A Subjective Theory of Compound Lotteries†

Haluk Ergin
Massachusetts Institute of Technology
and
Faruk Gul
Princeton University

February 2004

Abstract

We develop a Savage-type model of choice under uncertainty in which agents identify uncertain prospects with subjective compound lotteries. Our theory permits *issue preference*; that is, agents may not be indifferent among gambles that yield the same probability distribution if they depend on different issues. Hence, we establish subjective foundations for the Anscombe-Aumann framework and other models with two different types of probabilities. We define second-order risk as risk that resolves in the first stage of the compound lottery and show the equivalence of aversion to this risk with issue preference, the Ellsberg paradox, and uncertainty aversion.

† This research was supported by grants from the National Science Foundation.

1. Introduction

In the last ten years, the Ellsberg paradox has received more attention then any other experimentally observed violation of the expected utility hypothesis. What distinguishes the Ellsberg paradox from the Allais Paradox and other related violations of subjective expected utility theory that generated interest in non-expected utility models in the past is the fact that Ellsberg paradox type behavior cannot be explained within a model of choice among lotteries. That is, the Ellsberg paradox calls into question not only subjective expected utility theory but all models of choice under uncertainty that postulate behavior based on reducing uncertainty to risk.

The following "mini" version is useful for understand the issues raised by the Ellsberg paradox as well as our interpretation and resolution. An experimental subject is presented with an urn. He is told that the urn contains three balls, one of which is red. The remaining balls are either green or white. A ball will be drawn from the urn at random. The decision-maker is asked to choose between a bet that yields \$100 if a green ball is drawn and 0 dollars otherwise and a bet that yields \$100 if a red ball is drawn and 0 dollars otherwise.

Before the ball is drawn, the decision-maker is asked his preference over two other bets. In one bet he is to receive \$100 if either a green or a white ball is drawn and 0 if a red ball is drawn. With the other option, the decision-maker gets \$100 if either a red or a white ball is drawn and 0 if a green ball is drawn. These bets can be described as follows:

$$f = \begin{pmatrix} 100 & 0 & 0 \\ G & W & R \end{pmatrix} \quad \text{versus} \quad h = \begin{pmatrix} 0 & 0 & 100 \\ G & W & R \end{pmatrix}$$
$$f' = \begin{pmatrix} 100 & 100 & 0 \\ G & W & R \end{pmatrix} \quad \text{versus} \quad h' = \begin{pmatrix} 0 & 100 & 100 \\ G & W & R \end{pmatrix}$$

Consider a decision-maker who prefers h to f and f' to h'. Presumably, by preferring h to f the decision-maker is revealing his subjective assessment that there is a higher chance of a red ball being drawn than a green ball. Similarly, by revealing a preference for f' over h', the decision-maker is expressing his belief that the event "green or white" is more likely than "red or white". If the decision-maker's assessments of the likelihoods of G, W and R could be described by some probability μ, and if we assume that the decision-maker prefers a greater chance of winning \$100 to a smaller chance of winning \$100, we would conclude from the choice above that $\mu(R) > \mu(G)$ and $\mu(G \cup W) > \mu(R \cup W)$. Since, G, W, R are mutually exclusive events, no such probability exists, hence the paradox.

One intuitive explanation of the above behavior is the following: the decision-maker finds it difficult to associate unique probabilities with the events $G, W, R \cup G$ and $R \cup W$. In contrast, the probability of the events R is $\frac{1}{3}$ and hence the probability of the event $G \cup W$ is $\frac{2}{3}$. The ambiguity of the events G, W permits the agent to view each of these events as being less likely than R when they are associated with good prizes but more likely than R when associated with bad prizes. Consequently, the agent can prefer h to f and f' to h'. We call this interpretation of the Ellsberg paradox "the ambiguity aversion interpretation."

We provide an alternative interpretation of the Ellsberg paradox. In our model, behavior consistent with the Ellsberg paradox is identified as the agent's preference for betting on one issue over betting on another, even if the corresponding probabilities and prizes are the same. To understand how our model relates to the choice experiment above, imagine that the balls in the urn are numbered 1, 2, and 3. Without loss of generality, let ball 3 be the red ball. We can distinguish between two different kinds of uncertainty. First, there is the issue of which ball gets chosen. Second, there is the issue of the color of balls 1 and 2. With this description of the underlying type space, the events G, W, and R can be depicted as follows:

$$
\begin{array}{ccccl}
G & G & W & W & \text{ball 1} \\
G & W & G & W & \text{ball 2} \\
R & R & R & R & \text{ball 3} \\
gg & gw & wg & ww &
\end{array}
$$

In the matrix above, the last column describes the uncertainty regarding the ball that is chosen, while the bottom row depicts the uncertainty regarding the color of each ball. For example gg denotes the contingency in which both ball 1 and ball 2 are green, while wg describes the contingency where ball 1 is white and ball 2 is green. We refer to the uncertainty regarding the number of the ball that is drawn as issue a and the uncertainty regarding the color of the first two balls as issue b. Capital letters describe the color of the ball that is drawn.

Assume that the decision-maker considers every possible resolution of issue a equally likely and every possible resolution of issue b equally likely. Furthermore, assume that he considers the two issues to be statistically independent. Note that a bet on R corresponds to a bet on issue a. Hence, given our independence assumption, learning the outcome of issue b does not affect the decision-maker's belief regarding R. In contrast, the decision-maker's assessment of the probability of G conditional of ww is 0, while his assessment of G given gg is $\frac{2}{3}$. Hence, R is a bet on which ball gets chosen (issue a), while G entails both a bet on which ball gets chosen and on the number of green balls in the bag (issue b). Taking unions with W reverses the roles of G and R. Observe that the probability of $G \cup W$ does not change once issue b is resolved, while the probability of $R \cup W$ does. Hence a decision-maker who, *ceteris paribus* prefers lotteries that depend only on issue a to equivalent lotteries that depend both on issue b and issue a and has the subjective prior above, must prefer the act h (getting \$100 if and only if R occurs) to f (getting \$100 if and only if G occurs) but must prefer f' (getting \$100 if and only if G or W occurs) to h' (getting \$100 if and only if R or W occurs). We call this "the issue preference interpretation" of the Ellsberg paradox.

In the issue preference interpretation, the events associated with issue b are at least in principle, observable. We can imagine looking into the urn and verifying the number of green balls. In contrast, with the ambiguity aversion interpretation, the uncertainty regarding the probability of G depends on what Nau (2002) calls "credal" states. To understand the distinction between between observable and credal states consider a maxmin expected utility maximizer (Gilboa and Schmeidler (1989)) who has the the same four possible probability distributions over G, W, R that correspond to the columns of the matrix above. We would not expect uncertainty regarding this agent's beliefs to be resolved. We cannot identify the state space associated with this uncertainty, let alone construct the acts that are needed to verify the axioms of any particular theory or to elicit the agent's prior over these states. To put it differently, Savage's axioms provide a way of identifying the agent's subjective prior over *objective states*. Formulations of ambiguity based on credal states entail a subjective prior over *subjective* states. It is not clear how Savage's approach to probability and choice under uncertainty can be carried out when the state space itself is subjective.

In our model, we assume an objective state space that has a product structure. Hence, the two issues (a and b) are verifiable. We introduce assumptions that ensure that the agent has a subjective prior over this state space. Given his prior, the agent views any two acts that depend only on issue a, (i.e., are measurable with respect to issue a) as equivalent whenever they have the same probability distribution over prizes. Similarly, the agent identifies two acts that have the same probability distribution whenever they both depend only on issue b. However, the agent may not be indifferent between an act that depends only on issue a and one that depends on issue b even if they yield the same lottery.

Machina and Schmeidler (1992) define a probabilistically sophisticated decision-maker as one whose preference over acts (i.e., uncertain prospects) depends only on his subjective prior and the lottery that the act yields given his prior (i.e., the risky prospect). Thus, our agents are not probabilistically sophisticated. However, our assumptions also ensure that the (payoff) relevant uncertainty over the two issues can be described by a single, *compound* lottery. We call this, *second-order probabilistic sophistication*. Each stage of the compound lottery is identified with the resolution of one issue. The ordering of the issues is determined endogenously, from the agent's preferences.

In Theorem 1, we show that a suitable modification of Machina and Schmeidler's model of probabilistic sophistication on a state-space that has a product structure (i.e., with two distinct issues) yields second-order probabilistic sophistication; that is, preferences where the decision-maker's ranking of acts depends only on the corresponding compound lotteries. We call the preferences characterized by Theorem 1 second-order probabilistically sophisticated (SPS) preferences.

Like a probabilistically sophisticated agent, a second-order probabilistically sophisticated decision-maker has a well-defined subjective prior over the state space. However, unlike a probabilistically sophisticated decision-maker, for a second-order probabilistically sophisticated agent, the lotteries associated with the acts f, g given the prior μ do not provide sufficient information to rank all acts. A second-order probabilistically sophisticated agent needs to keep track not just of the risk associated with f and g but also the second-order risk.

In addition to providing a suitable framework for formulating the issue preference interpretation of the Ellsberg paradox, Theorem 1 provides Savage-type subjectivist foundations for the Anscombe-Aumann model. That is, it identifies features of preferences that lead to a model of uncertainty where the agent has two different types of subjective probabilities corresponding to the Anscombe-Aumann's horse race and roulette wheel probabilities.

The subjective model of compound lotteries provided by Theorem 1 is "robust" in the sense of Machina and Schmeidler. That is, it imposes no restrictions on how the decision-maker evaluates compound lotteries, other than stochastic dominance. Hence, Theorem 1 provides a framework for evaluating the interaction of the Ellsberg paradox with other known violations of the expected utility hypothesis.

Kreps and Porteus (1978) introduce a more general form of compound lotteries which they call temporal lotteries to study decision-makers who have preferences for when uncertainty resolves. In their theory, temporal lotteries are taken as given, that is they view objective, compound lotteries as the primitive description of the underlying uncertainty. Furthermore, their assumptions yield expected utility preferences over temporal lotteries. That is, both the decision-maker preferences over compound lotteries and their rankings of the prizes of compound lotteries (i.e., ordinary lotteries) satisfy the independence axiom.

Formally, the relationship between our Theorem 1 and the main theorem in Kreps and Porteus (1978) is analogous to the relationship between Machina and Schmeidler's main result and the von Neumann-Morgenstern theorem. We start with purely subjective uncertainty and provide necessary and sufficient conditions for a preference relation to be second-order probabilistically sophisticated.

In Theorem 3 we establish that assuming Savage's sure-thing principle on acts that depend only on one of the two issues characterizes SPS expected utility (SPS-EU) preferences. Hence, Theorem 3 provides Savage-type foundations for the Kreps and Porteus (1978) model.

Klibanoff, Marinacci and Mukerji (2002) (henceforth KMM) introduce an auxiliary state space S_2, which is a set of probability distributions over a basic state space S_1. They identify S_2 with the uncertainty regarding the agent's prior over S_1 and call a Savage act on S_2, a second order act. KMM assume that the agent is a subjective expected utility maximizer when evaluating second order acts and an expected utility maximizer when ranking objective lotteries over outcomes. They identify an ordinary act f on S_1, with a second order act that yields for each prior ν on S_1, the certainty equivalent of the lottery induced by f and ν, to obtain preferences with a Kreps-Porteus representation. While their utility function is formally analogous to the one describing SPS-EU preferences (Theorem 3), the underlying objects are different. In KMM the choice objects are second order acts and lotteries, while we consider ordinary, single stage acts over a state-space that has a product structure ($\Omega_a \times \Omega_b$). Our axioms enable a representation that evaluates acts *as if* issue $\omega_b \in \Omega_b$ is realized in the first stage and $\omega_a \in \Omega_a$ is realized in the second.

A second important distinction between the SPS-EU model and the model studied in KMM is in the interpretation. KMM compare their model to maxmin expected utility theory by identifying S_2 with uncertainty regarding the agent's prior. Hence, in the terminology of Nau (2002), these are credal states.

Segal (1987, 1990) uses preferences over compound lotteries (which he calls two-stage lotteries) to analyze the Ellsberg paradox and other related phenomena. In his model, second stage preferences satisfy the expected utility hypothesis while in the first stage, the decision-maker has anticipated utility preferences.[1] In Segal (1987), this particular model of choice over compound lotteries is assumed and Ellsberg paradox type behavior is related to the decision-maker's attitude towards second-order risk (which he calls ambiguity). Segal uses his model to address a large number of experimentally documented violations of probabilistic sophistication. Segal (1990) derives the model studied in Segal (1987) from assumptions on the preference relation over compound lotteries and investigates the relationships among various stochastic dominance and reduction (of compound lotteries) axioms.

In Theorem 4, we provide a definition of comonotonic sure thing principle for acts on a product set. We show that this leads to a characterization of SPS Choquet expected utility (SPS-CEU) preferences that maximize expected utility when evaluating acts that depend on issue a and maximize rank-dependent expected utility when evaluating acts that depend on issue b. Hence, Theorem 4 provides a subjectivist formulation of Segal's model.

[1] Quiggin (1982) introduces the theory and Yaari (1987) provides axiomatic foundations of the case with a linear utility index.

To see how second-order probabilistic sophistication relates to the Ellsberg paradox, consider a decision-maker who has a prior over the state-space describing the joint resolution of issues a and b above. Assume that this prior is uniform. Hence, $\mu(\omega_a, \omega_b) = \frac{1}{12}$ for all $\omega_a \in \{\text{ball } 1, \text{ball } 2, \text{ball } 3\}$ and all $\omega_b \in \{gg, gw, wg, ww\}$. Let p_α denote the lottery that yields \$100 with probability α and 0 dollars with probability $1 - \alpha$. For such a decision-maker, the act f can be described as a lottery that yields $p_{\frac{2}{3}}$ with probability $\frac{1}{4}$, $p_{\frac{1}{3}}$ with probability $\frac{1}{2}$ and p_0 with probability $\frac{1}{4}$. Hence, for each act, μ induces a lottery over lotteries, which we call a compound lottery. Let π_f denote the compound lottery associated with f. Computing the compound lotteries associated with the acts f, h, f', and h' above yields:

$$\pi_f = \frac{1}{4} \times p_{\frac{2}{3}} + \frac{1}{2} \times p_{\frac{1}{3}} + \frac{1}{4} \times p_0, \quad \pi_h = 1 \times p_{\frac{1}{3}}$$

$$\pi_{f'} = 1 \times p_{\frac{2}{3}}, \qquad \pi_{h'} = \frac{1}{4} \times p_{\frac{1}{3}} + \frac{1}{2} \times p_{\frac{2}{3}} + \frac{1}{4} \times p_1$$

Note that π_f and π_h both yield \$100 with probability $\frac{1}{3}$ and $\pi_{f'}$ and $\pi_{h'}$ both yield \$100 with probability $\frac{2}{3}$. However, π_f is a *mean preserving spread* of π_h and $\pi_{h'}$ is a mean preserving spread of $\pi_{f'}$. Our definition of second-order risk aversion requires that the agent prefers a compound lottery to a mean preserving spread. Hence, the Ellsberg paradox type behavior corresponds to second-order risk aversion.

Theorems 2 and 5 establish the relationship between issue preference, second-order risk aversion and uncertainty aversion (Schmeidler (1989)). We show that in general, uncertainty aversion implies second-order risk aversion which implies that among lotteries that have the same subjective probability distribution, the agent prefers ones that depend only on issue a to lotteries that depend on both issues and prefers the latter to lotteries that depend only on issue b. If the agent is a SPS-EU or a SPS-CEU maximizer, then second-order risk aversion and uncertainty aversion are equivalent.

Like Segal (1987), we relate Ellsberg paradox type behavior to a form of second-order risk aversion. However, we use a different notion of second-order risk. His notion is applicable to binary acts (i.e. act that yield the same two prizes) and is analogous to a notion of risk aversion based on comparing lotteries to their means. Our notion of second-order risk aversion is analogous to the standard notion of risk aversion based on mean preserving spreads and is applicable to all lotteries.

The notion of ambiguity aversion used in Klibanoff, Marinacci and Mukerji (2002) is formally similar to our concept of second-order risk aversion. Their definition entails comparing arbitrary lotteries to appropriate second-order riskless lotteries, i.e., it is analogous to a definition of risk aversion based on comparing lotteries to their means. Within their expected utility framework, risk aversion defined as not preferring the distribution to its mean is equivalent to risk aversion defined as preferring the distribution to a mean preserving spread.

In Kreps and Porteus, the "issue" is time and the decision-maker cares about how much of the uncertainty is resolved in period 1 versus in period 2. In our model, the fact that issue b resolves at the first stage of the compound lottery has no temporal significance. The particular representation is derived from the fact that the agent's preferences satisfy greater separability properties with respect to issue b than with respect to issue a. By taking temporal lotteries as the primitive of their model, Kreps and Porteus are implicitly making the same assumption with respect to lotteries that resolve earlier.

Nau (2002) also provides a model with two issues. His axioms ensure the existence of a representation that is additively separable when restricted to lotteries that depend only on issue a or only on issue b but permit state-dependent preferences. Since state-dependence makes it impossible to identify subjective probabilities, he offers a local measure of ambiguity aversion. Nau's is the closest to our interpretation of the Ellsberg paradox in that he places more emphasis on the agent's different treatment of two issues and less emphasis on identifying the more ambiguous issue.

Epstein and Zhang (2001) note that the notion of ambiguity requires a violation of probabilistic sophistication when formulated within the Savage setting but may be consistent with probabilistic sophistication within the Anscombe-Aumann framework. To the extent that the notion of ambiguity is motivated by the Ellsberg Paradox, our notion of issue preference resolves this apparent paradox. The Anscombe-Aumann model has two different types of uncertainty, corresponding to the objective probabilities (issue a in our model) and the subjective uncertainty (issue b) associated with the state. Hence, an agent may be probabilistically sophisticated over the subjective states and still prefer a particular objective lottery to the corresponding subjective one. In the Savage setting, both issues have to be specified in the description of the subjective state space. In this framework a preference for issue a versus issue b is synonymous with the failure of probabilistic sophistication.

2. Second-Order Probabilistic Sophistication

Let $\mathcal{Z}$ be a set of prizes and $\Omega := \Omega_a \times \Omega_b$ be the set of all states. We refer to a, b as *issues*. Let $\mathcal{A}$ be the algebra of all subsets of Ω_a and $\mathcal{B}$ be the algebra of all subsets of Ω_b. Let $\mathcal{E}_a$ denote the algebra of all sets of the form $A \times \Omega_b$ for some $A \in \mathcal{A}$, $\mathcal{E}_b$ denote the algebra of all sets of the form $\Omega_a \times B$ for some $B \in \mathcal{B}$, and $\mathcal{E}$ denote the algebra of all sets that can be expressed as finite unions of sets of the form $A \times B$ for $A \in \mathcal{A}$, $B \in \mathcal{B}$ (i.e., $\mathcal{E}$ is the algebra generated by $\mathcal{E}_a \cup \mathcal{E}_b$). A function $f : \Omega' \rightarrow \mathcal{Z}'$ is *simple* if $f(\Omega')$ is finite. For any algebra $\mathcal{E}'$, the function f is $\mathcal{E}'-$*measurable* if it is simple and $f^{-1}(z) \in \mathcal{E}'$ for all $z \in \mathcal{Z}$.

Let $\mathcal{F}$ denote the set of all (Savage) acts; that is, $\mathcal{F}$ is the set of $\mathcal{E}-$measurable functions from Ω to $\mathcal{Z}$. An individual is characterized by a binary relation $\succeq$ on $\mathcal{F}$. Our first assumption is that this binary relation is a preference relation.

Axiom 1: *(Preference Relation) $\succeq$ is complete and transitive.*

We use $\sim$ to denote the indifference relation associated with $\succeq$ and use $f \succ g$ to denote $f \succeq g$ and not $g \succeq f$. We identify each $z \in \mathcal{Z}$ with the corresponding constant prospect. Our second assumption ensures that the individual is not indifferent among all constant prospects.

Axiom 2: *(Nondegeneracy) There exists $x, y \in \mathcal{Z}$ such that $x \succ y$.*

Let E^c denote the complement of E in Ω. For any set $E \in \mathcal{E}$, we say that the uncertain prospects f and g agree on E if $f(s) = g(s)$ for all $s \in E$. We write $f = g$ on E to denote the fact that f agrees with g on E. An event $E \in \mathcal{E}$ is null if $f = g$ on E^c implies $f \sim g$. Otherwise, the event E is non-null. Our next assumption states that for all non-null events E and all uncertain prospects f, improving what the decision-maker gets if E

occurs, keeping what he gets in all other contingencies constant makes the decision-maker better off.

Axiom 3: *(Monotonicity) For all non-null E, $f = g$ on E^c, $f = z$ on E, $g = z'$ on E implies $z \succeq z'$ if and only if $f \succeq g$.*

Axioms $1-3$ are identical to their counter-parts in Savage's theory. The next assumption ensures that Ω can be divided into arbitrarily "small" events of the form $A \times \Omega_b$ and $\Omega_a \times B$:

Axiom 4: *(Continuity) For $f \succ g$, $z \in \mathcal{Z}$ there exist a partitions $E_1, \ldots, E_n \in \mathcal{E}_a$ of and partition $F_1, \ldots, F_n \in \mathcal{E}_b$ of Ω such that*

(a) $[f^i = f,\ g^i = g$ on E_i^c and $f^i = g^i = z$ on $E_i]$ implies $[f^i \succ g$ and $f \succ g^i]$

(b) $[f_j = f,\ g_j = g$ on F_j^c and $f_j = g_j = z$ on $F_j]$ implies $[f_j \succ g$ and $f \succ g_j]$

Most models that study uncertain prospects (i.e., the Savage setting) impose the assumptions above.[2] These models differ with respect to their comparative probability axiom and their separability axioms.

Note that for any $f \in \mathcal{F}$, there exists a partition $A_1, \ldots, A_n$ of $\mathcal{A}$ and a partition $B_1, \ldots, B_m$ of $\mathcal{B}$ such that the function f is constant on each $A_i \times B_j$, for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Hence, we can identify each $f \in \mathcal{F}$ with some $n+1$ by $m+1$ matrix. That is:

$$f = \begin{pmatrix} x_{11} & \ldots & x_{1m} & A_1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n1} & \ldots & x_{nm} & A_n \\ B_1 & \ldots & B_m & * \end{pmatrix}$$

Let $\mathcal{F}_a$ and $\mathcal{F}_b$ denote the set of $\mathcal{E}_a$ and $\mathcal{E}_b$ measurable acts respectively. For $f \in \mathcal{F}_a$ and $g \in \mathcal{F}_b$ we write

$$f = \begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix}, \ g = \begin{pmatrix} x_1 & \ldots & x_m \\ B_1 & \ldots & B_m \end{pmatrix}$$

[2] Axiom 4 is a slightly stronger than the usual continuity assumption since it requires that the event space can be partitioned both into small probability $\mathcal{E}_a$ and $\mathcal{E}_b$ events.

Axiom 5a: *($a-$Strong Comparative Probability) If $x \succ y$ and $x' \succ y'$ then*

$$\begin{pmatrix} x & A_1 \\ y & A_2 \\ z_3 & A_3 \\ \vdots & \vdots \\ z_n & A_n \end{pmatrix} \succ \begin{pmatrix} y & A_1 \\ x & A_2 \\ z_3 & A_3 \\ \vdots & \vdots \\ z_n & A_n \end{pmatrix} \text{ iff } \begin{pmatrix} x' & A_1 \\ y' & A_2 \\ z'_3 & A_3 \\ \vdots & \vdots \\ z'_n & A_n \end{pmatrix} \succ \begin{pmatrix} y' & A_1 \\ x' & A_2 \\ z'_3 & A_3 \\ \vdots & \vdots \\ z'_n & A_n \end{pmatrix}$$

Axiom 5a is the Machina-Schmeidler strong comparative probability axiom imposed on $\mathcal{E}_a-$measurable acts. Consider prizes x, y such that $x \succ y$ and an act f that yields x on event $A_1 \times \Omega_b$ and y on $A_2 \times \Omega_b$. If the decision-maker prefers f to the act that yields y on $A_1 \times \Omega_b$ and x on $A_2 \times \Omega_b$ agrees f outside of $(A_1 \cup A_2) \times \Omega_b$, this suggests that he considers $A_1 \times \Omega_b$ more likely than $A_2 \times \Omega_b$. The assumption asserts that prizes don't affect probabilities. That is, if we conclude that the decision-maker $A_1 \times \Omega_b$ strictly more likely than $A_2 \times \Omega_b$ using some act f, we should not be able to conclude the opposite using some other act f'. Our main assumption is the axiom below:

Axiom 5b: *($a|b-$Strong Comparative Probability) If*

$$\begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix} \succ \begin{pmatrix} y_1 & A_1 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix} \text{ and } \begin{pmatrix} x'_1 & A_1 \\ \vdots & \vdots \\ x'_n & A_n \end{pmatrix} \succ \begin{pmatrix} y'_1 & A_1 \\ \vdots & \vdots \\ y'_n & A_n \end{pmatrix}$$

Then,

$$\begin{pmatrix} x_1 & y_1 & z_{13} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & y_n & z_{n3} & \dots & z_{nm} & A_n \\ B_1 & B_2 & B_3 & \dots & B_m & * \end{pmatrix} \succ \begin{pmatrix} y_1 & x_1 & z_{13} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & x_n & z_{n3} & \dots & z_{nm} & A_n \\ B_1 & B_2 & B_3 & \dots & B_m & * \end{pmatrix}$$

iff

$$\begin{pmatrix} x'_1 & y'_1 & z'_{13} & \dots & z'_{1m} & A_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x'_n & y'_n & z'_{n3} & \dots & z'_{nm} & A_n \\ B_1 & B_2 & B_3 & \dots & B_m & * \end{pmatrix} \succ \begin{pmatrix} y'_1 & x'_1 & z'_{13} & \dots & z'_{1m} & A_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ y'_n & x'_n & z'_{n3} & \dots & z'_{nm} & A_n \\ B_1 & B_2 & B_3 & \dots & B_m & * \end{pmatrix}$$

Axiom 5b has three important implications. First, it implies the Machina-Schmeidler strong comparative probability axiom on $\mathcal{E}_b$. To see this, assume that all columns in the above statement of Axiom 5b are constants. That is, $x_1 = x_2 = \ldots = x_n$, $y_1 = y_2 = \ldots = y_n$, $z_{13} = z_{23} = \ldots = z_{n3}$, etc. Then, we obtain a symmetric version of Axiom 5a. This ensures that the agent is probabilistically sophisticated over $\mathcal{E}_b-$measurable acts. It follows from this observation that Axiom 5b is stronger than the symmetric counterpart of 5a. To see the other two implications consider the example below. (To simplify the presentation, we have represented the events with their probabilities):

$$f = \begin{pmatrix} 1 & .5 \\ 100 & .5 \end{pmatrix} \succ \begin{pmatrix} 100 & .5 \\ 0 & .5 \end{pmatrix} = g$$

Since the two rows of f and g are equiprobable, $f \succ g$ is simply a statement of the fact that 1 is preferred to 0. Then, Axiom 5b implies

$$f' = \begin{pmatrix} 1 & 100 & -500 & .5 \\ 100 & 0 & 1000 & .5 \\ .4 & .3 & .3 & * \end{pmatrix} \succ \begin{pmatrix} 100 & 1 & -500 & .5 \\ 0 & 100 & 1000 & .5 \\ .4 & .3 & .3 & * \end{pmatrix} = g'$$

But, the preference for f over g relies on the fact that the rows of f and g are equally likely and $f' \succ g'$ requires that conditional on each column, these rows are equally likely. Hence, Axiom 5b implies that equally likely issue a (i.e., row) events remain equally likely after conditioning on any issue b (column) event. Thus, the second implication of the axiom is that it renders the issues statistically independent.

To understand the final implication of Axiom 5b note that the second row of f' is better than the first row of f'. Switching the columns of f' (to obtain g') improves the first row at the expense of the second. That is, the utilities associated with the rows of g' are closer together than the corresponding utilities for the rows of f' and therefore g' has less risk with respect to issue a. Axiom 5b precludes aversion to such risk. It is this feature of Axiom 5b that permits the reduction of acts into compound lotteries where second-order risk is only associated with issue b.

Note that we can interpret Axiom 5b as a version of Machina-Schmeidler's comparative probability axiom applied to a richer set of prizes. To see this identify each $f \in \mathcal{F}_a$ with a function $f_a : \Omega_a \to \mathcal{Z}$ by setting $f_a(\omega_a) = f(\omega_a, \omega_b)$ for any $\omega_b \in \Omega_b$. Formally:

Definition: *Let $\mathcal{F}_a^0$ denote the set of all simple functions from Ω_a to $\mathcal{Z}$. Define the bijection*

$$\chi : \mathcal{F}_a \to \mathcal{F}_a^0$$

as follows:

$$\chi(f)(\omega_a) = f(\omega_a, \omega_b) \text{ for all } \omega_b \in \Omega_b$$

We let f_a denote $\chi(f)$.

Then, by identifying each $f \in \mathcal{F}$ with a function from Ω_b to $\mathcal{F}_a^0$, we can interpret Axiom 5b as Machina-Schmeidler's strong comparative probability axiom. To clarify this symmetry between Axioms 5a and 5b, we re-state Axiom 5b as follows:

Axiom 5b: *Suppose $f, \hat{f}, g, \hat{g}, h^3, \hat{h}^3, \dots h^m, \hat{h}^m \in \mathcal{F}_a$ and*

$$f \succ g \text{ and } \hat{f} \succ \hat{g}$$

Then,

$$\begin{pmatrix} f_a & g_a & h_a^3 & \dots & h_a^m \\ B_1 & B_2 & B_3 & \dots & B_m \end{pmatrix} \succ \begin{pmatrix} g_a & f_a & h_a^3 & \dots & h_a^m \\ B_1 & B_2 & B_3 & \dots & B_m \end{pmatrix}$$

iff

$$\begin{pmatrix} \hat{f}_a & \hat{g}_a & \hat{h}_a^3 & \dots & \hat{h}_a^m \\ B_1 & B_2 & B_3 & \dots & B_m \end{pmatrix} \succ \begin{pmatrix} \hat{g}_a & \hat{f}_a & \hat{h}_a^3 & \dots & \hat{h}_a^m \\ B_1 & B_2 & B_3 & \dots & B_m \end{pmatrix}$$

Henceforth, we refer to Axioms 5a and 5b together as Axiom 5. Let $\Delta(\mathcal{Z}') = \{p : \mathcal{Z}' \to [0,1] \,|\, p^{-1}((0,1])$ is finite and $\sum_z p(z) = 1\}$ be the set of all simple lotteries on some set $\mathcal{Z}'$. Let $p(G) = \sum_{z \in G} p(z)$ for $G \subset \mathcal{Z}'$. Let δ_z denote the degenerate lottery that yields z with probability 1. We say that $\phi : \Delta(\mathcal{Z}') \to I\!R$ satisfies stochastic dominance if for all $\alpha \in (0,1)$, $\phi(\alpha\delta_z + (1-\alpha)r) > \phi(\alpha\delta_{z'} + (1-\alpha)r)$ if and only if $\phi(\delta_z) > \phi(\delta_{z'})$.

We use p, p' and p'' to denote generic elements of $P := \Delta(\mathcal{Z})$ and π, π', π'' to denote generic elements of $\Delta(P)$. Hence, P is the set of simple lotteries on $\mathcal{Z}$ and $\Delta(P)$ is the set of simple lotteries on P. We call $\Delta(P)$ the set of *compound lotteries*.

Definition: *A function μ on $\mathcal{E}$ is a probability if (i) $\mu(E) \geq 0$ for all $E \in \mathcal{E}$, (ii) $\mu(\Omega) = 1$ and (iii) $E \cap E' = \emptyset$ implies $\mu(E \cup E') = \mu(E) + \mu(E')$.*

Given any probability μ on $\mathcal{E}$, we can associate with each $f \in \mathcal{F}$, a lottery $p_f \in P$. and a compound lottery $\pi_f \in \Delta(P)$. Since the underlying μ is clear we suppress the dependence on μ in the definitions below.

Definition: *Let μ be a probability on $\mathcal{E}$. For $f \in F$, define $p_f \in \Delta(\mathcal{Z})$ as follows:*

$$p_f(z) = \mu(f^{-1}(z)) \text{ for all } z \in \mathcal{Z}$$

Let μ_a and μ_b denote the marginals of the probability measure μ on the two issues, i.e. $\mu_a(A) = \mu(A \times \Omega_b)$ and $\mu_b(B) = \mu(\Omega_a \times B)$ for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$. We say that $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent if $\mu(A \times B) = \mu_a(A) \cdot \mu_b(B)$ for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$. When $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent, we also associate with each Savage act $f \in \mathcal{F}$, an Anscombe-Aumann act f^* and a compound lottery $\pi_f \in \Delta(P)$. Let $\mathcal{F}^*$ denote the set of all simple functions from Ω_b to P, i.e., $\mathcal{F}^*$ is the set of all Anscombe-Aumann acts.

Definition: *Let μ be a probability on $\mathcal{E}$ and let $\mathcal{E}_a, \mathcal{E}_b$ be $\mu-$independent. For $f \in F$ such that*

$$f = \begin{pmatrix} x_{11} & \dots & x_{1m} & A_1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n1} & \dots & x_{nm} & A_n \\ B_1 & \dots & B_m & * \end{pmatrix}$$

define f^* as follows:

$$p_j(z) = \sum_{i:x_{ij}=z} \mu_a(A_i)$$
$$f^*(s) = p_j \text{ for all } s \in B_j$$

Then, we can associate a compound lottery with each Anscombe-Aumann act by evaluating the probability of the event which yields p for each $p \in P$.

Definition: *Let μ be a probability on $\mathcal{E}$ and let $\mathcal{E}_a, \mathcal{E}_b$ be $\mu-$independent. For $f \in F$, define the compound lottery π_f as follows:*

$$\pi_f(p) = \mu_b(f^{*-1}(p))$$

We say that a measure μ is nonatomic on $\mathcal{E}_a$ if $E \in \mathcal{E}_a$ and $\lambda \in [0,1]$ implies there exists $E' \in \mathcal{E}_a$ such that $E' \subset E$ and $\mu(E') = \lambda\mu(E)$. A symmetric condition is required for μ to be nonatomic on $\mathcal{E}_b$. We say that μ is nonatomic if it is nonatomic on both $\mathcal{E}_a, \mathcal{E}_b$. Note that the three mappings described above are all onto when $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent and μ is nonatomic. That is, for each $p \in P$ there exists $f \in \mathcal{F}$ such that $p_f = p$, for each $h \in \mathcal{F}^*$ there exists $f \in \mathcal{F}$ such that $f^* = h$, and for each $\pi \in \Delta(P)$ there exists $f \in \mathcal{F}$ such that $\pi_f = \pi$.

Machina and Schmeidler call $\succeq$ probabilistically sophisticated if there exists a mixture continuous and stochastic dominance satisfying function $W : P \to I\!R$ and a subjective probability measure μ such $f \succeq g$ iff $W(p_f) \geq W(p_g)$. Hence, a probabilistically sophisticated decision-maker has a subjective prior over the state-space, considers all acts that yield the same lottery equivalent and satisfies stochastic dominance. The theorem below establishes a weaker version of probabilistic sophistication for decision-makers that satisfy Axioms $1-5$. According to this weaker notion of probabilistic sophistication, a decision-maker has a subjective prior over the state space, considers all acts that yield the same compound lottery as equivalent and satisfies the appropriate counter-parts of continuity and stochastic dominance.

For any function $W : \Delta(P) \to I\!R$, we will say that a sequence of compound lotteries π_n is W-bounded if there exist $z^*, z_* \in \mathcal{Z}$ such that $W(z^*) \geq W(x) \geq W(z_*)$ for any $x \in \bigcup_{n=1}^{\infty} \bigcup_{p \in \operatorname{supp} \pi_n} \operatorname{supp} p$. We endow P with the supremum metric: $d^{\infty}(p,p') = sup_{z \in Z}|p(z) - p'(z)|$ for $p, p' \in P$. Then the sequence of compound lotteries π_n weakly converges to a compound lottery π, if for every open subset G of P,

$$\text{liminf}_{n\to\infty}\pi_n(G) \geq \pi(G).$$

Definition: *A function $W : \Delta(P) \to I\!R$ is weakly continuous if for any W-bounded sequence π_n weakly converging to π, we have (i) if $W(\pi) > W(\pi')$ then there exists N such that $W(\pi_n) > W(\pi')$ for $n \geq N$, and (ii) if $W(\pi) < W(\pi')$ then there exists N such that $W(\pi_n) < W(\pi')$ for $n \geq N$.*

Theorem 1: *The binary relation $\succeq$ satisfies Axioms $1-5$ if and only if there exists a nonatomic probability μ on $\mathcal{E}$ and a weakly continuous function $W : \Delta(P) \to I\!R$, such that (i) $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent, (ii) $W(\pi_f) \geq W(\pi_g)$ iff $f \succeq g$, and (iii) Both W and the function $w : P \to I\!R$ defined by $w(p) = W(\delta_p)$ satisfy stochastic dominance.*

The details of the proof of Theorem 1, which relies on Theorem 2 of Machina and Schmeidler (1992), are in the appendix. Here, we give a sketch. Let Ω' be an arbitrary state space, $\mathcal{Z}'$ be any set of prizes and let $\mathcal{F}'$ be the set of all simple functions from Ω' to $\mathcal{Z}'$. Theorem 2 of Machina and Schmeidler (1992) establishes that if a binary relation $\succeq'$ over $\mathcal{F}'$ satisfies Axiom $1-3$ and continuity (i.e., replace Ω_a with Ω' in Axiom $4a$) and strong comparative probability (i.e., let the A_i's in Axiom $5a$ denote arbitrary subset of Ω') then $\succeq'$ is probabilistically sophisticated. Applying this theorem to $\mathcal{F}_a$ yields a probability measure μ_a and a stochastic dominance satisfying function $\tilde{w} : P \to I\!R$ such that $f \succeq g$ iff $\tilde{w}(p_f) \geq \tilde{w}(p_g)$ for all $f, g \in \mathcal{F}_a$, where the probability distributions p_f, p_g are derived from μ_a.

We next introduce a binary relation $\succeq^*$ on $\mathcal{F}^*$ by $f^* \succeq^* g^*$ if and only if $f \succeq g$ and show that $\succeq^*$ is well-defined. Applying Machina and Schmeidler's Theorem 2 again, we obtain a probability μ_b on Ω_b and a stochastic dominance satisfying function $W : \Delta(P) \to I\!R$ such that $f^* \succeq^* g^*$ iff $W(\pi^*_{f^*}) \geq W(\pi^*_{f^*})$ where for any $h \in \mathcal{F}^*$, the compound lottery π^*_h is defined by $\pi^*_h(p) = \mu_b(h^{-1}(p))$. Let μ be the product on $\mathcal{E}$ of μ_a and μ_b. That is, for any $\mathcal{E}-$set $E = \bigcup_{i=1}^n A_i \times B_i$, let $\mu(E) = \sum_{i=1}^n \mu_a(A_i) \cdot \mu_b(B_i)$. To complete to proof we show that W is weakly continuous.

We refer to preferences satisfying Axioms $1-5$ as second-order probabilistically sophisticated (SPS) preferences. Comparing Theorem 1 above to Theorem 2 of Machina and Schmeidler (1992), we note that Machina-Schmeidler decision-makers are indifferent between any two Savage acts that yield the same probability distribution over prizes. That is, they are indifferent between f and g whenever f and g imply the same risk $p_f = p_g$. In contrast, decision-makers characterized by our weaker notion of probabilistic sophistication are indifferent between f and g whenever f and g yield the same second-order risk on issue b. That is, $f \sim g$ whenever $\pi_f = \pi_g$. In section 3, we provide measures of second-order risk, relate second-order risk aversion to issue preference and the Ellsberg paradox.

3. Second-Order Risk Aversion and the Ellsberg Paradox

By Axiom 2, there are prizes $x, y \in \mathcal{Z}$ such that $x \succ y$. Let $x = 1$ and $y = 0$. Consider the following acts

$$f = \begin{pmatrix} 1 & 0 & A \\ 1 & 0 & \Omega_a \backslash A \\ B & \Omega_b \backslash B & * \end{pmatrix}, \; g = \begin{pmatrix} 1 & 1 & A \\ 0 & 0 & \Omega_a \backslash A \\ B & \Omega_b \backslash B & * \end{pmatrix}, \; h = \begin{pmatrix} 1 & 0 & A \\ 0 & 1 & \Omega_a \backslash A \\ B & \Omega_b \backslash B & * \end{pmatrix}$$

Note that if μ is any measure on $\mathcal{E}$ that renders the issues independent and $\mu(A) = \mu(B) = \frac{1}{2}$, then $\pi_g = \pi_h$. To see this note that both π_g and π_h assign probability 1 to the lottery p that yields 1 with probability $\frac{1}{2}$ and 0 with probability $\frac{1}{2}$. If $\succeq$ is an SPS decision-maker then $\pi_g = \pi_h$ implies $g \sim h$. Note also that $p_f = p_g = p_h = p$. Hence, if the decision-maker were probabilistically sophisticated he would be indifferent among all three acts. However, an SPS decision-maker may distinguish between acts that yield different second-order distributions. Observe that

$$\pi_f = \frac{1}{2}\delta_{\delta_1} + \frac{1}{2}\delta_{\delta_0} \neq \delta_{\frac{1}{2}\delta_1 + \frac{1}{2}\delta_0} = \pi_g = \pi_h$$

A decision-maker facing the bet f will know the outcome of his bet whenever he learns the resolution of issue b while a decision-maker holding the bet g or h will learn nothing when he learns the resolution of issue b. That is, a decision-maker confronting f faces significant second-order risk (with respect to issue b) while a decision-maker facing g or h faces no second-order risk.

This notion of second-order risk is analogous to the standard notion of risk. To see the similarity between the two concepts recall that when $\mathcal{Z}$ is an interval of real numbers, both $\mathcal{Z}$ and $\Delta(\mathcal{Z})$ are mixture spaces. Suppose there are two lotteries $p, p' \in \Delta(\mathcal{Z})$ such that $p = \alpha(\beta\delta_x + (1-\beta)\delta_y) + (1-\alpha)p''$ and $p' = \alpha\delta_{\beta x + (1-\beta)y} + (1-\alpha)p''$. Then, p is said to be a mean preserving spread of p'. A decision-maker who has preferences over lotteries is risk averse if he prefers p' to p whenever p is a mean preserving spread of p'. Hence, a risk averse decision-maker prefers it when a mixture in the space $\Delta(\mathcal{Z})$ is replaced by mixture in the space $\mathcal{Z}$. Since $\Delta(P)$ and P are also mixture spaces, we can use the same idea to define second-order risk aversion:

Definition: *The compound lottery π is a mean preserving spread of π' if there are $\alpha, \beta \in (0, 1)$, $p, q \in P$ and $\pi'' \in \Delta(P)$ such that*

$$\pi = \alpha(\beta\delta_p + (1-\beta)\delta_q) + (1-\alpha)\pi'' \text{ and}$$
$$\pi' = \alpha\delta_{\beta p+(1-\beta)q} + (1-\alpha)\pi''$$

We say that $\succeq$ is second-order risk averse if $f' \succeq f$ whenever π_f is a mean preserving spread of $\pi_{f'}$.

Theorem 2: *Let $\succeq$ be an SPS preference. Then, $(iii) \Rightarrow (i) \Leftrightarrow (ii) \Rightarrow (iv)$:*

(i) $\succeq$ is second order risk averse

(ii) $h^* = \alpha f^* + (1-\alpha)g^*$ and $\pi_f = \pi_g$ implies $h \succeq f$

(iii) $h^* = \alpha f^* + (1-\alpha)g^*$ and $f \sim g$ implies $h \succeq f$

(iv) $f \in \mathcal{F}_a, g \in \mathcal{F}, h \in \mathcal{F}_b$ and $p_f = p_g = p_h$ implies $f \succeq g \succeq h$.

Condition (iv) states that if f is a bet on issue a, h is a bet on issue b and g is any act that has the same subjective probability distribution as f and h, then, a decision-maker who is averse to second-order risk will prefer f to g and g to h. We call this condition issue preference. Theorem 2 establishes that in general, second-order risk aversion is a stronger condition than issue preference. Condition (iii) of Theorem 2 is exactly Schmeidler's definition of uncertainty aversion if we interpret his objective lotteries as issue a and the subjective uncertainty as issue b. The preferences studied by Schmeidler need not be SPS but satisfy what Schmeidler calls comonotonic independence. The above characterization of second-order risk aversion does not rely on comonotonic independence and hence can be adapted to any preference satisfying Axioms $1-5$. Finally, uncertainty aversion implies second-order risk aversion. Condition (ii) applies only to f and g that yield the same compound lottery. By Theorem 1 such acts are indifferent. Hence, (ii) is weaker than (iii). In the appendix, after the proof Theorem 2, we provide a counter-examples to (iv) implies (ii) and another counter-example to (ii) implies (iii). Hence, a result stronger than Theorem 2 cannot be proved.

In the following section, we show that conditions $(i) - (iv)$ are equivalent for SPS preferences that are EU within each issue or satisfy comonotonic independence in the sense of Schmeidler (1989).

In the introduction, we described how sensitivity to second-order risk can be utilized to interpret the Ellsberg paradox as a consequence of second-order risk aversion. Note however that by identifying issue a with the number of green balls and issue b with the number of the ball that gets chosen, we could have interpreted the Ellsberg paradox as a consequence of second-order risk loving behavior.

Which of the two possible ways of assigning issues is the right one? More generally, how can we distinguish issue a type uncertainty from issue b type uncertainty? In our approach, the choice of the issue b (i.e., the source of second-order risk) is, like the assignment of probabilities, is a subjective matter.

Recall the example at the beginning of this section. Regardless of which issue is issue a and which issue is issue b, the compound lottery associated with act h is $\delta_{p_{\frac{1}{2}}}$. Now, to verify which issue is the one associated with second-order uncertainty, i.e., issue b, we need to check if agent is indifferent between g and h or f and h. The former, means that the column events are issue b events, while the latter implies that the row events are issue b events.

3.1 Second-Order Risk with Expected Utility Preferences

Simpler and stronger characterization of second-order risk aversion are feasible for SPS preferences satisfying certain expected utility properties. Axiom $6a$ below is Savage's sure thing principle applied to acts in $\mathcal{F}_a$. In contrast, Axiom $6b$ is Savage's sure thing principle applied to all acts conditional on events in $\mathcal{E}_b$. We refer to Axioms $6a, b$ together as Axiom 6. Theorem 3 below establishes that imposing Axiom 6 on SPS preferences yields a version of the model introduced by Kreps and Porteus (1978). In this case, the agent is an expected utility maximizer with respect to both issue a and b lotteries but may not be indifferent between an issue a lottery and an equivalent issue b lottery. We refer to this type of preferences as SPS-EU preferences. Parts a and b of the following axiom impose Savage's sure thing principle on issue a and issue b dependent acts. Note that both parts of the axiom together are weaker than Savage's sure thing principle since neither part applies to acts that depend on both issues.

Axiom 6: *(Sure Thing Principles) Let $E_a \in \mathcal{E}_a, E_b \in \mathcal{E}_b$, $f, g, f', g' \in \mathcal{F}_a$, and $\tilde{f}, \tilde{g}, \tilde{f}', \tilde{g}' \in \mathcal{F}$. Then*

$$(a)\ f = f' \text{ on } E_a,\ g = g' \text{ on } E_a,\ f = g \text{ on } E_a^c,\ f' = g' \text{ on } E_a^c$$
$$\text{implies } f \succeq g \text{ if and only if } f' \succeq g'$$
$$(b)\ \tilde{f} = \tilde{f}' \text{ on } E_b,\ \tilde{g} = \tilde{g}' \text{ on } E_b,\ \tilde{f} = \tilde{g} \text{ on } E_b^c,\ \tilde{f}' = \tilde{g}' \text{ on } E_b^c$$
$$\text{implies } \tilde{f} \succeq \tilde{g} \text{ if and only if } \tilde{f}' \succeq \tilde{g}'$$

Axioms $1-6$ lead to SPS preferences with a different von Neuman-Morgenstern utility function for each issue. If we interpret issue b as the first period and issue a as the second period, Theorem 3 yields a subjective model of Kreps and Porteus temporal lotteries.[3]

Theorem 3: *The binary relation $\succeq$ satisfies Axioms $1-6$ if and only if there exists a nonatomic probability μ on $\mathcal{E}$ and a function $W : \Delta(P) \to I\!R$, such that (i) $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent, (ii) $W(\pi_f) \geq W(\pi_g)$ iff $f \succeq g$, and (iii) Both W and w defined by $w(p) = W(\delta_p)$ are expected utility functions. In particular,*

$$W(\pi) = \sum_{p \in P} v\left(\sum_{x \in \mathcal{Z}} u(x)p(x)\right)\pi(p)$$

for some continuous and strictly increasing $v : I\!R \to I\!R$ and $u : \mathcal{Z} \to I\!R$.

We refer to preferences that satisfy the hypothesis of Theorem 3 as SPS expected utility (SPS-EU) preferences.[4] We call the corresponding (v, u, μ) a representation of $\succeq$. It is easy to verify that a SPS-EU preference (v, u, μ) is second-order risk averse if and only if v is concave. Hence, second-order risk aversion of a SPS-EU preferences is formally equivalent to preference for late resolution of uncertainty as formulated by Kreps and Porteus (1978).

[3] Compound lotteries are simplified versions of Kreps-Porteus temporal lotteries. The latter allow for interim consumption and more importantly, multiple periods.

[4] Machina and Schmeidler's comparative probability axiom, which is analogous to our Axiom 5 is weaker than Savage's weak comparative probability axiom. In the presence of the sure thing principle, Savage's axiom is equivalent to the Machina-Schmeidler axiom. Therefore, in Theorem 3, we can replace Axiom 5 with suitable analogues of Savage's comparative probability axiom.

Nau (2002) considers the case where both Ω_a and Ω_b are finite and $\mathcal{Z}$ is an interval of real numbers. His assumptions yield additive separability but not state-independence. Hence, he characterizes preferences that can be represented by a function W of the form

$$W(f) = \sum_{i \in \Omega_b} v_i \left(\sum_{j \in \Omega_a} u_{ij}(f(i,j)) \right)$$

The state-dependence allows for a richer set of preferences but makes it impossible to identify subjective probabilities. Instead, Nau utilizes a differentiability assumption to define local probabilities. He notes that a finite state version of what we have called SPS-EU preferences is a special case of his preferences. He uses this model to discuss the Ellsberg Paradox and the Allais Paradox.

The Ellsberg paradox is often studied within the framework of Choquet expected utility or maxmin expected utility preferences. In the Anscombe-Aumann framework, Choquet expected utility preferences are defined as follows: Recall that $\mathcal{F}^*$ denotes the set of all simple functions from Ω_b to P. Let $u : \mathcal{Z} \to I\!R$ and define $U : P \to I\!R$ by $U(p) = \sum_{z \in \mathcal{Z}} u(z)p(z)$. A function $\nu : \mathcal{B} \to [0,1]$ is a capacity if $\nu(\Omega_b) = 1$, $\nu(\emptyset) = 0$ and $\nu(B) \geq \nu(B')$ whenever $B' \subset B$.

For any real-valued, $\mathcal{E}_b$ measurable function r define the Choquet integral of r as follows:

$$\int_{\Omega_b} r d\nu = \sum_{i=1}^{k} (\alpha_i - \alpha_{i+1}) \nu \left(\bigcup_{j \leq i} B_j \right)$$

where $\alpha_1 \geq \ldots \geq \alpha_k$, $\alpha_{k+1} = 0$ and $B_1, \ldots, B_k$ form a partition of Ω_b such that $r(s) = \alpha_i$ for all $s \in B_i$, $i = 1, \ldots, k$. Then, a preference relation $\succeq^*$ on $\mathcal{F}^*$ is a Choquet expected utility preference if there exists a capacity ν and a function $u : \mathcal{Z} \to I\!R$ such that the function W^* defined below represents $\succeq$.

$$W^*(f^*) = \int_{\Omega_b} u \circ f^* \, d\nu$$

Anscombe and Aumann (1963) choose P as their prizes. We do not assume the existence of objective lotteries. However, we utilize the agent's subjective probability μ_a on $\mathcal{A}$ to identify an Anscombe-Aumann act f^* with each f.

Schmeidler's axiomatization of Choquet expected utility relies on the comonotonic independence axiom. Gilboa (1987) provides a characterization of Choquet expected utility preferences in the Savage setting. The axiom below is a version of the comonotonic independence axiom that yields second-order probabilistically sophisticated Choquet expected utility preferences.

Definition: *Let f and g be the two acts below:*

$$f = \begin{pmatrix} x_{11} & \dots & x_{1m} & A_1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n1} & \dots & x_{nm} & A_n \\ B_1 & \dots & B_m & * \end{pmatrix} \quad g = \begin{pmatrix} y_{11} & \dots & y_{1m} & A_1 \\ \vdots & \ddots & \vdots & \vdots \\ y_{n1} & \dots & y_{nm} & A_n \\ B_1 & \dots & B_m & * \end{pmatrix}$$

These acts are comonotonic if for all j, k

$$\begin{pmatrix} x_{1j} & A_1 \\ \vdots & \vdots \\ x_{nj} & A_n \end{pmatrix} \succ \begin{pmatrix} x_{1k} & A_1 \\ \vdots & \vdots \\ x_{nk} & A_n \end{pmatrix} \textit{ implies } \begin{pmatrix} y_{1j} & A_1 \\ \vdots & \vdots \\ y_{nj} & A_n \end{pmatrix} \succeq \begin{pmatrix} y_{1k} & A_1 \\ \vdots & \vdots \\ y_{nk} & A_n \end{pmatrix}$$

The axiom below is the comonotonic sure thing principle for SPS preferences. It imposes the sure thing principle on all comonotonic acts conditional on $\mathcal{E}_a$ events.

Axiom $6c$**:** *(Comonotonic Sure Thing Principle) Let $E_a \in \mathcal{E}_a$ be nonnull and $f, g, f', g' \in \mathcal{F}$, be comonotonic acts. Then*

$$f = f' \text{ on } E_a,\ g = g' \text{ on } E_a,\ f = g \text{ on } E_a^c,\ f' = g' \text{ on } E_a^c$$
$$\textit{implies } f \succeq g \textit{ if and only if } f' \succeq g'$$

Since all acts in $\mathcal{F}_a$ are comonotonic, Axiom $6c$ implies Axiom $6a$. Of course Axiom $6c$ does not imply Axiom $6b$. Similarly, Axiom 6 does not imply Axiom Axiom $6c$. This last assertion can be verified by observing that the preferences characterized by Theorem 3 and 4 are not nested.

Theorem 4 describes preferences that are not probabilistically sophisticated. However, the associated $\succeq^*$ is probabilistically sophisticated. Hence, we characterize only a subset of the preferences studied in Gilboa (1987). This fact accounts for the relatively simpler axioms of Theorem 4.

Recall that for any SPS preference $\succeq$ there exists a non-atomic μ_a and $\succeq^*$ such that $f \succeq g$ if and only if $f^* \succeq^* g^*$. We call the preference $\succeq$ a second-order probabilistically sophisticated Choquet expected utility (SPS-CEU) preference if the corresponding $\succeq^*$ is a Choquet expected utility preference on $\mathcal{F}^*$ for some capacity $\nu = \gamma \circ \mu_b$ where μ_b is a non-atomic probability on $\mathcal{B}$ and $\gamma : [0,1] \to [0,1]$ is a strictly increasing bijection. We refer to (γ, u, μ) as a representation of $\succeq$. The theorem below identifies SPS-CEU preferences as the SPS preferences satisfying Axiom $6c$.

Theorem 4: *The binary relation $\succeq$ satisfies Axioms $1-5$ and $6c$ if and only if it is a SPS-CEU preference.*

Our final result provides a stronger characterization of second-order risk aversion for SPS-EU and SPS-CEU preferences. The theorem shows the equivalence of Schmeidler's notion of uncertainty aversion and our notion of second-order risk aversion for SPS-CEU and SPS-EU preferences. It follows that all of the characterizations of uncertainty aversion provided by Schmeidler are also characterizations second-order risk aversion.[5]

Theorem 5: *Let $\succeq$ be a SPS-EU preference or a SPS-CEU preference. Then, the following conditions are equivalent:*

(i) $\succeq$ *is second order risk averse*

(ii) $h^* = \alpha f^* + (1-\alpha) g^*$ *and* $\pi_f = \pi_g$ *implies* $h \succeq f$

(iii) $h^* = \alpha f^* + (1-\alpha) g^*$ *and* $f \sim g$ *implies* $h \succeq f$

(iv) $f \in \mathcal{F}_a, g \in \mathcal{F}, h \in \mathcal{F}_b$ *and* $p_f = p_g = p_h$ *implies* $f \succeq g \succeq h$.

Theorem 5 establishes that for SPS-CEU preferences and SPS-EU preferences, second-order risk aversion is equivalent to issue preference. Also, the theorem establishes the equivalence of Schmeidler's notion of uncertainty aversion and our notion of second-order risk aversion provided we identify Schmeidler's objective lotteries with issue a.

[5] For example, Schmeidler show that an Choquet expected utility preference is uncertainty averse if and only if ν is convex, that is $\nu(B) + \nu(B') \leq \nu(B \cap B') + \nu(B \cup B')$ for all $B, B' \in \mathcal{B}$. For SPS-CEU preferences, this condition is equivalent to γ being a convex function.

Definition of ambiguity aversion in Klibanoff, Marinacci and Mukerji (2002) is similar to (iv). They impose this condition only on binary acts. For SPS-EU preferences, this ensures that the condition holds for all acts.[6]

4. Conclusion

As in the first interpretation outlined in the introduction, the Ellsberg paradox is often intuitively identified with aversion to ambiguity. Recently, a number of authors have formalized this intuition by providing definitions of ambiguity and ambiguity aversion.

The approach taken by these authors is roughly the following: first, a set of unambiguous acts is defined. Then, an ambiguity neutral agent is defined. Finally, agent 1 is defined to be ambiguity averse if there is an ambiguity neutral agent 2 such that for any act g and any unambiguous act f, $f \succeq_2 g$ implies $f \succeq_1 g$. The notion of ambiguity/ambiguity aversion formalized in Epstein (1999) and Epstein and Zhang (2001) differs from the one in Ghirardato and Marinacci (2001) with respect to the underlying notion of ambiguity neutrality. Epstein (1999) identifies being ambiguity neutral with probabilistic sophistication while Ghirardato and Marinacci (2001) define ambiguity neutrality as expected utility maximization. Ghirardato and Marinacci (2001) seek a very broad notion that permits them to relate any departure from the expected utility model as either ambiguity aversion or ambiguity loving, while the Epstein/Epstein and Zhang formulation is tailored to the analysis of the Ellsberg paradox.

In contrast, both in Nau (2002) and in our model, the emphasis is on the agent having different preferences over uncertain prospects that depend on separate issues. Nau defines the agent's preferred issue as the unambiguous one. Like Ghirardato and Marinacci, Nau utilizes his model to provide a unified framework analyzing state dependent preferences, the Ellsberg paradox and the Allais' paradox. Like Epstein (1999) and Epstein and Zhang (2001), we have attempted to identify our central concept (issue preference or equivalently, second-order risk aversion) exclusively with the Ellsberg paradox.

[6] This can be verified by noting that in the proof of Theorem 5, only binary acts are used for establishing $(iv) \Rightarrow (iii)$ for SPS-EU preferences.

5. Appendix

5.1 Proof of Theorem 1

In the proofs we use the following notion of p-convergence.

Definition: *Given a nonatomic probability μ on Ω such that $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent, a sequence π_n of compound lotteries p-converges to a compound lottery π if there exist acts $f, f_n \in \mathcal{F}$ with $\pi_f = \pi$ and $\pi_{f_n} = \pi_n$, such that for any $\epsilon > 0$:*

$$\lim_{n\to\infty} \mu_b(\{s \in \Omega_b : d^\infty(f_n^*(s), f^*(s)) \geq \epsilon\}) = 0.$$

Lemma 0: *Weak convergence is equivalent to p-convergence.*

Proof: First note that p-convergence is the convergence of f_n^* to f^* in probability. It is well-known that convergence in probability implies weak convergence when the underlying probability measure is countably additive (see for example the corollary to Theorem 3.1 in Billingsley (1999)). Although in our case μ_b need not be countably additive, the standard proof can be adapted since the distribution of f has finite support.

For the converse, suppose π_n weakly converges to π. Since μ is nonatomic, there exists f such that $\pi_f = \pi$. Let $O_\epsilon(p)$ denote the open ball around p with radius ϵ. Choose $\delta > 0$ small enough such that $d^\infty(p, q) > 2\delta$ for any distinct $p, q \in \operatorname{supp} \pi$. Since μ_b is nonatomic, for any n, one can construct $f_n^{**} : \Omega_b \to P$ such that $\pi_{f_n^{**}} = \pi_n$ and

$$\mu_b\left[f_n^{**-1}(O_\delta(p)) \cap f^{*-1}(p)\right] = \min\left\{\pi_n\left(O_\delta(p)\right), \pi\left(p\right)\right\} \qquad \forall p \in \operatorname{supp} \pi$$

By nonatomicity of μ_a, we can find f_n such that $f_n^* = f_n^{**}$ for all n. For any $\epsilon \leq \delta$, $O_\epsilon(p)$'s are disjoint for $p \in \operatorname{supp} \pi$, hence weak convergence of π_n to π ensures that:

$$\lim_{n\to\infty} \pi_n(O_\epsilon(p)) = \pi(O_\epsilon(p)) = \pi(p) \qquad \forall p \in \operatorname{supp} \pi$$

In particular for $\epsilon = \delta$, we have

$$\lim_{n\to\infty} \mu_b\left[f_n^{**-1}(O_\delta(p)) \cap f^{*-1}(p)\right] = \pi(p) \qquad \forall p \in \operatorname{supp} \pi$$

therefore

$$\lim_{n\to\infty} \mu_b(B_n) = 0 \quad \text{where} \quad B_n = \Omega_b \setminus \left(\bigcup_{p\in \operatorname{supp}\pi} [f_n^{**-1}(O_\delta(p)) \cap f^{*-1}(p)] \right).$$

Let $\epsilon \leq \delta$. Since $O_\epsilon(p)$ is disjoint from $O_\delta(q)$ for distinct $p, q \in \operatorname{supp}\pi$, we have:

$$f_n^{**-1}(O_\epsilon(p)) = \left[f_n^{**-1}(O_\epsilon(p)) \cap f^{*-1}(p)\right] \cup \left[f_n^{**-1}(O_\epsilon(p)) \cap B_n\right]$$

therefore for any $p \in \operatorname{supp}\pi$,

$$\lim_{n\to\infty} \mu_b \left[f_n^{**-1}(O_\epsilon(p)) \cap f^{*-1}(p)\right] = \lim_{n\to\infty} \mu_b \left[f_n^{**-1}(O_\epsilon(p))\right] = \pi(p)$$

establishing the p-convergence of π_n to π. $\square$

With this alternative definition of convergence, it is straightforward to verify that if $\succeq$ has the representation in Theorem 1, then it satisfies Axioms $1-5$. For the other direction, assume that $\succeq$ satisfies Axioms $1-5$. Then $\succeq|_{\mathcal{F}_a}-$the restriction of $\succeq$ to $\mathcal{A}-$measurable acts, satisfies the Machina-Schmeidler axioms. Therefore by Theorem 2 of Machina and Schmeidler (1992), there is a nonatomic probability measure μ_a on $(\Omega_a, \mathcal{A})$ and a mixture continuous and monotonic (with respect to first order stochastic dominance) function $\tilde{w} : P \to I\!R$ such that $\tilde{w}(p_f) \geq \tilde{w}(p_g)$ if and only if $f \succeq g$ for all $f \in \mathcal{F}_a$. Hence, $\tilde{w}$ represents $\succeq|_{\mathcal{F}_a}$, the restriction of $\succeq$ to $\mathcal{F}_a$.

Lemma 1: *If B_1 is nonnull, then:*

$$\begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix} \succ \begin{pmatrix} y_1 & A_1 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix} \Leftrightarrow \begin{pmatrix} x_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix} \succ \begin{pmatrix} y_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix}$$

Proof: Let B_1 be nonnull. We prove the lemma in two steps. In Step 1, we show that strict preference $\succ$ on the left hand side implies strict preference $\succ$ on the right hand side. In the second step we show that indifference $\sim$ on the left hand side implies indifference $\sim$ on the right hand side.

Step 1: First assume that

$$\begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix} \succ \begin{pmatrix} y_1 & A_1 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix}.$$

By Axiom 2 (Nondegeneracy), there exist $x, y \in \mathcal{Z}$ such that $x \succ y$. By Axiom 3 (Monotonicity),

$$\begin{pmatrix} x & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix} \succ \begin{pmatrix} y & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix}.$$

Applying to the partition $\{B_1, \emptyset, B_2, \dots, B_m\}$, Axiom 5b ($a|b-$Strong Comparative Probability) yields

$$\begin{pmatrix} x_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix} \succ \begin{pmatrix} y_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix}.$$

Step 2: Assume that

$$\begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix} \sim \begin{pmatrix} y_1 & A_1 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix}.$$

Suppose that the acts

$$\begin{pmatrix} x_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix} \text{ and } \begin{pmatrix} y_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix}$$

are not indifferent. Without loss of generality, assume

$$\begin{pmatrix} x_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix} \succ \begin{pmatrix} y_1 & z_{12} & \dots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix} \qquad (*)$$

Let z_* be a $\succeq$-worst and z^* a $\succeq$-best outcome in $\{x_i : i = 1,\ldots,n\} \cup \{y_i : i = 1,\ldots,n\} \cup \{z_{ij} : i = 1,\ldots,n, j = 2,\ldots,m\}$. Note that it can not be that $z^* \sim z_*$, otherwise by iterated application of Axiom 3 (Monotonicity) we would have indifference in $(*)$. Therefore,

$$z^* \succ \begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix} \sim \begin{pmatrix} y_1 & A_1 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix} \text{ or } \begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix} \sim \begin{pmatrix} y_1 & A_1 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix} \succ z_*.$$

Suppose that z^* is as in above (The other case can be covered by a symmetric argument). By the representation obtained for $\succeq|_{\mathcal{F}_a}$, there exists $i^* \in \{1,\ldots,n\}$, such that $z^* \succ y_{i^*}$ and $\mu_a(A_{i^*}) > 0$. Without loss of generality let $i^* = 1$.

By Axiom $4a$ (Continuity), there is a partition $C_1,\ldots,C_k$ of Ω_a such that for any $i = 1,\ldots,k$:

$$\begin{pmatrix} x_1 & z_{12} & \ldots & z_{1m} & A_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & z_{n2} & \ldots & z_{nm} & A_n \\ B_1 & B_2 & \ldots & B_m & * \end{pmatrix} \succ \begin{pmatrix} z^* & z^* & \ldots & z^* & C_i \\ y_1 & z_{12} & \ldots & z_{1m} & A_1 \setminus C_i \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \ldots & z_{nm} & A_n \setminus C_i \\ B_1 & B_2 & \ldots & B_m & * \end{pmatrix}.$$

Since $\mu_a(A_1) > 0$, there is $i \in \{1,\ldots,k\}$ such that $\mu_a(C_i \cap A_1) > 0$. By iterated application of Axiom 3 (Monotonicity) we have

$$\begin{pmatrix} z^* & z^* & \ldots & z^* & C_i \\ y_1 & z_{12} & \ldots & z_{1m} & A_1 \setminus C_i \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \ldots & z_{nm} & A_n \setminus C_i \\ B_1 & B_2 & \ldots & B_m & * \end{pmatrix} \succeq \begin{pmatrix} z^* & z_{12} & \ldots & z_{1m} & A_1 \cap C_i \\ y_1 & z_{12} & \ldots & z_{1m} & A_1 \setminus C_i \\ y_2 & z_{22} & \ldots & z_{2m} & A_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \ldots & z_{nm} & A_n \\ B_1 & B_2 & \ldots & B_m & * \end{pmatrix}.$$

By transitivity

$$\begin{pmatrix} x_1 & z_{12} & \ldots & z_{1m} & A_1 \cap C_i \\ x_1 & z_{12} & \ldots & z_{1m} & A_1 \setminus C_i \\ x_2 & z_{22} & \ldots & z_{2m} & A_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & z_{n2} & \ldots & z_{nm} & A_n \\ B_1 & B_2 & \ldots & B_m & * \end{pmatrix} \succ \begin{pmatrix} z^* & z_{12} & \ldots & z_{1m} & A_1 \cap C_i \\ y_1 & z_{12} & \ldots & z_{1m} & A_1 \setminus C_i \\ y_2 & z_{22} & \ldots & z_{2m} & A_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \ldots & z_{nm} & A_n \\ B_1 & B_2 & \ldots & B_m & * \end{pmatrix} \qquad (**)$$

Since $\mu_a(C_i \cap A_1) > 0$ and $z^* \succ y_1$, by the representation for $\succeq|_{\mathcal{F}_a}$ we have

$$\begin{pmatrix} z^* & A_1 \cap C_i \\ y_1 & A_1 \setminus C_i \\ y_2 & A_2 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix} \succ \begin{pmatrix} y_1 & A_1 \cap C_i \\ y_1 & A_1 \setminus C_i \\ y_2 & A_2 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix} \sim \begin{pmatrix} x_1 & A_1 \cap C_i \\ x_1 & A_1 \setminus C_i \\ x_2 & A_2 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix}.$$

But then Step 1 implies that

$$\begin{pmatrix} z^* & z_{12} & \dots & z_{1m} & A_1 \cap C_i \\ y_1 & z_{12} & \dots & z_{1m} & A_1 \setminus C_i \\ y_2 & z_{22} & \dots & z_{2m} & A_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix} \succ \begin{pmatrix} x_1 & z_{12} & \dots & z_{1m} & A_1 \cap C_i \\ x_1 & z_{12} & \dots & z_{1m} & A_1 \setminus C_i \\ x_2 & z_{22} & \dots & z_{2m} & A_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & z_{n2} & \dots & z_{nm} & A_n \\ B_1 & B_2 & \dots & B_m & * \end{pmatrix}$$

a contradiction to $(**)$. □

Lemma 1 and the representation obtained for $\succeq|_{\mathcal{F}_a}$ permit us associate with $\succeq$ a preference $\succeq^*$ on Anscombe-Aumann acts $\mathcal{F}^*$ as follows. For any $f \in \mathcal{F}$ let $f^* \succ^* g^* \Leftrightarrow f \succ g$. We next note that $\succeq^*$ is well defined as a preference on $\mathcal{F}^*$.

Lemma 2:

(i) *For any $h \in \mathcal{F}^*$ there is $f \in \mathcal{F}$ such that $h = f^*$.*

(ii) *If $f^* = g^*$ then $f \sim g$.*

Proof: Part (i) follows from nonatomicity of μ_a, part (ii) follows from iterated application of Lemma 1. □

By construction $\delta_p \succ^* \delta_q \Leftrightarrow \tilde{w}(p) > \tilde{w}(q)$. The preference relation $\succeq^*$ inherits the Nondegeneracy axiom from $\succeq$. Note that a set $B \in \mathcal{B}$ is null with respect to $\succeq^*$ if and only if it is null with respect to $\succeq$. Therefore, Lemma 1 implies that $\succeq^*$ satisfies Statewise Monotonicity. By Axiom 3 (Monotonicity) and Axiom 4 (Continuity), $\succeq$ satisfies continuity on Ω_b. Finally, by Axiom 5b ($a|b-$Strong Comparative Probability), $\succeq^*$ also satisfies Strong Comparative Probability. Therefore, we can apply Theorem 2 of Machina and Schmeidler (1992) once again, in order to obtain a nonatomic probability measure μ_b

on $(\Omega_b, \mathcal{B})$ and a mixture continuous and monotonic with respect to first order stochastic dominance function $W\colon \Delta(P) \to I\!R$ such that $W(\pi^*_{f^*}) \geq W(\pi^*_{g^*})$ iff $f^* \succeq g^*$ for all $f, g \in \mathcal{F}$ where $\pi^*_{h^*} \in \Delta(P)$ is defined by $\pi^*_{h^*}(p) = \mu_b(h^{*-1}(p)))$. Since $\pi_f = \pi^*_{f^*}$ for any $f \in \mathcal{F}$, we have that

$$f \succ g \Leftrightarrow f^* \succ^* g^* \Leftrightarrow W(\pi^*_{f^*}) > W(\pi^*_{g^*}) \Leftrightarrow W(\pi_f) > W(\pi_g) \qquad\qquad \forall f, g \in \mathcal{F}$$

establishing that W represents $\succeq$. Define $w\colon P \to I\!R$ by $w(p) = W(\delta_p)$. Then

$$w(p) > w(q) \Leftrightarrow W(\delta_p) > W(\delta_q) \Leftrightarrow \delta_p \succ^* \delta_q \Leftrightarrow \tilde{w}(p) > \tilde{w}(q)$$

showing that w and $\tilde{w}$ are ordinally equivalent. In particular, w is also monotonic with respect to first order stochastic dominance. We conclude the proof by showing that W is weakly continuous.

Lemma 3: *W is weakly continuous.*

Proof: Assume that π_n is a W-bounded sequence that weakly converges to π. Then there exist $z^*, z_* \in \mathcal{Z}$ such that for any $x \in \bigcup_{n=1}^{\infty} \bigcup_{p \in \operatorname{supp} \pi_n} \operatorname{supp} p$, we have $W(z^*) \geq W(x) \geq W(z_*)$ implying $z^* \succeq x \succeq z_*$. Since weak convergence implies p-convergence, there exist acts $f, f_t \in \mathcal{F}$ with $\pi_f = \pi$ and $\pi_{f_t} = \pi_t$, such that for any $\epsilon > 0$:

$$\lim_{t \to \infty} \mu_b(\{s \in \Omega_b : d^{\infty}(f^*_t(s), f^*(s)) \geq \epsilon\}) = 0 \qquad (*)$$

We only consider the case where $W(\pi) > W(\pi')$, the argument for the other case is symmetric. By nonatomicity of μ there is $f' \in \mathcal{F}$ such that $\pi' = \pi_{f'}$, then $f \succ f'$.

Suppose without loss of generality that we can express f as:

$$\begin{pmatrix} x_{11} & \dots & x_{1m} & A_1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n1} & \dots & x_{nm} & A_n \\ B_1 & \dots & B_m & * \end{pmatrix}$$

where (i) $\mu_a(A_1) > 0$; (ii) $x_{1j} \succeq x_{ij}$, for $i = 1, \dots, n$ and $j = 1, \dots, m$. (If f does not have the above form, using nonatomicity of μ_a we can find $\bar{f} \in \mathcal{F}$ of the desired form such that $\bar{f}^* = f^*$.)

Let $\underline{x}$ be a $\succeq$-worst outcome in $\{z_*\} \cup \bigcup_{p \in \operatorname{supp} \pi} \operatorname{supp} p$. Now since $f \succ f'$, by Axiom 4a (Continuity) there is a partition $C_1, \ldots, C_k$ of Ω_a such that for any $i = 1, \ldots, k$:

$$g^i = \begin{pmatrix} \underline{x} & \dots & \underline{x} & C_i \\ x_{11} & \dots & x_{1m} & A_1 \setminus C_i \\ \vdots & \ddots & \vdots & \vdots \\ x_{n1} & \dots & x_{nm} & A_n \setminus C_i \\ B_1 & \dots & B_m & * \end{pmatrix} \succ f'.$$

Since $\mu_a(A_1) > 0$, there is $i \in \{1, \ldots, k\}$ such that $\mu_a(C_i \cap A_1) > 0$. Let $C = C_i \cap A_1$, then

$$g := \begin{pmatrix} \underline{x} & \dots & \underline{x} & C \\ x_{11} & \dots & x_{1m} & A_1 \setminus C \\ x_{21} & \dots & x_{2m} & A_2 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n1} & \dots & x_{nm} & A_n \\ B_1 & \dots & B_m & * \end{pmatrix} \succeq g^i \succ f',$$

where again the initial weak preference above follows from iterated application of Axiom 3 (Monotonicity).

Suppose without loss of generality that: (iii) $\mu_b(B_1) > 0$ and

$$(iv) \quad \begin{pmatrix} \underline{x} & C \\ x_{11} & A_1 \setminus C \\ \vdots & \vdots \\ x_{n1} & A_n \end{pmatrix} \succeq \begin{pmatrix} \underline{x} & C \\ x_{1j} & A_1 \setminus C \\ \vdots & \vdots \\ x_{nj} & A_n \end{pmatrix} \quad j = 1, \ldots, m.$$

(If g does not satisfy (iii) and (iv), we can reorder the columns by adjoining the null B_j's to a nonnull one and construct $\bar{g}$ satisfying (iii) and (iv) such that the equality $\bar{g}^* = g^*$ holds μ_b-almost surely.)

Since $g \succ f'$, by Axiom 4b (Continuity) there is a partition $D_1, \ldots, D_l$ of Ω_b such that for any $j = 1, \ldots, l$:

$$h^j = \begin{pmatrix} \underline{x} & \underline{x} & \dots & \underline{x} & C \\ \underline{x} & x_{11} & \dots & x_{1m} & A_1 \setminus C \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \underline{x} & x_{n1} & \dots & x_{nm} & A_n \\ D_j & B_1 \setminus D_j & \dots & B_m \setminus D_j & * \end{pmatrix} \succ f'.$$

Since $\mu_b(B_1) > 0$, there is $j \in \{1, \ldots, l\}$ such that $\mu_b(D_j \cap B_1) > 0$. Let $D = D_j \cap B_1$, then

$$h := \begin{pmatrix} \underline{x} & \underline{x} & \underline{x} & \cdots & \underline{x} & C \\ \underline{x} & x_{11} & x_{12} & \cdots & x_{1m} & A_1 \setminus C \\ \underline{x} & x_{21} & x_{22} & \cdots & x_{2m} & A_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \underline{x} & x_{n1} & x_{n2} & \cdots & x_{nm} & A_n \\ D & B_1 \setminus D & B_2 & \cdots & B_m & * \end{pmatrix} \succeq h^j \succ f',$$

where again the initial weak preference above follows from iterated application of Axiom 3 (Monotonicity).

Let $\epsilon = \min\{\mu_a(C)/n, \mu_b(D)/m\} > 0$. From $(*)$, choose N large enough such that

$$\mu_b(\{s \in \Omega_b : d^\infty(f_r^*(s), f^*(s)) \geq \epsilon\}) < \epsilon \quad \text{for any integer } r \geq N. \quad (**)$$

Fix any integer $r \geq N$. By nonatomicity of μ_a and $(**)$, there is $\bar{f}_r \in \mathcal{F}$ and events $E_1, \ldots, E_n \in \mathcal{A}$; $F_1, \ldots, F_m \in \mathcal{B}$ such that $\bar{f}_r^* = f_r^*$, $E_i \subset A_i$, $F_j \subset B_j$, $\mu_a(E_i) < \epsilon$, $\mu_b(F_j) < \epsilon$ and $\bar{f}_r$ gives x_{ij} on $A_i \setminus E_i \times B_j \setminus F_j$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

Then, by (iv), our choice of ϵ and iterated application of monotonicity with respect to stochastic dominance of W, we obtain:

$$h_r = \begin{pmatrix} \underline{x} & \underline{x} & \cdots & \underline{x} & \underline{x} & C \\ x_{11} & \underline{x} & \cdots & x_{1m} & \underline{x} & A_1 \setminus C \\ x_{21} & \underline{x} & \cdots & x_{2m} & \underline{x} & A_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{n1} & \underline{x} & \cdots & x_{nm} & \underline{x} & A_n \\ B_1 \setminus F_1 & F_1 & \cdots & B_m \setminus F_m & F_m & * \end{pmatrix} \succeq h.$$

Similarly, by (ii), our choice of ϵ and iterated application of monotonicity with respect to stochastic dominance of w and W, we obtain:

$$g_r = \begin{pmatrix} x_{11} & \underline{x} & \cdots & x_{1m} & \underline{x} & A_1 \setminus E_1 \\ \underline{x} & \underline{x} & \cdots & \underline{x} & \underline{x} & E_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{n1} & \underline{x} & \cdots & x_{nm} & \underline{x} & A_n \setminus E_n \\ \underline{x} & \underline{x} & \cdots & \underline{x} & \underline{x} & E_n \\ B_1 \setminus F_1 & F_1 & \cdots & B_m \setminus F_m & F_m & * \end{pmatrix} \succeq h_r.$$

Finally, by iterated application of Axiom 3 (Monotonicity) and the construction of $\bar{f}_r$, we have that $\bar{f}_r \succeq g_r$. Therefore $f_r \sim \bar{f}_r \succeq g_r \succeq h_r \succeq h \succ f'$, as desired. $\square$

5.2 Proof of Theorem 2

Part $(iii) \Rightarrow (ii)$ follows from $\pi_f = \pi_g \Rightarrow f \sim g$ for SPS preferences. To see the implication $(i) \Rightarrow (iv)$, it is enough to note that $f \in \mathcal{F}_a$, $g \in \mathcal{F}$, and $h \in \mathcal{F}_b$ induce the same lottery (i.e., $p_f = p_g = p_h$) then, there exists $\pi^1, \ldots, \pi^m$ and k such that π^{j+1} is a mean preserving spread of π^j for all $j = 1, \ldots, m-1$, $\pi^1 = \pi_f, \pi^k = \pi_g$ and $\pi^m = \pi_h$. We conclude the proof of the theorem by establishing the equivalence of (i) and (ii).
$(i) \Rightarrow (ii)$: Assume that $\succeq$ is second order risk averse. Let $f, g, h \in \mathcal{F}$ be such that $h^* = \alpha f^* + (1-\alpha) g^*$ and $\pi_f = \pi_g$. Then

$$\pi_h = \sum_{(p,q) \in P \times P} \mu_b \left(f^{*-1}(p) \cap g^{*-1}(q) \right) \delta_{\alpha p + (1-\alpha) q}.$$

Let

$$\pi = \sum_{(p,q) \in P \times P} \mu_b \left(f^{*-1}(p) \cap g^{*-1}(q) \right) (\alpha \delta_p + (1-\alpha) \delta_q).$$

Therefore there exists $\pi^1, \ldots, \pi^m$ such that $\pi^1 = \pi_h$, $\pi^m = \pi$ and π^{j+1} is a mean preserving spread of π^j for $j = 1, \ldots, m-1$. Hence, (i) implies $W(\pi_h) \geq W(\pi)$. We can rewrite π as

$$\begin{aligned}
\pi &= \sum_{p \in P} \left[\alpha \sum_{q \in P} \mu_b \left(f^{*-1}(p) \cap g^{*-1}(q) \right) + (1-\alpha) \sum_{q \in P} \mu_b \left(f^{*-1}(q) \cap g^{*-1}(p) \right) \right] \delta_p \\
&= \sum_{p \in P} \left[\alpha \mu_b \left(f^{*-1}(p) \right) + (1-\alpha) \mu_b \left(g^{*-1}(p) \right) \right] \delta_p \\
&= \sum_{p \in P} \left[\alpha \pi_f(p) + (1-\alpha) \pi_g(p) \right] \delta_p = \pi_f = \pi_g.
\end{aligned}$$

Since $\pi_f = \pi$ and W represents $\succeq$, we have $h \succeq f$. $\square$
$(ii) \Rightarrow (i)$: Assume that the SRS preference $\succeq$ satisfies condition (ii). It is enough to show that $f \succeq g$ whenever π_g is a mean preserving spread of π_f. Let π_g be a mean preserving spread of π_f, then there are $\alpha, \beta \in [0,1]$, $p, q \in P$ and $\pi' \in \Delta(P)$ such that

$$\begin{aligned}
&\pi_g = \alpha(\beta \delta_p + (1-\beta)\delta_q) + (1-\alpha)\pi' \text{ and} \\
&\pi_f = \alpha \delta_{\beta p + (1-\beta) q} + (1-\alpha)\pi'
\end{aligned}.$$

Let $B \subset \Omega_b$ be such that $\mu_b(B) = \alpha$ and $f^* = \beta p + (1-\beta)q$ on B. Without loss of generality, let $f^* = g^*$ outside of B.

Next, we define a sequence of partitions $\Pi^k = \{B_0^k, B_1^k, \ldots, B_{2^k-1}^k\}$ of B such that Π^{k+1} refines Π^k for $k \geq 1$. Let $B_0^1 = \{s \in B : g^*(s) = p\}$ and $B_1^1 = \{s \in B : g^*(s) = q\}$. Having defined the partition Π^k for some $k \geq 1$, inductively define Π^{k+1} as follows: For any $l \in \{0, \ldots, 2^k - 1\}$ by nonatomicity of μ_b, partition B_l^k into two subsets B_{2l}^{k+1} and B_{2l+1}^{k+1} such that $\mu_b(B_{2l}^{k+1}) = \beta\mu_b(B_l^k)$ and $\mu_b(B_{2l+1}^{k+1}) = (1-\beta)\mu_b(B_l^k)$.

Note that $\mu_b(B_l^k) = \alpha\beta^i(1-\beta)^{k-i}$ where i is the number of 0's in the k-digit binary expansion of l. (For example, if $k = 5$ and $l = 9$ then the 5-digit binary expansion of 9 is 01001 so $i = 3$.) By nonatomicity of μ_a, we can find a sequence of acts $g_k \in \mathcal{F}$ such that:

$$g_k^*(s) = \begin{cases} p & s \in B_l^k \text{ and } l \text{ is even} \\ q & s \in B_l^k \text{ and } l \text{ is odd} \\ g^*(s) & s \notin B \end{cases}$$

By definition $g_1^* = g^*$, implying that $g_1 \sim g$. By nonatomicity of μ_b, there exist acts h_m^n for $n = 0, 1, 2, \ldots$ and $m = 1, 2, \ldots$ such that:

$$h_m^{n*} = \sum_{k=(m-1)2^n+1}^{m2^n} \frac{1}{2^n} g_k^*$$

i.e., h_m^{n*} is the equal weight probability mixture (average) of the mth 2^n consecutive Anscombe-Aumann acts in the sequence g_k^*.

Note that on $B_l^{2^n}$, h_1^{n*} gives $\frac{i}{2^n}p + \frac{2^n-i}{2^n}q$ and on $B_l^{m2^n}$, h_m^{n*} gives $\frac{i}{2^n}p + \frac{2^n-i}{2^n}q$ where i is the number of 0's in the last 2^n digits in the $m2^n$-digit binary expansion of l. Therefore, we can write $\pi_{h_1^n}$ as:

$$\pi_{h_1^n} = \alpha \sum_{i=0}^{2^n} \binom{2^n}{i} \beta^i (1-\beta)^{2^n-i} \delta_{\frac{i}{2^n}p + \frac{2^n-i}{2^n}q} + (1-\alpha)\pi'$$

It is easy to verify that $\pi_{h_m^n} = \pi_{h_1^n}$ and therefore $h_m^n \sim h_1^n$ for all $m \geq 1$. Since $h_m^{n+1*} = \frac{1}{2}h_{2m-1}^{n*} + \frac{1}{2}h_{2m}^{n*}$ and $\pi_{h_{2m-1}^n} = \pi_{h_{2m}^n}$, by condition (ii), $h_m^{n+1} \succeq h_{2m}^n$. and therefore $h_1^n \succeq h_1^0$, by transitivity for any $n \geq 0$. Since $h_1^{0*} = g_1^*$ we have $h_1^0 \sim g_1$, thus $h_1^n \succeq h_1^0 \sim g_1 \sim g$ implying that $h_1^n \succeq g$ for any $n \geq 0$.

We next show that $\pi_{h_1^n}$ weakly converges to π_f. Let $\epsilon > 0$ be given. Let $\epsilon' > 0$ be such that $d^\infty(\beta' p + (1-\beta')q, \beta p + (1-\beta)q) < \epsilon$ for any $\beta' \in (\beta - \epsilon', \beta + \epsilon')$. By the Weak Law of Large Numbers, the average of i.i.d. Bernouilli random variables with mean β converges in probability to β. Therefore, there is N such that for any $n \geq N$:

$$\sum_{i=0}^{2^n} \binom{2^n}{i} \beta^i (1-\beta)^{2^n - i} 1_{\left\{\left|\frac{i}{2^n} - \beta\right| \geq \epsilon'\right\}} < \epsilon.$$

Then for any $n \geq N$,

$$\begin{aligned}
&\mu_b\left(\left\{s \in \Omega_b : d^\infty\left(h_1^{*n}(s), f^*(s)\right) \geq \epsilon\right\}\right) \\
=&\alpha \sum_{i=0}^{2^n} \binom{2^n}{i} \beta^i (1-\beta)^{2^n - i} 1_{\left\{d^\infty\left(\frac{i}{2^n} p + \frac{2^n - i}{2^n} q, \beta p + (1-\beta) q\right) \geq \epsilon\right\}} \\
\leq& \sum_{i=0}^{2^n} \binom{2^n}{i} \beta^i (1-\beta)^{2^n - i} 1_{\left\{\left|\frac{i}{2^n} - \beta\right| \geq \epsilon'\right\}} < \epsilon.
\end{aligned}$$

Thus $\pi_{h_1^n}$ weakly converges to π_f. Since $\bigcup_{n=0}^{\infty} \operatorname{supp}_{\mathcal{Z}} \pi_{h_1^n} = \operatorname{supp}_{\mathcal{Z}} \pi_f$ is finite and $W(\pi_{h_1^n}) \geq W(\pi_g)$, weak continuity of W implies that $W(\pi_f) \geq W(\pi_g)$. Therefore $\succeq$ is second order risk averse. □

Below, we provide counter-examples to (i) implies (iii) and (iv) implies (i). For both counter-examples assume that $\mathcal{Z} = \{0, 1\}$. Hence, P can be identified with the unit interval where $p \in P$ denotes the probability of getting 1. Also, each π can be identified with a simple probability distribution on the unit interval. Let μ_a be any nonatomic probability measure on the set of all subsets of some Ω_a. Similarly, let μ_b be any nonatomic probability measure on the set of all subsets of some Ω_b.

We first define a weakly continuous utility function W on $\Delta(P)$. Since each $f \in \mathcal{F}$ can be identified with a unique π_f, this utility function induces a preference $\succeq_W$ on $\mathcal{F}$. Define the function $m : \Delta(P) \to [0, 1]$ as follows:

$$m(\pi) = \sum_{x \in [0,1]} x \pi(x)$$

Hence, $m(\pi)$ is the mean of π. For any lottery π define η_π, the absolute error of π as follows:

$$\eta_\pi(z) = \sum_{x : |x - m(\pi)| = z} \pi(x)$$

Hence, $m(\eta_\pi)$ is the mean absolute error of π. Let $\psi(\alpha) = \frac{log(1+\alpha)}{3}$. Define

$$W(\pi) = m(\pi) - \psi(m(\eta_\pi))$$

The weak continuity of W is easy to verify. It is straightforward to check that for any $y > x$, an increase in $\pi(y)$ at the expense of $\pi(x)$ cannot increase $m(\eta_\pi)$ at a rate greater than $2(y-x)$ and hence $\psi(m(\eta_\pi))$ cannot increase at a rate greater than $2(y-x)/3$. On the other hand, $m(\pi)$ increases at a rate $y-x$. The overall effect is an increase in $W(\pi)$ establishing that W satisfies stochastic dominance. Since $W(\delta_x) = x$ we conclude that w satisfies stochastic dominance as well. We claim that $\succeq_W$ satisfies (i) but not (iii). To verify (i), note that W is risk averse since mean-preserving spreads leave $m(\pi)$ unchanged and (weakly) decrease $m(\eta_\pi)$.

To see that $\succeq_W$ does not satisfy (iii), let $\pi = .5\delta_1 + .5\delta_0$ and set $w = W(\pi) < .5$. Choose f, g such that $\pi_f = \delta_w$ and $\pi_g = \pi$. Let h be such that $h^* = .5f^* + .5g^*$, then $\pi_h = .5\delta_{.5w+.5} + .5\delta_{.5w}$ and $m(\eta_{\pi_h}) = .5m(\eta_\pi) + .5m(\eta_{\delta_w}) = .5m(\eta_\pi)$. Hence, the strict concavity of ψ ensures that $W(\pi_h) < .5W(\pi) + .5W(\delta_w) = w$ proving that $\succeq_W$ does not satisfy (iii).

For the second counter-example, let V be the nonexpected utility functional on $\Delta(P)$ defined as follows:

$$V(\pi) = \frac{\alpha m(\pi^1) + 2(1-\alpha)m(\pi^2)}{2-\alpha} \tag{$*$}$$

where $\pi = \alpha\pi^1 + (1-\alpha)\pi^2$, $\pi^1(x) > 0$ implies $x \geq V(\pi)$ and $\pi^2(x) > 0$ implies $x \leq V(\pi)$. The preference represented by this V belongs to the class introduced in Gul (1991). In particular, this preference is a disappointment averse preference with linear u and $\beta = 1$. Gul (1991) establishes that the function V is well-defined; that is a real number $V(\pi)$ satisfying $(*)$ always exists and that this number is the same for any α, π^1, π^2 satisfying the properties above. Define $W : \Delta(P) \to I\!R$ as follows:

$$W(\pi) = m(\pi) - \frac{1}{4}V(\eta_\pi)$$

Weak continuity of W follows from continuity of disappointment averse preferences in Gul (1991). Again, it can be verified that for $y > x$ a small increase in $\pi(y)$ at the expense of $\pi(x)$ increases $V(\eta_\pi)$ at a rate no greater than $3(y-x)$ and hence $V(\eta_\pi)/4$ increases at a rate no greater than $3(y-x)/4$, while $m(\pi)$ increases at a rate $y-x$, proving that W and the function w defined by $w(p) := W(\delta_p)$ both satisfy stochastic dominance. We claim that $\succeq_W$ satisfies (iv) but not (i).

Let $f \in \mathcal{F}_a, g \in \mathcal{F}$ such that $p_f = p_g$. Then, $W(\pi_f) = p_f = p_g \geq p_g - V(\eta_{\pi_g})/4 = W(\pi_g)$ and hence $f \succeq g$.

Next, take $h \in \mathcal{F}_b$ such that $p_h = p_g$. Then η_{π_g} can be expressed as $\sum_{j=1}^N \alpha_j \eta_j$ where each η_j has the form

$$\eta_j = \frac{w_j}{z_j + w_j}\delta_{z_j} + \frac{z_j}{z_j + w_j}\delta_{w_j}$$

for some α_j, z_j, w_j such that $1 \geq z_j \geq w_j \geq 0$, $\alpha_j \geq 0$, for $j = 1, \ldots, N$ and $\sum_{j=1}^N \alpha_j = 1$. Also,

$$\eta_{\pi_h} = \frac{w^*}{z^* + w^*}\delta_{z^*} + \frac{z^*}{z^* + w^*}\delta_{w^*}$$

for some z^*, w^* such that $1 \geq z^* \geq z_j$ and $z^* \geq w^* \geq w_j$ for all z_j, w_j. Gul (1991) shows that the nonexpected utility functional V satisfies betweenness; that is, $V(\nu) \geq V(\nu')$ implies

$$V(\nu) \geq V(\alpha\nu + (1-\alpha)\nu') \geq V(\nu')$$

Thus, to prove that $W(\pi_g) \geq W(\pi_h)$ it is enough to establish that

$$V(\frac{w^*}{z^* + w^*}\delta_{z^*} + \frac{z^*}{z^* + w^*}\delta_{w^*}) \geq V(\frac{w}{z + w}\delta_z + \frac{z}{z + w}\delta_w)$$

whenever $z^* \geq z$, $w^* \geq w$, $z^* \geq w^*$ and $z \geq w$. This is verified easily by noting that $V(\frac{z}{z+w}\delta_w + \frac{w}{z+w}\delta_z) = \frac{3zw}{2z+w}$ is increasing both in z and in w.

To show that W does not satisfy (i), we construct π and π' such that π' is a mean-preserving spread of π and $V(\eta_\pi) > V(\eta_{\pi'})$. For example $\pi = .4\delta_{5/6} + .6\delta_0$ and $\pi' = .2\delta_1 + .2\delta_{2/3} + .6\delta_0$ and hence $V(\eta_\pi) = 3/8$, $V(\eta_{\pi'}) = 10/27$ satisfy the desired inequality.

5.3 Proof of Theorem 3

Let W be the representation of $\succeq$ guaranteed by Theorem 1. Let $\mu = \mu_a \times \mu_b$ be the associated probability measure. Define $\succeq^*$ on $\mathcal{F}^*$ as follows $f^* \succeq^* g^*$ if and only if $W(\pi_f) \succeq W(\pi_g)$. Since μ_b and μ_a are nonatomic, $\succeq^*$ is well-defined. It follows from Axiom

$6b$ (The Sure Thing Principle) and Savage's Theorem that $\succeq^*$ has an expected utility representation W^* such that $W^*(f^*) = \sum_p U^*(p)\mu(f^{*-1}(p))$. By Axiom $6a$ (The Sure Thing Principle), the preference on $\mathcal{F}_a$ defined by $f \succeq' g$ if and only if $U^*(p_f) \geq U^*(p_g)$ satisfies all of the Savage axioms and therefore there exists an expected utility function $U : P \to I\!R$ such that if $U(p_f) \geq U(p_g)$ if and only if $U^*(p_f) \geq U^*(p_g)$. Since U^* and U represent the same preference $\succeq$ there exist a strictly increasing function $v : U(P) \to I\!R$ such that $U^* = v \circ U$. Let $u(z) = \delta_z$ for all $z \in \mathcal{Z}$. Define W' by

$$W'(\pi) = \sum_{p \in P} v\left(\sum_{x \in \mathcal{Z}} u(x)p(x)\right)\pi(p)$$

Note that $W^*(f^*) = W'(\pi_f)$ for all f and $f \succeq g$ iff $f^* \succeq^* g^*$ iff $W^*(f^*) \geq W^*(g^*)$ iff $W(\pi_f) \geq W(\pi_g)$. Since U is an expected utility function, $U(P)$ is an interval. Hence, if v is continuous it can easily be extended to a continuous, strictly increasing function on $I\!R$. Therefore, to conclude the proof, we need only to show that v is continuous. Since, v is strictly increasing, there are only two possible types of discontinuities it can have: There exists $t = U(p)$ and $\varepsilon > 0$ such that either $v(t) \geq v(t') + \varepsilon$ for all $t' < t$, $t' \in U(P)$ or $v(t') \geq v(t) + \varepsilon$ for all $t' > t$, $t' \in U(P)$. Suppose, the former holds for some t (the argument for the other case is symmetric and omitted).

Choose $t' < t$ such that $v(t') > v_- - \varepsilon$ where v_- is the left limit of v at t. Let $p' \in P$, $f \in \mathcal{F}_a$, $g \in \mathcal{F}$ and $B \in \mathcal{B}$ be such that $U(p') = t'$, $\mu_b(B) = .5$, $f^*(\omega_b) = p$ for all $\omega_b \in \Omega_b$, $g^*(\omega_b) = p$ for all $\omega_b \in B$ and $g^*(\omega_b) = p'$ for all $\omega_b \in \Omega_b \backslash B$. Then $W'(\pi_f) = v(t) > .5v(t) + .5v(t') = W'(\pi_g) > v_-$ and hence $f \succ g$. Let x minimize $U(\delta_z)$ among z in the support of p'. There exists $A \in \mathcal{A}$ with $\mu_a(A) > 0$ such that f gives a prize strictly better than x on A. Then, for any $A' \subset A$ such that $\mu_a(A') > 0$ $\hat{f}(\omega_a, \omega_b) = x$ for all $\omega_a \in A'$ and $\hat{f}(\omega_a, \omega_b) = f(\omega_a, \omega_b)$ otherwise implies $W'(\pi_{\hat{f}}) \leq v_-$. So, $g \succ \hat{f}$, contradicting Axiom $4a$ (Continuity). $\square$

5.4 Proof of Theorem 4

Two acts $f^*, g^* \in \mathcal{F}^*$ such that

$$f^* = \begin{pmatrix} p_1 & p_2 & \dots & p_m \\ B_1 & B_2 & \dots & B_m \end{pmatrix},\ g^* = \begin{pmatrix} q_1 & q_2 & \dots & q_m \\ B_1 & B_2 & \dots & B_m \end{pmatrix}$$

are comonotonic if $p_i \succ^* p_j$ implies $q_i \succeq^* q_j$ for all i, j. Three acts are comonotonic if each pair is comonotonic.

A preference relation $\succeq^*$ on $\mathcal{F}^*$ satisfies vNM continuity if $f^* \succ^* g^* \succ^* h^*$ implies that there exist $\alpha, \beta \in (0,1)$ such that $\alpha f^* + (1-\alpha)h^* \succ^* g^* \succ^* \beta f^* + (1-\beta)h^*$. The preference $\succeq^*$ satisfies comonotonic independence, if f^*, g^*, h^* are comonotonic, $f^* \succ^* g^*$ and $\alpha \in (0,1)$ implies $\alpha f^* + (1-\alpha)h^* \succ^* \alpha g^* + (1-\alpha)h^*$.

By the theorem on page 578 of Schmeidler (1989) if a preference relation $\succeq^*$ on $\mathcal{F}^*$ satisfies vNM continuity, weak stochastic dominance (i.e., $f^*(\omega_b) \succeq^* g^*(\omega_b)$ for all $\omega_b \in \Omega_b$ implies $f^* \succeq^* g^*$), weak nondegeneracy (i.e., there exists f^*, g^* such that $f^* \succ^* g^*$) and comonotonic independence then it is has a Choquet expected utility representation.

By Theorem 1, there exists a preference $\succeq^*$ on $\mathcal{F}^*$ such that $f \succeq g$ iff $f^* \succeq^* g^*$ and a weakly continuous, stochastic monotonicity satisfying W such that $W(\pi_f) \geq W(\pi_g)$ iff $f^* \succeq^* g^*$ for all $f, g \in \mathcal{F}$. Then, Axiom 2 (Nondegeneracy) implies that $\succeq^*$ satisfies weak nondegeneracy and the fact that W is satisfies stochastic dominance implies $\succeq^*$ satisfies weak monotonicity. Also, it follows from the weak continuity of W that $\succeq^*$ satisfies vNM continuity. We show that $\succeq^*$ is a Choquet expected utility by proving that $\succeq^*$ satisfies vNM comonotonic independence.

Observe that since Axiom 6c (Comonotonic Sure Thing Principle) implies Savage's sure thing principle on $\mathcal{F}_a$. Then, by Savage's theorem, there exists some expected utility function U such that $f \succeq g$ iff $U(p_f) \geq U(p_g)$ for all $f, g \in \mathcal{F}_a$.

Consider comonotonic f^*, g^* such that

$$f^* = \begin{pmatrix} p_1 & p_2 & \cdots & p_m \\ B_1 & B_2 & \ldots & B_m \end{pmatrix}, \; g^* = \begin{pmatrix} q_1 & q_2 & \cdots & q_m \\ B_1 & B_2 & \ldots & B_m \end{pmatrix}$$

Then, for any natural number n construct $\hat{f}, \hat{g}$ such that

$$\hat{f} = \begin{pmatrix} p_1 & \ldots & p_m & A_1 \\ \vdots & \ddots & \vdots & \vdots \\ p_1 & \ldots & p_m & A_n \\ B_1 & \ldots & B_m & * \end{pmatrix} \; \hat{g} = \begin{pmatrix} q_1 & \ldots & q_m & A_1 \\ \vdots & \ddots & \vdots & \vdots \\ q_1 & \ldots & q_m & A_n \\ B_1 & \ldots & B_m & * \end{pmatrix}$$

That is, $\hat{f}$ conditional on $A_i \times B_j$ has distribution p_j and $\hat{g}$ conditional on $A_i \times B_j$ has distribution q_j for all i, j.

For any $M \subset \{1, \ldots, m\}$, let $\hat{f}^M$ denote the act obtained from f by replacing each row $j \in M$ with the corresponding row of $\hat{g}$. Hence, $\hat{f}^{\emptyset} = \hat{f}$ and $\hat{f}^{\{1,\ldots,m\}} = \hat{g}$ etc. Note that $U(p_i) \geq U(p_j)$ and $U(q_i) \geq U(q_j)$ implies $U(\alpha p_i + (1-\alpha)q_i) \geq U(\alpha p_j + (1-\alpha)q_j)$. Hence, $\hat{f}^M$ and $\hat{f}^{M'}$ are comonotonic for all M, M'.

Then, by Axiom 6*c* (Comonotonic Sure Thing Principle), $g^* \succeq \frac{1}{n}f^* + \frac{n-1}{n}g^*$ implies

$$\frac{1}{n}f^* + \frac{n-1}{n}g^* \sim \begin{pmatrix} q_1 & \ldots & q_m & A_1 \\ p_1 & \ldots & p_m & A_2 \\ q_1 & \ldots & q_m & A_3 \\ \vdots & \ddots & \vdots & \vdots \\ q_1 & \ldots & q_m & A_n \\ B_1 & \ldots & B_m & * \end{pmatrix} \succeq \begin{pmatrix} p_1 & \ldots & p_m & A_1 \\ p_1 & \ldots & p_m & A_2 \\ q_1 & \ldots & q_m & A_3 \\ \vdots & \ddots & \vdots & \vdots \\ q_1 & \ldots & q_m & A_n \\ B_1 & \ldots & B_m & * \end{pmatrix} \sim \frac{2}{n}f^* + \frac{n-2}{n}g^*$$

Repeating the argument with other rows and using transitivity implies $g^* \succeq f^*$. It follows that $f^* \succ g^*$ implies $f^* \succ \alpha f^* + (1-\alpha)g^* \succ g^*$ for every rational $\alpha \in (0,1)$. It then follows from the weak continuity of W that the same holds for every $\alpha \in (0,1)$.

Suppose $f^* \succ g^*$ and

$$h^* = \begin{pmatrix} r_1 & r_2 & \ldots & r_m \\ B_1 & B_2 & \ldots & B_m \end{pmatrix}$$

is also comonotonic with f^* and g^*. For any $\alpha \in (0,1)$ choose $A \in \mathcal{A}$ such that $\mu_a(A) = \alpha$ and note that by the argument above

$$f^* \sim \begin{pmatrix} p_1 & \ldots & p_m & A \\ p_1 & \ldots & p_m & \Omega_a \backslash A \\ B_1 & \ldots & B_m & * \end{pmatrix} \succ \begin{pmatrix} q_1 & \ldots & q_m & A \\ p_1 & \ldots & p_m & \Omega_a \backslash A \\ B_1 & \ldots & B_m & * \end{pmatrix} \sim \alpha f^* + (1-\alpha)g^*$$

Applying Axiom 6*c* (Comonotonic Sure Thing Principle) again yields

$$\alpha f^* + (1-\alpha)h^* \sim \begin{pmatrix} p_1 & \ldots & p_m & A \\ r_1 & \ldots & r_m & \Omega_a \backslash A \\ B_1 & \ldots & B_m & * \end{pmatrix} \succ \begin{pmatrix} q_1 & \ldots & q_m & A \\ r_1 & \ldots & r_m & \Omega_a \backslash A \\ B_1 & \ldots & B_m & * \end{pmatrix} \sim \alpha g^* + (1-\alpha)h^*$$

Proving comonotonic independence.

Therefore $\succeq^*$ is a Choquet expected utility preference, let W^* be the Choquet expected utility that represents $\succeq^*$. Without loss of generality let $W^*(p) = U(p)$ for any constant act $p \in \mathcal{F}^*$.

By Theorem 1, $\succeq^*$ is probabilistically sophisticated. It follows that the associated capacity ν can be written as $\gamma \circ \mu_b$ for strictly increasing $\gamma : [0,1] \to [0,1]$ such that

$\gamma(0) = 0, \gamma(1) = 1$ and probability μ_b. To conclude the proof we show that γ is continuous. Since γ is strictly increasing, there are only two possible types of discontinuities it can have: Either $\gamma(t) \geq \gamma(t') + \varepsilon$ for all $t' < t$ or $\gamma(t') \geq \gamma(t) + \varepsilon$ for all $t' > t$. Suppose, the former holds for some t (the argument for the other case is symmetric and omitted).

Choose B such that $\mu_b(B) = t$. Such a B exists by Theorem 1. Choose p, q such that $U(p) > U(q)$ and $\alpha \in (\gamma(t) - \varepsilon, \gamma(t))$. Such p, q exits by nondegeneracy. Define $f, g \in \mathcal{F}$, such that $f^*(\omega_b) = p$ for all $\omega_b \in B$, $f^*(\omega_b) = q$ for all $\omega_b \in \Omega_b \backslash B$ and $g^*(\omega_b) = \alpha p + (1-\alpha)q$ for all $\omega_b \in \Omega_b$. Note that $W^*(f^*) = \gamma(t)U(p) + (1-\gamma(t))U(q) > \alpha U(p) + (1-\alpha)U(q) = W^*(g^*)$ and hence $f \succ g$.

Let x minimize $U(\delta_z)$ among z in the support of q. Then, for any $B' \subset B$ such that $\mu_b(B') > 0$ $\hat{f}(\omega_a, \omega_b) = x$ for all $\omega_b \in B'$ and $\hat{f}(\omega_a, \omega_b) = f(\omega_a, \omega_b)$ otherwise implies $W^*(\hat{f}^*) \leq (\gamma(t) - \varepsilon)U(p) + (1 - \gamma(t) + \varepsilon)U(q) < \alpha U(p) + (1-\alpha)U(q) = W^*(g^*)$. So, $g \succ \hat{f}$, contradicting Axiom 4a (Continuity). $\square$

5.5 Proof of Theorem 5

In view of Theorem 5, we need only show that $(iv) \Rightarrow (iii)$. Let $\succeq = (v, u, \mu)$ be a SPS-EU preference that satisfies (iv). Let t, t' be in the convex hull of $u(\mathcal{Z})$ and $\beta \in [0,1]$. Then, there exist lotteries $p, q \in P$ such that $u(p) = t$ and $u(q) = t'$. By nonatomicity of μ, there are acts $f, g, h \in \mathcal{F}_a$, and $B \in \mathcal{B}$ such that $p_f = p$, $p_g = p'$, $p_h = \beta p + (1-\beta)p'$ and $\mu_b(B) = \beta$. Define $h' \in \mathcal{F}$ as follows: $h'(\omega_a, \omega_b) = f(\omega_a, \omega_b)$ for all $\omega_b \in B$ and $h'(\omega_a, \omega_b) = g(\omega_a, \omega_b)$ for all $\omega_b \notin B$. Suppose (v, u, μ) is a representation of $\succeq$. Then, $W(h) = v[\beta u(p) + (1-\beta)u(p')] = v(\beta t + (1-\beta)t')$ and $W(h') = \beta v(u(p) + (1-\beta)v(u(p')) = \beta v(t) + (1-\beta)v(t')$. By condition (iv), $h \succeq h'$ and hence v is concave. Suppose $h^* = \alpha f^* + (1-\alpha)g^*$ for some $f, g, h \in \mathcal{F}$ such that $f \sim g$. Then, it follows from Theorem 3 that $W(h) = \sum_{i=1}^k v[\alpha u(p^i) + (1-\alpha)u(q^i)]\beta_i$ where $W(f) = \sum_{i=1}^k v[u(p^i)]\beta_i = W(g) = \sum_{i=1}^k v[u(q^i)]\beta_i$ for some $p^1, \ldots, p^k, q^1 \ldots, q^k$ and $\beta_i > 0$. It follows from the concavity of v that W viewed as a function of $(u(p^1), \ldots, u(p^k))$ [and hence $(u(q^1), \ldots, u(q^k))$] is concave. Hence, $W(h) \geq W(f)$ as desired.

Next, assume that $\succeq$ is a SPS-CEU preference. Let (γ, u, μ) be a representation of $\succeq$. Without loss of generality, assume $u(z^*) = 1, u(z_*) = 0$ for some $z^*, z_* \in \mathcal{Z}$. Let $\alpha \in (0,1)$ and $t, t' \in [0,1]$. Assume without loss of generality that $t \leq t'$. Choose $A \in \mathcal{A}$

and $B, B' \in \mathcal{B}$ such that $\mu_a(A) = \alpha$, $B \cap B' = \emptyset$, $\mu_b(B) = t$, $\mu_b(B') = t' - t$. Also, choose $B'' \in \mathcal{B}$ such that $\mu_b(B'') = t + \alpha(t' - t)$. Let $f(\omega_a, \omega_b) = z^*$ if ($\omega_b \in B$ or $\omega_a \in A, \omega_b \in B'$) and $f(\omega_a, \omega_b) = z_*$ otherwise. Also, let $g(\omega_a, \omega_b) = z^*$ if $\omega_b \in B''$ and $g(\omega_a, \omega_b) = z_*$ otherwise. Then, $W(f) = (1 - \alpha)\gamma(t) + \alpha\gamma(t')$ while $W(g) = \gamma(\alpha t' + (1 - \alpha)t)$. Since $g \in \mathcal{F}_b$ and $p_g = p_f$, (iv) establishes that γ is convex which implies that the capacity $\nu = \gamma \circ \mu_b$ is convex. That is:

$$\nu(B \cup B') + \nu(B \cap B') \geq \nu(B) + \nu(B')$$

Then, (iii) follows from the characterization of uncertainty aversion (the proposition on page 582) in Schmeidler (1989). $\square$

References

Allais, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21, 503–546.

Anscombe F.J. and R.J. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205

Billingsley, P. (1999): *Convergence of Probability Measures*, Wiley, New York.

Ellsberg, D. (1961): "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

Epstein, L. G., (1999): "A Definition of Ambiguity Aversion," *Review of Economic Studies*, 66, 579–608.

Epstein, L. G., and J. Zhang (2001): "Subjective Probabilities on Subjectively Unambiguous Events," *Econometrica*, 69, 265–306.

Ghirardato, P., and M. Marinacci (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251–289.

Gilboa, I. (1987): "Expected Utility with Purely Subjective Non-additive Probabilities," *Journal of Mathematical Economics*, 16, 65-88.

Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility with a Non-unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59, 667–686.

Klibanoff, P. M. Marinacci and S. Mukerji (2002): "A Smooth Model of Decision Making Under Ambiguity," mimeo, Northwestern University.

Kreps, D. M., and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

Machina, M. J. (1979): "Temporal von Neumann-Morgenstern and Induced Preference," *Journal of Economic Theory*, 20, 81–109.

Machina, M. J. and D. Schmeidler (1992): "A More Robust Definition of Subjective Probability,"*Econometrica*, 60, 745–780.

Nau, R. F. (2002): "Uncertainty Aversion with Second-Order Probabilities and Utilities," mimeo, Duke University.

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343.

Savage, L. J. (1954): The Foundations of Statistics, Wiley, New York.

Segal, U. (1987): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28, 175–202.

——— (1990): "Two-Stage Lotteries without the Reduction Axiom," *Econometrica*, 58, 349–377.

Schmeidler, D. (1989): " Subjective Probability and Expected Utility without Additivity,"*Econometrica*, 57, 571–587.

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