Discussion Paper No. 1212

Stochastic Independence and Uncertainty Aversion[1]

by

Peter Klibanoff
Department of Managerial Economics and Decision Sciences
J. L. Kellogg Graduate School of Management
Northwestern University, Evanston, IL 60208
peterk@nwu.edu
(847)491-5153

January 1998
revised: April 1998

[1] I thank Massimo Marinacci, Paolo Ghirardato, Bart Lipman, and Eddie Dekel for helpful discussions and also thank seminar audiences at the Canadian Economic Theory Conference, U. Toronto, the Conference on Rationality in Game Theory and Utility Theory, U. Caen, and the Math Center Bag Lunch, Northwestern U. All errors are my own.

Abstract

This paper proposes a preference-based condition for stochastic independence of a randomizing device in a product state space. This condition, when imposed on Choquet Expected Utility preferences in a Savage framework displaying uncertainty aversion, results in a collapse to Expected Utility (EU). In contrast, Maxmin EU with multiple priors preferences continue to allow for a very wide variety of uncertainty averse preferences when stochastic independence is imposed. These points are used to reexamine recent arguments against preference for randomization with uncertainty averse preferences. In particular, these arguments are shown to rely on preferences that do not treat randomization as a stochastically independent event.

Journal of Economic Literature Classification Number: D81.

1 Introduction

An example seminal to interest in uncertainty (or ambiguity) aversion is Ellsberg's [8] "two-color" problem. There is a "known urn" which contains 50 red balls and 50 black balls, and an "unknown urn" which contains a mix of red and black balls about which no information is given. Ellsberg observed (as did many afterwards, more carefully) that a substantial fraction of individuals were indifferent between the colors in both urns, but preferred to bet on either color in the "known urn" than the corresponding color in the "unknown urn". This violates not only expected utility, but probabilistically sophisticated behavior more generally. One contemporary criticism of the displayed behavior was put forward by Raiffa [21] who pointed out that flipping a coin to decide which color to bet on in the unknown urn should be viewed as equivalent to betting on the "known" 50-50 urn. One can think of such preferences as displaying a preference for randomization.

Jumping ahead to more contemporary work, there is a burgeoning literature attempting to model uncertainty (or ambiguity) aversion in decision makers. Some of this work (e.g. Lo [18], Klibanoff [16]) accepts this preference for mixture or randomization as a facet of uncertainty aversion, while other work (e.g. Dow and Werlang [5], Eichberger and Kelsey [6]) does not. This has led to several papers, most directly Eichberger and Kelsey [7], but also Ghirardato [10] and Sarin and Wakker [22], related to this difference. In particular, all three papers observe that the choice of a "one-stage" or Savage model as opposed to a "two-stage" or Anscombe-Aumann model can lead to different preferences when modeling uncertainty aversion. In Eichberger and Kelsey [7] the authors set out to "show that while individuals with non-additive beliefs may display a strict preference for randomization in an Anscombe-Aumann framework they will not do so in a Savage-style decision theory."[1]

This paper was motivated in part by the intuition that the one-stage/two-stage modeling distinction is largely a red herring, at least as it relates to preference for randomization. In particular, while appreciating that there can be differences between the frameworks, one goal of this paper is to relate these differences to violations of stochastic independence and to point out that they have essentially no role to play in the debate over preference for randomization in uncertainty aversion. In making this point, the related finding of the restrictiveness of Choquet expected utility preferences in allowing for randomizing devices is key.

An additional contribution of the paper is to provide preference based conditions to describe a stochastically independent randomizing device in a non-Bayesian environment. Section 2 sets out some preliminaries and notation. Section 3 describes two frameworks in which a randomizing

[1] [7, Abstract].

device can be modeled. Section 4 provides the key preference conditions and contains the main results on the restrictiveness of Choquet expected utility when stochastic independence is required and the relative flexibility of Maxmin expected utility with multiple priors. Section 5 concludes.

2 Preliminaries and Notation

We will consider two representations of preferences, each of which generalizes expected utility and allows for uncertainty aversion. The first model is Choquet Expected Utility (CEU). CEU was axiomatized first in an Anscombe-Aumann framework by Schmeidler [23], and then in a Savage framework by Gilboa [13] and Sarin and Wakker [22]. In a Savage framework, but with a finite state space, Wakker [25], Nakamura [19], and Chew and Karni [4] have axiomatized CEU. The second model is Maxmin Expected Utility with non-unique prior (MMEU). MMEU was first axiomatized in an Anscombe-Aumann framework by Gilboa and Schmeidler [14]. In a Savage framework, but with a finite state space, MMEU has recently been axiomatized by Casadesus-Masanell, Klibanoff, and Ozdenoren [3].

Consider a finite set of *states of the world* S. Let X be a set of consequences. An *act* f is a function from S to X. Denote the set of acts by F. A function $v : 2^S \rightarrow [0,1]$ is a *capacity* or *non-additive probability* if it satisfies,

(i) $v(\emptyset) = 0$,

(ii) $v(S) = 1$, and

(iii) $A \subseteq B$ implies $v(A) \leq v(B)$.
It is *convex* if, in addition,

(iv) For all $A, B \subseteq S, v(A) + v(B) \leq v(A \cup B) + v(A \cap B)$.

Now define the (finite) *Choquet integral* of a real-valued function a to be:
$\int a dv = \alpha_1 v(E_1) + \Sigma_{i=2}^{n} \alpha_i [v(\bigcup_{j=1}^{i} E_j) - v(\bigcup_{j=1}^{i-1} E_j)]$, where α_i is the ith largest value that a takes on, and $E_i = a^{-1}(\alpha_i)$.

Let $\succeq$ be a binary relation on acts, F, that represents (weak) preferences. A decision maker is said to have CEU preferences if there exists a utility function $u : X \rightarrow \Re$ and a non-additive probability $v : 2^S \rightarrow \Re$ such that, for all $f, g \in F$, $f \succeq g$ if and only if $\int u \cdot f dv \geq \int u \cdot g dv$. CEU preferences are said to display (weak) *uncertainty aversion* if v is convex.[2] A decision maker is

[2]This characterization of uncertainty aversion for the CEU model stems from an axiom of Schmeidler's [23] of

said to have MMEU preferences if there exists a utility function $u : X \to \Re$ and a non-empty, closed, convex set B of additive probability measures on S such that, for all $f, g \in F$, $f \succeq g$ if and only if $\min_{p \in B} \int u \cdot f dp \geq \min_{p \in B} \int u \cdot g dp$. All MMEU preferences display (weak) uncertainty aversion. Finally, note that the set of MMEU preferences strictly contains the set of CEU preferences with convex capacities.

3 Modeling a Randomizing Device

Corresponding to the two standard frameworks for modeling uncertainty (Anscombe-Aumann and Savage) there are at least two alternative ways to model a randomizing device. In an Anscombe-Aumann setting, a randomizing device is incorporated in the structure of the consequence space. Specifically the "consequences" X, are often taken to be the set of all simple probability distributions over some more primitive set of outcomes, Z. In this set-up, a randomization over two acts f and g with probabilities p and $1-p$ respectively is modeled by an act h where $h(s)(z) = pf(s)(z) + (1-p)g(s)(z)$, for all $s \in S, z \in Z$. Observe that h is, indeed, a well-defined act because the set of simple probability distributions is closed under mixture.

Returning to the "unknown urn" of the introduction, Figure 1 shows the three acts (a) "bet on red," (b) "bet on black," and (c) "randomize 50-50 over betting on red or on black" as modeled in this setting.

Figure 1: Unknown urn with randomization in the consequence space (Anscombe-Aumann)

	R(ed)	B(lack)
a	\$100	\$0
b	\$0	\$100
c	$\frac{1}{2}\$100 \oplus \frac{1}{2}\$0$	$\frac{1}{2}\$100 \oplus \frac{1}{2}\$0$

Alternatively, consider a Savage-style setting with a finite state space (e.g. Wakker [24], Nakamura [19], or Gul [15]). Here a convex combination of two elements of the consequence space X need not be an element of X. Therefore, to model a randomization, we may instead expand the original state space, S, by forming the cross product of S with the possible outcomes (or "states") of the randomizing device. For example, Figure 2 shows the acts (a) "bet on red," (b) "bet on black," (c) "bet on red if heads, black if tails," and (d) "bet on black if heads, red if tails" in the case of the unknown urn with a coin used to randomize.

the same name. This notion of uncertainty aversion has been by far the most common in the literature. Recently, Epstein [9] and Ghirardato and Marinacci [12] have proposed alternative notions of uncertainty aversion.

Figure 2: Unknown urn with randomization in the state space only (Savage)

	R(ed), H(eads)	B(lack), H(eads)	R(ed), T(ails)	B(lack), T(ails)
a	\$100	\$100	\$0	\$0
b	\$0	\$0	\$100	\$100
c	\$100	\$0	\$0	\$100
d	\$0	\$100	\$100	\$0

In comparing the two models, observe that the Anscombe-Aumann setting builds in several key properties that a randomizing device should satisfy while the Savage setting does not. In particular, the probabilities attached to the outcomes of the randomizing device should be unambiguous and the device should be stochastically independent from the (rest of the) state space. Arguably these two properties capture the essence of what is meant by a randomizing device. Both properties are automatically satisfied in an Anscombe-Aumann setting. In a Savage setting, as we will see below, these properties require additional restrictions on preferences.[3]

Several recent papers (including Eichberger and Kelsey [7], Ghirardato [10], and Sarin and Wakker [22]), have noted that CEU need not give identical results in the two frameworks. Specifically, they suggest that the choice of a one-stage (Savage) or two-stage (Anscombe-Aumann) model can lead to different behavior. To see this in the unknown urn example, consider the case where the decision maker's marginal capacity over the colors is $v(R) = v(B) = \frac{1}{3}$. In the Anscombe-Aumann setting this is enough to pin down preferences as $c \succ a \sim b$, (i.e., the Raiffa preferences or preference for randomization).

In the Savage setting, consider the capacity given by

$$\begin{aligned} v(R \times \{H,T\}) = v(B \times \{H,T\}) &= \frac{1}{3}, \\ v(H \times \{R,B\}) = v(T \times \{R,B\}) &= \frac{1}{2}, \\ v(R \times H) = v(R \times T) = v(B \times H) = v(B \times T) &= \frac{1}{6}, \\ v((R \times H) \cup (B \times T)) = v((R \times T) \cup (B \times H)) &= \frac{1}{3}, \\ v(\text{any 3 states}) &= \frac{2}{3}. \end{aligned}$$

This capacity yields the preferences $a \sim b \sim c \sim d$, and thus does not provide a preference

[3]A randomizing device *could* be modeled in an Anscombe-Aumann setting by expanding the state space in exactly the same way as illustrated for the Savage setting. In this case, the same additional restrictions on preferences as in the latter setting would be required to ensure that the randomizing device was unambiguous and stochastically independent.

for randomization as in the Anscombe-Aumann setting. Why can this occur despite the fact that the marginals are identical in the two cases and the product capacity is equal to the product of the marginals on all rectangles? Mathematically, as Ghirardato [10] explains, the source is a failure of the usual Fubini Theorem to hold for Choquet integrals. Intuitively, however, it is not clear what is going "wrong" in the example.

To gain some insight, it is useful to examine the weights applied to each state when evaluating the randomized acts. For example, as Figure 3 shows, "Bet on Red if Heads, Black if Tails" is evaluated using *non-product* weights. The fact that such non-product weights can be applied suggests that the CEU preferences with the capacity above reflect ambiguity not only about the color of the ball drawn from the urn but also about the correlation between the randomizing device and the color of the ball. While such ambiguity is certainly possible, it runs directly counter to the stochastic independence we would expect of a randomizing device. In the next section, therefore, I propose conditions on preferences that ensure this independence.

Figure 3: Non-product weights for randomized act

	R, H	R, T	B, H	B, T
c	\$100	\$0	\$0	\$100
weights	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

4 Stochastic Independence and Preferences

Here I propose conditions on preferences that are designed to reflect two properties of a randomizing device: unambiguous probabilities and stochastic independence. These two properties are essential to what is meant by a randomizing device.

Formally, consider preferences, $\succeq$, over acts, $F : S \to X$, on a finite product state space, $S = S_1 \times S_2 \times \ldots \times S_N$. Denote by F_{S_i} the subset of acts that are measurable with respect to 2^{S_i}. For $f, g \in F$ and $A \subseteq S$, denote by $f_A g$ the act which equals $f(s)$ for $s \in A$ and equals $g(s)$ for $s \notin A$. We now state some useful definitions concerning preferences.

Definition 1 *$\succeq$ satisfies* **solvability** *on S_i if, for $f \in F_{S_i}$, $x, y, z \in X$ and $A_i \subseteq S_i$, $x_{A_i \times S_{-i}} z \succeq f \succeq y_{A_i \times S_{-i}} z$ implies $f \sim w_{A_i \times S_{-i}} z$ for some $w \in X$.*

Solvability should be seen as a joint richness condition on $\succeq$ and X. It is satisfied in all axiomatizations of which we are aware of EU, CEU, or MMEU over Savage acts on a finite state-space. For example, Nakamura [19] imposes solvability directly, while Wakker [24, 25], Gul [15] and Casadesus-Masanell, Klibanoff, and Ozdenoren [3] ensure it is satisfied through topological assumptions on X and continuity assumptions on $\succeq$.

Definition 2 *$\succeq$ satisfies* **unambiguous probabilities (UP)** *on S_i if $\succeq$ restricted to F_{S_i} can be represented by expected utility where the utility function is unique up to a positive affine transformation and the probability measure on the set of all subsets of S_i is unique.*

While the definition is intentionally stated somewhat flexibly, it could easily be made more primitive/rigorous by assuming that preferences restricted to F_{S_i} satisfy the axioms in one of the existing axiomatizations of expected utility over Savage acts on a finite state space such as Wakker [24], Nakamura [19], Gul [15], or Chew and Karni [4]. This definition is intended to capture the fact that the decision-maker associates a unique probability distribution with S_i. Note that the uniqueness requirement on the probability measure entails the existence of consequences $x, y \in X$ such that $x \succ y$ (where preferences over X are derived from preferences over the associated constant-consequence acts in the usual way). Furthermore, any of the axiomatizations cited will imply solvability on S_i as well. For this reason, it may be best to think of solvability as part of the foundation for UP rather than as an independent requirement.

Definition 3 *$\succeq$ satisfies* **stochastic independence (SI)** *on S_i if, for any $\hat{s}_{-i} \in S_{-i}$, any $f \in F$ and any $g, h \in F_{S_i}$,*

$$g \sim h$$

implies,

$$g_{S_i \times \hat{s}_{-i}} f \sim h_{S_i \times \hat{s}_{-i}} f.$$

The idea here is that if the realization of $s_i \in S_i$ is perceived to be stochastically independent from the realization of $s_{-i} \in S_{-i}$, then fixing any s_{-i} and replacing the payoffs across S_i with payoffs across S_i that are deemed equivalent (when considered as elements of F_{S_i}) should not affect preferences. Another way to think about this is to think about conditional preferences: if the outcome on S_{-i} does not affect the perceived probabilities over S_i, then overall evaluation of S_i-measurable acts should agree with evaluation conditional on any given s_{-i}. Further supporting the idea that this axiom reflects stochastic independence is the observation that if preferences are EU and non-trivial, then SI is equivalent to requiring that the representing probability measure be a product measure on $S_i \times S_{-i}$.

Definition 4 *$s_i \in S_i$ is* **null** *if $f_{s_i \times S_{-i}} h \sim g_{s_i \times S_{-i}} h$ for all $f, g, h \in F_{S_i}$.*

Definition 5 *If $\succeq$ satisfies solvability, UP and SI on S_i and S_i contains at least two non-null states, then S_i is a* **candidate randomizing device (CRD)**.

Note that given UP, the requirement of two non-null states is equivalent to at least two elements of S_i being assigned non-zero probability.

4.1 MMEU and Randomizing Devices

This section develops the implications for MMEU preferences of one ordinate of the state space being a CRD. MMEU will be found to be flexible enough to easily incorporate both a CRD and uncertainty aversion.

Theorem 1 *Assume $\succeq$ are MMEU preferences satisfying solvability for some S_i that contains at least two non-null states. Then the following are equivalent:*

(i) S_i is a CRD;

(ii) There exists a probability measure on 2^{S_i}, $\hat{p}$, such that all probability measures, p, in the closed, convex set of measures, B, of the MMEU representation satisfy $p(s) = \hat{p}(s_i)p(S_i \times s_{-i})$, for all $s \in S$.

Proof: ($(i) \Rightarrow (ii)$) We first show that all $p \in B$ must have the same marginal on S_i. CRD implies that $\succeq$ restricted to F_{S_i} may be represented by $\sum_{s_i \in S_i} u(f(s_i))\hat{p}(s_i)$ where $\hat{p}$ is the unique representing probability measure on 2^{S_i}, and u is normalized so that $u(x) = 1$ and $u(y) = 0$. Using the MMEU representation of $\succeq$ yields,

$$\min_{p \in B} \sum_{s_i \in S_i} u(f(s_i))p(s_i \times S_{-i}) = \sum_{s_i \in S_i} u(f(s_i))\hat{p}(s_i) \text{ for all} f \in F_{S_i}. \tag{1}$$

Suppose there is some $p' \in B$ such that $p'(s_i \times S_{-i}) \neq \hat{p}(s_i)$ for some $s_i \in S_i$. Without loss of generality, assume that $\hat{p}(s_i') > p'(s_i' \times S_{-i})$ for an $s_i' \in S_i$. Consider the act $f = x_{s_i' \times S_{-i}}y$. By MMEU and our assumption,

$$\begin{aligned} \min_{p \in B} \sum_{s_i \in S_i} u(f(s_i))p(s_i \times S_{-i}) &\leq u(x)p'(s_i' \times S_{-i}) \\ &< u(x)\hat{p}(s_i') \\ &= \sum_{s_i \in S_i} u(f(s_i))\hat{p}(s_i), \end{aligned}$$

contradicting equation 1. Therefore it must be that $p \in B$ implies $p(s_i \times S_{-i}) = \hat{p}(s_i)$ for all $s_i \in S_i$. In other words, all the marginals on S_i agree.

Now we show that each $p \in B$ is a product measure on $S_i \times S_{-i}$. This part of the argument proceeds by contradiction. Suppose that $p \in B$ does not imply that $p(s) = \hat{p}(s_i)p(S_i \times s_{-i})$, for all $s \in S$. Then there must exist a $p_0 \in B$, non-null $s_i, s_i' \in S_i$ and an $s_{-i}' \in S_{-i}$ such that

$$\frac{p_0(s_i, s_{-i}')}{\hat{p}(s_i)} < \frac{p_0(s_i', s_{-i}')}{\hat{p}(s_i')}.$$

According to p_0, the probability of s_{-i}' conditional on s_i' is higher than the probability of s_{-i}' conditional on s_i. We now show that this difference in conditional probabilities is inconsistent with stochastic independence on S_i. There are two cases to consider: the case where $\hat{p}(s_i) \geq \hat{p}(s_i')$ and the case where $\hat{p}(s_i) < \hat{p}(s_i')$. Assume the former is true and consider the act $f \in F_{S_i}$ such that $f = x_{s_i' \times S_{-i}} y$. Since $x_{s_i \times S_{-i}} y \succeq f \succeq y$, solvability on S_i implies there exists a $w \in X$ such that $w_{s_i \times S_{-i}} y \sim f$. Define $g = w_{s_i \times S_{-i}} y$. Observe that our normalization of u and the preference representation imply $u(w) = \frac{\hat{p}(s_i')}{\hat{p}(s_i)}$.

Define the act $h = g_{S_i \times s_{-i}'} f$. By SI, $f \sim g$ implies $f \sim h$. We have the following contradiction:

$$\begin{aligned}
\min_{p \in B} \sum_{s \in S} u(f(s))p(s) &= \hat{p}(s_i') \\
&= \min_{p \in B} \sum_{s \in S} u(h(s))p(s) \\
&\leq \sum_{s \in S} u(h(s))p_0(s) \\
&= u(w)p_0(s_i, s_{-i}') + u(x)(p_0(s_i' \times S_{-i}) - p_0(s_i', s_{-i}')) \\
&= \hat{p}(s_i') \left[\frac{p_0(s_i, s_{-i}')}{\hat{p}(s_i)} + 1 - \frac{p_0(s_i', s_{-i}')}{\hat{p}(s_i')} \right] \\
&< \hat{p}(s_i').
\end{aligned}$$

A similar argument yields a contradiction for the case $\hat{p}(s_i) < \hat{p}(s_i')$ using acts $f = x_{s_i \times S_{-i}} y \sim w_{s_i' \times S_{-i}} y = g \sim h = f_{S_i \times s_{-i}'} g$, where $u(w) = \frac{\hat{p}(s_i)}{\hat{p}(s_i')}$. Therefore each $p \in B$ must in fact be a product measure on $S_i \times S_{-i}$ and (ii) is proved.

($(ii) \Rightarrow (i)$) That (ii) implies UP is satisfied on S_i is clear because $\hat{p}$ is the unique representing probability measure. To see that SI is satisfied on S_i, consider acts $g, h \in F_{S_i}$ such that $g \sim h$ and any act $f \in F$. Fix $\hat{s}_{-i} \in S_{-i}$ and define acts $d = g_{S_i \times \hat{s}_{-i}} f$ and $e = h_{S_i \times \hat{s}_{-i}} f$. By (ii),

$$\min_{p \in B} \sum_{s \in S} u(d(s))p(s) = \min_{p \in B} \sum_{s_{-i} \in S_{-i}} p(S_i \times s_{-i}) \left[\sum_{s_i \in S_i} u(d(s_i, s_{-i}))\hat{p}(s_i) \right]$$

and,

$$\min_{p\in B}\sum_{s\in S} u(e(s))p(s) = \min_{p\in B}\sum_{s_{-i}\in S_{-i}} p(S_i\times s_{-i})\left[\sum_{s_i\in S_i} u(e(s_i,s_{-i}))\hat{p}(s_i)\right].$$

Since $g\sim h$ and $f\sim f$,

$$\sum_{s_i\in S_i} u(d(s_i,s_{-i}))\hat{p}(s_i) = \sum_{s_i\in S_i} u(e(s_i,s_{-i}))\hat{p}(s_i) \text{ for all } s_{-i}\in S_i.$$

Therefore the two minimization problems are the same and $d\sim e$. *QED*

Thus, we get quite a natural representation in the MMEU framework:

- All the marginals on the randomizing device agree, reflecting the lack of ambiguity about the device.
- All the measures in B are product measures on $S_i\times S_{-i}$, reflecting the independence of the randomizing device.

Remark 1 This result may be seen as additional confirmation that SI is a reasonable condition; in particular that it does not confound stochastic independence with the violations of independence inherent in uncertainty aversion. In fact, the theorem shows that one ordinate being a CRD places *no* restrictions on the preferences over acts measurable with respect to the non-CRD ordinates of the state space.

Remark 2 It is not hard to see from the theorem that, in the Ellsberg "unknown urn" example, if "bet on red" is indifferent to "bet on black" then *any* MMEU preferences that are not EU and for which the coin is a CRD lead to the "Raiffa" preference for randomization. Thus, at least in this setting, there is a strong argument that uncertainty aversion implies a preference for randomization in the sense discussed here.

Remark 3 The set of product measures that emerges from the MMEU characterization is consistent with a notion of independent product of two sets of measures proposed by Gilboa and Schmeidler [14]. Specifically, the set B is trivially the independent product (in their sense) of the (unique) marginal on S_i and the set of marginals on S_{-i} used in representing preferences over $F_{S_{-i}}$. It is worth noting that no purely preference based justification for their broader notion is known.

4.2 CEU, Uncertainty Aversion, and Randomizing Devices

This section examines uncertainty averse CEU preferences on a product state space where one of the ordinates is assumed to be a candidate randomizing device. In stark contrast to the results of the previous section, this class is shown to include only expected utility (EU) preferences. This suggests that CEU preferences with a convex capacity are incapable of modelling both a randomizing device and uncertainty aversion simultaneously.

Theorem 2 *If CEU preferences, $\succeq$, display uncertainty aversion and, for some i, S_i is a CRD then $\succeq$ must be EU preferences.*

Proof:

Recall that the state space is $S = S_1 \times S_2 \times \ldots \times S_N$. Without loss of generality, let S_1 be a CRD. Define $K \equiv \#S_1$ and $L \equiv \#S_{-1}$. Again, without loss of generality, assume $L > 1$. Fix outcomes $x, y \in X$ such that $x \succ y$. Normalize the von Neumann-Morgenstern utility function u given in the CEU representation so that $u(x) = 1$ and $u(y) = 0$. Uncertainty aversion implies that the capacity v in the CEU representation must be convex.

The proof strategy is to assume convex v, S_1 a CRD and *not* EU and derive a contradiction. The following lemmas will be needed.

Lemma 1 *If CEU preferences satisfy UP with respect to S_1 then the capacity, v, in the CEU representation of these preferences must satisfy*

$$v(\{s_1\} \times S_{-1}) = p(s_1), \ \forall s_1 \in S_1, \tag{2}$$

where p is a probability measure on 2^{S_1}.

Proof: Omitted.

The next lemma is due to Eichberger and Kelsey [7].

Lemma 2 *Let $A, B \subseteq S_{-1}$ and let $Y \subseteq S_1$ and Y_1, Y_2 be a partition of Y. Let $P = (Y_1 \times A) \cup (Y_2 \times B)$. If v is a convex capacity on 2^S and equation 2 holds, then*

$$v(P) = v(Y_1 \times A) + v(Y_2 \times B). \tag{3}$$

Proof: See [7, Lemma 3.1].

Lemma 3 *If S_1 is a CRD and v is a convex capacity on 2^S, then*

$$v(A_1 \times B_{-1}) = p(A_1)v(S_1 \times B_{-1}), \tag{4}$$

where $A_1 \subseteq S_1$, $B_{-1} \subseteq S_{-1}$ and p is a probability measure on 2^{S_1}.

Proof: Fix a non-null $s_1 \in S_1$ and consider $f \in F_{S_1}$ such that $f = x_{s_1 \times S_{-1}}y$. By lemma 1, the CEU representation assigns value $p(s_1)u(x) + (1 - p(s_1))u(y) = p(s_1)$ to f. Since $x \succeq f \succeq y$, solvability implies that there exists a $w \in X$ such that $w \sim f$. It must be that $u(w) = p(s_1)$.

Fix any $B_{-1} \subseteq S_{-1}$. By SI on S_1, $w \sim f$ implies $w_{S_1 \times B_{-1}}y \sim f_{S_1 \times B_{-1}}y$. The CEU representation assigns value $v(S_1 \times B_{-1})u(w) = v(S_1 \times B_{-1})p(s_1)$ to $w_{S_1 \times B_{-1}}y$ and value $v(s_1 \times B_{-1})u(x) = v(s_1 \times B_{-1})$ to $f_{S_1 \times B_{-1}}y$. Therefore, it must be that

$$v(s_1 \times B_{-1}) = p(s_1)v(S_1 \times B_{-1}). \tag{5}$$

Fixing some $A_1 \subseteq S_1$ and summing across equation 5 yields,

$$\sum_{s_1 \in A_1} v(s_1 \times B_{-1}) = p(A_1)v(S_1 \times B_{-1}).$$

Using lemma 2, $v(A_1 \times B_{-1}) = \sum_{s_1 \in A_1} v(s_1 \times B_{-1})$. Therefore,

$$v(A_1 \times B_{-1}) = p(A_1)v(S_1 \times B_{-1}).$$

QED

Now let $s_{-11}, \ldots, s_{-1L}$ be an enumeration of the elements of S_{-1} and let $s_{11}, \ldots, s_{1K}$ be an enumeration of the elements of S_1. Assume without loss of generality that $p(s_{11}) \geq p(s_{12}) > 0$.

Consider the act $h_0 \in F_{S_1}$ defined by $h_0 = x_{s_{12} \times S_{-1}}y$. The CEU representation assigns h_0 value $p(s_{12})$. Since $x_{s_{11} \times S_{-1}}y \succeq h_0 \succeq y$, solvability implies the existence of $w \in X$ such that $w_{s_{11} \times S_{-1}}y \sim h_0$. Therefore it must be that $u(w) = \frac{p(s_{12})}{p(s_{11})}$. Let $g_0 = w_{s_{11} \times S_{-1}}y$ and let $f = w_{s_{11} \times s_{-11}}y$. Observe that $g_0 \sim h_0$ by construction.

Define acts $h_1 = h_{0_{S_1 \times s_{-12}}}f$ and $g_1 = g_{0_{S_1 \times s_{-12}}}f$. By SI on S_1, $g_0 \sim h_0$ implies $g_1 \sim h_1$. The CEU representation assigns h_1 value
$u(x)v(s_{12} \times s_{-12}) + u(w)\left[v\left((s_{11} \times s_{-11}) \cup (s_{12} \times s_{-12})\right) - v(s_{12} \times s_{-12})\right]$. By lemma 2,

$$v\left((s_{11} \times s_{-11}) \cup (s_{12} \times s_{-12})\right) = v(s_{11} \times s_{-11}) + v(s_{12} \times s_{-12}).$$

By lemma 3,

$$v(s_{11} \times s_{-11}) = p(s_{11})v(S_1 \times s_{-11})$$

and,

$$v(s_{12} \times s_{-12}) = p(s_{12})v(S_1 \times s_{-12}).$$

Therefore, h_1 gets value
$u(x)p(s_{12})v(S_1 \times s_{-11}) + u(w)p(s_{11})v(S_1 \times s_{-11}) = p(s_{12}) \sum_{i=1}^{2} v(S_1 \times s_{-1i})$. The CEU representation assigns g_1 value $u(w)v(s_{11} \times \{s_{-11}, s_{-12}\})$. By lemma 3,

$$v(s_{11} \times \{s_{-11}, s_{-12}\}) = p(s_{11})v(S_1 \times \{s_{-11}, s_{-12}\}).$$

Therefore g_1 gets value $p(s_{12})v(S_1 \times \{s_{-11}, s_{-12}\})$. Since $g_1 \sim h_1$ it follows that

$$v(S_1 \times \{s_{-11}, s_{-12}\}) = \sum_{i=1}^{2} v(S_1 \times s_{-1i}). \quad (6)$$

Now define acts $h_2 = h0_{S_1 \times s_{-13}} g_1$ and $g_2 = g0_{S_1 \times s_{-13}} g_1$. The CEU representation gives h_2 value $u(x)v(s_{12} \times s_{-13}) + u(w)\left[v\left((s_{11} \times \{s_{-11}, s_{-12}\}) \cup (s_{12} \times s_{-13})\right) - v(s_{12} \times s_{-13})\right]$. Using equation 6 along with lemma 2 and lemma 3 as above, it can be shown that this value is $p(s_{12}) \sum_{i=1}^{3} v(S_1 \times s_{-1i})$. As the CEU representation gives g_2 value $p(s_{12})v(S_1 \times \{s_{-11}, s_{-12}, s_{-13}\})$, and SI implies $h_2 \sim g_2$, it must be that $v(S_1 \times \{s_{-11}, s_{-12}, s_{-13}\}) = \sum_{i=1}^{3} v(S_1 \times s_{-1i})$.

We continue this iterative construction for acts,

$$h_j = h0_{S_1 \times s_{-1(j+1)}} g_{j-1}, \; j = 2, \ldots, L-1$$

with value,

$$p(s_{12}) \sum_{i=1}^{j+1} v(S_1 \times s_{-1i})$$

and acts,

$$g_j = g0_{S_1 \times s_{-1(j+1)}} g_{j-1}, \; j = 2, \ldots, L-1$$

with value,

$$p(s_{12})v(S_1 \times \left\{s_{-11}, \ldots, s_{-1(j+1)}\right\}).$$

By SI on S_1, $g_0 \sim h_0$ implies $g_j \sim h_j$, $j = 1, \ldots, L-1$. Observe, however, that $g_{L-1} = w_{s_{11} \times S_{-1}} y = g_0$. Therefore g_{L-1} is assigned value $p(s_{12})$ by the CEU representation. Since $g_{L-1} \sim h_{L-1}$, this implies that

$$\sum_{i=1}^{L} v(S_1 \times s_{-1i}) = 1.$$

In other words, the capacity v has an additive marginal on S_{-1}, which we denote by p_v. We now show that v must in fact be additive.

By lemma 3 and the above,

$$\begin{aligned}\sum_{s\in S} v(s) &= \sum_{s\in S} p(s_1)p_v(s_{-1}) \\ &= \sum_{s_1\in S_1}\left(p(s_1)\sum_{s_{-1}\in S_{-1}} p_v(s_{-1})\right) \\ &= \sum_{s_1\in S_1} p(s_1) \\ &= 1.\end{aligned}$$

Since v is convex and $v(S)=1$, $\sum_{s\in S} v(s)=1$ implies v is additive. *QED*

Remark 4 This theorem shows that CEU with a convex capacity is a *very* restrictive class of preferences in a Savage-like setting. In particular a decision maker with such preferences must be either uncertainty neutral (i.e. an expected utility maximizer) or must not view any ordinate of the state space as a candidate randomizing device. Note that this fact is disguised in an Anscombe-Aumann setting because there the randomizing device is built into the outcome space and thus automatically separated from the uncertainty over the rest of the world.

Remark 5 The theorem allows us to better understand the result of Eichberger and Kelsey [7], who find that convexity of v, a symmetric additive marginal on S_1, and a requirement that relabeling the states in S_1 not affect preference, together imply no preference for randomization. Note that they do not impose a condition like SI. Thus, the result shown here makes clear that the lack of preference for randomization in their paper comes from the fact that decision makers having preferences in this class (with v somewhere strictly convex) *cannot* act as if the randomizing device is stochastically independent. In other words, the uncertainty averse preferences they consider rule out *a priori* the possibility of a stochastically independent device and thus of true randomization. Once they admit preferences like MMEU, which, as shown above, can reflect a CRD as well as uncertainty aversion, preference for randomization reappears.

4.3 Further Discussion of the SI Condition

The key to these results is the SI condition. I argued above that it is quite natural to accept SI as reflecting stochastic independence of S_i from S_{-i}. However, it is worth elaborating a bit on why SI seems appropriate in the sense that it is neither too strong nor too weak an axiom.

How might SI be too strong? Since stochastic independence does concern *independence*, and uncertainty aversion fundamentally involves violations of the independence axiom/sure thing principle of subjective expected utility theory, it is fair to ask whether imposing SI unnecessarily restricts uncertainty aversion. Theorem 1 answers this question in the negative and suggests that uncertainty aversion is not restricted at all by imposing SI. Specifically, any MMEU preferences[4] over $F_{S_{-i}}$ are compatible with S_i being a CRD.

How might SI be too weak? A natural question to raise here is whether the logic supporting SI also supports imposing a similar axiom with the roles of S_i and S_{-i} reversed. The key to understanding why it may not is to think of S_i as an ordinate for which UP applies, whereas preferences over acts in $F_{S_{-i}}$ display some uncertainty aversion. With uncertainty averse preferences, the effective probabilities or decision weights over S_{-i} depend on how payoffs vary across these states – this is the crux of the matter. Therefore, even if the realization of $s_i \in S_i$ is perceived to be stochastically independent from that of $s_{-i} \in S_{-i}$, there is no reason to think that the evaluation of a number of conditional payoffs over S_{-i} should aggregate in a way that agrees with overall preferences over $F_{S_{-i}}$. Fundamentally, the weight attached to any particular s_{-i} depends on the way payoffs vary over the whole space. Thus, imposing SI seems to be appropriate only on ordinates over which the agent is uncertainty neutral, as in UP, in the sense of using fixed probabilities when evaluating acts measurable with respect to that ordinate. The following theorem formalizes some of this intuition by showing that when S_i is a CRD and we impose SI on S_{-i}[5] as well, even MMEU preferences collapse to EU.

Theorem 3 *Assume $\succeq$ are MMEU preferences satisfying solvability for some S_i and S_{-i} that each contain at least two non-null states. If S_i is a CRD and $\succeq$ satisfy SI on S_{-i} then $\succeq$ are EU preferences. Furthermore there exist probability measures $\hat{p}$ on 2^{S_i} and $\hat{q}$ on $2^{S_{-i}}$ such that the representing probability measure p satisfies $p(s) = \hat{p}(s_i)\hat{q}(s_{-i})$, for all $s \in S$.*

Proof: From the assumption that S_i is a CRD, it follows by theorem 1 that $p \in B$ implies $p(s) = \hat{p}(s_i)p(S_i \times s_{-i})$, for all $s \in S$. Thus, we already know that B contains only product measures, and all that remains to be shown is that B is a singleton set. We use a proof by contradiction.

Suppose that B contains at least two distinct probability measures. Then, for some $\hat{s}_{-i} \in S_{-i}$ it must be that

$$\overline{p} \equiv \max_{p \in B} p(S_i \times \hat{s}_{-i}) > \min_{p \in B} p(S_i \times \hat{s}_{-i}) \equiv \underline{p}.$$

The idea of the proof is to construct a pair of indifferent acts, g and h, that are measurable with respect to S_{-i} such that one act is evaluated using weight $\overline{p}$ on $\hat{s}_{-i}$ while the other is evaluated

[4] Recall that this includes any uncertainty averse CEU preferences as well.

[5] To define "SI on S_{-i}" simply replace i with $-i$ and $-i$ with i everywhere in definition 3

using weight $\underline{p}$. Then we can show that, for a non-null $\hat{s}_i \in S_i$, $g_{\hat{s}_i \times S_{-i}} h \succ h$ which violates SI on S_{-i} and thus generates the contradiction. The only slight complication is that a single pair of g and h will not work for all values of $\overline{p}$ and $\underline{p}$. Thus we must repeat essentially the same argument for four possible cases with a different g, h pair for each case:

Fix $x, y \in X$ such that $x \succ y$. Normalize $u(x) = 1$ and $u(y) = 0$.

Case 1: Assume $1 > \overline{p} > \underline{p} > 0$ and $1 - \overline{p} \geq \underline{p}$. Let $g = x_{S_i \times \hat{s}_{-i}} y$. According to the MMEU representation, g has value $u(x)\underline{p} = \underline{p}$. Let $h = y_{S_i \times \hat{s}_{-i}} w$ where w is chosen so that $h \sim g$. Note that solvability on S_{-i} and our assumption on $\overline{p}$ and $\underline{p}$ guarantee that such a w exists. Furthermore $0 < u(w) = \frac{\underline{p}}{1-\overline{p}} < 1$. Pick a non-null $\hat{s}_i \in S_i$. Since $g \sim h$, SI on S_{-i} implies $g_{\hat{s}_i \times S_{-i}} h \sim h$. (To see this set $f = h$ in the statement of SI.) According to the MMEU representation, $g_{\hat{s}_i \times S_{-i}} h$ has value:

$$\min_{p \in B} \hat{p}(\hat{s}_i) p(S_i \times \hat{s}_{-i}) + (1 - \hat{p}(\hat{s}_i))(1 - p(S_i \times \hat{s}_{-i})) u(w)$$

Note that the minimand is linear in $p(S_i \times \hat{s}_{-i})$, so the minimum is attained at $p(S_i \times \hat{s}_{-i}) = \overline{p}$ or $\underline{p}$ (or both). Substituting in $u(w) = \frac{\underline{p}}{1-\overline{p}}$ it is straightforward to calculate that the value assigned to $g_{\hat{s}_i \times S_{-i}} h$ is strictly greater than $\underline{p}$ which is the value assigned to g (and thus h, as well). Therefore $g_{\hat{s}_i \times S_{-i}} h \succ h$ in violation of SI on S_{-i}. Thus we have a contradiction.

Case 2: Assume $1 > \overline{p} > \underline{p} > 0$ and $1 - \overline{p} < \underline{p}$. For this case, let $h = y_{S_i \times \hat{s}_{-i}} x$ and $g = w_{S_i \times \hat{s}_{-i}} y$, where w is chosen to make $g \sim h$. Again, solvability and the assumptions on $\overline{p}$ and $\underline{p}$ guarantee such a w exists and we can calculate $0 < u(w) = \frac{1-\overline{p}}{\underline{p}} < 1$. Now the argument proceeds exactly as in case 1except we use the case 2 definitions of h and g. In particular it is again true that $g_{\hat{s}_i \times S_{-i}} h \succ h$. Thus we have a contradiction in this case as well.

Case 3: Assume $\overline{p} > \underline{p} = 0$. Let $h = y_{S_i \times \hat{s}_{-i}} x$ and $g = x_{S_i \times \hat{s}_{-i}} w$, where $u(w) = 1 - \overline{p} < 1$ exists and guarantees $g \sim h$. Again it follows from calculations similar to those in the above cases that $g_{\hat{s}_i \times S_{-i}} h \succ h$.

Case 4: Assume $1 = \overline{p} > \underline{p} > 0$. Let $h = w_{S_i \times \hat{s}_{-i}} x$ and $g = x_{S_i \times \hat{s}_{-i}} y$, where $u(w) = \underline{p} < 1$ exists and guarantees $g \sim h$. Again it follows from calculations similar to those in the above cases that $g_{\hat{s}_i \times S_{-i}} h \succ h$.

These four cases cover all the possible configurations of $\overline{p} > \underline{p}$. Thus it must be that $\overline{p} = \underline{p}$. Therefore all $p \in B$ must assign the same weight to $\hat{s}_{-i}$. However, $\hat{s}_{-i}$ was arbitrary, so the same must be true for all $s_{-i} \in S_{-i}$. This proves that B is a singleton set (since MMEU implies B cannot be empty). *QED*

Remark 6 The above theorem shows that imposing SI in both directions is in conflict with uncertainty aversion as embodied in MMEU. One might wonder whether, as was the case with uncertainty averse CEU preferences, it is something about the MMEU representation apart from uncertainty aversion that is causing the conflict with "bi-directional" SI. However, a simple example in the context of the Ellsberg "two-color" problem shows that the conflict is truly between the stronger SI condition and uncertainty aversion itself. Recall from the introduction, and assume for the sake of simplicity only, that an individual is indifferent between betting on red or betting on black in the 50-50 urn and is also indifferent between betting on red or betting on black in the "unknown urn." Uncertainty aversion then manifests itself in the urn problem by a strict preference for betting on the color red (black) in the 50-50 urn over betting on the color red (black) in the "unknown urn." However, one application of SI to the ordinate representing the "unknown urn" implies betting on red in the "unknown urn" is indifferent to a bet which wins only if the same color is drawn from both urns (i.e. either red and red or black and black). But by an application of SI to the ordinate representing the 50-50 urn, betting on red in the 50-50 urn is also indifferent to betting that the same color will be drawn from both urns. By transitivity then, betting on red in the 50-50 urn is indifferent to betting on red in the "unknown urn," contradicting uncertainty aversion.

5 Conclusion

This paper has provided preference-based conditions that a randomizing device should satisfy. When these conditions are applied to the class of CEU preferences with convex capacities in a product state-space model a collapse to expected utility results. This does not occur with MMEU preferences in the same setting. In particular, it appears that some previous results on the absence of preference for randomization were driven not by some deep difference in Anscombe-Aumann and Savage style models as they relate to uncertainty aversion, but by the restrictiveness of the CEU functional form as it relates to stochastic independence which is exacerbated in Savage style models. When stochastic independence is properly accounted for, preference for randomization by uncertainty averse decision makers arises in both one- and two- stage models.

To my knowledge, Blume, Brandenburger and Dekel [2] are the only others to have developed a preference axiom for stochastic independence. Their work is in the context of preferences satisfying the decision-theoretic independence axiom. This leads their condition to be unsatisfactory in the setting of this paper. In particular, their axiom asks more of conditional preferences than is reasonable in the presence of uncertainty aversion and does not need to address the consistency of conditional with unconditional preferences.

Some other recent work on shortcomings of the CEU model is Nehring [20]. In the context of inequality measurement under uncertainty, Ben-Porath, Gilboa and Schmeidler [1] advocate MMEU type functionals and show that they are closed under iterated application while CEU functionals are not.

Differences between CEU and MMEU are also discussed in Klibanoff [17] and Ghirardato, Klibanoff and Marinacci [11].

It is worth noting that whether or not actual decision makers act as if randomizing devices are independent, and when and where these preferences for randomization (or mixture) come into play are questions which cannot be answered here. In addressing these issues in the future, there is scope for both tackling important theoretical issues relating to dynamic flow of utility for decision makers and for doing careful experimental examinations of behavior relating to randomization in decision making.

References

[1] BEN-PORATH, E., I. GILBOA, AND D. SCHMEIDLER (1997) "On the Measurement of Inequality under Uncertainty," *Journal of Economic Theory* **75**, 194–204.

[2] BLUME, L., A. BRANDENBURGER, AND E. DEKEL (1991) "Lexicographic Probabilities and Choice Under Uncertainty," *Econometrica* **59**, 61–79.

[3] CASADESUS-MASANELL, R., P. KLIBANOFF, AND E. OZDENOREN (1998) "Maxmin Expected Utility over Savage Acts with a Set of Priors," mimeo, Northwestern University.

[4] CHEW, S.H. AND E. KARNI (1994) "Choquet Expected Utility with a Finite State Space: commutativity and Act-Independence," *Journal of Economic Theory* **62**, 469–479.

[5] DOW, J. AND S. WERLANG (1994) "Nash Equilibrium under Knightian Uncertainty: Breaking Down Backward Induction," *Journal of Economic Theory* **64**, 305–324.

[6] EICHBERGER, J. AND D. KELSEY (1995) "Non-additive Beliefs and Game Theory," Mimeo, University of Saarbrucken.

[7] ——(1996) "Uncertainty Aversion and Preference for Randomization," *Journal of Economic Theory* **71**, 31–43.

[8] ELLSBERG, D. (1961) "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Economics* **75**, 643–669.

[9] EPSTEIN, L. (1997) "Uncertainty Aversion," mimeo, U. Toronto and Hong Kong University of Science and Technology.

[10] GHIRARDATO, P. (1997) "On Independence for Non-Additive Measures with a Fubini Theorem," *Journal of Economic Theory* **73**, 261–291.

[11] GHIRARDATO, P., P. KLIBANOFF, AND M. MARINACCI (forthcoming) "Additivity with Multiple Priors," *Journal of Mathematical Economics* .

[12] GHIRARDATO, P. AND M. MARINACCI (1997) "Ambiguity Made Precise: A Comparative Foundation and Some Implications," Social Science Working Paper 1026, California Institute of Technology.

[13] GILBOA, I. (1987) "Expected utility theory with purely subjective non-additive probabilities," *Journal of Mathematical Economics* **16**, 141–153.

[14] GILBOA, I. AND D. SCHMEIDLER (1989) "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics* **18**, 141–153.

[15] GUL, F. (1992) "Savage's Theorem with a Finite Number of States," *Journal of Economic Theory* **57**, 99–110.

[16] KLIBANOFF, P. (1994) "Uncertainty, Decision, and Normal Form Games," mimeo, Northwestern University.

[17] ——(1996) "Characterizing Uncertainty Aversion through Preference for Mixtures," CMS-EMS Discussion Paper 1159, Northwestern University.

[18] LO, K.C. (1996) "Equilibrium in Beliefs under Uncertainty," *Journal of Economic Theory* **71**, 443–484.

[19] NAKAMURA, Y. (1990) "Subjective Expected Utility with Non-additive Probabilities on Finite State Spaces," *Journal of Economic Theory* **51**, 346–366.

[20] NEHRING, K. (1997) "Capacities and Probabilistic Beliefs: A Precarious Coexistence," Department of Economics Working Paper Series 97-08, University of California, Davis.

[21] RAIFFA, H. (1961) "Risk, ambiguity, and the Savage axioms: comment," *Quarterly Journal of Economics* **75**, 690–694.

[22] SARIN, R. AND P. WAKKER (1992) "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica* **60**, 1255–1272.

[23] SCHMEIDLER, D. (1989) "Subjective probability and expected utility without additivity," *Econometrica* **57**, 571–587.

[24] WAKKER, P. (1984) "Cardinal coordinate independence for expected utility," *Journal of Mathematical Psychology* **28**, 110–117.

[25] ——(1989) "Continuous subjective expected utility with non-additive probabilities," *Journal of Mathematical Economics* **18**, 1–27.