A Theory of Subjective Compound Lotteries[†]

Haluk Ergin

Washington University in St. Louis

and

Faruk Gul

Princeton University

September 2008

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 that *uncertainty aversion* implies aversion to second-order risk which implies issue preference and behavior consistent with *the Ellsberg paradox*.

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

1. Introduction

What distinguishes the Ellsberg paradox from *the Allais Paradox* and other related violations of subjective expected utility theory 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 understanding our interpretation and resolution of the Ellsberg Paradox. 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. We depict these bets 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 then 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 choices 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 makes the agent behave as if each of these events is 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."

The literature on the Ellsberg paradox has considered two other related interpretations of the above choices. To understand the "second-order uncertainty aversion" interpretation, consider the matrix below:

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

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. Assume that the decision-maker considers every possible column equally likely, every possible row equally likely, and the row and column events independent. Furthermore, assume that the uncertainty is revealed in two stages; first, the appropriate column is revealed, then the row. In this setting, a bet on R yields (for sure) a gamble that wins with probability $\frac{1}{3}$ while a bet on W (or G) yields a lottery over gambles: with probability $\frac{1}{4}$ the decision maker gets a gamble that wins with probability 0; with probability $\frac{1}{2}$, she gets a gamble that wins with probability $\frac{1}{3}$; and with probability $\frac{1}{4}$ she gets a gamble that wins with probability $\frac{2}{3}$. Furthermore, assume that the decision maker does not "reduce" this uncertainty and equate her chance of winning to $\frac{1}{3}$, but rather, is averse to it. That is, he would rather have a "sure" $\frac{1}{3}$ chance of winning. Such preferences would justify Ellsberg behavior.

Finally, to understand the closely related "issue (or source) preference interpretation," call the uncertainty regarding the number of the ball that is drawn issue a and the uncertainty regarding the color of the first two balls issue b, and let capital letters describe the color of the ball that is drawn. Note that a bet on R is 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 on 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: the probability of $G \cup W$ does not change once issue b is resolved, while the probability of $R \cup W$ does. Therefore, 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 (i) prefer getting \$100 if and only if R occurs getting \$100 if and only if G occurs and (ii) prefer getting \$100 if and only if G or W occurs to getting \$100 if and only if R or W occurs.

Note that both of the last two interpretations suggest the existence of two different sets of probabilities over two different collections of events (i.e., the row events and the column events above). Given the probabilities associated with the rows, by conditioning on each column, we can obtain a lottery. Then, the column probabilities yield a lottery over these lotteries. We describe decision-makers that identify each act with such a lottery over lotteries as *second-order probabilistic sophisticated*. The last two interpretations above suggest that the Ellsberg paradox is a consequence of greater aversion to the risk associated with the column events versus row events, or equivalently, aversion to second-order risk.

The goal of this paper is to provide a theory of *second-order probabilistic sophistication (SPS)*, (Theorem 1), and to relate the Ellsberg paradox and various formulations of uncertainty/ambiguity aversion to each other within that theory (Theorems 2 and 5). Our main result, Theorem 2, reveals that in general, uncertainty aversion implies second-order risk aversion which implies preference for issue a. Theorems 3 and 4 provide characterization of two special classes of SPS preferences: *SPS expected utility (SPS-EU)* and *SPS Choquet expected utility (SPS-CEU)*. We show that if the agent is a SPS-expected utility maximizer or a SPS-Choquet expected utility maximizer, then second-order risk aversion, preference for issue a, and uncertainty aversion are all equivalent (Theorem 5). However, we provide examples proving that in general, uncertainty aversion is stronger than second-order risk aversion which is stronger than preference for issue a.

1.1 Related Literature

Kreps and Porteus [16] introduce a more general form of compound lotteries, which they call temporal lotteries, to study decision-makers who have a preference for when uncertainty resolves. In their theory, temporal lotteries are taken as given and their assumptions yield expected utility preferences over temporal lotteries.

We start with purely subjective uncertainty and provide necessary and sufficient conditions for a preference relation to be second-order probabilistically sophisticated. The derived subjective prior enables us to view every act as a mapping that associates a lottery with each resolution of one issue. Hence our model provides subjectivist foundations for the framework of the Anscombe and Aumann (1963) model. We then use the same prior to reduce each Anscombe-Aumann act to a compound lottery. That is, Theorem 3 establishes Savage-type foundations for a simple version of the Kreps and Porteus [16] model.

By introducing the notion of a temporal lottery and hence abandoning the often implicit reduction of compound lotteries axiom, Kreps and Porteus develop a novel tool for the analysis of a variety of behavioral phenomena. One such phenomenon is ambiguity and ambiguity aversion. Given the amount and variety of work that has been done in this field, our attempt to relate our analysis to the relevant literature is bound to be incomplete. In particular, we focus our review of the literature on models that suggest a multi-stage resolution of uncertainty or model multiple issues. A more detailed list and discussion of ambiguity and ambiguity aversion can be found in Wakker [28].

Schmeidler [22] provides the first axiomatic model of choice geared at analyzing the Ellsberg Paradox. In that paper, he models a decision maker with preferences over Anscombe-Aumann acts who has expected utility preferences over objective (i.e., constant) acts and Choquet expected utility preferences over general acts. He introduces the notion of uncertainty aversion and relates it to the Ellsberg Paradox. Theorem 4 provides (purely) subjectivist foundations for the subclass of SPS preferences that fall in Schmeidler's model and Theorem 5 proves the equivalence of uncertainty aversion to issue preference and second order risk aversion for that subclass.

Segal [23], [24] introduces the idea of using preferences over compound lotteries (which he calls two-stage lotteries) to analyze the Ellsberg paradox and other issues related to ambiguity. In Segal [23], 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 [24] derives a subclass of the preferences studied in Segal [23] from assumptions on the preference relation over compound lotteries and investigates the relationships among various stochastic dominance and reduction (of compound lotteries) axioms.

Like Segal [23], we relate Ellsberg paradox type behavior to a form of second-order risk aversion (Theorem 2). However, we use a different notion of second-order risk. His notion is applicable to binary acts (i.e., acts 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.

Our analysis of second-order risk aversion is closely related to Grant, Kajii and Polak [12]'s analysis of information aversion (or information loving). The Grant, Kajii and Polak [12] notion of an elementary linear bifurcation is equivalent to our definition of a mean-preserving spread. Hence, their notion of single-action information aversion corresponds to our second-order risk aversion. In Grant, Kajii and Polak [12], the set of all two-stage lotteries is taken as primitive. Nevertheless, once we establish second-order probabilistic sophistication, we can easily compare their results to ours. In particular, parts of their Proposition 1 are equivalent to parts of our Theorem 2.

Klibanoff, Marinacci and Mukerji [15] also have a model that describes the resolution of uncertainty in two stages. Some interpretational and modeling differences notwithstanding, their utility function is formally analogous to the one describing SPS-EU preferences (Theorem 3). The most significant interpretational difference between our model and that of Klibanoff, Marinacci and Mukerji is that they permit both the possibility that the first stage of uncertainty might be identified with a particular observable and the possibility that the first stage uncertainty is simply a part of the individual's preferences that can be pinned-down only with "cognitive data or thought experiments."

The notion of ambiguity aversion used in Klibanoff, Marinacci and Mukerji [15] is formally similar to our concept of second-order risk aversion. Their definition entails comparing arbitrary second-order 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 (and hence the setting of Theorem 3), 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.

Seo [25] considers preferences over the following choice objects: a lottery, P, is a probability distribution over acts f which map states to probability distributions, p, over prizes x. Hence, this is Anscombe-Aumann's original framework. Seo characterizes preferences that can be represented by an expected utility function, U, such that its von Neumann-Morgenstern utility, u, is analogous to the SPS-EU utility function of our Theorem 3. The richer setting permits Seo to derive this representation from an ingenious dominance assumption.

Tversky and Fox [26] introduce the notion of source (i.e., issue preference) and provide evidence showing that, ceteris paribus, basketball fans prefer gambles on the outcomes of basketball games to objective gambles. Abdellaoui, Baillon, and Wakker [1] expand on the Tversky and Fox approach and provide more evidence of issue preference. They show that decision makers prefer betting on the temperature in their own city (Paris) to betting on the temperature in foreign cities.

Nau [19] provides a theoretical model with two issues. Nau assumes a finite state space $\Omega = \Omega_a \times \Omega_b$, provides axioms that 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. He also derives the state-independent representation by imposing Wakker [27]'s trade-off consistency axiom. Unlike Nau, we do not allow for state dependent preferences. However, we permit general, nonseparable (i.e., nonexpected utility) preferences and probabilistic sophistication.

2. Second-Order Probabilistic Sophistication

Let $\mathcal{Z}$ be the 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 $f : \Omega' \to \mathcal{Z}'$ 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 $\mathcal{Z}$. A binary relation $\succeq$ on $\mathcal{F}$ characterizes an individual. 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 act. Our second assumption ensures that the individual is not indifferent among all constant acts.

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 acts 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 acts 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. Hence, this axiom ensures that the ordinal ranking of prizes is state independent.

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 all $f, g \in \mathcal{F}$ and $z \in \mathcal{Z}$, if $f \succ g$ then, there exist a partition $E_1, \ldots, E_n \in \mathcal{E}_a$ of and a partition $F_1, \ldots, F_n \in \mathcal{E}_b$ of Ω such that*

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

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

Most models that study acts (i.e., the Savage setting) impose the assumptions above.[1] These models differ in 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:

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

Definition ($\mathcal{F}_a$, $\mathcal{F}_b$): *Let $\mathcal{F}_a$ and $\mathcal{F}_b$ denote the set of $\mathcal{E}_a$–measurable and $\mathcal{E}_b$–measurable acts respectively.*

For $f \in \mathcal{F}_a$ and $g \in \mathcal{F}_b$ we write

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

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

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

[1] 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 $5a$ 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 $A_2 \times \Omega_b$ and agrees with 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 new 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

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

Henceforth, we refer to Axioms $5a$ and $5b$ together as Axiom 5. 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: suppose

$$f = \begin{pmatrix} 100 & A_1 \\ 1 & A_2 \end{pmatrix} \succ \begin{pmatrix} 0 & A_1 \\ 100 & A_2 \end{pmatrix} = g \text{ and}$$
$$\hat{f} = \begin{pmatrix} 1 & A_1 \\ 100 & A_2 \end{pmatrix} \succ \begin{pmatrix} 100 & A_1 \\ 0 & A_2 \end{pmatrix} = \hat{g}$$

suggesting that both $\mu_a(A_1)$ and $\mu_a(A_2)$ are close to $\frac{1}{2}$. Then, Axiom 5b implies

$$f' = \begin{pmatrix} 100 & 0 & -500 & A_1 \\ 1 & 100 & 1000 & A_2 \\ B_1 & B_2 & B_3 & * \end{pmatrix} \succ \begin{pmatrix} 0 & 100 & -500 & A_1 \\ 100 & 1 & 1000 & A_2 \\ B_1 & B_2 & B_3 & * \end{pmatrix} = g'$$

iff

$$f'' = \begin{pmatrix} 1 & 100 & -500 & A_1 \\ 100 & 0 & 1000 & A_2 \\ B_1 & B_2 & B_3 & * \end{pmatrix} \succ \begin{pmatrix} 100 & 1 & -500 & A_1 \\ 0 & 100 & 1000 & A_2 \\ B_1 & B_2 & B_3 & * \end{pmatrix} = g''$$

But, the preferences for f over g and $\hat{f}$ over $\hat{g}$ rely on the fact that the rows of f and g are (almost) equally likely and $f' \succ g'$ and $f'' \succ g''$ require that conditional on each column, these rows remain (almost) 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, suppose that B_1 is more likely than B_2. Then, the first row of g' is worse than the first row of f' but the second row of g' is better than the second row of f'. But, assuming $\mu_b(B_3)$ is not very small, this means that the rows of f' yield closer payoffs than those of g', while the rows of g'' yield closer payoffs than those of f''. Hence, f' has less (second-order) issue a risk than g', while f'' has more second order issue a risk than g''. By demanding that $f' \succ g'$ implies $f'' \succ g''$, Axiom 5b precludes aversion to such risk. It is this feature of Axiom 5b that permits the translation of acts into compound lotteries where second-order risk is only associated with issue b.

Let $\Delta(\mathcal{Z}') = \{p : \mathcal{Z}' \to [0,1] : p^{-1}((0,1]) \text{ is finite and } \sum_z p(z) = 1\}$ be the set of all simple lotteries on some set $\mathcal{Z}'$. Let δ_z denote the degenerate lottery that yields z with probability 1.

Definition: *The function* $\phi : \Delta(\mathcal{Z}') \to I\!R$ *satisfies stochastic dominance*[2] *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 (μ, μ_a, μ_b): *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* $\mu(E \cup E') = \mu(E) + \mu(E')$. *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 $B \in \mathcal{B}$. The next definition describes how to associate an Anscombe-Aumann act f^* with each Savage act $f \in \mathcal{F}$ whenever $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent. Since the underlying μ is clear we suppress the dependence on μ in the definitions below.

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

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

Also, for $f \in \mathcal{F}$ *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 $f^* : \Omega_b \to P$ *as follows:*

$$p_j(z) = \sum_{i : x_{ij} = z} \mu_a(A_i)$$

$$f^*(s) = p_j \text{ for all } s \in B_j$$

and let $\mathcal{F}^*$ *be the set of all simple functions from* Ω_b *to* P.

[2] When Z' itself is a set of lotteries, this condition is sometimes called *compound independence* (Segal [24]) or *recursivity* (Grant, Kajii and Polak [12]).

Next, we describe how to associate a compound lottery with each Anscombe-Aumann act. Note that $\mathcal{F}^*$ is a mixture space (with the usual) Anscombe-Aumann mixture operation which defines $\alpha f^* + (1-\alpha)g^*$ as the act that yields in state s, the von Neumann-Morgenstern mixture $\alpha f^*(s) + (1-\alpha)g^*(s)$.

Definition (π_f): *Let μ be a probability on $\mathcal{E}$ and let $\mathcal{E}_a, \mathcal{E}_b$ be $\mu-$independent. For $f \in \mathcal{F}$, 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 $\alpha \in [0,1]$ implies there exists $E' \in \mathcal{E}_a$ such that $E' \subset E$ and $\mu(E') = \alpha\mu(E)$. 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 $\mathcal{E}_a, \mathcal{E}_b$. 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 $\pi_f = \pi$.

We endow P with the supremum metric and $\Delta(P)$ with the Prohorov metric.[3] That is, let $d^\infty(p,p') = \sup_{z \in Z}|p(z) - p'(z)|$ for $p, p' \in P$ and for any finite set $T \subset P$, let $T^\epsilon = \{p' \in P : \min_{p \in T} d^\infty(p', p) < \epsilon\}$. Then, let $d_\Delta(\pi, \pi')$ be the infimum of ϵ that satisfy

$$\pi(T) \leq \pi'(T^\epsilon) + \epsilon \text{ and } \pi'(T) \leq \pi(T^\epsilon) + \epsilon$$

for all finite $T \subset P$.[4] For any compound lottery π, define $\operatorname{supp}_{\mathcal{Z}} \pi$ to be the union of all z in the support of p over p in the support of π. That is, $\operatorname{supp}_{\mathcal{Z}} \pi = \bigcup_{p \in \operatorname{supp} \pi} \operatorname{supp} p$.

Definition: *A function $W : \Delta(P) \to I\!R$ is weakly continuous if π_n converges to π and $\bigcup_{n=1}^\infty \operatorname{supp}_{\mathcal{Z}} \pi_n$ is finite implies* $\lim W(\pi_n) = W(\pi)$.[5]

[3] The Prohorov metric is usually defined more generally on the set of all probability measures on the Borel sets of a metric space. Here, we are considering the subspace, $\Delta(P)$, of that space.

[4] In the standard definition of the Prohorov metric (see for example Billingsley [4]), the infimum is taken over ϵ such that the above inequalities are satisfied for all Borel measurable $T \subset P$. Note that the two definitions coincide for simple probability measures.

[5] Since we require $\bigcup_{n=1}^\infty \operatorname{supp}_{\mathcal{Z}} \pi_n$ to be finite, our results would not change if we endow P with any metric topology such that the relative topology of $\Delta(\mathcal{Z}')$ in P coincides with the Euclidean topology of $\Delta(\mathcal{Z}')$, for all finite $\mathcal{Z}' \subset \mathcal{Z}$. For example, our results go not change if we replace the supremum metric with the metric $d^\alpha(p,q) = \sum_{z \in \mathcal{Z}} |p(z) - q(z)|^\alpha$ for all $p, q \in P$, where $\alpha > 0$.

It is well known that the Prohorov metric topology and the topology of weak convergence are equivalent when P is a separable metric space. Therefore, for finite $\mathcal{Z}$, our weak continuity indeed becomes continuity in the topology of weak convergence.

Machina and Schmeidler call a decision-maker probabilistically sophisticated if he has a subjective prior over the state-space, considers all acts that yield the same lottery equivalent and satisfies stochastic dominance. We refer to preferences satisfying Axioms $1-5$ as *second-order probabilistically sophisticated (SPS)* preferences. Theorem 1 shows decision-makers with SPS preferences are indifferent between f and g whenever f and g yield the same compound lottery; that is, $f \sim g$ whenever $\pi_f = \pi_g$. But, they need not be indifferent between f and g if $p_f = p_g$ because these decision-makers also care about second-order risk.

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, nonconstant $W : \Delta(P) \to I\!R$ 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.*

Familiar arguments ensure that the μ of Theorem 1 is unique and W is unique up to a continuous monotone transformation. The details of the proof of Theorem 1, which rely on Theorem 2 of Machina and Schmeidler [17], are in the appendix. Here, we give a sketch. The main idea is to apply Machina and Schmeidler's proof of probabilistic sophistication twice, first to $\mathcal{F}_a$, then to $\mathcal{F}^*$. 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 [17] 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. Next, we introduce the binary relation, $\succeq^*$, on $\mathcal{F}^*$ the set of all Anscombe-Aumann acts.

Definition ($\succeq^*$): *Let $f^* \succeq^* g^*$ if and only if $f \succeq g$.*

In the proof of Theorem 1, we show that $\succeq^*$ is well-defined; that is, $f^* = g^*$ implies $f \sim g$. 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 $W(\pi^*_{f^*}) \geq W(\pi^*_{g^*})$ 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.

To see how we can derive a subjective version of the Anscombe-Aumann model for our framework and use it to relate Axiom 5b to existing comparative probability and replacement axioms, we 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 ($\mathcal{F}_a^0$, χ): *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 applied to acts on Ω_b with prizes in $\mathcal{F}_a^0$. To clarify this symmetry between Axioms 5a and 5b, we re-state Axiom 5b as follows:

Axiom 5b: *Suppose $f, \hat{f}, g, \hat{g}, h^1, \hat{h}^1, \ldots h^m, \hat{h}^m \in \mathcal{F}_a$ and*

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

Then,

$$\begin{pmatrix} f_a & g_a & h_a^1 & \ldots & h_a^m \\ B_1 & B_2 & B_3 & \ldots & B_m \end{pmatrix} \succ \begin{pmatrix} g_a & f_a & h_a^1 & \ldots & h_a^m \\ B_1 & B_2 & B_3 & \ldots & B_m \end{pmatrix}$$

iff

$$\begin{pmatrix} \hat{f}_a & \hat{g}_a & \hat{h}_a^1 & \dots & \hat{h}_a^m \\ B_1 & B_2 & B_3 & \dots & B_m \end{pmatrix} \succ \begin{pmatrix} \hat{g}_a & \hat{f}_a & \hat{h}_a^1 & \dots & \hat{h}_a^m \\ B_1 & B_2 & B_3 & \dots & B_m \end{pmatrix}$$

Hence, we impose the Machina-Schmeidler comparative probability assumption on $\mathcal{E}_a$–measurable acts with prizes $\mathcal{Z}$ and on $\mathcal{E}_b$–measurable acts with prizes in $\mathcal{F}_a^0$ but never on arbitrary $\mathcal{E}$–measurable acts. Therefore, in one sense, Axiom 5 is stronger than the corresponding assumption in Machina and Schmeidler [17], it is applied to a richer set of prizes. On the other hand, there is a sense in which Axiom 5 is weaker; it is applied to a smaller set of acts.[6]

The Machina-Schmeidler comparative probability assumption applied to all acts in $\mathcal{F}$ would not yield a μ that is a product measure while our assumptions do not ensure that acts can be reduced to lotteries; only that they can be reduced to compound lotteries. More specifically, our assumptions enable us to identify Savage acts on $\Omega_a \times \Omega_b$ with Anscombe-Aumann acts on Ω_b. The implied preferences on these Anscombe-Aumann acts are weaker than the preferences considered in Machina and Schmeidler [18] and Grant and Polak [13]). These two papers define preferences over Anscombe-Aumann acts and impose assumptions (a replacement axiom in the former paper and two axioms, betting neutrality and two-outcome independence in the latter) that our model does not satisfy. By circumventing the calibration of subjective probabilities with objective probabilities, or equivalently, the $b-$probabilities with $a-$probabilities that the additional Machina and Schmeidler [18] or Grant and Polak [13] assumptions enable, our model permits issue preference and related violations of probabilistic sophistication such as the Ellsberg paradox.

An alternative approach to the one we have taken here would be to impose more structure on $\mathcal{Z}$, for example, assume that it is an interval of real numbers, and enough continuity to ensure that conditional certainty equivalents are well defined. Then, we could strengthen Axiom 3 to cover a richer set of prizes such as the f_a's in the second version of Axiom 5b. Given this stronger version of monotonicity, we could weaken Axiom 5b to apply only to acts in $\mathcal{F}_b$ (i.e., make it just like 5a) and still obtain the representation of Theorem 1.

[6] In fact, it is not difficult to see that an SPS preference satisfies the comparative probability axiom on all acts in $\mathcal{F}$ if and only if it is an expected utility preference. Therefore, the intersection of the current model with Machina Schmeidler [17] is the subjective expected utility model with independent μ.

3. Second-Order Risk Aversion and the Ellsberg Paradox

In this section, we provide measures of second-order risk, relate second-order risk aversion to issue preference 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.

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 issue b (i.e., the source of second-order risk), like the assignment of probabilities, is a subjective matter.

Regardless of which issue is issue a and which issue is issue b, the compound lottery associated with act h is $\delta_{\frac{1}{2}\delta_1 + \frac{1}{2}\delta_0}$. Now, to verify which issue is the one associated with second-order risk, 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.

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 $f \succeq g \succeq h$.

The equivalence of (i) and (ii) is the most difficult argument in the proof of Theorem 2 and is closely related to Proposition 1(i) in Grant, Kajii and Polak [12]. 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. Condition (ii) applies only to f and g that yield the same compound lottery. By Theorem 1, such acts are indifferent. Hence, (iii) implies (ii). 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, the decision-maker will prefer f to g and g to h. We call this condition issue preference. It is not difficult to verify that if $f \in \mathcal{F}_a$, $g \in \mathcal{F}$ and $h \in \mathcal{F}_b$ such that $p_f = p_g = p_h$, then there is a finite sequence of compound lotteries $\pi_1, \ldots, \pi_n$ such that $\pi_1 = \pi_f$, $\pi_j = \pi_g$ for some j, $\pi_n = \pi_h$, and π_{i+1} is a mean preserving spread of π_i for all $i = 1, \ldots, n-1$. Hence, condition (ii) implies condition (iv).

In the appendix, after the proof Theorem 2, we provide counter-examples to (iv) implies (ii) and to (ii) implies (iii). Hence, a result stronger than Theorem 2 cannot be proved. To see why (iv) does not imply (ii), note that (iv) provides no criterion for comparing f and g when neither is an element of $\mathcal{F}_a \cup \mathcal{F}_b$ even if π_g is a mean-preserving spread of π_f. To get intuition as to why (ii) does not imply (iii), consider the simplest case with only two prizes and two equally likely states in Ω_b. Then, each act f^* can be identified with a point in $[0,1]^2$. In this setting, (iii) is the statement that the implied preferences on $[0,1]^2$ are quasiconcave while (ii) is the statement that $\alpha(a_1, a_2)+(1-\alpha)(a_2, a_1) \succeq^* (a_1, a_2)$ for all $(a_1, a_2) \in [0,1]^2$.

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 [22].

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 [16]. 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 permits conditioning on events $E \notin \mathcal{E}_a \cup \mathcal{E}_b$.

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,*

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

Nau [19] deals with a finite state space that also has a product structure; that is, he too has two issues. His Axiom 2 is analogous to Axiom 6(b) above and his Axiom (3) is a stronger version of Axiom 6(a). Then, he imposes additional axioms to derive a version of Theorem 3 that does not require the independence of the two issues.

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.[7,8]

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) W is given by*

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

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

We refer to preferences that satisfy the hypothesis of Theorem 3 as SPS expected utility (SPS-EU) preferences. We call the corresponding (v, u, μ) a representation of $\succeq$.

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

[8] Machina and Schmeidler's comparative probability axiom, which is analogous to our Axiom 5 is stronger than Savage's 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.

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 [16].

The Ellsberg paradox is often studied within the framework of Choquet expected utility or maxmin expected utility preferences. To define a Choquet expected utility preference, we first state the definition of the Choquet integral in the Anscombe-Aumann setting: Recall that $\mathcal{F}^*$ denotes the set of all simple functions from Ω_b to P.

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 $B' \subset B$. Given any capacity ν, for any real-valued, $\mathcal{E}_b$–measurable simple 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 an expected utility function $U : P \to I\!R$ such that the function W^* defined below represents $\succeq$.

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

Schmeidler's axiomatization of Choquet expected utility relies on the comonotonic independence axiom. Gilboa [10] 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:*

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

These acts are comonotonic if for all j, k

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

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 f. Then*

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

Since all acts in $\mathcal{F}_a$ are comonotonic, Axiom $6c$ implies Axiom $6a$. Of course, Axiom $6c$ does not imply Axiom $6b$ because $6c$ only permits conditioning on events in $\mathcal{E}_a$ and is restricted to comonotonic acts. Similarly, Axiom $6a$ does not imply $6c$ because $6a$ only applies to acts in $\mathcal{F}_a$ while $6c$ applies to arbitrary comonotonic acts in $\mathcal{F}$. In fact, the combination of $6a$ and $6b$ do not imply $6c$ as verified by observing that the preferences characterized by Theorem 3 and 4 are not nested.

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 an SPS 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 $\nu = \gamma \circ \mu_b$. 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 there exists (i) a nonatomic probability μ on $\mathcal{E}$ such that $\mathcal{E}_a, \mathcal{E}_b$ are $\mu-$independent, (ii) $W^* : \mathcal{F}^* \to I\!R$ such that $f \succeq g$ if and only if $W^*(f^*) \geq W^*(g^*)$, and (iii) a nonconstant $u : \mathcal{Z} \to I\!R$ and a strictly increasing and continuous bijection $\gamma : [0,1] \to [0,1]$ such that W^* is defined by the Choquet integral*

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

where $U(p) = \sum_{z \in \mathcal{Z}} u(z)p(z)$ *and* $\nu = \gamma \circ \mu_b$.

Our final result provides a stronger characterization of second-order risk aversion for SPS-EU and SPS-CEU preferences. The theorem establishes that for such preferences, second-order risk aversion is equivalent to issue preference and to Schmeidler's notion of uncertainty aversion provided we identify Schmeidler's objective lotteries with issue a. It follows that all of the characterizations of uncertainty aversion provided by Schmeidler are also characterizations second-order risk aversion.[9]

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* $f \succeq g \succeq h$.

Note that by Theorem 2, to prove Theorem 5, it is enough to show that (iv) implies (iii). Consider the following weakening of condition (iv):

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

For SPS-EU preferences, (v) is equivalent to (iv). This fact is formally related to the observation that for expected utility preferences, defining risk aversion by comparing lotteries to their means versus by comparing lotteries to their mean preserving spreads leads to identical formulations of risk aversion and comparative risk aversion. Furthermore, it is known that (v) implies (i) (see for example Grant, Kajii and Polak [12]). Klibanoff, Marinacci and Mukerji (2005) define ambiguity aversion as (v) applied to binary acts. For SPS-EU preferences, this ensures that the condition holds for all acts.[10] For SPS-CEU

[9] 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.

[10] 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.

preferences, (v) is not equivalent to (iv): while (iv) is equivalent to the convexity of γ, (v) is equivalent to $\gamma(t) \leq t$ for all $t \in [0,1]$.[11]

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 [6] and Epstein and Zhang [7] differs from the one in Ghirardato and Marinacci [9] with respect to the underlying notion of ambiguity neutrality. Epstein [6] identifies being ambiguity neutral with probabilistic sophistication while Ghirardato and Marinacci [9] define ambiguity neutrality as expected utility maximization. Ghirardato and Marinacci [9] 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 [19] and in our model, the emphasis is on the agent having different preferences on uncertain acts that depend on separate issues. Nau defines the agent's preferred issue as the unambiguous one. Like Ghirardato and Marinacci, Nau uses his model to provide a unified framework analyzing state dependent preferences, the Ellsberg paradox and the Allais' paradox. Like Epstein [6] and Epstein and Zhang [7], we have attempted to identify our central concept (issue preference or equivalently, second-order risk aversion) exclusively with the Ellsberg paradox.

[11] We are grateful to the associate editor for this last observation.

5. Appendix

5.1 Proof of Theorem 1

We start by proving a useful result about the Prohorov metric on $\Delta(P)$.

Lemma 0: *Suppose that μ_b is a nonatomic probability measure on $\mathcal{E}_b$. Let $\pi, \pi' \in \Delta(P)$, $f^* \in \mathcal{F}^*$, and $\pi = \pi_{f^*}$. Then, $d_\Delta(\pi, \pi')$ is the infimum of $\epsilon > 0$ for which there exist $g^* \in \mathcal{F}^*$ such that $\pi_{g^*} = \pi'$ and*

$$\mu_b(\{s \in \Omega_b : d^\infty(f^*(s), g^*(s)) \geq \epsilon\}) \leq \epsilon \qquad (1)$$

Proof: It is easy to see that if ϵ satisfies Equation (1) for some $g^* \in \mathcal{F}^*$ with $\pi_{g^*} = \pi'$, then $d_\Delta(\pi, \pi') \leq \epsilon$. To see the other direction, suppose that $d_\Delta(\pi, \pi') < \epsilon$. Let $\mathcal{O}_\epsilon(p)$ denote the open ball around p with radius ϵ and $\mathcal{D} = \text{supp}(\pi)$. Let $\mathcal{S}$ denote the partition of $\mathcal{D}^\epsilon$ generated by $\{\mathcal{O}_\epsilon(p) : p \in \mathcal{D}\}$. Consider a supply-demand network where the demand nodes are lotteries in $\mathcal{D}$, and the supply nodes are cells of the partition $\mathcal{S}$ and $\emptyset$. Each $p \in \mathcal{D}$ has demand $\pi(p)$, each $Q \in \mathcal{S}$ has supply $\pi'(Q)$, and $\emptyset$ has supply ϵ. The supplier $\emptyset$ is connected to all demand nodes. The supplier Q is connected to p if and only if $Q \subset \mathcal{O}_\epsilon(p)$, in which case we write $Q \to p$. Note that for all $T \subset \mathcal{D}$

$$\sum_{p \in T} \pi(p) = \pi(T) \leq \pi'(T^\epsilon) + \epsilon = \sum_{Q \,:\, Q \to p \text{ for some } p \in T} \pi'(Q) + \epsilon.$$

Therefore, Gale [8]'s Feasibility theorem implies that there exist flows $\lambda_{Qp}, \lambda_{\epsilon p} \geq 0$ for all $p \in \mathcal{D}$ and $Q \to p$ such that: (i) $\forall p \in \mathcal{D} : \sum_{Q \,:\, Q \to p} \lambda_{Qp} + \lambda_{\epsilon p} \geq \pi(p)$, (ii) $\forall Q \in \mathcal{S}$: $\sum_{p \,:\, Q \to p} \lambda_{Qp} \leq \pi'(Q)$, and (iii) $\sum_{p \in \mathcal{D}} \lambda_{\epsilon p} \leq \epsilon$. For each $p \in \mathcal{D}$, by (i) and nonatomicity of μ_b, the event ${f^*}^{-1}(p)$ can be partitioned into events $\{B_{Qp}\}_{Q \,:\, Q \to p} \cup \{B_{\epsilon p}\}$ such that $\mu_b(B_{Qp}) \leq \lambda_{Qp}$ and $\mu_b(B_{\epsilon p}) \leq \lambda_{\epsilon p}$. By (ii) and nonatomicity of μ_b, there exists $g^* \in \mathcal{F}^*$ such that $\pi_{g^*} = \pi'$ and $g^*(s) \in Q$ for all $Q \in \mathcal{S}$ and $s \in \bigcup_{p \,:\, Q \to p} B_{Qp}$. By (iii), Equation (1) is satisfied. $\square$

To see the necessity of the axioms, suppose that the μ and W that the theorem specifies exist. Then, Axioms 1, 3, and 5 are obviously satisfied. Note that if all constant acts were indifferent, then the fact that w and W satisfy stochastic dominance would imply that both w, W are constant functions, a contradiction. Hence, Axiom 2 is also satisfied. To see that Axiom 4 is satisfied, consider any f, g such that $W(\pi_f) > W(\pi_g)$ and let z be any element of Z. Without loss of generality that we can express f as:

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

Define the partition $\mathcal{P}_a^k$ by dividing Ω into k equiprobable events $E_1^k, \dots, E_k^k \in \mathcal{E}_a$. (This is possible since μ_a is nonatomic.)

Let f_E be the act obtained from f by replacing the outcome of f with z at $s \in E \in \mathcal{E}$. Consider the sequence $(f_n) = (f_{E_1^1}, f_{E_1^2}, f_{E_2^2}, f_{E_1^3}, f_{E_2^3}, f_{E_3^3}, \dots)$. Since f^* and $f^*_{E_l^k}$ satisfy Equation (1) for $\epsilon = {}^1\!/_k$, by Lemma 0, π_{f_n} converges to π_f. Since $\bigcup_{n=1}^{\infty} \operatorname{supp}_{\mathcal{Z}} \pi_{f_n} = \operatorname{supp}_{\mathcal{Z}} \pi_f \cup \{z\}$ and W is weakly continuous, for some k large enough, $W(\pi_{f_E}) > W(\pi_g)$ for all $E \in \mathcal{P}_a^k$. Hence, $f_E \succ g$. A symmetric argument establishes that for k large enough, $f \succ g_E$ for all $E \in \mathcal{P}_a^k$, proving Axiom $4a$. Dividing Ω into $k-$equiprobable $\mathcal{E}_b$–measurable events and repeating the same argument proves $4b$.

Next, assume that $\succeq$ satisfies Axioms $1-5$. Then $\succeq|_{\mathcal{F}_a}-$the restriction of $\succeq$ to $\mathcal{E}_a$–measurable acts, satisfies the Machina-Schmeidler axioms. Therefore by Theorem 2 of Machina and Schmeidler [17], 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 $f, g \in \mathcal{F}_a$. 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, there exist $x, y \in \mathcal{Z}$ be such that $x \succ y$. By Axiom 3,

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

Applying to the partition $\{B_1, \emptyset, B_2, \dots, B_m\}$, Axiom 5b 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

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

Let $\underline{z}$ be a $\succeq$-worst and $\overline{z}$ a $\succeq$-best outcome in $\{x_i : i = 1, \ldots, n\} \cup \{y_i : i = 1, \ldots, n\} \cup \{z_{ij} : i = 1, \ldots, n, j = 2, \ldots, m\}$. Then, by Axioms 2 and 3, we can find z^* such that

$$z^* \succ \begin{pmatrix} x_1 & A_1 \\ \vdots & \vdots \\ x_n & A_n \end{pmatrix} \sim \begin{pmatrix} y_1 & A_1 \\ \vdots & \vdots \\ y_n & A_n \end{pmatrix} \text{ and } z^* \succeq \overline{z}$$

or z_* such that

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

Suppose we have a z^* 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 $4a$, 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 $\mu_a(C_i \cap A_1) > 0$. Since $z^* \succeq \overline{z}$, by iterated application of Axiom 3, 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 $(*)$. $\square$

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. $\square$

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 and Axiom 4, $\succeq^*$ satisfies continuity on Ω_b. Finally, by Axiom 5b, $\succeq^*$ also satisfies Strong Comparative Probability. Therefore, we can apply Theorem 2 of Machina and Schmeidler [17] once again, in order to obtain a nonatomic probability measure μ_b on $(\Omega_b, \mathcal{B})$ and a function $W\colon \Delta(P) \to I\!R$ such that (i) W is mixture continuous and monotonic with respect to first order stochastic dominance and (ii) $W(\pi^*_{f^*}) \geq W(\pi^*_{g^*})$ iff $f^* \succeq g^*$ for all $f, g \in \mathcal{F}$ where $\pi^*_{h^*} \in \Delta(P)$ is defined by $\pi^*_{h^*}(p) = \mu_b(h^{*-1}(p)))$. Note also that the proof in Machina and Schmeidler [17] yields W such that $\{W(\pi^*) : \pi^* \in \Delta(P)\}$ is convex.

Since $\pi_f = \pi^*_{f^*}$ for any $f \in \mathcal{F}$, we have that

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

establishing that W represents $\succeq$. Define w: $P \to I\!R$ 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 the sequence π_t converges to π and $\bigcup_{t=1}^{\infty} \mathrm{supp}_{\mathcal{Z}}\, \pi_t$ is finite. Suppose that $W(\pi) > \liminf W(\pi_t)$. Note that Theorem 2 in Machina and Schmeidler [17] ensures that W has a convex range. Therefore, there exists π' such that $W(\pi) > W(\pi') > \liminf W(\pi_t)$. By nonatomicity of μ, there exists f' such that $\pi_{f'} = \pi'$. We can assume without loss of generality that f can be expressed 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 $(\bigcup_{t=1}^{\infty} \operatorname{supp}_{\mathcal{Z}} \pi_t) \bigcup \operatorname{supp}_{\mathcal{Z}} \pi$. Now since $f \succ f'$, by Axiom 4a, there is a partition $C_1, \ldots, C_I$ of Ω_a such that for any $i = 1, \ldots, I$:

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

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

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

where the weak preference above follows from iterated application of Axiom 3.

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

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

(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, there is a partition $D_1, \ldots, D_J$ of Ω_b such that for any $j = 1, \ldots, J$:

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

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

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

where again the weak preference above follows from iterated application of Axiom 3.

Let $\epsilon = \min\{\mu_a(C)/n, \mu_b(D)/m\} > 0$ and let t be such that $d_\Delta(\pi, \pi_t) < \epsilon$. By Lemma 0, there exists $f_t^* \in \mathcal{F}^*$ such that $\pi_t = \pi_{f_t^*}$ and $\mu_b(\{s \in \Omega_b : d^\infty(f_t^*(s), f^*(s)) \geq \epsilon\}) < \epsilon$. By nonatomicity of μ_a, there is $\bar{f}_t \in \mathcal{F}$ and events $E_1, \dots, E_n \in \mathcal{A}$; $F_1, \dots, F_m \in \mathcal{B}$ such that $\bar{f}_t^* = f_t^*$, $E_i \subset A_i$, $F_j \subset B_j$, $\mu_a(E_i) < \epsilon$, $\mu_b(F_j) < \epsilon$ and $\bar{f}_t$ gives x_{ij} on $A_i \setminus E_i \times B_j \setminus F_j$ for $i = 1, \dots, n$ and $j = 1, \dots, m$.

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

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

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

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

Finally, by iterated application of Axiom 3, we have that $\bar{f}_t \succeq g_t$. Therefore, $f_t \sim \bar{f}_t \succeq g_t \succeq h_t \succeq h \succ f'$, implying that $W(\pi_t) > W(\pi')$ for all $t \geq N$, a contradiction to $W(\pi') > \liminf W(\pi_t)$. Therefore, $W(\pi) \leq \liminf W(\pi_t)$. A symmetric argument shows that $\limsup W(\pi_t) \leq W(\pi)$, implying that $W(\pi) = \lim W(\pi_t)$. □

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 $1 \leq k \leq m$ 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 $\pi^m = \pi_h$. To conclude the proof we will show 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). We will 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 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), $h_m^{n+1} \succeq h_{2m-1}^n$. Therefore, by transitivity, $h_1^n \succeq h_1^0$ 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 $n \geq 0$.

Next, we show that $\pi_{h_1^n}$ converges to π_f. Let $\epsilon > 0$ be given. Define the probability measure μ^* on $\mathcal{B}$ by $\mu^*(B') = \frac{\mu_b(B' \cap B)}{\mu_b(B)}$. Also, define $T^k(s) = 1$ if $s \in B_l^k$ for some even l

and $T^k(s) = 0$ otherwise. 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[12] (applied to T^k on the probability space $(\Omega_b, \mathcal{B}, \mu^*)$), 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 $n \geq N$:

$$\mu_b(\{s \in \Omega_b : |\sum_{i=0}^{2^n} T^i(s)/2^n - \beta| \geq \epsilon'\}) < \epsilon.$$

Then, for any $n \geq N$, $\mu_b\left(\{s \in \Omega_b : d^\infty\left((h_1^n)^*(s), f^*(s)\right) \geq \epsilon\}\right) < \epsilon$, which, by Lemma 0, implies that $d_\Delta(\pi_{h_1^n}, \pi_f) \leq \epsilon$. Therefore, $\pi_{h_1^n}$ 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. $\square$

Below, we provide counter-examples to (iv) implies (i) and (i) implies (iii). 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 mean 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 $\psi(\alpha) = \frac{log(1+\alpha)}{4}$. Define,

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

[12] Despite the fact that μ and hence μ^* are finitely additive, the Weak Law of Large Numbers still applies since Bernoulli random variables are simple.

The weak continuity of W is easy to verify. It is straightforward to check that for any $y > x$, a small increase in $\pi(y)$ at the expense of $\pi(x)$ cannot increase $m(\eta_\pi)$ at a rate greater than $2(y-x)$. To see this, note that such an increase will have two kinds of effects. First, it could change $m(\pi)$ and second it will increase the weight on the term $|y-m(\pi)|$ at the expense of the weight on the term $|x-m(\pi)|$. Each of these effects can increase $m(\eta_\pi)$ at a rate no more than $y-x$ and hence $\psi(m(\eta_\pi))$ cannot increase at a rate greater than $(y-x)/2$. On the other hand, such an increase in $\pi(y)$ increases $m(\pi)$ at a rate $y-x$ and therefore, the overall effect is an increase in $W(\pi)$. It follows that W satisfies stochastic dominance. Since $w(x) = W(\delta_x) = x$, 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) increase $m(\eta_\pi)$.

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 $\pi_g = \pi$. Then, straightforward calculations yield $\pi_{.5f^*+.5g^*} = .5\delta_{.5w+.5} + .5\delta_{.5w}$ and $m(\eta_{\pi_{.5f^*+.5g^*}}) = .5m(\eta_\pi) + .5m(\eta_{\delta_w}) = .5m(\eta_\pi)$. Hence, the strict concavity of ψ ensures that $W(\pi_{.5f^*+.5g^*}) < .5W(\pi) + .5W(\delta_w) = w$ 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 [14].[13] In that paper, it is shown 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:

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

The function V described in $(*)$ is neither risk averse nor risk loving and since the rest of W is linear, this means that W is neither risk averse nor risk loving. This feature

[13] In particular, this preferences is a disappointment averse preference with linear u and $\beta = 1$.

of W permits the preferences that W represents to fail (ii) but still satisfy (iv). Again, it can be verified that W is weakly continuous and that for $y > x$ a small increase in $\pi(y)$ at the expense of $\pi(x)$ has two effects. First, it can increase $V(\eta_\pi)$ by changing $m(\pi)$ and increasing each $|z - m(\pi)|$. This effect is at most $y - x$. The second effect is the one associated with increasing the weight of $|y - m(\pi)|$ and decreasing the weight of $|x - m(\pi)|$ (while keeping $m(\pi)$ fixed) in the definition of V. Note that the function V is piecewise differentiable. Considering each case separately ($|y-m(\pi)| > |x-m(\pi)| \geq V(\eta_\pi)$; $|x - m(\pi)| > |y - m(\pi)| \geq V(\eta_\pi)$; $|y - m(\pi)| \geq V(\eta_\pi) \geq |x - m(\pi)|$ etc.) reveals that this derivative of V is at most $2(y - x)$. Hence, the total effect of a small increase in $\pi(y)$ together with the same small decrease in $\pi(x)$ is no greater than $3(y - x)$ and therefore the total effect of this change on W is at least $(y - x)/4 > 0$, 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 $\gamma := p_h = p_g$. Note that $W(\pi_g) \geq W(\pi_h)$ if and only if $V(\eta_{\pi_h}) \geq V(\eta_{\pi_g})$. By $(*)$, if $\gamma < \frac{1}{2}$, then

$$V(\eta_{\pi_h}) = \frac{\gamma(1-\gamma) + 2(1-\gamma)\gamma}{2-\gamma}$$

On the other hand, for $V(\eta_{\pi_g})$, $(*)$ yields

$$V(\eta_{\pi_g}) = \frac{\alpha|y-\gamma| + 2(1-\alpha)|x-\gamma|}{2-\alpha}$$

for x, y, α such that $|y - \gamma| \geq |x - \gamma|$, $\alpha y + (1-\alpha)x = \gamma$, $\alpha, x, y \in [0, 1]$. Somewhat tedious but straightforward calculations reveal that setting $y = 1, x = 0$, and $\alpha = \gamma$ maximizes the right-hand side of the display equation above among all α, x, y satisfying these constraints. A symmetric argument shows that $V(\eta_{\pi_h}) \geq V(\eta_{\pi_g})$ for the $\gamma \geq \frac{1}{2}$ case as well.

To show that W does not satisfy (ii), 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. Hence, the preference that W represents fails (i) and by Theorem 2, it fails (ii).

5.3 Proof of Theorem 3

Verifying that if the desired representation exists, then Axioms 1-6 are satisfied is straightforward and omitted. Suppose Axioms 1-6 are satisfied and 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 $W(\pi_f) \geq W(\pi_g)$. Since μ_b and μ_a are nonatomic, $\succeq^*$ is well-defined. It follows from Axiom 6b 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 6$a$, 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, the restriction of $\succeq$ to $\mathcal{F}_a$, 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 $p \in P$ such that $U(p) = t$ and $t' < t$ 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. $\square$

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 [22], 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.

The proof that the existence of the W in the statement of Theorem 3 implies that $\succeq^*$ is a CEU preference is standard. The proof that $\succeq$ satisfies Axiom 6c whenever $\succeq^*$ is a CEU preference is straightforward and omitted. We conclude by proving that the axioms guarantee the desired representation.

Assume that $\succeq$ satisfies Axioms $1-5$ and 6c. By Theorem 1, there exists a preference $\succeq^*$ on $\mathcal{F}^*$ such that $f \succeq g$ iff $f^* \succeq^* g^*$ and a nonconstant, weakly continuous, stochastic monotonicity satisfying W such that $W(\pi_f) \geq W(\pi_g)$ iff $f^* \succeq^* g^*$ for all $f, g \in \mathcal{F}$. Since W is nonconstant, $\succeq^*$ satisfies weak nondegeneracy and since W is satisfies stochastic dominance, $\succeq^*$ satisfies weak stochastic dominance. 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 comonotonic independence.

Observe that since Axiom 6c 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

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

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

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

for $A_1, \dots, A_n$ such that $\mu_a(A_i) = \frac{1}{n}$ for all i. That is, $\hat{f}$ conditional on $A_i \times B_j$ is $\mathcal{E}_a$–measurable and has distribution p_j, and $\hat{g}$ conditional on $A_i \times B_j$ is also $\mathcal{E}_a$–measurable and has distribution q_j for all i, j.

For any $N \subset \{1, \dots, n\}$, let $\hat{f}^N$ denote the act obtained from f by replacing each row $j \in N$ with the corresponding row of $\hat{g}$. Hence, $\hat{f}^{\emptyset} = \hat{f}$ and $\hat{f}^{\{1,\dots,n\}} = \hat{g}$ etc. Note that since U is an expected utility function, $U(p_i) \geq U(p_j)$ 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, $\hat{f}^N$ and $\hat{f}^{N'}$ are comonotonic for all N, N'.

Then, by Axiom 6c, $g^* \succeq^* \frac{1}{n} f^* + \frac{n-1}{n} g^*$ implies

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

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

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

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

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

Applying Axiom 6c again yields

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

Proving that $\succeq^*$ satisfies 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 $p \in \mathcal{F}^*$.

It follows Theorem 1 that the capacity of every event depends only on its μ_b probability. That is, the associated capacity ν can be written as $\gamma \circ \mu_b$ for strictly increasing $\gamma : [0,1] \to [0,1]$ such that $\gamma(0) = 0, \gamma(1) = 1$. 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. $\square$

5.5 Proof of Theorem 5

By Theorem 2, 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 $p, p' \in P$ 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 [22]. $\square$

References

[1] M. Abdellaoui, A. Baillon, P. P. Wakker Combining Bayesian Beliefs and Willingness to Bet to Analyze Attitudes towards Uncertainty, Working Paper (2007).

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

[3] F.J. Anscombe, R.J. Aumann, A Definition of Subjective Probability, Annals of Mathematical Statistics, 34 (1963), 199–205.

[4] P. Billingsley, Convergence of Probability Measures, Wiley, New York (1999).

[5] D. Ellsberg, Risk, Ambiguity and the Savage Axioms, Quarterly Journal of Economics, 75 (1961), 643–669.

[6] L. G. Epstein, A Definition of Ambiguity Aversion, Review of Economic Studies, 66 (1999), 579–608.

[7] L. G. Epstein, J. Zhang, Subjective Probabilities on Subjectively Unambiguous Events, Econometrica, 69 (2001), 265–306.

[8] D. Gale, A Theorem on Flows in Networks, Pacific Journal of Mathematics, 7 (1957), 1073–1082.

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

[10] I. Gilboa, Expected Utility with Purely Subjective Non-Additive Probabilities, Journal of Mathematical Economics, 16 (1987), 65–88.

[11] I. Gilboa, D. Schmeidler, Maxmin Expected Utility with a Non-unique prior, Journal of Mathematical Economics, 18 (1989), 141–153.

[12] S. Grant, A. Kajii, B. Polak, Intrinsic Preference for Information, Journal of Economic Theory, 83 (1998), 233–259.

[13] S. Grant, A., B. Polak Bayesian Beliefs with Stochastic Monotonicity: An Extension of Machina and Schmeidler, Journal of Economic Theory, 130 (2006), 264–82.

[14] F. Gul, A Theory of Disappointment Aversion, Econometrica, 59 (1991), 667–686.

[15] P. Klibanoff, M. Marinacci, S. Mukerji, A Smooth Model of Decision Making Under Ambiguity, Econometrica, 73 (2005), 1849–92.

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

[17] M. J. Machina, D. Schmeidler, A More Robust Definition of Subjective Probability, Econometrica, 60 (1992), 745–780.

[18] M. J. Machina, D. Schmeidler, Bayes without Bernoulli, Journal of Economic Theory, 67 (1995), 106–28.

[19] R. F. Nau, Uncertainty Aversion with Second-Order Probabilities and Utilities, Management Science, 52 (2006), 136–145.

[20] J. Quiggin, A Theory of Anticipated Utility, Journal of Economic Behavior and Organization, 3 (1982), 323–343.

[21] L. J. Savage, The Foundations of Statistics, Wiley, New York (1954).

[22] D. Schmeidler, Subjective Probability and Expected Utility without Additivity, Econometrica, 57 (1989), 571–587.

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

[24] U. Segal, Two-Stage Lotteries without the Reduction Axiom, Econometrica, 58 (1990), 349–377.

[25] Seo, K. Ambiguity and Second-Order Belief,' mimeo (2007), University of Rochester.

[26] A. Tversky and C. Fox, Weighing Risk and Uncertainty, Psychological Review, 102 (1995), 269–283.

[27] P. P. Wakker, Additive Representations of Preferences, Kluwer Academic Publishers, Dordrecht (1989).

[28] P. P. Wakker, Uncertainty, In Lawrence Blume and Steven N. Durlauf (Eds.), The New Palgrave: A Dictionary of Economics, 6780-6791. The MacMillan Press, London (2008).

[29] M.E. Yaari, The Dual Theory of Choice under Risk, Econometrica, 55 (1987), 95–115.