CIRJE-F-219

A Simple Axiomatization of Iterated Choquet Objectives

Kiyohiko G. Nishimura
The University of Tokyo

Hiroyuki Ozaki
Tohoku University

May 2003

CIRJE Discussion Papers can be downloaded without charge from:

http://www.e.u-tokyo.ac.jp/cirje/research/03research02dp.html

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.

A Simple Axiomatization of Iterated Choquet Objectives*

by

Kiyohiko G. Nishimura

Faculty of Economics
The University of Tokyo

and

Hiroyuki Ozaki

Faculty of Economics
Tohoku University

May 4, 2003

Abstract

A set of axioms which characterizes a preference representable by the iterated Choquet expected utility is presented. This objective function is attractive since it possesses a feature of dynamical consistency. Furthermore, we show that under the same axioms the conditional preference is represented by the Choquet expected utility with respect to the capacity which is updated according to the Dempster-Shafer rule. We do this by weakening Schmeidler's axiom of comonotonic independence to our axiom of constrained comonotonic independence and by adding the axiom of dynamical consistency.

*The work reported here is partially supported by a grant from the Economic and Social Research Insitute, the Cabinet Office, the Government of Japan.

1. Introduction

This paper provides a set of axioms under which a preference relation is represented by the iterated Choquet expected utility with respect to some probability capacity in a two-period dynamic setup.[1] The iterated Choquet expected utility axiomatized in the current paper is attractive since it can incorporate atemporal theory of Choquet expected utility, which has been motivated by huge literature on Knightian uncertainty or ambiguity, into a dynamic setup while still retaining a feature of dynamical consistency. Here, the dynamical consistency means that observing a state realized in the first period should not give the decision-maker any incentive to revise her initial plan for the remaining period optimally chosen before the observation, which is a desirable feature for any tractable economic model dealing with a choice over time.[2]

Furthermore, we show that under our axioms the conditional preference given the first-period's observation, that is, the restriction of the preference over the realized state, is also represented by the Choquet expected utility and that the Choquet integral here is defined with respect to the probability capacity which is obtained by updating some probability capacity according to the Dempster-Shafer rule, the updating rule which is extensively studied in the statistics literature (see, for example, Shafer, 1976 and Dempster, 1967, 1968). This provides one justification of a usage of the Dempster-Shafer rule in the literature on learning under Knightian uncertainty or ambiguity.[3]

A similar objective function is axiomatized by Wang (2002), who employs a rather complicated hierarchical domain of preferences in order to incorporate preferences on the information filtration in the Savege-act framework.[4] In contrast, we assume that the information filtration is exogenously given and that the domain of the preferences are lottery acts along the line developed by Anscombe and Aumann (1963). With these sacrifices in generality, however, our axioms

[1]For the definitions of the probability capacity, the Choquet integral and other related concepts, see Section 2.

[2]For example, Nishimura and Ozaki (2001) study a job-search behavior of an unemployed worker whose preference is given by a general-state-space and infinite-horizon extension of the preferences axiomatized in the current paper. The dynamical consistency allows them to show that the optimal strategy for the worker has a reservation-wage property.

[3]Nishimura and Ozaki (2002) study learning behavior under Knightian uncertainty by assuming dynamically consistent preferences which are similar to the one axiomatized in the current paper and by using the Dempster-Shafer rule as an updating rule.

[4]One of the novelty in Wang's approach which is absent from here is that it can explicitly analyze the decision-maker's preference on the timing of uncertainty resolution.

are much easier to state and to interpret than those by Wang. In particular, our axioms can be easily compared with those of Schmeidler's (1989, first appeared in 1982 as a working paper) pioneering work in the literature of Choquet expected utility. Roughly, we weaken Schmeidler's axiom of comonotonic independence to what we call the axiom of constrained comonotonic independence, which seems to be new to the literature, and then add the axiom of dynamical consistency.

We see that if both the (unconditional) preference and the conditional preference are represented by the *non*iterated Choquet expected utilities with respect to some probability capacity, which would be the case if we maintain the comonotonic independence, then the dynamical consistency implies that the capacity must be additive. (Similar observations are made by Epstein and Le Breton, 1993; and Gilboa and Schmeidler, 1993. For more details, see Section 4.) In contrast, we require that the unconditional preference should be represented by the *iterated* Choquet expected utility by weakening the comonotonic independence to the constrained comonotonic independence. Thus, the dynamical consistency is retained and the conditional preference is still represented by the Choquet expected utility with respect to a *nonadditive* probability capacity.

The organization of the paper is as follows. The next section provides some preliminary definitions which are necessary for the following analysis. Section 3 presents our axioms and the main results of the paper: the representation theorem and a corollary which states that the updating rule in the theorem coincides with naive Bayes' rule and the Dempster-Shafer rule. The proof of the theorem is offered in Section 5. Section 4 discusses the axioms and relates our results to the existing literature.

2. Preliminaries

Suppose that there are two periods. Let $m, n \in \mathbb{N}$ and let the first and the second period's state space be given by $S = \{s_1, \ldots, s_m\}$ and $T = \{t_1, \ldots, t_n\}$, respectively. Therefore, the whole state space is given by $\Omega \equiv S \times T$. A generic element of Ω is denoted by ω or (s, t). Let Y be a mixture space. We call an element of Y a *lottery*. For example, if we let X be a set of prizes and if we let Y be the set of simple probability measures on $(X, 2^X)$, then Y will be clearly a mixture space with the operation in a vector space. Given $y, y' \in Y$ and $\lambda \in [0,1]$, we denote by $\lambda y + (1-\lambda)y'$ the *compound lottery*.[5]

We follow Anscombe and Aumman's (1963) framework and define a *simple lottery act* as a Y-valued function on Ω whose range is a finite subset of Y. We henceforth call it a *lottery act*, or more simply, an act. The set of simple lottery acts is denoted by L_0. Given $f, g \in L_0$ and $\lambda \in [0,1]$, a *compound lottery act* $\lambda f + (1-\lambda)g \in L_0$ is defined by $(\forall\omega)\ (\lambda f + (1-\lambda)g)(\omega) = \lambda f(\omega) + (1-\lambda)g(\omega)$. A lottery act whose range is a singleton is referred to as a *constant act* and the set of constant acts is denoted by L_c. We say that a lottery act f is *1st-period-measurable* if $(\forall s)(\forall t, t')\ f(s,t) = f(s,t')$. As its name suggests, the outcome of the 1st-period-measurable act is determined by the state of the first period only. We sometimes write the outcome of a 1st-period-measurable act f at (s,t) as $f(s)$ rather than as $f(s,t)$.

The decision-maker's preference is given by a *class* of binary relations, $\{\succ_i\}_{i=0,1,\ldots,m}$, on L_0. We understand that $\succ_0$ denotes the decision-maker's *un*conditional preference and $\succ_i$ denotes her *conditional* preference after she knows that $s_i \in S$ has been realized in the first period. The two classes of binary relations, $\{\succeq_i\}_i$ and $\{\sim_i\}_i$, are defined from $\{\succ_i\}_i$ by: $(\forall i)\ \succeq_i \Leftrightarrow \not\prec_i$ and $\sim_i \Leftrightarrow [\not\succ_i$ and $\not\prec_i]$. In general, a binary relation $\succ$ is a *preference order* by definition if it is asymmetric and negatively transitive.[6] For each $i \in \{0, 1, \ldots, m\}$, we define a binary relation over Y by restricting $\succ_i$ on L_c and denote it by the same symbol $\succ_i$, that is,

$$(\forall y, y' \in Y) \quad y \succ_i y' \ \Leftrightarrow\ (\exists f, g \in L_c)\ \ (\forall\omega \in \Omega)\ f(\omega) = y,\ g(\omega) = y' \text{ and } f \succ_i g\,.$$

A pair of acts, f and g, are *comonotonic with respect to* $\succ_i$ if $(\forall\omega, \omega')\ f(\omega) \succ_i f(\omega') \Rightarrow g(\omega) \not\prec_i g(\omega')$. Note that the comonotonicity is defined in terms of the preference induced on Y from $\succ_i$.

Let Ω' be a generic finite set. A real-valued set function θ on Ω' is a *probability capacity* if it satisfies $\theta(\phi) = 0$, $\theta(\Omega') = 1$ and $A \subseteq B \Rightarrow \theta(A) \leq \theta(B)$. If in addition θ is additive, that is, if it satisfies that $A \cap B = \phi \Rightarrow \theta(A \cup B) = \theta(A) + \theta(B)$, then θ is a probability *measure*. Let

[5] Here, $\lambda y + (1-\lambda)y'$ should be understood as the element of Y into which (y, y', λ) is mapped by the operation which defines Y as a mixture space, and hence, it does not necessarily mean the convex combination in a vector space. Accidentally, it does when Y is the set of simple probability measures on $(X, 2^X)$ as in the example of the main text.

[6] A binary relation $\succ$ is *asymmetric* if $(\forall f, g \in L_0)\ f \succ g \Rightarrow g \not\succ f$, and it is *negatively transitive* if $(\forall f, g, h \in L_0)\ [f \not\succ g$ and $g \not\succ h] \Rightarrow f \not\succ h$.

a be a real-valued function on Ω' and suppose that a is representable by

$$a = \sum_{i=0}^{k} a_i \chi_{A_i}$$

where $a_0 < a_1 < \cdots < a_k$, $(\forall i)$ $A_i = \{\, \omega' \in \Omega' \,|\, a(\omega') = a_i \,\}$ and χ_A is the indicator function for a set A.[7] Such a representation is always possible and unique. Then, the *Choquet integral* of a with respect to a probability capacity θ is defined by

$$\int a\, d\theta \equiv \int_{\Omega'} a(\omega')\, \theta(d\omega') \equiv a_0 + \sum_{i=1}^{k} (a_i - a_{i-1})\theta \left(\cup_{j=i}^{k} A_j \right) .$$

3. Axioms and Main Results

We take as a primitive a class of binary relations, $\{\succeq_i\}_{i=0,1,\ldots,m}$, that the decision-maker possesses over L_0, and we consider the following axioms which may be imposed on that class of binary relations. In the axioms, f, g and h denote arbitrary elements in L_0 and λ denotes an arbitrary real number such that $\lambda \in (0, 1]$.

A1 (Ordering) For each $i \in \{0, 1, \ldots, m\}$, the binary relation $\succ_i$ is a preference order.

A2(0) (Constrained Comonotonic Independence) If f, g, h are 1st-period-measurable and pairwise comonotonic with respect to $\succ_0$, then $f \succ_0 g \Rightarrow \lambda f + (1-\lambda)h \succ_0 \lambda g + (1-\lambda)h$.

A2(1) (Conditional Comonotonic Independence) For each $i \in \{1, \ldots, m\}$, if f, g, h are pairwise comonotonic with respect to $\succ_i$, then $f \succ_i g \Rightarrow \lambda f + (1-\lambda)h \succ_i \lambda g + (1-\lambda)h$.

A3 (Continuity) For each $i \in \{0, 1, \ldots, m\}$, if $f \succ_i g$ and $g \succ_i h$, then

$$(\exists \alpha, \beta \in (0,1)) \;\; \alpha f + (1-\alpha)h \succ_i g \;\text{ and }\; g \succ_i \beta f + (1-\beta)h\,.$$

A4 (Monotonicity) For each $i \in \{0, 1, \ldots, m\}$, if $(\forall \omega \in \Omega)$ $f(\omega) \succeq_i g(\omega)$, then $f \succeq_i g$.

[7]That is, $\chi_A : \Omega' \to \{0, 1\}$ is a function defined by

$$(\forall \omega') \quad \chi_A(\omega') = \begin{cases} 1 & \text{if} \quad \omega' \in A \\ 0 & \text{if} \quad \omega' \notin A\,. \end{cases}$$

A5 (Non-degeneracy) $(\exists f, g \in L_0)\ f \succ_0 g$.

A6 (Independence of Unrealized Events) $(\forall i \in \{1, \dots, m\})\ f(s_i, \cdot) = g(s_i, \cdot) \ \Rightarrow\ f \sim_i g$.

A7 (Ordinal Preference Consistency) $(\forall i \in \{1, \dots, m\})(\forall y, y' \in Y)\ y \succ_0 y' \ \Leftrightarrow\ y \succ_i y'$.

A8 (Dynamical Consistency) $[(\forall i \in \{1, \dots, m\})\ f \succeq_i g] \ \Rightarrow\ f \succeq_0 g$.

We discuss each axiom in the next section.

The main results of this paper are the following theorem and corollary. The theorem shows that under Axioms A1-A8, the unconditional preference $\succ_0$ is represented by an iterated Choquet expected utility with respect to some class of probability capacities (see (1) in the Theorem) and each conditional preference $\succ_i$ is represented by its restriction over $\{s_i\} \times T$ (see (2) in the Theorem). The proof of the Theorem is relegated to Section 5.

Theorem. *A class of binary relations, $\{\succ_i\}_{i=0}^m$, satisfies A1-A8 if and only if there exist a unique probability capacity θ_0 on S, a unique class of probability capacities $\langle\theta_{s_i}\rangle_{i=1}^m$ on T and a nonconstant affine function $u : Y \to \mathbb{R}$, which is unique up to a positive affine transformation, such that*

$$f \succ_0 g \ \Leftrightarrow\ \int_S \int_T u(f(s,t))\, \theta_s(dt)\, \theta_0(ds) > \int_S \int_T u(g(s,t))\, \theta_s(dt)\, \theta_0(ds) \tag{1}$$

$$\textit{and} \quad (\forall i \in \{1, \dots, m\}) \quad f \succ_i g \ \Leftrightarrow\ \int_T u(f(s_i,t))\, \theta_{s_i}(dt) > \int_T u(g(s_i,t))\, \theta_{s_i}(dt)\,. \tag{2}$$

We now turn to the corollary. Let $\hat\theta_0$ be any probability capacity on S and let $\langle\hat\theta_{s_i}\rangle_{i=1}^m$ be any class of probability capacities on T. Define a real-valued set function $\hat\theta$ on Ω by

$$(\forall A \in 2^\Omega) \quad \hat\theta(A) = \int_S \int_T \chi_A(s,t)\, \hat\theta_s(dt)\, \hat\theta_0(ds)\,. \tag{3}$$

It follows immediately that $\hat\theta(\phi) = 0$, $\hat\theta(\Omega) = 1$ and $A \subseteq B \Rightarrow \hat\theta(A) \le \hat\theta(B)$. Therefore, $\hat\theta$ is a probability capacity on Ω.

First, we observe[8] that $(\forall E \subseteq S)(\forall F \subseteq T)$

$$\hat\theta(E \times F) = \int_S \int_T \chi_{E\times F}(s,t)\, \hat\theta_s(dt)\, \hat\theta_0(ds) = \int_S \hat\theta_s(F) \chi_E(s)\, \hat\theta_0(ds)\,,$$

[8]Similar observations to those in this and the next paragraphs are also made by Wang (2002).

from which it follows that

$$(\forall E) \;\; \hat{\theta}(E \times T) = \int_S \hat{\theta}_s(T)\chi_E(s)\,\hat{\theta}_0(ds) = \hat{\theta}_0(E)$$
$$\text{and} \quad (\forall i)(\forall F) \;\; \hat{\theta}(\{s_i\} \times F) = \int_S \hat{\theta}_s(F)\chi_{\{s_i\}}(s)\,\hat{\theta}_0(ds) = \hat{\theta}_0(\{s_i\})\hat{\theta}_{s_i}(F)\,.$$

We thus conclude that

$$\hat{\theta}_{s_i}(F) = \frac{\hat{\theta}(\{s_i\} \times F)}{\hat{\theta}_0(\{s_i\})} = \frac{\hat{\theta}(\{s_i\} \times F)}{\hat{\theta}(\{s_i\} \times T)}$$

as far as the denominators are non-zero. This is an update rule which would be obtained if we simply applied Bayes' rule to $\hat{\theta}$ by regarding $\hat{\theta}$ as a probability measure and may be called *naive Bayes' rule*.

Second, we observe that $(\forall i)(\forall F)$

$$\begin{aligned}
\hat{\theta}\left((\{s_i\} \times F) \cup (S\backslash\{s_i\} \times T)\right) &= \int_S \int_T \chi_{(\{s_i\}\times F)\cup(S\backslash\{s_i\}\times T)}(s,t)\,\hat{\theta}_s(dt)\,\hat{\theta}_0(ds) \\
&= \int_S \left(\hat{\theta}_{s_i}(F)\chi_{\{s_i\}}(s) + \chi_{S\backslash\{s_i\}}(s)\right)\,\hat{\theta}_0(ds) \\
&= (1-\hat{\theta}_{s_i}(F))\hat{\theta}_0(S\backslash\{s_i\}) + \hat{\theta}_{s_i}(F) \\
&= (1-\hat{\theta}_0(S\backslash\{s_i\}))\hat{\theta}_{s_i}(F) + \hat{\theta}_0(S\backslash\{s_i\})\,,
\end{aligned}$$

from which we conclude that

$$\begin{aligned}
\hat{\theta}_{s_i}(F) &= \frac{\hat{\theta}\left((\{s_i\} \times F) \cup (S\backslash\{s_i\} \times T)\right) - \hat{\theta}_0(S\backslash\{s_i\})}{1-\hat{\theta}_0(S\backslash\{s_i\})} \\
&= \frac{\hat{\theta}\left((\{s_i\} \times F) \cup (S\backslash\{s_i\} \times T)\right) - \hat{\theta}(S\backslash\{s_i\} \times T)}{1-\hat{\theta}(S\backslash\{s_i\} \times T)}
\end{aligned}$$

as far as the denominators are non-zero. This is an update rule for capacities which is known as the *Dempster-Shafer rule* in the statistics literature (see, for example, Shafer, 1976 and Dempster, 1967, 1968).

By taking, as $\hat{\theta}_0$ and $\langle\hat{\theta}_{s_i}\rangle_i$, θ_0 and $\langle\theta_{s_i}\rangle_i$ whose existence is guaranteed by the Theorem under A1-A8 and by defining $\hat{\theta}$ from θ_0 and $\langle\theta_{s_i}\rangle_i$ via (3), we have the next corollary.

Corollary. *Suppose that a class of binary relations, $\{\succ_i\}_{i=0}^m$, satisfies A1-A8. Then, there exist a probability capacity $\hat{\theta}$ on Ω and an affine function $u : Y \to \mathbb{R}$, which is unique up to a* *positive affine transformation, such that for any i satisfying $\hat{\theta}(\{s_i\}\times T) \neq 0$ and $\hat{\theta}(S\backslash\{s_i\}\times T) \neq 1$, it holds that*

$$f \succ_i g \;\Leftrightarrow\; \int_T u(f(s_i,t))\,\theta_{s_i}(dt) > \int_T u(g(s_i,t))\,\theta_{s_i}(dt)\,,$$

where θ_{s_i} is derived from $\hat{\theta}$ by

$$(\forall F) \quad \theta_{s_i}(F) = \frac{\hat{\theta}(\{s_i\}\times F)}{\hat{\theta}(\{s_i\}\times T)} = \frac{\hat{\theta}\left((\{s_i\}\times F)\cup(S\backslash\{s_i\}\times T)\right) - \hat{\theta}(S\backslash\{s_i\}\times T)}{1-\hat{\theta}(S\backslash\{s_i\}\times T)}\,.$$

This corollary shows that under Axioms A1-A8, each conditional preference is represented by the Choquet expected utility with respect to the probability capacity which is updated from some probability capacity $\hat{\theta}$ according to the Dempster-Shafer rule. Therefore, the Corollary provides one justification of a usage of the Dempster-Shafer rule in the literature on learning under Knightian uncertainty or ambiguity (see, for example, Nishimura and Ozaki, 2002).

Furthermore, the Corollary also shows that for the probability capacity $\hat{\theta}$, naive Bayes' rule and the Dempster-Shafer rule coincide, which is not always the case for general probability capacities. This and that the updating rule must be the Dempster-Shafer rule are among strong implications of the dynamical consistency imposed on the class of (un)conditional preferences.

4. Discussion of Axioms

This section discusses the axioms in the Theorem with relation to those in the existing literature. The whole set of the axioms are divided into two groups, that is, Axioms A1 through A5 and Axioms A6 through A8.

Except for Axiom A2, each axiom in the first group, A1 and A3-A5, requires that all of the binary relations, $\{\succ_i\}_{i=0}^m$, should satisfy the axiom of Schmeidler (1989) with the same name. Note that, while Axiom A5 requires the non-degeneracy only of $\succ_0$, Axioms A4 and A5 applied to $\succ_0$ and Axiom A7 imply that $\succ_i$ also satisfies the non-degeneracy for all i (see Step 1 of the proof in Section 5).[9]

[9]To be precise, Schmeidler's axioms of ordering and non-degeneracy are stated in terms of the weak order while ours are stated in terms of its asymmetric part. Of course, his and ours are equivalent.

Axiom A2(1) (conditional comonotonic independence) requires that all *conditional* preferences should satisfy Schmeidler's comonotonic independence. The motivation for the axiom of comonotonic independence can be found in Schmeidler's (1989) original work.

In contrast, Axiom A2(0) (constrained comonotonic independence) is concerned with the *un*conditional preference. On the one hand, when applied to the unconditional preference, the axiom of comonotonic independence would require

A2S(0) (Comonotonic Independence) If f, g, h are pairwise comonotonic with respect to $\succ_0$, then $f \succ_0 g \;\Rightarrow\; \lambda f + (1-\lambda)h \succ_0 \lambda g + (1-\lambda)h$.

On the other hand, Axiom A2(0) requires the comonotonic independence to hold only among acts which are 1st-period-measurable and does not say anything about a triplet of acts at least one of which is not 1st-period-measurable. Clearly, Axiom A2(0) is implied by Axiom A2S(0). Actually, it is substantially weaker than A2S(0). We come back to this point later.

Among the second group of the axioms, Axiom A6 (independence of unrealized events) applies only to the conditional preferences. It is well-known also as the axiom of *consequentialism*[10] and requires that if two acts behave exactly in the same manner after the realization of state s_i, the conditional preference given s_i should evaluate these two acts indifferently. Axiom A6 forces the representation of $\succ_i$ to be independent of unrealized states, s_j $(j \neq i)$ (see (2) in the Theorem).

The last two axioms are concerned with the connection between the unconditional preference and the conditional preferences. The former, Axiom A7 (ordinal preference consistency), is also well-known[11] and requires that all the preferences should evaluate constant acts in the same manner. This axiom implies that the von-Neumann-Morgenstern (vNM) utility index, u, in the Theorem can be taken to be common for all representations (see (1) and (2) in the Theorem).

The latter, Axiom A8 (dynamical consistency), is a version of a well-known axiom of

[10] See, for example, Axiom 7 of Ghirardato (2002), which, in the Savage-act framework, axiomatizes the class of (un)conditional preferences which can be represented by an expected utility with respect to a probability measure P and conditional expected utilities with respect to the conditional probability measures updated from P by Bayes' rule.

[11] See, for example, Axiom 3 of Ghirardato (2002), cited in the previous footnote.

dynamical consistency.[12] To see an implication of this axiom, suppose that the axiom is now violated so that there exists a pair of acts, f and g, such that $(\forall i)\ f \succeq_i g$ and $g \succ_0 f$. Then, there are two possible cases: (a) $(\forall i)\ f \sim_i g$ and $g \succ_0 f$; and (b) $(\forall i)\ f \succeq_i g$, $(\exists i)\ f \succ_i g$ and $g \succ_0 f$, in the latter case of which there exists some state such that the decision-maker has a definite incentive to revise her initial plan after observing that state. Axiom A8 requires that there should be no such pair of acts that either (a) or (b) is the case.

We now turn to a discussion of implications of the axioms as a whole. In particular, we argue that although the dynamical consistency (A8) seems to be a mild requirement, its implication is fairly strong in the presence of other axioms. To see this, assume that all the axioms of the Theorem is satisfied. Further, assume that Axiom A2(0) is now strengthened to Axiom A2S(0). Then, by Axioms A1, A2S(0) and A3-A5, Schmeidler's (1989) theorem implies that there exists a unique probability capacity θ on Ω and an affine function $u : Y \to \mathbb{R}$ such that

$$f \succ_0 g \ \Leftrightarrow\ \int_\Omega u(f(\omega))\, \theta(d\omega) > \int_\Omega u(g(\omega))\, \theta(d\omega)\,, \tag{4}$$

where u may be assumed, without loss of generality, to be the same as the one in the Theorem.[13] Furthermore, the Theorem shows that (1) holds with some θ_0 and $\langle \theta_{s_i} \rangle_i$. Therefore, it follows that[14]

$$(\forall f \in L_0) \qquad \int_\Omega u(f(\omega))\, \theta(d\omega) = \int_S \int_T u(f(s,t))\, \theta_s(dt)\, \theta_0(ds)\,. \tag{5}$$

By considering an act f^A satisfying $u(f^A(\cdot)) = \chi_A$ for each A,[15] equation (5) implies that $(\forall A)\ \theta(A) = \hat{\theta}(A)$, where $\hat{\theta}$ is derived from θ_0 and $\langle \theta_{s_i} \rangle_i$ by (3) right after the statement of the Theorem in the previous section (set $\hat{\theta}_0$ and $\langle \hat{\theta}_{s_i} \rangle_i$ there to be equal to θ_0 and $\langle \theta_{s_i} \rangle_i$ here).

[12] Another version of the dynamical consistency, which is conceptually pretty close to ours but adapted to a different framework, appears as Axiom 5 (consistency) in Wang (2002). For the difference between his framework and ours, see the Introduction of the current paper.

[13] From (4), it follows that u is an affine function which represents $\succ_0$ on Y. The vNM utility index in the Theorem is also an affine function representing $\succ_0$ on Y. Therefore, one index is an affine transformation of the other, and hence, we can take u in (4) to be the same as the one in the Theorem.

[14] To see this, note that the both sides of equation (5) coincide when f is a constant act. Step 1 of the proof in Section 5 proves that for any $f \in L_0$, there exists a constant act which is indifferenct to f with respect to $\succ_0$ (see (9)). Therefore, the both sides of equation (5) must *always* coincide since they both represent $\succ_0$.

[15] Such an act certainly exists. See (8) in Step 1 of the proof in Section 5 (let f^A be such that $f^A(\omega) = y^*$ if $\omega \in A$ and $f^A(\omega) = (1/2)y^* + (1/2)y_*$ if $\omega \notin A$).

Therefore, the discussion there and the fact that $\hat{\theta} = \theta$ show that

$$\theta_0 = \hat{\theta}(\cdot \times T) = \theta(\cdot \times T) \text{ and } \theta_s = \frac{\hat{\theta}(\{s\} \times \cdot)}{\hat{\theta}(\{s\} \times T)} = \frac{\theta(\{s\} \times \cdot)}{\theta(\{s\} \times T)} . \tag{6}$$

It is well-known that the probability capacity θ which satisfies both (5) and (6) must be additive.[16] Therefore, a class of (un)conditional preferences which satisfies Axioms A1, A2S(0), A2(1) and A3-A8 can be represented by an expected utility with respect to a unique probability *measure* and conditional expected utilities with respect to the conditional probability *measures* updated by Bayes' rule.

Epstein and Le Breton (1993) observe that in the Savage-act framework, if the unconditional preference is represented by using a unique probability measure P (but not necessarily in a form of expected utility), the axiom of dynamical consistency implies that each conditional preference is represented by using the conditional probability measure updated from P by Bayes' rule (again not necessarily in a form of conditional expected utility).[17] The discussion in the previous paragraph shows that if both the unconditional and conditional preferences are represented by the (noniterated) Choquet expected utilities (under Axioms A1, A2S(0), A2(1) and A3-A5), the axiom of dynamical consistency (as well as Axioms A6 and A7) implies that the representation of the preferences must be the (un)conditional expected utilities with respect to a probability measure and the conditional probability measures updated by Bayes' rule. This is a variant of the observation made by Epstein and Le Breton in our lottery-act framework.

Furthermore, in the lottery-act framework as ours, Gilboa and Schmeidler (1993) show that if both the unconditional and conditional preferences are represented by the (noniterated) Choquet expected utilities, the dynamical consistency must be violated except for the trivial case where the capacity is additive.[18] In contrast, we require that the unconditional preference should be represented only by an *iterated* Choquet expected utility by substantially weakening A2S(0) to A2(0). By this, the class of (un)conditional preferences restores the dynamical consistency while still allowing the conditional preferences to be represented by the Choquet expected utility with respect to a probability capacity which is not necessarily reduced to a probability measure.

[16] See, for example, Yoo (1991).

[17] Their work is largely motivated by Machina and Schmeidler's (1992) theory of probabilistical sophistication.

[18] To be more precise, they show that both the unconditional and conditional preferences are represented by the (noniterated) Choquet expected utilities if and only if the class of (un)conditional preferences satisfies what they call an *f-Bayesian rule* with f being an act which takes on the best and the worst outcomes only.

5. Proof of Theorem

Showing that the axioms in the Theorem are necessary for the representation is straightforward in view of the comonotonic additivity of Choquet integrals (Schmeidler, 1986), and hence, it is omitted. We prove the sufficiency of them in several steps.

Step 1. This step shows that there exists a function $U : L_0 \to \mathbb{R}$ which represents $\succ_0$, that is,

$$(\forall f, g \in L_0) \quad f \succ_0 g \Leftrightarrow U(f) > U(g), \tag{7}$$

and whose restriction on L_c is an affine function. Note that L_c is a mixture space and that by A1, A2(0) and A3, $\succ_0$ restricted on L_c satisfies all the axioms of the mixture-space theorem of Herstein and Milnor (1953). Therefore, it follows that there exists an affine function J on L_c which represents $\succ_0$ restricted on L_c. Define a function u on Y by $(\forall y)$ $u(y) = J(f)$ where $f \in L_c$ is such that $(\forall \omega)$ $f(\omega) = y$. Clearly, u is an affine function on Y. By A4 and A5, there exist $y^*, y_* \in Y$ such that $y^* \succ_0 y_*$. Therefore, by making a suitable positive affine transformation on u, we may assume without loss of generality that

$$u(y^*) = 1 \quad \text{and} \quad u(y_*) = -1\,. \tag{8}$$

We claim that for any $f \in L_0$, there exist $\bar{y}, \underline{y} \in Y$ and $\alpha \in [0, 1]$ such that

$$f \sim_0 \alpha\bar{y} + (1 - \alpha)\underline{y}\,. \tag{9}$$

This holds because it follows from A4 that there exist $\bar{y}, \underline{y} \in Y$ such that $\bar{y} \succeq_0 f \succeq_0 \underline{y}$ and because it follows from A2(0) and A3 that there exists $\alpha \in [0, 1]$ such that $f \sim_0 \alpha\bar{y} + (1 - \alpha)\underline{y}$.

Then, define $U : L_0 \to \mathbb{R}$ by

$$(\forall f \in L_0) \quad U(f) = u\left(\alpha\bar{y} + (1 - \alpha)\underline{y}\right)\,, \tag{10}$$

where $\bar{y}, \underline{y} \in Y$ and $\alpha \in [0, 1]$ are such that (9) holds. Such $\bar{y}$, $\underline{y}$ and α certainly exist as shown in the previous paragraph. It is then immediate that U is well-defined and represents $\succ_0$ on L_0. Furthermore, its restriction on L_c is an affine function since u is an affine function on Y.

Step 2. Define L_0' by

$$L_0' = \left\{ f' : T \to Y \;\middle|\; f'(T) \text{ is a finite subset of } Y \right\} .$$

That is, L_0' is the space of simple lottery acts whose domain is T. Given any $f \in L_0$ and any $s \in S$, it holds obviously that $f(s, \cdot) \in L_0'$.

For each $i \in \{1, \ldots, m\}$, we define a binary relation $\succ_i'$ on L_0' by

$$(\forall f', g' \in L_0') \quad f' \succ_i' g' \;\Leftrightarrow\; (\forall f, g \in L_0) \;\; \big[f(s_i, \cdot) = f' \;\text{ and }\; g(s_i, \cdot) = g' \;\Rightarrow\; f \succ_i g \big] . \tag{11}$$

We derive $\succeq_i'$, $\sim_i'$ and $\succ_i'$ on Y, from $\succ_i'$ as usual. Two acts, f' and g', are *comonotonic with respect to* $\succ_i'$ if $(\forall t, t' \in T)$ $f'(t) \succ_i' f'(t') \Rightarrow g'(t) \not\prec_i' g'(t')$.

The rest of this step shows that the following holds:

$$(\forall f', g' \in L_0') \quad f' \succ_i' g' \;\Leftrightarrow\; (\exists f, g \in L_0) \;\; f(s_i, \cdot) = f', g(s_i, \cdot) = g' \;\text{ and }\; f \succ_i g . \tag{12}$$

To show ($\Rightarrow$), suppose that $f' \succ_i' g'$. Then, the right-hand side of (12) clearly holds true by the definition (11) since we can always find f and g such that $f(s_i, \cdot) = f'$ and $g(s_i, \cdot) = g'$ for any f' and g'. To show ($\Leftarrow$), suppose that the right-hand side of (12) holds, that is, assume that there exist $\hat{f}$ and $\hat{g}$ such that $\hat{f}(s_i, \cdot) = f'$, $\hat{g}(s_i, \cdot) = g'$ and $\hat{f} \succ_i \hat{g}$. Let f, g be any pair of acts such that $f(s_i, \cdot) = f'$ and $g(s_i, \cdot) = g'$. Then, A6 implies that $\hat{f} \sim_i f$ and $\hat{g} \sim_i g$, and hence, it follows that $f \succ_i g$ by A1. Therefore, $f' \succ_i' g'$ holds by the definition (11).

Step 3. This step proves that $\succ_i'$ defined in the previous step satisfies all the axioms of Schmeidler's Theorem (1989, p.578). In the rest of this step, we fix $i \in \{1, \ldots, n\}$ arbitrarily.

(Ordering) We need to show that $\succ_i'$ is asymmetric and negatively transitive. The asymmetry is immediate from the definition (11) of $\succ_i'$ and the asymmetry of $\succ_i$ (A1). To show the negative transitivity, let $f', g', h' \in L_0'$ be such that $f' \not\succ_i' g'$ and $g' \not\succ_i' h'$. Then, by the definition (11) of $\succ_i'$, there exist $f, g \in L_0$ such that $f(s_i, \cdot) = f'$, $g(s_i, \cdot) = g'$ and $f \not\succ_i g$ and there exist $\hat{g}, h \in L_0$ such that $\hat{g}(s_i, \cdot) = g'$, $h(s_i, \cdot) = h'$ and $\hat{g} \not\succ_i h$. Since A6 implies that $g \sim_i \hat{g}$ and since $\succ_i$ is a preference order (A1), it follows that $f \not\succ_i h$. Therefore, we have $f' \not\succ_i' h'$ by the definition (11) of $\succ_i'$, which completes the proof of the negative transitivity.

(Comonotonic Independence) Let $f', g', h' \in L_0'$ be pairwise comonotonic with respect to $\succ_i'$ and such that $f' \succ_i' g'$. We need to show that for any $\lambda \in (0,1)$,

$$\lambda f' + (1-\lambda)h' \succ_i' \lambda g' + (1-\lambda)h' \,. \tag{13}$$

Let $\lambda \in (0,1)$ and let $f, g, h \in L_0$ be such that $(\forall s)\ f(s,\cdot) = f'$, $g(s,\cdot) = g'$ and $h(s,\cdot) = h'$. This paragraph proves that f, g, h thus defined are pairwise comonotonic with respect to $\succ_i$. To see this, let $s, s' \in S$ and $t, t' \in T$ be such that $f(s,t) \succ_i f(s',t')$. Since $f(s,t) = f'(t)$ and $f(s',t') = f'(t')$, it follows that $f'(t) \succ_i' f'(t')$ by (12) and that $g'(t) \not\prec_i' g'(t')$ by the comonotonicity of f' and g' with respect to $\succ_i'$. Since $g(s,t) = g'(t)$ and $g(s',t') = g'(t')$, (12) also implies that $g(s,t) \not\prec_i g'(s',t')$. The same argument applies to the other pairs of acts.

Note that $f \succ_i g$ by the definition (11) of $\succ_i'$ and the assumption that $f' \succ_i' g'$. Therefore, A2(1) and the pairwise comonotonicity of f, g, h proven in the previous paragraph imply that $\lambda f + (1-\lambda)h \succ_i \lambda g + (1-\lambda)h$. Finally, (13) follows from (12) because $(\lambda f + (1-\lambda)h)(s_i,\cdot) = \lambda f(s_i,\cdot) + (1-\lambda)h(s_i,\cdot) = \lambda f' + (1-\lambda)h'$ and $(\lambda g + (1-\lambda)h)(s_i,\cdot) = \lambda g(s_i,\cdot) + (1-\lambda)h(s_i,\cdot) = \lambda g' + (1-\lambda)h'$.

(Continuity) Let $f', g', h' \in L_0'$ be such that $f' \succ_i' g'$ and $g' \succ_i' h'$. We need to show the existence of $\alpha, \beta \in (0,1)$ such that $\alpha f' + (1-\alpha)h' \succ_i' g'$ and $g' \succ_i' \beta f' + (1-\beta)h'$. To do this, let $f, g, h \in L_0$ be such that $f(s_i,\cdot) = f'$, $g(s_i,\cdot) = g'$ and $h(s_i,\cdot) = h'$. Then, the definition (11) of $\succ_i'$ shows that $f \succ_i g$ and $g \succ_i h$. Therefore, A3 implies that there exists $\alpha \in (0,1)$ such that $\alpha f + (1-\alpha)h \succ_i g$. Finally, (12) shows that $\alpha f' + (1-\alpha)h' \succ_i' g'$ because $(\alpha f + (1-\alpha)h)(s_i,\cdot) = \alpha f(s_i,\cdot) + (1-\alpha)h(s_i,\cdot) = \alpha f' + (1-\alpha)h'$. The existence of β can be proven similarly.

(Monotonicity) We first show that for any $y, y' \in Y$, $y \succ_i' y'$ if and only if $y \succ_i y'$. To do this, suppose that $y \succ_i' y'$ and let $f, g \in L_0$ be constant acts such that $(\forall \omega \in \Omega)\ f(\omega) = y$ and $g(\omega) = y'$. Then, the definition (11) of $\succ_i'$ immediately shows that $f \succ_i g$, and hence, that $y \succ_i y'$. Next, suppose that $y \succ_i y'$. Then, it follows that $y \succ_i' y'$ from (12) by letting f and g there be the constant acts defined above.

We now turn to the proof of monotonicity. Let $f', g' \in L_0'$ be such that $(\forall t \in T)\ f'(t) \succeq_i' g'(t)$. We need to show that $f' \succeq_i' g'$. To do this, let $f, g \in L_0$ be such that $(\forall s \in S)\ f(s,\cdot) = f'$

and $g(s,\cdot) = g'$. Then, from the assumption that $(\forall t)\ f'(t) \succeq_i' g'(t)$, it follows that $(\forall \omega \in \Omega)\ f(\omega) \succeq_i' g(\omega)$, and hence, that $(\forall \omega \in \Omega)\ f(\omega) \succeq_i g(\omega)$ by the claim proven in the previous paragraph. Therefore, A4 implies that $f \succeq_i g$, which in turn implies that $f' \succeq_i' g'$ by the definition (11) of $\succ_i'$.

(Non-degeneracy) We need to show that there exist $f', g' \in L_0'$ such that $f' \succ_i' g'$. To do this, note that there exist $y^*, y_* \in Y$ such that $y^* \succ_0 y_*$ (Step 1). Then, A7 implies that $y^* \succ_i y_*$, and hence, the first paragraph of the proof of monotonicity shows that $y^* \succ_i' y_*$, which completes the proof.

Step 4. By the previous step, we may invoke Schmeidler's Theorem (1989, p.578) to conclude that for each $i \in \{1, \ldots, m\}$, there exist a unique probability capacity θ_{s_i} on T and an affine function $u_i : Y \to \mathbb{R}$, which is unique up to a positive affine transformation, such that

$$(\forall f, g \in L_0') \quad f' \succ_i' g' \ \Leftrightarrow\ \int_T u_i(f'(t))\, \theta_{s_i}(dt) > \int_T u_i(g'(t))\, \theta_{s_i}(dt)\,. \tag{14}$$

In the rest of this step, we prove (2) in the Theorem. Fix $i \in \{1, \ldots, n\}$ arbitrarily. First, we show that

$$(\forall f, g \in L_0) \quad f \succ_i g \ \Leftrightarrow\ \int_T u_i(f(s_i, t))\, \theta_{s_i}(dt) > \int_T u_i(g(s_i, t))\, \theta_{s_i}(dt)\,. \tag{15}$$

To show ($\Leftarrow$), assume that the right-hand side of (15) holds. Then, by (14), it holds that $f(s_i,\cdot) \succ_i' g(s_i,\cdot)$. Then, by the definition (11) of $\succ_i'$, it follows that $f \succ_i g$. To show ($\Rightarrow$), assume that the right-hand side of (15) does *not* hold. Then, by (14), it holds that $f(s_i,\cdot) \not\succ_i' g(s_i,\cdot)$. If $f \succ_i g$ holds, it must hold that $f(s_i,\cdot) \succ_i' g(s_i,\cdot)$ by (12), which contradicts the asymmetry of $\succ_i'$ which was established in (Ordering) of Step 3.

Second, note that $y^* \succ_i y_*$ since $y^* \succ_0 y_*$ by the definitions of y^* and y_* (Step 1) and since $y^* \succ_0 y_* \ \Leftrightarrow\ y^* \succ_i y_*$ by A7. Therefore, by making a suitable positive affine transformation on u_i, we may assume without loss of generality that $u_i(y^*) = 1$ and $u_i(y_*) = -1$. Since two affine functions which intersect at two distinct points coinside, it follows that $(\forall y \in Y)\ u_i(y) = u(y)$ by (8), where $u : Y \to \mathbb{R}$ is an affine function defined in Step 1. Therefore, we conclude that there exist a unique class of probability capacity $\langle \theta_{s_i} \rangle_{i=1}^m$ on T and an affine function $u : Y \to \mathbb{R}$

such that

$$(\forall f, g \in L_0) \quad f \succ_i g \;\Leftrightarrow\; \int_T u(f(s_i, t))\, \theta_{s_i}(dt) > \int_T u(g(s_i, t))\, \theta_{s_i}(dt)\,, \tag{16}$$

which completes the proof of (2) in the Theorem.

Step 5. Let $K \subseteq \mathbb{R}$ be defined by $K \equiv u(Y)$. Note that K is convex by the affinity of u and $[-1, 1] \subseteq K$ by (8). We denote by $B_0(K)$ the space of K-valued simple functions on S. Two elements, a and b, of $B_0(K)$ are said to be *comonotonic* if $(\forall s, s' \in S)$ $(a(s)-a(s'))(b(s)-b(s')) \geq 0$. Given $f \in L_0$, a mapping defined by

$$s \mapsto \int_T u(f(s, t))\, \theta_s(dt)$$

is an element of $B_0(K)$. Furthermore, for any element a of $B_0(K)$, there exists a 1st-period-measurable act $f \in L_0$ such that

$$(\forall s \in S) \quad a(s) = \int_T u(f(s, t))\, \theta_s(dt) = u(f(s)) \tag{17}$$

(recall that we may write the outcome of a 1st-period-measurable act as $f(s)$ instead of $f(s, t)$).

We define a functional $I : B_0(K) \to \mathbb{R}$ by

$$(\forall a \in B_0(K)) \quad I(a) = U(f)\,,$$

where $f \in L_0$ is an act which satisfies

$$(\forall s \in S) \quad a(s) = \int_T u(f(s, t))\, \theta_s(dt)\,.$$

Such an act certainly exists by (17). In the rest of this paragraph, we show that I is well-defined. To this end, let $a \in B_0(K)$ and let $f, g \in L_0$ be any pair of acts such that

$$(\forall i) \quad a(s_i) = \int_T u(f(s_i, t))\, \theta_{s_i}(dt) = \int_T u(g(s_i, t))\, \theta_{s_i}(dt)\,.$$

Then, (16) implies that $(\forall i)$ $f \sim_i g$, which in turn implies that $f \sim_0 g$ by A8. We thus conclude that $U(f) = U(g)$ by (7).

By the definition of I, we have

$$(\forall f \in L_0) \quad U(f) = I\left(\int_T u(f(\cdot, t))\, \theta_{\cdot}(dt)\right)\,. \tag{18}$$

In particular, when f is 1st-period-measurable, we have

$$U(f) = I\left(u(f(\cdot))\right)\,. \tag{19}$$

Step 6. In this step, we show that the functional I defined in the previous step satisfies all the assumptions of Corollary of Schmeidler (1986, p.258), which proves that I can be represented as

$$(\forall a \in B_0(K)) \quad I(a) = \int_S a(s)\, \theta_0(ds)$$

with some probability capacity θ_0 on S. This, (18) and (7) complete the proof of (1) in the Theorem.

(Positive Homogeneity) Let $\lambda \in K$. We need to show that $I(\lambda\chi_S) = \lambda$. To do this, let $y \in Y$ be an outcome such that $u(y) = \lambda$ and let $f \in L_0$ be a 1st-period-measurable act such that $(\forall s)\ f(s) = y$. Then,

$$I(\lambda\chi_S) = I\left(u(f(\cdot))\right) = U(f) = u(y) = \lambda\,,$$

where the second equality holds by (19) and the third equality holds by (10).

(Comonotonic Independence) Let $a, b, c \in B_0(K)$ be pairwise comonotonic (see the first paragraph of Step 5). We need to show that for any $\alpha \in (0, 1)$,

$$I(a) > I(b) \;\Rightarrow\; I(\alpha a + (1-\alpha)c) > I(\alpha b + (1-\alpha)c)\,. \tag{20}$$

To do this, let f, g, h be 1st-period-measurable acts such that $(\forall s)\ u(f(s)) = a(s)$, $u(g(s)) = b(s)$ and $u(h(s)) = c(s)$. Such f, g and h certainly exist by (17). Then, f and g are comonotonic because for any $s, s' \in S$,

$$\begin{aligned} & f(s) \succ_0 f(s') \;\Leftrightarrow\; f(s) \succ_i f(s') \;\Leftrightarrow\; u(f(s)) > u(f(s')) \;\Leftrightarrow\; a(s) > a(s') \\ \Rightarrow\;\; & b(s) \not< b(s') \;\Leftrightarrow\; u(g(s)) \not< u(g(s')) \;\Leftrightarrow\; g(s) \not\prec_i g(s') \;\Leftrightarrow\; g(s) \not\prec_0 g(s')\,, \end{aligned}$$

where the first and last equivalences hold by A7; the second and fifth equivalences hold by (16); and the implication holds by the comonotonicity between a and b. Similarly, the other pairs among f, g and h are comonotonic. Therefore, (20) holds because

$$\begin{aligned} I(a) > I(b) \;&\Rightarrow\; I\left(u(f(\cdot))\right) > I\left(u(g(\cdot))\right) \\ &\Rightarrow\; U(f) > U(g) \\ &\Rightarrow\; U(\alpha f + (1-\alpha)h) > U(\alpha g + (1-\alpha)h) \end{aligned}$$

$$\Rightarrow \quad I\left(u(\alpha f(\cdot)+(1-\alpha)h(\cdot))\right) > I\left(u(\alpha g(\cdot)+(1-\alpha)h(\cdot))\right)$$

$$\Rightarrow \quad I\left(\alpha u(f(\cdot))+(1-\alpha)u(h(\cdot))\right) > I\left(\alpha u(g(\cdot))+(1-\alpha)u(h(\cdot))\right)$$

$$\Rightarrow \quad I\left(\alpha a+(1-\alpha)c\right) > I\left(\alpha a+(1-\alpha)c\right)\ ,$$

where the second and fourth implications hold by (19); the third implication holds by (7), the assumption that f, g, h are 1st-period-measurable, the fact that they are pairwise comonotonic (proven above) and A2(0); and the fifth implication holds by the affinity of u.

(Monotonicity) Let $a, b \in B_0(K)$ be such that $a \geq b$. We need to prove that $I(a) \geq I(b)$. To do this, let f and g be 1st-period-measurable acts such that $(\forall s)\ u(f(s)) = a(s)$ and $u(g(s)) = b(s)$. Such f and g certainly exist by (17). Then, $I(a) \geq I(b)$ holds because

$$a \geq b \ \Leftrightarrow\ (\forall s)\ u(f(s)) \geq u(g(s)) \ \Leftrightarrow\ (\forall s)\ f(s) \succeq_0 g(s)$$

$$\Rightarrow \quad f \succeq_0 g \Leftrightarrow U(f) \geq U(g) \ \Leftrightarrow\ I(u(f(\cdot))) \geq I(u(g(\cdot))) \ \Leftrightarrow\ I(a) \geq I(b)\,,$$

where the second equivalence holds by (10) and (7); the implication holds by A4; the third equivalence holds by (7); and fourth equivalence holds by (19). ■

REFERENCES

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

Dempster, A. P. (1967): "Upper and Lower Probabilities Induced by A Multivalued Map," *Annals of Mathematical Statistics* 38, 325-339.

Dempster, A. P. (1968): "A Generalization of Bayesian Inference," *Journal of Royal Statistics Society, Series B* 30, 205-247.

Epstein, L. G. and M. Le Breton (1993): "Dynamically Consistent Beliefs Must Be Bayesian," *Journal of Economic Theory* 61, 1-22.

Ghirardato, P. (2002): "Revisiting Savage in A Conditional World," *Economic Theory* 20, 83-92.

Gilboa, I. and D. Schmeidler (1993): "Updating Ambiguous Beliefs," *Journal of Economic Theory* 59, 33-49.

Herstein, I. N. and J. Milnor (1953): " An Axiomatic Approach to Measurable Utility," *Econometrica* 21, 291-297.

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

Nishimura, K. G., and H. Ozaki (2001): "Search and Knightian Uncertainty," Discussion Paper, University of Tokyo.

Nishimura, K. G., and H. Ozaki (2002): "Economics of Self-Feeding Fear," Discussion Paper, University of Tokyo.

Schmeidler, D. (1986): "Integral Representation without Additivity," *Proceedings of the American Mathematical Society* 97, 255-261.

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

Shafer G. (1976): *A Mathematical Theory of Evidence*, Princeton University Press, Princeton.

Wang, T. (2002): "Conditional Preferences and Updating," mimeo, University of British Columbia, forthcoming in *Journal of Economic Theory*.

Yoo, K-R. (1991): "The Iterative law of Expectation and Non-additive Probability Measure," *Economics Letters* 37, 145-149.