# SUBJECTIVE ERROR MEASURE

by Hiroyuki Ozaki*

Faculty of Economics
Keio University

May 9, 2013

**Abstract**

The decision-maker who complies with Savage's axioms can be regarded as evaluating each act by solving a statistical inference problem, in which the estimation error is measured by the squared-error loss function. However, it is more desirable to derive a loss function, as well as the utility index and the subjective probability, from the decision-maker's own preference. We weaken Savage's axioms to characterize the preference which is based on the loss-minimization in the sense that the decision-maker evaluates each act by solving a statistical inference problem, in which the estimation error is measured by *some* loss function which may be different from the squared-error loss function. Our results build on the representation theorem proved by Grant, Kajii and Polak (2000) which characterizes the preferences which satisfy the betweenness property. Also, we provide two examples of preferences which are based on the loss-minimization with the loss function more general than the squared-error loss functions and we discuss conditional preferences for a class of preferences considered in this paper.

*e-mail address: ozaki@econ.keio.ac.jp

## 1. Introduction

A celebrated theorem of Savage (1954) has been underlying much of economic modeling and analysis. It states that if a decision-maker (DM)'s preference relation complies with some set of axioms, she behaves as if she chooses an act which maximizes the expected utility calculated by means of some utility index and some subjective probability. As we will see closely in Section 2.4, such a behavior of the DM may be also described as that of evaluating each act by solving a statistical inference problem where she estimates the "true" utility value without knowing which state has actually occurred. Most importantly, in this inference problem, she is supposed to seek the "best" estimate, which is the best in the sense of minimizing the estimation error measured by the *squared-error* loss function. Savage's theorem then claims that such a behavior of the DM is rational in that it is completely characterized by some reasonable axioms.

While Savage's theory certainly has an appeal as such, an analogy to an inference problem reveals its limitation also. The theory takes it as granted that the DM uses the estimate which is the best in the sense of minimizing the estimation error defined via the squared-error loss function. However, the loss function may be naturally understood to measure her "psychological" distance between the estimand and the estimate. Therefore, similarly to the utility index and the subjective probability, the loss function itself *is* a part of her preference, and hence, it should be derived from the axioms. Various loss functions are proposed in the statistics literature.[1] Which of these or other loss functions will be employed by the DM should be determined from her attitude toward uncertainty, rather than is assumed to be given from the very outset.

Thus motivated, this paper weakens Savage's axioms (mainly, his sure-thing principle) to derive a new class of preferences, which we call the *preference based on the loss-minimization*. The DM with such a preference evaluates each act by solving a statistical inference problem defined by means of the utility index, the subjective probability and the loss function which are all derived from her own preference. It is quite often the case that an inference problem in the statistics literature is described with a general loss function which may be different from the squared-error loss function. Therefore, we may conclude that many inference problems in

[1]These includes the absolute-error loss function and the weighted squared-error loss function. For these and others, see, for example, Lehmann and Casella (1998).

the literature are relevant to the behavior of the DM whoes preference is based on the loss-minimization.

In the next section, we first show that some set of axioms is sufficient for the DM's preference to be based on the loss-minimization, that is, the DM evaluates each act by solving a statistical inference problem in which the estimation error is measured by *some* loss function. We then show that its converse also holds under an additional condition on the loss function: Any preference which is based on the loss-minimization must satisfies the above axioms. These results constitute a main contribution of the paper. Our proofs builds on the representation theorem proved by Grant, Kajii and Polak (2000), which characterizes the preferences under uncertainty which satisfy the betweenness property.[2]

By allowing loss functions which differ from the squared-error loss function, our representation makes it possible to model some aspects of a preference which are difficult to be analyzed within Savage's framework of expected utility. Section 3 provides two examples of loss functions, each of which defines a loss-minimization problem on which the preference is based in the abvoe sense. The first example is the $L^q$-loss function, which, together with a subjective probability $\mu$ and a utility index $v$, generates a one-parameter family of preferences. We argue that the parameter determines whether the DM makes much of probability values governed by $\mu$ or utility values governed by $v$. The second example is the asymmetric loss function, which, together with $\mu$ and $v$, generates a two-parameter family of preferences. The asymmetric loss function measures the estimation error asymmetrically between the underestimate and the overestimate. We then argue that under some specification of parameter values, the preference in this family could exhibit the disappointment aversion. The preferences in the both families violate Savage's sure-thing principle and thus are unable to be represented by the expected utility.

Section 4 turns to the inference problem with an observation opportunity. We argue that solving such an inference problem leads to a natural definition of a class of conditional preferences. When the inference problem is defined in terms of the squared-error loss function, this procedure of updating *ex-ante* preference into a class of conditional preferences exactly

[2]The preferences under risk which satisfy the betweenness property were studied by Chew (1983, 1989) and Dekel (1986).

corresponds to updating the expected utility to the conditional expected utility with the prior subjective probability revised according to Bayes' rule. Therefore, we may argue that our definition of conditional preferences substantially extends the conditional expected utility in Savage's framework because it is derived by solving statistical inference problems in which a loss function may be quite different from the squared-error loss function.

Finally, Section 5 contains lemmas and proofs.

## 2. Axiomatization

The Introduction mentioned that the DM who complies with Savage's axioms can be seen as solving a suitably-defined inference problem, in which the estimation error is measured by the squared-error loss function. This section weakens Savage's axioms to those which are sufficient for the DM's preference to be based on the loss-minimization, in which the estimation error is measured by *some* loss function which may be distinct from the squared-error loss function.

### 2.1. The Axioms

This subsection introduces a set of axioms. Following Savage's framework, let $S$ be the set of states, let $\mathcal{E} \equiv 2^S$ be the set of events and let $X$ be the set of outcomes. Throughout the paper, we assume that $X$ is a compact metric space which is also connected. An *act* is a function from $S$ into $X$. We denote by $F_0$ the set of all simple acts. Here, we say that an act $f$ is *simple* if the image of $S$ under $f$, $f(S) \equiv \{\, x \in X \mid (\exists s \in S)\ f(s) = x \,\}$, is a finite set. We consider only simple acts in this paper, and call them "acts" without the adjective. We take as the primitive the DM's preference relation $\succeq$ on $F_0$, and axioms will be imposed on $\succeq$.

To state the axioms, we introduce some notations, which are standard in the literature. Derive the asymmetry part $\succ$ and the indifference part $\sim$ from $\succeq$.[3] An event $E \in \mathcal{E}$ is *null* (with respect to $\succeq$) if any pair of acts which differ only on $E$ are indifferent to each other. We use the same symbol, $\succeq$, to denote the preference relation induced on $X$ from $\succeq$, that is, $(\forall x, y \in X)\ x \succeq y$ if $f \succeq g$ where $f$ and $g$ are *constant acts* such that $(\forall s)\ f(s) = x$ and $g(s) = y$.

[3] $(\forall f, g \in F_0)\ f \succ g \Leftrightarrow g \not\succeq f$; $f \sim g \Leftrightarrow [f \succeq g$ and $g \succeq f]$.

We thus identify an outcome with the constant act taking on that outcome. We frequently employ the notation such as

$$\begin{bmatrix} x & \text{on} & A \\ f & \text{on} & A^c \end{bmatrix}$$

to denote the act which takes on $x \in X$ on an event $A$ and equals $f \in F_0$ when restricted to $A^c$, the complement of $A$ in $S$. Similar notations apply in obvious manners.

The axioms we consider are as follows:

**P1** (Ordering) *The preference relation $\succeq$ is complete and transitive.*

**P3** (Eventwise Monotonicity) *For any outcome $x$ and $y$, any non-null event $E$, and any act $f$,*

$$\begin{bmatrix} x & \text{on} & E \\ f & \text{on} & E^c \end{bmatrix} \succeq \begin{bmatrix} y & \text{on} & E \\ f & \text{on} & E^c \end{bmatrix} \Leftrightarrow x \succeq y$$

**P4$^{\mathbf{c}}$** (Conditional Weak Comparative Probability) *For any outcome $x, x', y, y'$ such that $x \succ x'$ and $y \succ y'$, any event $A$, $B$ and $T$ such that $A \cup B \subseteq T$, and any act $g$,*

$$\begin{bmatrix} x & \text{on} & A \\ x' & \text{on} & T \backslash A \\ g & \text{on} & T^c \end{bmatrix} \succeq \begin{bmatrix} x & \text{on} & B \\ x' & \text{on} & T \backslash B \\ g & \text{on} & T^c \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} y & \text{on} & A \\ y' & \text{on} & T \backslash A \\ g & \text{on} & T^c \end{bmatrix} \succeq \begin{bmatrix} y & \text{on} & B \\ y' & \text{on} & T \backslash B \\ g & \text{on} & T^c \end{bmatrix}$$

**P5** (Nondegeneracy) *There exist constant acts $x$ and $x'$ such that $x \succ x'$.*

**P6** (Small Event Continuity) *For any act $f$ and $g$ such that $f \succ g$ and any outcome $x$, there exists a finite partition of $S$, $\langle E_i \rangle_{i=1}^n$, such that*

$$(\forall i, j \in \{1, 2, \ldots, n\}) \quad f \succ \begin{bmatrix} x & \text{on} & E_i \\ g & \text{on} & (E_i)^c \end{bmatrix} \text{ and } \begin{bmatrix} x & \text{on} & E_j \\ f & \text{on} & (E_j)^c \end{bmatrix} \succ g$$

**A1** (Best and Worst Acts) *There exist constant acts $x^*$ and $x_*$ such that* $(\forall f \in F_0)$ $x^* \succeq f \succeq x_*$.

**A2** (Weak Decomposability) *For any event $A$ and any act $f$ and $g$,*

$$\begin{bmatrix} g & \text{on} & A \\ f & \text{on} & A^c \end{bmatrix} \succ f \text{ and } \begin{bmatrix} f & \text{on} & A \\ g & \text{on} & A^c \end{bmatrix} \succ f \Rightarrow g \succ f$$

**A3** (Continuity on $X$) *For any outcome $x$, $\{ y \in X \mid y \succeq x \}$ and $\{ y \in X \mid x \succeq y \}$ are closed.*

Note that each of Axioms P1, P3, P5 and P6 is identical to each of Machina and Schmeidler's (1992) axioms with the same name. Axiom P4$^c$ is introduced by Epstein and Le Breton (1993) in the context of conditional preferences and dynamic consistency and it is closely related to the axiom of strong comparative probability (so-called Axiom P4*) by Machina and Schmeidler (1992). Axiom A1 postulates that there should exist the best and worst constant acts. Axiom A2 is called weak decomposability and introduced by Grant, Kajii and Polak (2000). Finally, Axiom A3 is a weak form of the continuity assumption and it only requires that $\succeq$ restricted to $X$ be continuous in the given topology on $X$.

### 2.2. Main Theorems

In this subsection, we characterize a class of preferences whose representation is given by a unique solution to some suitably-defined (no-data) inference problem. We first define an error function formally. That is, a function $\phi : [0,1] \times [0,1] \to \mathbb{R}$ is an *error function* if it satisfies

E1. $(\forall u, v)\ \ \phi(v,u) \geq 0$ ,

E2. $(\forall u, v)\ \ \phi(v,u) = 0 \ \Leftrightarrow\ u = v$ ,

E3. $(\forall v)\ \ \phi(v,\cdot)$ is continuously differentiable ,

E4. $(\forall u)\ \ \phi_2(\cdot,u)$ is strictly decreasing and

E5. $(\forall v)\ \ \phi_2(v,v) = 0$ ,

where $\phi_2$ denotes $\phi$'s partial derivative with respect to its second argument. Intuitively, an error function measures a distance between two numbers with zero distance occurring if and only if two numbers coincide (E1 and E2). In this regard, it may be considered as an analogue of a metric (in a metrizable topological space). We also impose on an error function three technical assumptions (E3-E5).

We say that a preference relation $\succeq$ on $F_0$ is *based on the loss-minimization* if there exist a unique finitely-additive, non-atomic[4] probability measure $\mu$ on $(S, \mathcal{E})$, a continuous utility

[4]According to Machina and Schmeidler (p.751, footnote 12), we call a finitely-additive probability measure $\mu$ *non-atomic* if $(\forall B \in \mathcal{E})(\forall \rho \in [0,1])(\exists C \subseteq B)\ \mu(C) = \rho\mu(B)$.

index $v : X \to [0,1]$ satisfying $(\exists x^*, x_* \in X)\ v(x^*) = 1,\ v(x_*) = 0$ and $(\forall x, x' \in X)\ x \succeq x' \Leftrightarrow v(x) \geq v(x')$ and an error function $\phi : [0,1] \times [0,1] \to \mathbb{R}$, that are altogether such that $(\forall f, g \in F_0)\ f \succeq g \Leftrightarrow U(f) \geq U(g)$, where[5]

$$(\forall f \in F_0) \quad U(f) = \operatorname*{argmin}_{u \in [0,1]} \int_S \phi(v \circ f(s), u)\, d\mu(s)\,. \tag{1}$$

We now claim that the next theorem holds.

**Theorem 1** *Suppose that a preference relation $\succeq$ on $F_0$ satisfies Axioms P1, P3, P4$^c$, P5, P6, A1, A2 and A3. Then, $\succeq$ is based on the loss-minimization.*

The theorem shows that a DM whose preference complies with the given set of axioms evaluates each act by solving an inference problem in which estimation errors are measured by the loss function which itself is derived from her own preference.

A "converse" of Theorem 1 is given by the next theorem, which is not exactly the converse because the assumption E6 imposed on $\phi$ in the theorem is not derived in Theorem 1.

**Theorem 2** *Suppose that the preference relation $\succeq$ on $F_0$ is based on the loss-minimization with an error function which satisfies*

E6. $(\forall v)\ \phi(v, \cdot)$ *is strictly convex*

*as well as E1-E5. Then, $\succeq$ satisfies P1, P3, P4$^c$, P5, P6, A1, A2 and A3.*

Consider a DM who, in evaluating an act $f$, solves a statistical inference problem composed of a subjective probability $\mu$, a utility index $v$ and an error function $\phi$, which are primitives characterizing the DM's preference. Also, suppose that $\phi$ satisfies E6. Then, Theorem 2 shows that the DM's behavior must be based upon a preference which complies with the given set

[5]Formally, the right-hand side of (1) is a set. When it is a singleton, we write as in the theorem, instead of writing as

$$U(f) \in \operatorname*{argmin}_{u \in [0,1]} \int_S \phi(v \circ f(s), u)\, d\mu(s)\,.$$

The same convention applies in what follows. The right-hand side of (1) may be suggestively written as $\operatorname{argmin}_{u \in [0,1]} E^{\mu}\left[\phi\left(v \circ f, u\right)\right]$.

of axioms. Note that the strict convexity of an error function in its second argument is often assumed in the statistics literature to guarantee the existence of estimators.[6] We call an error function satisfying E1-E6 a *convex* error function.

Both Theorems 1 and 2 show that the preference relation represented by the solution to some statistical inference problem has a sound behavioral foundation based on the set of axioms given in the theorems.

The proofs of our theorems build on the representation theorem proved by Grant, Kajii and Polak (2000), which characterizes the preference under uncertainty which satisfies what is called the betweenness property. We turn to this subject in the next subsection.

### 2.3. Betweenness-satisfying Preferences

We say that a preference relation $\succeq$ on $F_0$ is *betweenness-satisfying* if there exist a unique finitely-additive, non-atomic probability measure $\mu$ on $(S, \mathcal{E})$, a (not necessarily continuous) utility index $v : X \to [0,1]$ satisfying $(\exists x^*, x_* \in X)\ v(x^*) = 1,\ v(x_*) = 0$ and $(\forall x, x' \in X)\ x \succeq x' \Leftrightarrow v(x) \geq v(x')$, and a function $\varphi : v(X) \times [0,1] \to \mathbb{R}$ satisfying

B1. $(\forall u \in [0,1])\ \ \varphi(\cdot, u)$ is strictly increasing and

B2. $(\forall v \in v(X))\ \ \varphi(v, \cdot)$ is continuous,

that are altogether such that $(\forall f, g \in F_0)\ f \succeq g \ \Leftrightarrow\ U(f) \geq U(g)$, where for any $f \in F_0$, $u = U(f)$ is the unique solution to the following equation in $u$ :

$$\int_S \varphi(v \circ f(s), u)\, d\mu(s) = 0\,. \tag{2}$$

A function $\varphi$ which satisfies B1 and B2 in the definition is called a *betweenness function*. The preferences under risk which satisfy the betweenness property are studied and axiomatized by Chew (1983, 1989) and Dekel (1986). Grant, Kajii and Polak (2000) studied the betweenness-satisfying preferences under uncertainty and proved the following representation theorem.

**Theorem 3 (Grant, Kajii and Polak, 2000)** *A preference relation $\succeq$ on $F_0$ satisfies Axioms P1, P3, P4$^c$, P5, P6, A1 and A2 if and only if it is betweenness-satisfying.*

[6]For example, see Lehmann and Casella (1998, p.88, Theorem 1.11) which establishes the existence of the uniform minimum variance unbiased (UMVU) estimator.

Let $\succeq$ be a preference based on the loss-minimization with a convex error function. Then, $\succeq$ satisfies all the axioms of Theorem 3 by Theorem 2, and hence, it is betweenness-satisfying.[7] Such a preference further specifies the betweenness-satisfying preference by imposing an additional axiom of A3, the continuity of the preference on $X$.

To see the role played by Axiom A3, suppose that all the axioms in Theorem 3 are satisfied. Then, a seemingly mild axiom of A3 guarantees the existence of a betweenness function which satisfies the condition B3:

$$\text{B3.} \quad (\forall u, v) \;\; \varphi(v, u) = 0 \;\Leftrightarrow\; u = v$$

as well as B1 and B2 (Lemma 1 in Section 5), which in turn allows us to construct an error function satisfying E1-E5 (see the proof of Theorem 1 in Section 5).

It follows from (2) (see also the equation (9) in the Appendix) that the DM's preference over the probability measures induced on $X$ by $\mu$ and $f$ is the same as the one axiomatized by Chew (1983, 1989) and Dekel (1986) in the framework of risk. They relax the independence axiom of von Neumann-Morgenstern's expected-utility theory and only require that the indifference sets be convex, the so-called *betweenness property*. It is well-known that such a preference admits Allais-type behavior.

### 2.4. The Squared-error Loss Function: Savage's Theory Revisited

Consider the following axioms which are introduced by Savage (1954).

**P2** (Sure-thing Principle) *For any act $f, f', g, g'$ and any event $A$,*

$$\begin{bmatrix} f & \text{on} & A \\ g & \text{on} & A^c \end{bmatrix} \succeq \begin{bmatrix} f' & \text{on} & A \\ g & \text{on} & A^c \end{bmatrix} \Leftrightarrow \begin{bmatrix} f & \text{on} & A \\ g' & \text{on} & A^c \end{bmatrix} \succeq \begin{bmatrix} f' & \text{on} & A \\ g' & \text{on} & A^c \end{bmatrix}$$

**P4** (Weak Comparative Probability) *For any outcome $x, x', y, y'$ such that $x \succ x'$ and $y \succ y'$ and any event $A$ and $B$,*

$$\begin{bmatrix} x & \text{on} & A \\ x' & \text{on} & A^c \end{bmatrix} \succeq \begin{bmatrix} x & \text{on} & B \\ x' & \text{on} & B^c \end{bmatrix} \Leftrightarrow \begin{bmatrix} y & \text{on} & A \\ y' & \text{on} & A^c \end{bmatrix} \succeq \begin{bmatrix} y & \text{on} & B \\ y' & \text{on} & B^c \end{bmatrix}$$

[7] There may exist a preference which is based on the loss-minimization with a *non*-convex error function and is *not* betweenness-satisfying.

Savage (1954) proved essentially the following result (also, see Fishburn, 1970):

**Theorem 4 (Savage, 1954)** *A preference relation $\succeq$ on $F_0$ satisfies Axioms P1, P2, P3, P4, P5, P6, A1 and A3 if and only if there exist a unique finitely-additive, non-atomic probability measure $\mu$ on $(S, \mathcal{E})$ and a continuous utility index $v : X \to [0,1]$ satisfying $(\exists x^*, x_* \in X)\ v(x^*) = 1$ and $v(x_*) = 0$ such that $f \succeq g \Leftrightarrow U(f) \geq U(g)$, where*

$$(\forall f \in F_0) \quad U(f) = \operatorname*{argmin}_{u \in [0,1]} \int_S (v \circ f(s) - u)^2 \, d\mu(s)\,. \tag{3}$$

The minimization problem (3) can be explicitly solved for $u$ to obtain

$$(\forall f \in F_0) \quad U(f) = \int_S v \circ f(s) \, d\mu(s)\,. \tag{4}$$

Therefore, the representation in Theorem 4 turns out to be equivalent to the one by the *expected utility*, which gives a more familiar form of Savage's theorem.

Theorem 4 provides a further specification of the representation in Theorems 1 and 2 since it specifies the error function $\phi$ in those theorems by $\phi^2$, which is defined by $(\forall u, v)\ \phi^2(v, u) \equiv (v - u)^2$. Note that $\phi^2$ satisfies E6 as well as E1-E5. Savage's theorem (Theorem 4) thus characterizes the DM's behavior who, in evaluating each act, solves a statistical inference problem by using the *squared-error* loss function.

The preference based on the loss-minimization characterized by Theorems 1 and 2 substantially extends the expected utility characterized by Theorem 4. Since the requirement that $(\forall x, x' \in X)\ x \succeq x' \Leftrightarrow v(x) \geq v(x')$ holds because $(\forall x)\ v(x) = U(x)$ by (4), the preference represented by the expected utility is based on the loss-minimization by definition. On the other hand, the next section provides some examples of preferences based on the loss-minimization which do *not* satisfy the sure-thing principle (Axiom P2). Therefore, the preferences based on the loss-minimization define a strictly larger class of preferences than the expected utility, each of which can be described in terms of statistical inference problems defined with some (not necessarily squared-error) loss function.

## 3. Examples

This section considers some preferences based on the loss-minimization which violate the sure-thing principle. They are defined by means of loss functions which are different from the squared-error loss function and seem to be interesting in that they make it possible to model some aspects of preferences which are difficult to be analyzed within the framework of expected utilities.

### 3.1. The $L^q$-Error Function

Let $\mu$ be a finitely-additive, non-atomic probability measure on $(S, \mathcal{E})$ and let $v : X \to [0,1]$ be a continuous function such that $(\exists x^*, x_* \in X)\ v(x^*) = 1$ and $v(x_*) = 0$. For any $q \in (1, +\infty)$, define the *$L^q$-error function* $\phi^q : [0,1] \times [0,1] \to \mathbb{R}$ by $(\forall u, v)\ \phi^q(v, u) = |v - u|^q$. Note that $\phi^q$ satisfies E1-E6.

Now, define the preference relation $\succeq^q$ on $F_0$ by $(\forall f, g)\ f \succeq^q g \Leftrightarrow U(f) \geq U(g)$, where for any $f \in F_0$, $U(f)$ is defined as a unique solution $u$ to

$$\min_{u \in [0,1]} \int_S \phi^q \left(v \circ f(s), u\right)\, d\mu(s)\,. \tag{5}$$

Since $U(f)$ is well-defined and the requirement that $(\forall x, x' \in X)\ x \succeq^q x' \Leftrightarrow v(x) \geq v(x')$ holds because $(\forall x)\ v(x) = U(x)$, $\succeq^q$ is well-defined and based on the loss-minimization by definition.

When $q = 2$, $\succeq^q$ $(=\succeq^2)$ satisfies Axiom P2 and it is represented by the expected utility. However, $\succeq^q$ need *not* satisfy P2 unless $q = 2$ as shown by the next example.

**Example 1** Let $q > 1$, let $z \in X$ be an outcome such that $v(z) = 1/2$ and let $\langle E_1, E_2, E_3 \rangle$ be a partition of $S$ such that $\mu(E_1) = \mu(E_2) = 1/4$ and $\mu(E_3) = 1/2$. Define an act $f$ by

$$f = \begin{bmatrix} x_* & \text{on} & E_1 \\ x^* & \text{on} & E_2 \cup E_3 \end{bmatrix}.$$

Then, it follows that

$$\begin{bmatrix} z & \text{on} & E_1 \cup E_2 \\ z & \text{on} & E_3 \end{bmatrix} \sim^q \begin{bmatrix} f & \text{on} & E_1 \cup E_2 \\ z & \text{on} & E_3 \end{bmatrix}$$

and that

$$g_1 \equiv \begin{bmatrix} z & \text{on} & E_1 \cup E_2 \\ x^* & \text{on} & E_3 \end{bmatrix} \succ^q \begin{bmatrix} f & \text{on} & E_1 \cup E_2 \\ x^* & \text{on} & E_3 \end{bmatrix} \equiv g_2 \;\Leftrightarrow\; q > 2\,.$$

Thus, Axiom P2 is violated whenever $q \neq 2$.

In this example, the act $g_2$ is "better" than the act $g_1$ in the sense that $g_2$ has a higher probability of getting the best outcome than $g_1$. On the other hand, $g_2$ is "worse" than $g_1$ in the sense that $g_2$ contains in its range the worst outcome while $g_1$ does not. This suggests that the parameter $q$ determines whether the DM makes much of probability values governed by $\mu$ or utility values governed by $v$. When $q$ becomes smaller and closer to 1, the probability values become relatively more important to her, and hence $g_2$ becomes more preferred to $g_1$. On the contrary, when $q$ becomes larger and closer to $+\infty$, the utility values become relatively more important, and hence $g_2$ becomes *less* preferred to $g_1$.

The point made in the previous paragraph will be highlighted if we consider an extreme case. Let $\langle E_1, E_2\rangle$ be a partition of $S$ such that $(\forall i)\ \mu(E_i) > 0$, let $x_1, x_2 \in X$ be outcomes, let $f$ be an act defined by

$$f = \begin{bmatrix} x_1 & \text{on} & E_1 \\ x_2 & \text{on} & E_2 \end{bmatrix},$$

and let $U$ be a representation of $\succeq^q$ as defined by (5). Then, when $q = 1$, $U(f) = v(x_i)$, where $i = \arg\max_{j=1,2} \mu(E_j)$, and the probability values are decisive in evaluating the act.[8] On the other hand, when $q = +\infty$, $U(f) = (v(x_1) + v(x_2))/2$ regardless of $\mu(E_i)$'s, and the utility values are decisive in it.[9]

### 3.2. The Asymmetric Error Function

Let $\mu$ be a finitely-additive, non-atomic probability measure on $(S, \mathcal{E})$ and let $v : X \to [0, 1]$ be a function such that $(\exists x^*, x_* \in X)\ v(x^*) = 1$ and $v(x_*) = 0$. For any $q \in (1, +\infty)$ and any $\gamma > 0$, define the *asymmetric error function* $\phi^{q,\gamma} : [0, 1] \times [0, 1] \to \mathbb{R}$ by

$$(\forall u, v) \quad \phi^{q,\gamma}(v, u) = \begin{cases} \gamma|v-u|^q & \text{if} \quad v \geq u \\ |v-u|^q & \text{if} \quad v < u\ . \end{cases}$$

Note that $\phi^{q,\gamma}$ satisfies E1-E6.

[8]When $\mu(E_1) = \mu(E_2)$, the solution to (5) could be any number between $v(x_1)$ and $v(x_2)$, and hence, $\succeq^q$ is not well-defined. This is why we exclude the case where $q = 1$ at the outset. Hence, the argument in this paragraph stands only heuristically.

[9]When $q = +\infty$, we think of (5) as

$$\min_{u\in[0,1]} \lim_{q\to+\infty} \left( \int_S |v \circ f(s) - u|^q \, d\mu(s) \right)^{1/q}.$$

Again, the argument remains only to be heuristic.

Now, define the preference relation $\succeq^{q,\gamma}$ on $F_0$ by $(\forall f, g)\ f \succeq^{q,\gamma} g \ \Leftrightarrow\ U(f) \geq U(g)$, where for any $f \in F_0$, $U(f)$ is defined as a unique solution $u$ to

$$\min_{u\in[0,1]} \int_S \phi^{q,\gamma}\left(v \circ f(s), u\right)\, d\mu(s)\,.$$

Since $U(f)$ is well-defined and the requirement that $(\forall x, x' \in X)\ x \succeq^{q,\gamma} x' \Leftrightarrow v(x) \geq v(x')$ holds because $(\forall x)\ v(x) = U(x)$, $\succeq^{q,\gamma}$ is well-defined and based on the loss-minimization by definition.

When $\gamma = 1$, $\phi^{q,\gamma}$ is reduced to $\phi^q$ in the previous subsection. Otherwise, $\phi^{q,\gamma}$ treats an underestimate $(u < v)$ and an overestimate $(u > v)$ asymmetrically. When $\gamma < 1$, an overestimate is more penalized than an underestimate since the error caused by the latter is tolerated by $\gamma$ while the error caused by the former is not. On the other hand, when $\gamma > 1$, the underestimate is more penalized than the overestimate.

When $\gamma = 1$ *and* $q = 2$, $\succeq^{q,\gamma}$ $(= \succeq^{2,1} = \succeq^2)$ satisfies Axiom P2 and it is represented by the expected utility. However, $\succeq^{q,\gamma}$ need *not* satisfy P2 unless $\gamma = 1$, *even if* $q = 2$, as shown by the next example.

**Example 2** Let $\gamma > 0$, let $z \in X$ be an outcome such that $v(z) = \gamma/(1+\gamma)$ and let $\langle E_1, E_2, E_3\rangle$ be a partition of $S$ such that $\mu(E_1) = \mu(E_2) = 1/4$ and $\mu(E_3) = 1/2$. Define an act $f$ by

$$f = \begin{bmatrix} x_* & \text{on} & E_1 \\ x^* & \text{on} & E_2 \cup E_3 \end{bmatrix}.$$

Then, it follows that

$$\begin{bmatrix} z & \text{on} & E_1 \cup E_2 \\ z & \text{on} & E_3 \end{bmatrix} \sim^{2,\gamma} \begin{bmatrix} f & \text{on} & E_1 \cup E_2 \\ z & \text{on} & E_3 \end{bmatrix}$$

and that

$$g_1 \equiv \begin{bmatrix} z & \text{on} & E_1 \cup E_2 \\ x^* & \text{on} & E_3 \end{bmatrix} \succ^{2,\gamma} \begin{bmatrix} f & \text{on} & E_1 \cup E_2 \\ x^* & \text{on} & E_3 \end{bmatrix} \equiv g_2 \ \Leftrightarrow\ \gamma > 1\,.$$

Thus, Axiom P2 is violated whenever $\gamma \neq 1$, even if $q = 2$.

Given the act $g_1$, the DM would be "disappointed" with a probability 1/2 if she expected any utility value between two utility values $v \circ g_1$ might assume. Similarly, given the act $g_2$, she would be "disappointed" with a probability 1/4 if she expected any utility value between two utility values $v \circ g_2$ might assume. When $\gamma < 1$, she prefers the act under which she incurs a disappointment with a less probability. Therefore, the example suggests that when $\gamma < 1$, $\succeq^{2,\gamma}$ exhibits a disappointment aversion (also, see Gul, 1991).

## 4. Conditional Preferences

The discussion in Section 2 leads us to a natural definition of conditional preferences. Formally, let $\langle T_i \rangle_{i=1}^n$ be a finite partition of $S$ and let an *estimator* be a function from $S$ into $[0,1]$ which is $\langle T_i \rangle_i$-measurable.[10] We denote the set of all estimators by $B(\langle T_i \rangle_i)$ and its generic element by $\hat{u}$. Suppose that the DM's (unconditional) preference is based on the loss-minimization with some $\mu$, $v$ and $\phi$ such that $(\forall i)\ \mu(T_i) > 0$ (as is characterized by Theorem 1). When the DM observes an event $T_i$, the preference after observing it is called a *conditional preference* and denoted by $\succeq_{T_i}$. We define a class of conditional preferences, $\langle \succeq_{T_i} \rangle_{i=1}^n$, as follows. Let $\hat{U}(f) = (U_1(f), \ldots, U_n(f))$ be such that

$$\hat{U}(f) = \operatorname*{argmin}_{\hat{u} \in B(\langle T_i \rangle_i)} \int_S \phi(v \circ f(s), \hat{u}(s))\, d\mu(s)\,. \tag{6}$$

Here, $f$ is an act to be evaluated and the minimum is assumed to exist uniquely.[11] Then, for each $i$, define $\succeq_{T_i}$ by $f \succeq_{T_i} g \Leftrightarrow U_i(f) \geq U_i(g)$.[12]

The problem (6) can be decomposed into sub-problems, which is formally stated in the next theorem.

**Theorem 5** *Let $\hat{U}(f)$ be the unique solution to the problem* (6)*. Then, for each $i$,*

$$U_i(f) = \operatorname*{argmin}_{u_i \in [0,1]} \int_S \phi(v \circ f(s), u_i)\, d\mu_{T_i}(s)\,, \tag{7}$$

*where $\mu_{T_i}$ is a probability measure on $(S, \mathcal{E})$ which is derived from $\mu$ according to Bayes' rule: $(\forall E)\ \mu_{T_i}(E) = \mu(T_i \cap E)/\mu(T_i)$.*

In the theorem, $\mu_{T_i}$ is well-defined by the assumption that $(\forall i)\ \mu(T_i) > 0$. We may naturally think of (7) to correspond the inference problem in which the DM estimates the utility

[10] We say that a function is $\langle T_i \rangle_i$-*measurable* if it is measurable with respect to the algebra generated by $\langle T_i \rangle_i$.
[11] For the unique existence, it suffices to assume that $\phi$ satisfies E6.
[12] For an axiomatic treatment of conditional preferences, see Epstein and Le Breton (1993).

value *after* observing the occurrence of $T_i$ with $\mu$ updated into $\mu_{T_i}$. In particular, the DM minimizes the error function which defines her original (or unconditional) preference.

When the error function is specified by $\phi^2$, it turns out that the solution $\hat{U}(f)$ to (6) equals the *conditional* expectation of $v \circ f$ with respect to the probability measure $\mu$ given the information partition $\langle T_i \rangle_i$ and that the solution $U_i(f)$ to (7) equals the expectation of $v \circ f$ with respect to the *posterior* probability measure $\mu_{T_i}$. Hence, we may suggestively write as $\hat{U}(f) = E^{\mu}[\, v \circ f \,|\, \langle T_i \rangle_i \,]$ and $U_i(f) = E^{\mu_{T_i}}[\, v \circ f \,]$ in such a case. It is well-known that in Savage's framework the preference represented by the expected utility is naturally revised to the conditional preferences represented by the conditional expected utilities with the prior belief updated according to Bayes' rule (Kreps, 1988, Chapter 10; Ghirardato, 2002). Hence, the conditional preferences defined through (6) (or (7)), as well as the unconditional preference defined through (1), extend Savage's framework by replacing the squared-error loss function by a more general loss function in corresponding inference problems.

## 5. Lemmas and Proofs

In order to prove Theorem 1, we first prove a couple of lemmas.

**Lemma 1** *Suppose that the preference relation $\succeq$ on $F_0$ satisfies all the axioms in Theorem 3 and hence that it is betweenness-satisfying. Also, assume that $\succeq$ satisfies Axiom A3 as well. Then, the utility index $v$ is continuous and $v(X) = [0, 1]$. Furthermore, the betweenness function $\varphi$ can be taken so as to satisfy*

$$\text{B3.} \quad (\forall u, v \in [0, 1]) \;\; \varphi(v, u) = 0 \;\Leftrightarrow\; u = v \; .$$

**Proof** Let $v$ be a utility index and let $\varphi$ be a betweenness function. Their existence is guaranteed under Axioms P1, P3, P4$^c$, P5, P6, A1 and A2 by Theorem 3. Also, for each $f \in F_0$, let $U(f)$ denote a unique solution to (2). The continuity of $v$ follows from A3 and the fact that $v$ represents $\succeq$ restricted to $X$. The continuity of $v$ and the assumption that $X$ is connected imply that $v(X)$ is connected (Munkres, 1975, p.149, Theorem 1.5), which shows that $v(X) = [0, 1]$ by $v(x^*) = 1$ and $v(x_*) = 0$. In the rest of proof, we show the existence of a betweenness function

$\hat{\varphi}$ which satisfies B3. First, note that

$$(\forall x, x' \in X) \quad v(x) \geq v(x') \Leftrightarrow x \succeq x' \Leftrightarrow U(x) \geq U(x')\,, \tag{8}$$

which implies that there exists a strictly increasing function $g : [0,1] \to \mathbb{R}$ such that $(\forall x \in X)\ U(x) = g \circ v(x)$. Since $U$ is continuous on $X$ by A3 and (8), $U(X)$ is connected. Hence, the strict increase of $g$ implies that $g$ is continuous. Define $\hat{\varphi} : [0,1] \times [0,1] \to \mathbb{R}$ by $(\forall \hat{u}, v)\ \hat{\varphi}(v, \hat{u}) = \varphi(v, g(\hat{u}))$. Then, $\hat{\varphi}$ satisfies B1 since $\varphi$ satisfies B1 and it satisfies B2 since $g$ is continuous and $\varphi$ satisfies B2. Furthermore, both $\hat{\varphi}$ and $\varphi$ generate the identical betweenness-satisfying preference by the strict increase of $g$. Finally, we show that $\hat{\varphi}$ satisfies B3. To show ($\Leftarrow$), let $\bar{v} \in [0,1]$ be arbitrary and let $x \in X$ be such that $v(x) = \bar{v}$. Then, $\hat{\varphi}(\bar{v}, \bar{v}) = \varphi(\bar{v}, g(\bar{v})) = \varphi(v(x), g \circ v(x)) = \varphi(v(x), U(x)) = 0$, where the last equality holds by (2). To show ($\Rightarrow$), it suffices to note that $(\forall \bar{v})\ \hat{\varphi}(\bar{v}, \bar{v}) = 0$ (just proven) and $\hat{\varphi}$ satisfies B1. ■

Let $P_0(X)$ be the set of all simple probability measures on $X$. Here, we say that a probability measure $p$ on $X$ is *simple* if it has a finite support. Given an act $f$ and a probability measure $\mu$ on $(S, \mathcal{E})$, define the simple probability measure $p_f \in P_0(X)$ by $(\forall x)\ p_f(x) = \mu(f^{-1}(\{x\}))$. Note that with this notation, the equation (2) can be rewritten as

$$\int_X \varphi(v(x), u)\, dp_f(x) = 0\,. \tag{9}$$

**Lemma 2** *Let $\mu$ be a finitely-additive, non-atomic probability measure, let $v$ be a utility index, let $\varphi$ be a betweenness function and for each $f \in F_0$, let $U(f)$ be a unique solution to* (2)*. Define a map $V : P_0(X) \to \mathbb{R}$ by $(\forall p)\ V(p) = U(f)$, where $f$ is such that $p = p_f$. Then, $V$ is well-defined and, for any $p, q \in P_0(X)$, the map defined by*

$$\lambda \mapsto V(\lambda p + (1 - \lambda) q)\,,$$

*is continuous on* $[0,1]$.

**Proof** (Well-definition) First, note that for any $p \in P_0(X)$, there exists an act $f$ such that $p = p_f$ by the non-atomicity of $\mu$. Second, suppose that $p_f = p_g$ for some pair of acts, $f, g \in F_0$. Then, $U(f) = U(g)$ since both are the unique solution to (9).[13] Hence, $V(p)$ does not depend

[13]This property is called *probabilistic sophistication* according to Machina and Schmeidler (1992).

on the choice of $f \in F_0$ such that $p = p_f$.

(Continuity) Let $p, q \in P_0(X)$ and let $\langle \lambda_n \rangle_{n=1}^{\infty} \subseteq [0,1]$ converge to $\lambda_0$. We show that $\lim_{n\to\infty} V(\lambda_n p + (1-\lambda_n)q) = V(\lambda_0 p + (1-\lambda_0)q)$, which completes the proof. In the rest of proof, $V(\lambda_n p + (1-\lambda_n)q)$ is abbreviated as $V_n$ for each $n$. Note that it follows from (9) that

$$(\forall n \geq 1) \quad \lambda_n \int_X \varphi(v(x), V_n)\, dp(x) + (1-\lambda_n) \int_X \varphi(v(x), V_n)\, dq(x) = 0\,. \tag{10}$$

Let $\langle \lambda_{n_i} \rangle_{i=1}^{\infty}$ be a subsequence of $\langle \lambda_n \rangle_{n=1}^{\infty}$ such that $\langle V_{n_i} \rangle_{i=1}^{\infty}$ converges to $\overline{\lim}_{n\to\infty} V_n$. Then, (10), the dominated convergence theorem and B2 imply

$$\begin{aligned} 0 &= \lim_{i\to\infty} \lambda_{n_i} \int_X \varphi(v(x), V_{n_i})\, dp(x) + \lim_{i\to\infty} (1-\lambda_{n_i}) \int_X \varphi(v(x), V_{n_i})\, dq(x) \\ &= \lambda_0 \int_X \lim_{i\to\infty} \varphi(v(x), V_{n_i})\, dp(x) + (1-\lambda_0) \int_X \lim_{i\to\infty} \varphi(v(x), V_{n_i})\, dq(x) \\ &= \lambda_0 \int_X \varphi(v(x), \overline{\lim_{n\to\infty}} V_n)\, dp(x) + (1-\lambda_0) \int_X \varphi(v(x), \overline{\lim_{n\to\infty}} V_n)\, dq(x)\,. \end{aligned}$$

By a symmetric argument, it also holds that

$$\lambda_0 \int_X \varphi(v(x), \underline{\lim_{n\to\infty}} V_n)\, dp(x) + (1-\lambda_0) \int_X \varphi(v(x), \underline{\lim_{n\to\infty}} V_n)\, dq(x) = 0\,.$$

Therefore, we conclude that $V(\lambda_0 p + (1-\lambda_0)q) = \overline{\lim}_{n\to\infty} V_n = \underline{\lim}_{n\to\infty} V_n$, and hence, that $V(\lambda_0 p + (1-\lambda_0)q) = \lim_{n\to\infty} V_n$. ■

**Proof of Theorem 1** Suppose that the preference relation $\succeq$ on $F_0$ satisfies Axioms P1, P3, P4$^c$, P5, P6, A1, A2 and A3. Then, Theorem 3 shows that $\succeq$ is betweenness-satisfying with some finitely-additive, non-atomic probability measure $\mu$, some utility index $v$ and some betweenness function $\varphi$. In particular, we may assume that $\varphi$ satisfies B1, B2 and B3 by Lemma 1. Define a map $\phi : [0,1] \times [0,1] \to \mathbb{R}$ by

$$(\forall u, v) \quad \phi(v, u) = \int_0^v \varphi(v, t)\, dt - \int_0^u \varphi(v, t)\, dt\,.$$

Note that $\phi$ is well-defined by B2. Also, note that $\varphi(v, u) > 0$ if $u < v$. To see this, suppose that $\varphi(v, u) \leq 0$ and $u < v$. Then, it follows that $\varphi(u, u) < 0$ by B1, which contradicts B3. Similarly, it holds that $\varphi(v, u) < 0$ if $u > v$. These properties show that $\phi$ thus constructed satisfies E1 and E2. Furthermore, $\phi$ satisfies E3, E4 and E5 by B2, B1 and B3, respectively. Therefore, $\phi$ is an error function.

For each $f \in F_0$, let $U(f)$ denote a unique solution to (2). Then, define a map $\xi : F_0 \times [0,1] \to \mathbb{R}$ by

$$(\forall f, u) \quad \xi(f,u) = -\int_S \varphi(v \circ f(s), u)\, d\mu(s)\,.$$

It follows by definition that $\xi(f,u) = 0$ if and only if $u = U(f)$. The rest of this paragraph proves (a) $\xi(f,u) < 0$ if $u < U(f)$ and (b) $\xi(f,u) > 0$ if $u > U(f)$. To prove (a), suppose that $f$ and $u$ are such that $0 \le u < U(f)$. Define a map $V : P_0(X) \to \mathbb{R}$ by $(\forall p)\ V(p) = U(g)$, where $g$ is such that $p = p_g$. The well-definition of $V$ is among the conclusions of Lemma 2. Then, it follows that $U(f) = V(p_f)$ and that $0 = V(\delta_{x_*})$ because $V(\delta_{x_*}) = U(x_*) = v(x_*) = 0$, where the second equality holds by (2) and B3. Therefore, we have $V(\delta_{x_*}) \le u < V(p_f)$. Then, Lemma 2 and the intermediate value theorem imply that there exists $\lambda \in (0,1]$ such that $V(\lambda\delta_{x_*} + (1-\lambda)p_f) = u$. Therefore, it turns out that

$$\begin{aligned} 0 &= -\int_X \varphi(v(x), u)\, d[\lambda\delta_{x_*} + (1-\lambda)p_f](x) \\ &= -\lambda\varphi(v(x_*), u) - (1-\lambda)\int_X \varphi(v(x), u)\, dp_f(x) \\ &> -\int_X \varphi(v(x), u)\, dp_f(x) = \xi(f,u)\,, \end{aligned}$$

where the first and last equalities hold by (9) and the strict inequality follows from $\lambda \neq 0$ and the fact that

$$\varphi(v(x_*), u) < \int_X \varphi(v(x), u)\, dp_f(x)\,. \tag{11}$$

To see (11), note that $p_f$ assigns a positive probability to some $x' \in X$ such that $x' \succ x_*$ since otherwise (11) would hold with an equality for all $u$ because $v$ represents $\succeq$ restricted to $X$ and $x_*$ is the worst outcome, which would contradict the assumption that $V(\delta_{x_*}) < V(p_f)$. This, that $v$ represents $\succeq$ restricted to $X$ and B1 show (11), which completes the proof of (a). By a symmetric argument which uses $x^*$, (b) can be also established.

Let $f \in F_0$. By the definition of $\phi$, $u \in [0,1]$ minimizes

$$\int_S \phi(v \circ f(s), u)\, d\mu(s) \tag{12}$$

if and only if it minimizes

$$\begin{aligned} & \int_S \left( -\int_0^u \varphi(v \circ f(s), t)\, dt \right) d\mu(s) \\ = \; & \int_0^u \left( -\int_S \varphi(v \circ f(s), t)\, d\mu(s) \right) dt \\ = \; & \int_0^u \xi(f, t)\, dt \end{aligned}$$

where the first equality holds by Fubini's theorem and the second equality holds by the definition of $\xi$. The previous paragraph ((a) and (b)) shows that $u = U(f)$ uniquely minimizes the last expression, and hence, it also uniquely minimizes (12), which completes the proof. ■

**Proof of Theorem 2** Axiom A3 follows immediately by the assumption that $v$ is continuous on $X$ and it represents $\succeq$ restricted to $X$. We show that the preference based on the loss-minimization with a convex error function is betweenness-satisfying, which completes the proof by Theorem 3.

To show that $\succeq$ is betweenness-satisfying, let $f$ be an act and suppose that $u^* = U(f)$ solves

$$\min_{u \in [0,1]} \int_S \phi(v \circ f(s), u)\, d\mu(s)\,.$$

First, assume that $u^* \in (0, 1)$. Then, by E3, it must hold that

$$\int_S \phi_2(v \circ f(s), u^*)\, d\mu(s) = 0\,, \tag{13}$$

where $\phi_2$ denotes $\partial\phi/\partial u$. Next, assume that $u^* = 0$. Then, by E3, it must hold that

$$\int_S \phi_2(v \circ f(s), u^*)\, d\mu(s) \geq 0\,. \tag{14}$$

Note that $(\forall s)\ v \circ f(s) \geq u^*$ by the fact that $u^* = 0$ and $v(X) \subseteq [0, 1]$. Hence, the inequality in the opposite direction also holds in (14) by E4 and E5, which shows that the equation (13) holds even when $u^* = 0$. A symmetric argument shows that the equation (13) holds also when $u^* = 1$. Note that any other $u$ except for $u^*$ does not satisfy (13) by E6. Therefore, if we set $\varphi = -\phi_2$, $u^* = U(f)$ will be a unique solution to the equation (2). Furthermore, $\varphi$ thus defined satisfies B1 by E4 and it satisfies B2 by E3. ■

**Proof of Theorem 5** Let $p(i|s)$ be the probability that $T_i$ occurs given that the true state is $s$. Then, $(\forall s)$ $\sum_{i=1}^n p(i|s) = 1$. Therefore, for any $f \in F_0$ and any $\hat{u} \in B(\langle T_i \rangle_i)$, it holds that

$$\begin{aligned}
\int_S \phi(v \circ f(s), \hat{u}(s))\, d\mu(s) &= \int_S \phi(v \circ f(s), \hat{u}(s)) \sum_{i=1}^n p(i|s)\, d\mu(s) \\
&= \sum_{i=1}^n \int_S \phi(v \circ f(s), \hat{u}(s)) p(i|s)\, d\mu(s) \\
&= \sum_{i=1}^n \int_S \phi(v \circ f(s), \hat{u}(s)) \frac{p(i|s) d\mu(s)}{\mu(T_i)} \cdot \mu(T_i) \\
&= \sum_{i=1}^n \int_S \phi(v \circ f(s), \hat{u}(s)) \frac{\mu(T_i \cap ds)}{\mu(T_i)} \cdot \mu(T_i) \\
&= \sum_{i=1}^n \int_S \phi(v \circ f(s), \hat{u}(s))\, d\mu_{T_i}(s) \cdot \mu(T_i)\,.
\end{aligned}$$

Since $\hat{u}$ is constant over each $T_i$ and the integrals in the last line can be minimized one by one, the proof is complete. ■

## REFERENCES


Chew, S. H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica* 51, 1065-1092.

Chew, S. H. (1989): "Axiomatic Utility Theories with the Betweenness Property," *Annals of Operations Research* 19, 273-298.

Dekel, E. (1986): "An Axiomatic Characterization of Preferences under Uncertainty," *Journal of Economic Theory* 40, 304-318.

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

Fishburn, P. C. (1970): *Utility Thoery for Decision Making*. Wiley, New York.

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

Grant, S., A. Kajii and B. Polak (2000): "Decomposable Choice under Uncertainty," *Journal of Economic Theory* 92, 169-197.

Gul, F. (1991): “A Theory of Disappointment Aversion,” *Econometrica* 59, 667-686.

Kreps, D. (1988): *Notes on the Theory of Choice*. Boulder, Colorado: Westview Press.

Lehmann, E. L. and G. Casella (1998): *Theory of Point Estimation* (Second Edition), Springer-Verlag.

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

Munkres, J. R. (1975): *Topology: A First Course*. Prentice-Hall, New Jersey.

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