ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208269
27 Hillhouse Avenue
New Haven, CT 06520-8269

CENTER DISCUSSION PAPER NO. 774

AN OUTCOME-ORIENTED THEORY OF CHOICE AND EMPIRICAL PARADOXES IN EXPECTED UTILITY THEORY

Fumihiro Kaneko

Yale University

May 1997

Note: Center Discussion Papers are preliminary materials circulated to stimulate discussions and critical comments. Dr. Kaneko is a Ph.D. candidate in the Department of Economics.

An Outcome-Oriented Theory of Choice and Empirical Paradoxes in Expected Utility Theory

Fumihiro Kaneko*

Department of Economics

Yale University

12 May 1997

Abstract

I analyze observed choice between lotteries from an outcome-oriented point of view in the framework of choice between random variables. I characterize a choice maker, who faces a choice situation between lotteries, by (1) a surmising process that associates, with a pair of lotteries, a set of well-defined choice situations between random variables, and (2) a choice set that is a collection of well-defined choice situations. I give a partial axiomatic foundation of the theory. The theory is applied to explain the well-known paradoxes in expected utility theory.

Keywords: Expected Utility Theory, SSA Representation, SSB Representation, Prospect Theory, Choice under Uncertainty

---

*The author thanks Professors Donald Brown, Truman Bewley, David Pearce, Dilip Abreu, and other participants of the seminar in Yale University for very helpful comments.

1 Introduction

In this paper, I develop a theory of choice among lotteries where choices satisfy neither transitivity nor the independence axiom of expected utility theory. The theory developed explains well-known empirical paradoxes in classical expected utility theory. Many researchers have developed other theories that extend expected utility theory and are compatible with either the violation of the independence axiom or the violation of the transitivity.[1] However, most of them are not compatible with both, and fail to identify the origins of paradoxes.

In this paper, I start with a consideration of reasoning that results in a sensitive decision process. I construct an axiomatic theory that yields the decision process. Then, I show that the theory identifies the origins of well-known paradoxes in expected utility theory.

To illustrate a choice-maker's reasoning that I believe to be natural, consider experiments in which subjects are asked the following question.

> Lottery 1 guarantees a probability p of winning a prize x. Lottery 2 guarantees a probability q of winning a prize y. You can take one of these but not both. Which one do you choose?

Suppose that a choice maker answered, "I choose lottery 1." Should we take this answer at face value without second thought? My answer is no. It is not clear what subjects really choose. Traditional expected utility theory assumes that they choose among probability measures on the outcome space, but careful examination reveals that this may be inadequate. Subjects may be suspicious of probabilities as sound objects of choice unless the way they are generated is common knowledge to both subjects and the experimenter. Since subjects are not offered any particular kind of random device, they do not know how the experimenter generates the probabilities. Since all random devices that generates a pair of lotteries define pairs of random variables, I assume that the objects of choice are random variables defined on a state space. I also assume that subjects believe that they know all kinds of random devices which the experimenter use to generate lotteries. Hence, though such a state space is not given in the experimenter's description, I can assume that there

[1] A list of such representation results is in [Machina, 87].

exists a universal state space, Ω, that is common knowledge for both subjects and the experimenter. Subjects must be informed of a probability on the universal state space so that they can compute distributions induced by random variables. Hence, from subjects' point of view, a well-described choice situation for subjects is given by $(f, g; P)$, where f and g are random variables defined on Ω and P is a probability on Ω. (Subjects face choice between f and g, given probabilistic information P.) How subjects can obtain such random variables and probabilities when the experimenter does not give any of them directly? My answer is, by guessing. The way this guessing is done is part of subjects' characteristics. Each subject conceives of a set of well-described choice situations from the description given by the experimenter. Subjects are sure that they are facing one of these choice situations, but they do not know which one. Finally, I interpret the statement, "I choose lottery 1" as "For all well-described choice situations I can think of, I will choose the random variable that induces the distribution on outcomes, p." I summarize my points in the following monologue of a subject.

> I am confused. I do not think that choosing a probability measure is a good idea because I do not know how it is generated. Assuming we both agree on a state space, I believe that the objects of choice should be random variables on the state space. For me, a well-described choice situation consistent with the description is given by $(f, g; P)$ where f and g are random variables and P is a probability on the state space, that satisfy $p = P \circ f^{-1}$ and $q = Q \circ g^{-1}$. In this way, I can see how you generate your "lotteries." I can think of a set H of such well-defined choice situations. I am sure that you are suggesting that I am facing one of choice situations in H, but I do not know which one I am really facing. But I can say, "I choose the lottery 1" in the sense that, for any well-described choice situation in H, I will choose the random variable which induces the distribution on outcomes, p.

The theory developed in this paper describes a choice maker exactly the way the monologue suggests. Formally, I characterize a choice maker in the following way. Let Ω be the universal state space and Z be a set of outcomes. Also let the notation Λ refer, "the probability space on." I characterize a choice maker by (1) a correspondence,

$$\Gamma : \Lambda(Z) \times \Lambda(Z) \to \{(f, g; P) \,|\, f, g \text{ are random variables on } \Omega, \text{ and } P \in \Lambda(\Omega)\}.$$

that satisfies, for all $(f, g; P) \in \Gamma(p, q)$, $P \circ f^{-1} = p$ and $P \circ g^{-1} = q$, and (2) a subset $\succ$ of $\{(f, g; P) \,|\, f, g \text{ are random variables on } \Omega, \text{ and } P \in \Lambda(\Omega)\}$ that is called a choice set. The characteristic (1) governs the selection of well-defined choice situations from the description of lotteries in an obvious way. The characteristic (2) governs the choice when a choice maker is sure that she is facing one well-defined choice situation. $(f, g; P) \in \succ$ if and only if f is chosen over g by a choice maker when she is informed of a probability P on the state space Ω.

My trial for axiomatic foundation of the theory is not complete because I do not give any axiomatic foundation for Γ. However I give a set of axioms on $\succ$ that yields an "outcome-oriented" counterpart of SSA theory developed by Fishburn in [Fishburn, 89]. By "outcome-orientation", I mean generally that information that is directly associated with outcomes has priority over other information in decision process. In this paper, "outcome-orientation" means that a choice maker concentrates on information about distributions on pairs of outcomes in the decision process. I introduce five axioms and prove, under a domain restriction, that $\succ$ allows an additive representation with a skew-symmetric function. Then, I introduce two additional axioms and extend the result to the general case without domain restriction.

In application of the theory, I consider the following empirical paradoxes in expected utility theory; (a) the Allais Paradox, (b) the common consequence effect, (c) the common ratio effect, (d) the utility evaluation effect, and (e) preference reversal. Traditionally (a), (b), (c) and (d) are explained as violations of the independence axiom in expected utility theory, while (e) is explained as a violation of transitivity.[2] My theory is not only consistent with all of these paradoxes, but also is useful to identify the origins of these paradoxes. My analysis may be summarized as follows.

[2]The survey paper, [Machina, 87], strictly follows this viewpoint.

1. Under some regularity on experiments, (a) and (c) happen if Γ selects choice situations in which two random variables are negatively correlated. The analysis of (c) also requires that $\succ$ exhibits increasing risk aversion (see 3).

2. (b) happens if the outcome space is the probability space $\Lambda(R)$, a set of all probabilities on money-prizes, and $\succ$ exhibits a change of attitude toward risk. (The precise meaning of "change of attitude toward risk" is given in the section 5.) For this explanation to make sense, probabilities on money-prizes must be common knowledge to subjects and an experimenter. When the outcome space is $\Lambda(R)$, we can no longer see (b) as a violation of the independence axiom. The analysis of (b) with money-prizes as outcomes can be done in a simple case. I show that, if the configuration of experiments satisfies some regularity conditions, (b) happens when Γ selects choice situations in which the two random variables are negatively correlated.

3. I regard (d) and (e) as violations of transitivity for $\succ$. They happen if $\succ$ exhibits increasing risk aversion. (In the section 5, I define a notion of "increasing risk aversion" for $\succ$.) The fact that certainty equivalents are used in both experiments (d) and (e) is crucial for my argument. The analysis of (e) requires that the configuration of experiments satisfies some regularity.

All regularity conditions are derived in forms that can be checked by asking subjects.

I also apply my theory for criticizing properties imposed on a prospect function in [Kahneman & Tversky, 79]. I argue that it is hard to justify such properties, and the existence of a prospect function itself, because there would be no "fixed base outcome" in the decision process of a choice maker, which is essential for their construction of a prospect function.

I comment briefly on other works about the extension of the expected utility theory related to the one in this paper. Most researches on non-expected utility theory may be classified into three categories; (A) objection to the reduction of compound lotteries, (B) objection to the independence axiom, and (C) objection to transitivity. My theory belongs to both (B) and (C).

Though I take seriously researches about theory of choice between compounded lotteries in the category (A) (cf. [Kreps & Porteus, 78], [Krantz, Luce, Suppes & Tversky, 90]), I leave them out of scope in this paper, for there is no clear relationship between them and the theory developed here.[3]

The first systematic approach for constructing non-expected utility theory that is consistent with violation of the independence axiom, other than the pioneering work by Allais [Allais, 79], appeared in [Kahneman & Tversky, 79]. In this paper, the authors advocated use of a prospect function and a probability modification function. They call their theory prospect theory. The use of prospect functions has been very controversial. On the other hand, the use of a probability modification function has been welcomed by many researchers as the first solution for violation of the independence axiom. The authors show that the "subcertainty" of the probability modification function causes the Allais paradox while its "subproportionality" causes the common ratio effect.[4] However, these properties are not derived from axioms. It is also hard to find some natural interpretation that gives these properties intuitive meaning. The authors also considered other empirical paradoxes that are out of scope in this paper and tried to identify their causes.[5] Unfortunately, their analysis depends heavily on properties of prospect functions that have been very controversial. Based on my theory, I deny such properties of prospect functions. For their credit, prospect theory characterizes a choice maker partly by an "editing phase." In my theory, Γ plays the role of their "editing phase."

Several researchers have worked on extension of prospect theory. Most of them conjectured that the weighting function may depend on outcomes. The first axiomatic theory that exhibits this property is expected utility theory with rank-dependent probabilities by Quiggin. [Quiggin, 82]. He calls his theory anticipated utility theory. In this theory, a distribution modification function, instead of a probability modification function, is used. For each lottery with a finite support, a

[3] Here I just note that the argument for the common consequence effect in this paper requires a careful consideration about the reduction of lotteries.

[4] They also derive other properties of the probability modification function that are necessary to explain empirical results. For example, the condition that the probability modification function overweights small probabilities has been proposed to explain risk-taking behavior with small probabilities.

[5] (1) Preference for regular insurance over probabilistic insurance, (2) Increase of risk seeking by a negative translation of a choice. For details, see [Kahneman & Tversky, 79].

rank-dependent probability is constructed by the distribution function corresponding to the lottery and a distribution modification function. Then, lotteries are ordered by "expected" utilities with respect to their rank-dependent probabilities. On axiomatic foundation, he introduces the first-order stochastic dominance principle and a weak version of the independence axiom.[6] Properties on the distribution modification function that are necessary for several paradoxes considered in this paper can be found. I do not regard the first-order stochastic dominance as a compelling axiom, since I design my theory for completely general outcome spaces.

In [Chew, 83], Chew introduces a weighting function on outcomes and uses it to describe a modification process of lotteries in choice maker's mind. He calls his theory weighted utility theory. He introduces the weak substitution axiom, that replaces the independence axiom, and derives a representation result such that the utility function in expected utility theory is replaced by a weighted utility function. A weighted utility function is defined as a product of two functions defined on outcomes. One of them is considered as a usual utility function. The other is called a weight-of-utility function and used to describe a psychological modification process of lotteries. It has been shown that this theory is consistent with violations of the independence axiom, including the Allais paradox. It is not difficult to show that certain properties on weight-of-utility functions imply those of the probability modification function imposed by Kahneman and Tversky. However, they are not derived from axioms, and not particularly intuitive.

A very different approach was proposed by [Dekel, 86]. By introducing a very weak independence axiom, called the betweenness axiom, he derived a representation function whose function form is exactly the same as the one derived in this paper. Unfortunately, he assumes transitivity, as do all works I have mentioned so far. This results in the use of certainty equivalents to order lotteries. But, this procedure has an obvious conflict with the preference reversal, one of the well-known paradoxes.[7] He assumed that a pair of lotteries induces a distribution over pairs of outcomes by its product measure. This I do not assume. In my theory, the surmising process includes the consideration of correlation among random variables, so that my theory is more general than his.

[6] The former axiom forces distributions to be sufficient statistics for choice.

[7] It could be the case that discarding the transitivity from his set of axioms may result in a similar representation result to the one in this paper.

The first contribution to axiomatic theory of choice without transitivity is SSB (Skew-Symmetric Bilinear) theory developed by Fishburn [Fishburn, 82]. This theory is the first work which suggests that, given a choice situation, a choice maker may not treat each alternative in it separately. The theory assumes that a pair of lotteries in a choice situation induces a distribution over pairs of outcomes by its product measure. This nature restricts the application of the theory. In particular, it seems to me that SSB theory does not give a clear explanation of the Allais paradox. Fishburn [Fishburn, 89] refined SSB theory in the Savage setup. He called the resulting theory SSA (Skew-Symmetric Additive) theory. It is a minimal additive extension of Savage's theory without transitivity. It allows consideration of correlated random variables. On the other hand, since, in SSA theory, a probability on the state space is a part of choice maker's characteristics, we cannot apply SSA theory to solve empirical paradoxes in expected utility theory. My theory repairs this shortcoming. My theory fills a gap between two different formations and solve paradoxes in expected utility theory. In this sense, I regard the representation result for $\succ$ in this paper as an "outcome-oriented" version of SSA theory.

This paper is organized as follows. Section 2 introduces the model. Section 3 introduces five axioms on the choice maker's characteristics. Section 4 presents and proves the representation result for finite-valued acts. The result is applied to solve various paradoxes in expected utility theory in section 5. The representation result is extended for general acts in section 6. Section 7 concludes the paper.

2 Model

The outcome space is given by a measurable space, which I denote by (Z, Ξ). I assume that there is a state space that is common knowledge for both experimenter and subjects. It is given by a continuum measurable space. I denote it by $\left(\Omega, 2^{\Omega}\right)$. I call each random variable, a measurable function from the state space to the outcome space, an *act*. I denote the set of all acts by A.

Definition 2.1 *A* <u>*well-described choice situation*</u> *for subjects is an ordered pair of acts,* (f, g)*, together with an atomless probability measure,* P*, on the state space. I denote it by* $(f, g; P)$.

I denote the set of all well-described choice situations by B.

Definition 2.2 *An* <u>*intended choice situation*</u> *by experimenter is an ordered pair of probabilities on the outcome space,* (p, q).

I denote the set of all intended choice situations by L.

Objects of choice for choice makers are acts. A choice maker is characterized by

1. $\Gamma : L \to 2^B$ such that, for all $(f, g; P) \in \Gamma(p, q)$, $P \circ f^{-1} = p$ and $P \circ g^{-1} = q$,

2. $\succ \subset B$.

I call Γ a surmising process and $\succ$ a choice set.

A surmising process describes a modelling process by a choice maker who is given an intended choice situation.

A well-described choice situation $(f, g; P)$ is in the choice set $\succ$ if and only if the choice maker chooses the act f over the act g when the probability measure P is given.

Definition 2.3 *Given an intended choice situation* $(p, q) \in L$*, I say that a choice maker chooses* p *over* q *if* $\Gamma(p, q) \subset \succ$.

In my axiomatic foundation, I leave the surmising process out of scope and concentrate on the choice set. It is clearly desirable to give axioms that determines a (unique) structure of the surmising process. This task is left for my future research.

3 Axioms for the Additive Representation

In this section, I introduce a simple set of axioms on the choice set $\succ$ that guarantees the existence and the uniqueness of its skew-symmetric additive representation.

The first axiom is called outcome-orientation. In general, outcome-orientation means that, in a decision process, information directly associated with outcomes has priority over other information. In this paper, it means that, given a choice situation, a choice maker cares only the induced distribution over pairs of outcomes. To state this assumption formally, I take a quotient space of B by equivalence of the induced distribution over ordered pairs of outcomes. I denote the resulting quotient space by B^* and an equivalence class with representative element $(f, g; P)$ by $< f, g; P >$. I call each class in B^* an consequential class of choice situations. The first axiom essentially says that $\succ$ determines a unique subset $\succ^*$ of B^*. Since B^* is isomorphic to the space of all distributions over pairs of outcomes, we can regard $\succ^*$ as a subset of the latter space. I will use this identification freely in the rest of this paper.

Axiom 3.1 (Outcome-Orientation) *If two well-described choice situations,* $(f, g; P)$ *and* $(f', g'; P')$, *are such that* $P \circ (f, g)^{-1} = P' \circ (f', g')^{-1}$, *then* $(f, g; P) \in \succ$ *if and only if* $(f', g'; P') \in \succ$.

For the convenience of exposition, I define $\sim^*$ by

$$\sim^* \equiv \{< f, g; P > \mid < f, g; P > \notin \succ^* \text{ and } < g, f; P > \notin \succ^*\}.$$

It will be used later in the proof.

The next axiom is asymmetricity.

Axiom 3.2 (Asymmetricity) *If* $(f, g; P) \in \succ$, *then* $(g, f; P) \notin \succ$.

If the outcome-orientation axiom is satisfied, we can restate this axiom by replacing choice situations by consequential classes of choice situations and $\succ$ by $\succ^*$. Suppose that is the case. In terms of distributions over pairs of outcomes, this axiom simply claims that changing the order of coordinates does not preserve the membership to $\succ^*$.

The other three axioms are essentially restrictions on the mixture operation on consequential classes of choice situations. Since B^* is isomorphic to the space of probabilities over ordered pairs of outcomes, I can define a mixture operation on B^* by the familiar weighted sum operation on probabilities. For any given $\alpha \in [0,1]$, the mixture operation $\oplus_\alpha : B^* \times B^* \longrightarrow B^*$ is defined as follows.

$$< f, g; P > \oplus_\alpha < f', g'; P' > \equiv < f'', g''; P'' >.$$

where

$$P'' \circ (f'', g'')^{-1} = \alpha P \circ (f, g)^{-1} + (1 - \alpha) P' \circ (f', g')^{-1}.$$

A simple way to create a probability measure P'' and a pair of acts (f'', g'') is as follows.

1. Take any set $E \in 2^\Omega$ and assign the probability α to it.

2. On $\left(E, 2^E\right)$, construct a pair of random variables (f_E, g_E) and an atomless probability measure P_E such that $P_E \circ (f_E, g_E)^{-1} = P \circ (f, g)^{-1}$.

3. On $\left(\Omega - E, 2^{\Omega - E}\right)$, construct a pair of random variables $(f'_{\Omega-E}, g'_{\Omega-E})$ and an atomless probability measure $P'_{\Omega-E}$ such that $P'_{\Omega-E} \circ (f'_{\Omega-E}, g'_{\Omega-E})^{-1} = P' \circ (f', g')^{-1}$.

4. Construct a pair of acts (f'', g'') by the rule by

$$f''(\omega) = \begin{cases} f_E(\omega), & \text{if } \omega \in E, \\ f'_{\Omega-E'}(\omega) & \text{if } \omega \in \Omega - E, \end{cases}$$

   and

$$g''(\omega) = \begin{cases} g_E(\omega), & \text{if } \omega \in E, \\ g'_{\Omega-E}(\omega), & \text{if } \omega \in \Omega - E. \end{cases}$$

5. Construct an atomless probability measure P'' by

$$P''(F) = \alpha P_E(F \cap E) + (1 - \alpha) P'_{\Omega-E}(F \cap (\Omega - E)).$$

Then $< f'', g''; P'' >$ is exactly $< f, g; P > \oplus_\alpha < f', g'; P' >$. Note that this operation associates, with a pair of well-described choice situations, a set of well-described choice situations. I use the same notation $\oplus_\alpha$ to denote this correspondence. Since the state space is a continuum, the second and the third step of this procedure are feasible.[8] I call these steps *replication by miniatures.*

The third axiom suggests that, if a choice maker cannot make a choice in choice situations, $(f, g; P)$ and $(f', g'; P')$, then she should not be able to make a choice in any mixture of these two because a mixture of two choice situations is regarded as more complex than original two choice situations. To state the axiom, I introduce an additional notation. I denote $\{(f, g; P) | (f, g; P) \notin \succ \text{ and } (g, f; P) \notin \succ\}$ by $\sim$.

Axiom 3.3 (Betweenness) *If* $(f, g; P) \in \sim$ *and* $(f', g'; P') \in \sim$, *then* $(f, g; P) \oplus_\alpha (f', g'; P') \subset \sim$ *for all* $\alpha \in [0, 1]$.

The name of this axiom comes from the following observation. Suppose that the outcome-orientation axiom is satisfied. Then I can restate the axiom by replacing choice situations by consequential classes of choice situations and $\sim$ by $\sim^*$. I can regard that mixture operations $\{\oplus_\alpha | \alpha \in [0, 1]\}$, applied to two consequential classes of choice situations, determine all consequential classes *between* them. Then, the axiom is equivalent to the preservation of the membership to $\sim^*$ *between* two consequential classes of choice situations.

The forth axiom guarantees that there is no choice situation such that:

1. A choice maker cannot make a choice in it;

2. It affects the choice if it is mixed with some particular choice situation.

Axiom 3.4 (Substitution) *If* $(f, g; P) \oplus_\alpha (g', f'; P') \not\subset \succ$ *for some* $\alpha \in (0, 1)$, *then* $(f, g; P) \in \succ$ *implies* $(f', g'; P') \in \succ$.

Finally I introduce an axiom that is a key to find a measurement tool in my theory. This axiom is a natural counterpart of the Archimedian axiom in expected utility theory. When choice is not necessarily transitive, a unitary measurement device (frequently called a "rod") implied by the

[8] This is the reason why I restricted the probability on the state space to an atomless one.

Archimedian axiom is replaced by a binary measurement device (which I call a "balance") implied by the following axiom.

Axiom 3.5 (Archimedian) *If* $(f, g; P) \in \succ$ *and* $(f', g'; P') \in \succ$, *then there is an* $\alpha \in (0, 1)$ *such that* $(f, g; P) \oplus_\alpha (g', f'; P') \subset \succ$.

These are all axioms that I need to prove my representation result for consequential classes with finite support.

Before going to the next section, I introduce an operation on B (and B^*) other than the mixture operation. The twisting operation, $(.)^-$, on B is defined by $(f, g; P)^- \equiv (g, f; P)$. It introduces a unique operation on B^* as follows. Let $T : Z \times Z \to Z \times Z$ be defined by $T(x, y) \equiv (y, x)$. Then, the twisting operation, $(.)^-$, on B^* is defined by $P^- \equiv P \circ T^{-1}$. (I used the identification of B^* and the space of probabilities on ordered pairs of outcomes.) The twisting operation is used often in the proof.

4 Representation of Choice over Finite-valued Acts

The following result is the representation formula for all consequential classes with finite support.

Theorem 4.1 (Skew-Symmetric Additive Representation : Finite Supports) *A choice set $\succ$ restricted on $B^F \equiv \{(f,g;P) | P \circ (f,g)^{-1}$ has a finite support. $\}$ satisfies axiom 3.1, 3,2, 3.3, 3.4 and 3.5 if and only if there is a $\Phi : (Z \times Z, \Xi \otimes \Xi) \to R$ such that:*

1. *$\Phi(x,y) + \Phi(y,x) = 0$ for any $x, y \in Z$;*
2. *$(f,g;P) \in \succ \cap B^F$ if and only if $P\Phi(f,g) > 0$.*

If the two function Φ and Φ' satisfies 1 and 2, then there is an $a > 0$ such that $\Phi'(x,y) = a\Phi(x,y)$ for all $x, y \in Z$.

The intuition behind this theorem is as follows. By the outcome-orientation axiom, I can translate the statement of the theorem into the equivalent statement in $\Lambda^F(Z \times Z, \Xi \otimes \Xi)$ where $\Lambda(.)$ means "the set of all probabilities on" and the superscript "F" indicates finite support property. Let $\succ^{*-}$ be the set $\{Q^- \in \Lambda(Z \times Z, \Xi \times \Xi) | Q \in \succ^*\}$. The asymmetricity axiom says $\succ^* \cap \succ^{*-} = \emptyset$. The substitution axiom says that $\succ^*$ is closed with respect to the mixture operation (or, "convex"), and that $\sim^*$ is not "thick". At this stage, $\sim^*$ starts to look like a part of hyperplane. The betweenness axiom reinforces this intuition by saying that $\sim^*$ is "convex." The Archimedian axiom says that the mixture operation is "$\succ^*$-continuous," and that a "balance" can be constructed. This balance is a measurement device that, for each well-described choice situation and a choice of an act by a choice maker, outputs a positive number if the choice is supported, a negative number if the opposite choice is supported, and zero otherwise. The additivity of the balance is guaranteed by the substitution axiom and the betweenness axiom.

I note the similarity of the intuition behind my result and that behind expected utility theory. In the expected utility theory, each lottery is "plotted" on the real line. The procedure to do this is as follows. At first, take two lotteries. They determine a unit rod. Then, measure each lottery by the replication coefficient as a mixture of two lotteries in the unit rod. The choice order is translated to the usual order on R. In my theory, each consequential class of choice situations is "plotted" on the real line. The choice set $\succ^*$ is translated to the positive orthant of R and $\sim^*$ is translated to zero.[9]

I also note the relation between my theory and Fishburn's theories. Fishburn's SSB theory [Fishburn, 82] suggests that, without transitivity, a measurement tool should be a "balance". Fishburn's SSA theory [Fishburn, 89] proves skew-symmetric additive representation of chqice in the Savage's setup. In this setup, a probability on the state space is not given as information in any choice situation. It is uniquely determined by characteristics of a choice maker. To do that, it is assumed that a choice maker possesses detailed knowledge about the nature of the state space when she makes a choice. In a choice between lotteries, the description of lotteries gives information about possible probabilities on the state space. Moreover, since the state space is not explicitly mentioned in the description of lotteries, a choice maker is likely to pay her attention to outcomes. This suggests that a choice maker would not consider detailed information about states. Therefore, Fishburn's SSA theory cannot analyze paradoxes in expected utility theory.[10] The formula of the skew-symmetric additive representation in my theory is the same as the one in SSA theory. But, my theory is made to explain choice between lotteries. In other words, my theory generalize the expected utility theory of Milnor, not the Savage theory.

Now I start a formal proof. At first, I rewrite the statement of the theorem with $\succ \in B$ to the equivalent statement with $\succ^* \in \Lambda(Z \times Z, \Xi \otimes \Xi)$. It is possible, due to the outcome-orientation axiom. To make statements simple, I use the usual mixture notation, instead of $\oplus_\alpha$. Since the computation rule of mixture on $\Lambda(Z \times Z, \Xi \otimes \Xi)$ corresponds to the usual mixture notation, there will be no confusion.

[9] For a general notion of representation as a construction of geometry, I refer to [Beja & Gilboa, 92].

[10] It was reported in [Machina, 87] that some researchers applied SSA theory to explain several paradoxes in expected utility theory. Readers should convince themselves that this cannot be done. [Loomes & Sugden, 82] gives the formula equivalent to the one in this paper and analyze some paradoxes in expected utility theory. Since the paper does not give any axiomatic foundation, the validity of the representation formula cannot be tested. In other words, it does not conform a "theory."

Axioms 3.2, 3.3, 3.4 and 3.5 is rewritten as follows.

Axiom 4.1 (Asymmetricity) *If* $Q \in \succ^*$, *then* $Q^- \notin \succ^*$.

Axiom 4.2 (Betweeness) *If* $Q, Q' \in \sim^*$, *then* $\alpha Q + (1-\alpha)Q' \in \sim^*$ *for all* $\alpha \in [0,1]$.

Axiom 4.3 (Substitution) *If there is an* $\alpha \in (0,1)$ *such that* $\alpha Q + (1-\alpha)Q'^- \notin \succ^*$, *then* $Q \in \succ^*$ *implies* $Q' \in \succ^*$.

Axiom 4.4 (Archimedian) *If* $Q \in \succ^*$ *and* $Q' \in \succ^*$, *then there is an* $\alpha \in (0,1)$ *such that* $\alpha Q + (1-\alpha)Q'^- \in \succ^*$.

The theorem 4.1 is equivalent to the following proposition on $\Lambda^F(Z \times Z, \Xi \otimes \Xi)$.

Proposition 4.1 (Skew-Symmetric Additive Representation on $\Lambda^F(Z \times Z, \Xi \otimes \Xi)$) *Axiom 4.1, 4.2, 4.3 and 4.4 are satisfied on* $\Lambda^F(Z \times Z, \Xi \otimes \Xi)$ *if and only if there is a function* $\Phi : (Z \times Z, \Xi \otimes \Xi) \to R$ *such that:*

1. $\Phi(x,y) + \Phi(y,x) = 0$ *for any* $x, y \in Z$;

2. $Q \in \succ^* \cap \Lambda^F(Z \times Z, \Xi \otimes \Xi)$ *if and only if* $Q\Phi > 0$.

If the two function Φ *and* Φ' *satisfies 1 and 2, then there is an* $a > 0$ *such that* $\Phi'(x,y) = a\Phi(x,y)$ *for all* $x, y \in Z$.

It is easy to prove that the proposition and the theorem are equivalent. I omit a simple proof.

To prove the proposition, I prove a key lemma at first. This lemma suggests that a balance can be constructed from $\succ$. I also derive several simple corollaries. In proofs, I use the idempotent property, $P^{--} = P$.

Lemma 4.1 (Cancellation) *Assume that axioms 4.1, 4.2, 4.3 and 4.4 hold. If $P, Q \in \succ^*$, then there exists a unique θ such that $\theta P + (1-\theta)Q^- \in \sim^*$.*

Proof:

Let $\bar{\Theta} \equiv \{\alpha \in [0,1] | \alpha P + (1-\alpha)Q^- \in \succ^*\}$. Since it contains 1, it is not empty. By the Archimedian axiom, it also contains an $\alpha \in (0,1)$. If $\alpha \in \bar{\Theta}$ and $\beta > \alpha$, then $\beta \in \bar{\Theta}$ because, otherwise, it violates the substitution axiom applied to P and $[\alpha P + (1-\alpha)Q^-]^-$. (I used the idempotent property $Q'^{--} = Q'$ here.) Also, for any $\alpha \in \bar{\Theta}$, the Archimedian axiom applied to $\alpha P + (1-\alpha)Q^-$ and Q guarantees the existence of a $\beta < \alpha$ such that $\beta \in \bar{\Theta}$. Hence $\bar{\Theta}$ is an open interval in $[0,1]$ that contains 1 and does not contain 0. Similar argument proves that $\underline{\Theta} \equiv \{\alpha \in [0,1] | [\alpha P + (1-\alpha)Q^-]^- \in \succ^*\}$ is an open interval on $[0,1]$ that contains 0 and does not contain 1. By the asymmetricity axiom, $\bar{\Theta} \cap \underline{\Theta} = \emptyset$. Hence there must be a $\theta \in (0,1)$ such that θ is contained in neither $\bar{\Theta}$ nor $\underline{\Theta}$. Clearly $\theta P + (1-\theta)Q^- \in \sim^*$. For any $\alpha > \theta$, the substitution axiom applied to P and $\theta P + (1-\theta)Q^-$, and the asymmetricity axiom implies that $\alpha \in \bar{\Theta}$. By using the distributive law $[\alpha Q' + (1-\alpha)Q'']^{-1} = \alpha Q'^- + (1-\alpha)Q''^{-1}$, a similar argument shows that, if $\beta < \theta$, then $\beta \in \underline{\Theta}$. Hence θ is uniquely determined.

♠

Corollary 4.1 (Replication) *Assume that axioms 4.1, 4.2, 4.3 and 4.4 hold. Then, for all Q, $\frac{1}{2}Q + \frac{1}{2}Q^- \in \sim^*$.*

Proof:

If $Q \in \sim^*$, the statement is obviously true by the betweenness axiom. Suppose that $Q \in \succ^*$. If $\frac{1}{2}Q + \frac{1}{2}Q^- \in \succ^*$, then, as we showed in the proof of the lemma, there must be an $\alpha < \frac{1}{2}$ such that $\alpha Q + (1-\alpha)Q^- \in \succ^*$. Also, since $1-\alpha > \frac{1}{2}$, the argument used in the proof of the lemma shows that $(1-\alpha)Q + \alpha Q^- = [\alpha Q + (1-\alpha)Q^-]^- \in \succ^*$. This contradicts the asymmetricity axiom. A similar argument shows that assuming $[\frac{1}{2}Q + \frac{1}{2}Q^-]^- \in \succ^*$ leads to a contradiction to the asymmetricity axiom. Hence it must be the case that $\frac{1}{2}Q + \frac{1}{2}Q^- \in \sim^*$.

♠

Corollary 4.2 (Reduction of Effective Support) *Assume that axioms 4.1, 4.2, 4.3 and 4.4 hold. Fix $z \in Z$. Let $x, y \in Z$ to be such that $\delta_{(x,y)} \in \succ^*$. Then, for any $\alpha \in [0.1]$, $\alpha\delta_{(x,y)} + (1-\alpha)\delta_{(y,x)} \in \succ^*$ if and only if $\alpha > \frac{1}{2}$ and $(2\alpha - 1)\delta_{(x,y)} + 2(1-\alpha)\delta_{(z,z)} \in \succ^*$.*

Proof:

In the proof of the replication corollary, it was proven that $Q \equiv \alpha\delta_{(x,y)} + (1-\alpha)\delta_{(y,x)} \in \succ^*$ if and only if $\alpha > \frac{1}{2}$. For such α, $Q' \equiv (2\alpha-1)\delta_{(x,y)} + 2(1-\alpha)\delta_{(z,z)}$ is well defined and $\frac{1}{2}Q + \frac{1}{2}Q'^- \in \sim^*$ by the betweenness axiom. Hence, by the substitution axiom, $Q' \in \succ^*$.

♠

Proof of the proposition:

If $\delta_{(x,y)} \in \sim^*$ for all $x, y \in Z$, a simple induction using the betweenness axiom says that $\sim^* \cap \Lambda^F(Z \times Z, \Xi \otimes \Xi) = \Lambda^F(Z \times Z, \Xi \otimes \Xi)$. Hence $\Phi(x, y) \equiv 0$ is the unique function that satisfies the claim of the proposition.

Hereafter I assume that there is a $x, y \in Z$ such that $\delta_{(x,y)} \in \succ^*$. I fix such a pair and denote it by (x_0, y_0).

I prove the existence of the representation function Φ by actually constructing it. For each $(x, y) \in Z \times Z$ such that $\delta_{(x,y)} \in \succ^*$, the cancellation lemma guarantees the existence of unique $\theta(x, y) \in (0, 1)$such that $\theta(x, y)\delta_{(x,y)} + (1-\theta(x, y))\delta_{(y_0,x_0)} \in \sim^*$. Then, I define $\Phi(x, y)$ by $\Phi(x, y) \equiv \frac{1-\theta(x,y)}{\theta(x,y)}$. For each $(x, y) \in Z \times Z$ such that $\delta_{(x,y)} \in \sim^*$, I define $\Phi(x, y)$ by $\Phi(x, y) \equiv 0$. For each $(x, y) \in Z \times Z$ such that $(y, x) \in \succ^*$, I define $\Phi(x, y)$ by $\Phi(x, y) \equiv -\Phi(y, x)$. By the asymmetricity axiom and the nature of the construction, it is clear that Φ is skew-symmetric. Hence the only matter I have to prove is the additive representation property.

Let $Q \in \Lambda^F(Z \times Z, \Xi \otimes \Xi)$. Then I can express Q as

$$Q \equiv \sum_{i=1}^{N} \alpha_i \delta_{(x_i,y_i)},$$

where $\alpha_i > 0$ for all i and $\sum_{i=1}^{N} \alpha_i = 1$. I partition N to three sets in the following way.

$$N_+ \equiv \left\{ i \in N | \delta_{(x_i,y_i)} \in \succ^* \right\},$$
$$N_- \equiv \left\{ i \in N | \delta_{(y_i,x_i)} \in \succ^* \right\},$$
$$N_0 \equiv \left\{ i \in N | \delta_{(x_i,y_i)} \in \sim^* \right\}.$$

If $N_+ = N_- = \emptyset$, then an induction argument using the betweenness axiom shows that $Q \in \sim^*$. In this case, the construction of Φ implies that $\sum_{i=1}^{N} \alpha_i \Phi(x_i, y_i) = 0$. If $N_+ \neq \emptyset$ and $N_- = \emptyset$, then an induction argument using the substitution axiom and the asymmetricity axiom shows that $Q \in \succ^*$. Also, by the construction of Φ, $\sum_{i=1}^{N} \alpha_i \Phi(x_i, y_i) > 0$. If $N_+ = \emptyset$ and $N_- \neq \emptyset$, then an induction argument using the distributive law of the operator $^-$, the asymmetricity axiom and the substitution axiom proves that $Q^- \in \succ^*$. Again, by the construction of Φ, $\sum_{i=1}^{N} \alpha_i \Phi(x_i, y_i) < 0$.[11]

The only case left to prove is $N_+ \neq \emptyset$ and $N_- \neq \emptyset$. Pick any arbitrary $i \in N_+$ and $j \in N_-$. I rewrite the expression of Q as follows.

$$Q = (\alpha_i + \alpha_j) \left[\frac{\alpha_i}{\alpha_i + \alpha_j} \delta_{(x_i,y_i)} + \frac{\alpha_j}{\alpha_i + \alpha_j} \delta_{(x_j,y_j)} \right] + (1 - \alpha_i - \alpha_j) \sum_{k \neq i,j} \alpha_k \delta_{(x_k,y_k)}.$$

By the cancellation lemma, there is a unique $\theta^* \in (0,1)$ such that $\theta^* \delta_{(x_i,y_i)} + (1 - \theta^*) \delta_{(x_j,y_j)} \in \sim^*$. I distinguish three different cases.

[11] The steps in these induction arguments are almost exactly the same as the arguments used in the proof of the cancellation lemma. Hence I do not reproduce the details here.

<u>**Case 1**</u> $\frac{\alpha_i}{\alpha_i+\alpha_j} > \theta^*$.

By the substitution axiom, $\frac{\alpha_i}{\alpha_i+\alpha_j}\delta_{(x_i,y_i)} + \frac{\alpha_j}{\alpha_i+\alpha_j}\delta_{(x_j,y_j)} \in \succ^*$. Fix any $z \in Z$. I will show that there exists a $\beta \in (0,1)$ such that, if I define Q_β by $Q_\beta \equiv (\alpha_i + \alpha_j)[\beta\delta_{(x_i,y_i)} + (1-\beta)\delta_{(z,z)}] + (1-\alpha_i-\alpha_j)\sum_{k\neq i,j}\alpha_k\delta_{(x_k,y_k)}$, then:

1. $Q \in \succ^*$ if and only if $Q_\beta \in \succ^*$, $Q^- \in \succ^*$ if and only if $Q_\beta^- \in \succ^*$ and $Q \in \sim^*$ if and only if $Q_\beta \in \sim^*$;
2. $Q\Phi = Q_\beta\Phi$.

The task to find a β satisfying the first claim reduces to the task to find a $\beta \in (0,1)$ such that

$$\frac{1}{2}\left[\frac{\alpha_i}{\alpha_i+\alpha_j}\delta_{(x_i,y_i)} + \frac{\alpha_j}{\alpha_i+\alpha_j}\delta_{(x_j,y_j)}\right] + \frac{1}{2}\left[\beta\delta_{(y_i,x_i)} + (1-\beta)\delta_{(z,z)}\right] \in \sim^* .$$

The reason is as follows. By using the replication corollary of the lemma, it is easy to show that, for such β, $\frac{1}{2}Q + \frac{1}{2}Q_\beta^-$ is a mixture of two probabilities in $\Lambda^F(Z \times Z, \Xi \otimes \Xi) \cap \sim^*$ with coefficient $(\alpha_i + \alpha_j, 1-\alpha_i-\alpha_j)$. Hence $\frac{1}{2}Q + \frac{1}{2}Q_\beta^- \in \sim^*$ by the betweenness axiom. Then the substitution axiom and the asymmetricity axiom implies 1 of the claim. To prove 2 of the claim, I note at first the following simple fact. Since $\delta_{(z,z)} \in \sim^*$ by the asymmetricity axiom, the substitution axiom implies that finding β satisfying the restriction stated above is equivalent to finding a $\beta \in (0,1)$ such that

$$\frac{1}{1+\beta}\left[\frac{\alpha_i}{\alpha_i+\alpha_j}\delta_{(x_i,y_i)} + \frac{\alpha_j}{\alpha_i+\alpha_j}\delta_{(x_j,y_j)}\right] + \frac{\beta}{1+\beta}\delta_{(y_i,x_i)} \in \sim^* .$$

I determine β in this claim in the following way. By the Archimedian axiom, there exists a unique $\theta \in (0,1)$ such that
$R \equiv \theta\left[\frac{\alpha_i}{\alpha_i+\alpha_j}\delta_{(x_i,y_i)} + \frac{\alpha_i}{\alpha_i+\alpha_j}\delta_{(x_j,y_j)}\right] + (1-\theta)\delta_{(y_i,x_i)} \in \sim^*$. Then I define β by $\beta \equiv \frac{1-\theta}{\theta}$. By using the replication corollary, I can show that $\theta > \frac{1}{2}$. Hence $\beta \in (0,1)$. Clearly my claim is satisfied with this β.

Finally I have to show that $Q\Phi = Q_\beta\Phi$. To prove this, at first I derive a formula of θ^* in terms of θ.

I note that the expression of R can be rewritten as

$$R \equiv (\alpha'\theta + (1-\theta))\left[\frac{\alpha'\theta}{\alpha'\theta+(1-\theta)}\delta_{(x_i,y_i)} + \frac{1-\theta}{\alpha'\theta+(1-\theta)}\delta_{(y_i,x_i)}\right] + \theta(1-\alpha')\delta_{(x_j,y_j)},$$

where $\alpha' \equiv \frac{\alpha_i}{\alpha_i+\alpha_j}$. Since $R \in \sim^*$, the substitution axiom implies that $P \equiv \frac{\alpha'\theta}{\alpha'\theta+(1-\theta)}\delta_{(x_i,y_i)} + \frac{1-\theta}{\alpha'\theta+(1-\theta)}\delta_{(y_i,x_i)} \in \succ^*$. By applying the "reduction of effective support" corollary, this is equivalent to the statement that $\alpha'\theta > 1-\theta$ and $P' \equiv \frac{\alpha'\theta-(1-\theta)}{\alpha'\theta+(1-\theta)}\delta_{(x_i,y_i)} + \frac{2(1-\theta)}{\alpha'\theta+(1-\theta)}\delta_{(z,z)} \in \succ^*$. Moreover, by the betweenness axiom and the replication corollary, $\frac{1}{2}P + \frac{1}{2}P'^{-} \in \sim^*$. Let $R' \equiv (\alpha'\theta+(1-\theta))P' + \theta(1-\alpha')\delta_{(x_j,y_j)}$. The betweenness axiom says that $\frac{1}{2}R + \frac{1}{2}R'^{-} \in \sim^*$. Since $R \in \sim^*$, this implies, by the substitution axiom, that $R' \in \sim^*$. I can rewrite the expression of R' as

$$R' = (2\theta-1)\left[\frac{\alpha'\theta-(1-\theta)}{2\theta-1}\delta_{(x_i,y_i)} + \frac{\theta(1-\alpha')}{2\theta-1}\delta_{(x_j,y_j)}\right] + 2(1-\theta)\delta_{(z,z)}.$$

Since $R' \in \sim^*$ and $\delta_{(z,z)} \in \sim^*$, the substitution axiom says that

$$\frac{\alpha'\theta-(1-\theta)}{2\theta-1}\delta_{(x_i,y_i)} + \frac{\theta(1-\alpha')}{2\theta-1}\delta_{(x_j,y_j)} \in \sim^*,$$

or, $\theta^* = \frac{\alpha'\theta-(1-\theta)}{2\theta-1}$.

Next I compute a formula of θ^* in terms of $\Phi(x_i,y_i)$ and $\Phi(x_j,y_j)$. Let θ_i and θ_j to be balancing coefficients such that

$$\begin{aligned} \theta_i\delta_{(x_i,y_i)} + (1-\theta_i)\delta_{(y_0,x_0)} &\in \sim^*, \\ \theta_j\delta_{(x_j,y_j)} + (1-\theta_j)\delta_{(x_0,y_0)} &\in \sim^*. \end{aligned}$$

I take a mixture of these probability measures with coefficient $\left(\frac{1-\theta_j}{(1-\theta_i)+(1-\theta_j)}, \frac{1-\theta_i}{(1-\theta_i)+(1-\theta_j)}\right)$. By the betweenness axiom, the resulting probability must be in $\sim^*$, and is expressed as

$$\frac{\theta_i(1-\theta_j)}{(1-\theta_i)+(1-\theta_j)}\delta_{(x_i,y_i)} + \frac{\theta_j(1-\theta_i)}{(1-\theta_i)+(1-\theta_j)}\delta_{(x_j,y_j)} + \frac{2(1-\theta_i)(1-\theta_j)}{(1-\theta_i)+(1-\theta_j)}\left[\frac{1}{2}\delta_{(x_0,y_0)} + \frac{1}{2}\delta_{(y_0,x_0)}\right].$$

Since $\frac{1}{2}\delta_{(x_0,y_0)} + \frac{1}{2}\delta_{(y_0,x_0)} \in \sim^*$ by the replication corollary, the substitution axiom implies that

$$\frac{\theta_i(1-\theta_j)}{\theta_i(1-\theta_j)+\theta_j(1-\theta_i)}\delta_{(x_i,y_i)} + \frac{\theta_j(1-\theta_i)}{\theta_i(1-\theta_j)+\theta_j(1-\theta_i)}\delta_{(x_j,y_j)} \in \sim^*.$$

or, $\theta^* = \frac{\theta_i(1-\theta_j)}{\theta_i(1-\theta_j)+\theta_j(1-\theta_i)}$. By using the construction of Φ, the right hand side of this equation is rewritten as

$$\frac{\theta_i(1-\theta_j)}{\theta_i(1-\theta_j)+\theta_j(1-\theta_i)} = \frac{1}{1+\frac{1-\theta_i}{\theta_i}\frac{\theta_j}{1-\theta_j}} = \frac{1}{1-\frac{\Phi(x_i,y_i)}{\Phi(x_j,y_j)}}.$$

Now I can compare two formulas for θ^*. After some computation, I get

$$\theta = \frac{\Phi(x_i,y_i)}{(1+\alpha')\Phi(x_i,y_i)+(1-\alpha')\Phi(x_j,y_j)}.$$

Since $\beta = \frac{1-\theta}{\theta}$, a tedious computation shows that

$$\beta = \frac{\alpha'\Phi(x_i,y_i)+(1-\alpha')\Phi(x_j,y_j)}{\Phi(x_i,y_i)}.$$

It is obvious that this implies $Q\Phi = Q_\beta\Phi$.

<u>**Case 2**</u> $\frac{\alpha_i}{\alpha_i+\alpha_j} < \theta^*$.

By considering Q^- instead of Q, I can apply the same argument as the one for the previous case. Hence, I can find a $\beta \in (0,1)$ such that

1. $Q \in \succ^*$ if and only if $Q_\beta \in \succ^*$, $Q^- \in \succ^*$ if and only if $Q_\beta^- \in \succ^*$ and $Q \in \sim^*$ if and only if $Q_\beta \in \sim^*$;
2. $Q\Phi = Q_\beta\Phi$,

where $Q_\beta \equiv (\alpha_i+\alpha_j)[\beta\delta_{(x_j,y_j)} + (1-\beta)\delta_{(z,z)}] + (1-\alpha_i-\alpha_j)\sum_{k\neq i,j}\alpha_k\delta_{(x_k,y_k)}$.

Case 3 $\frac{\alpha_i}{\alpha_i+\alpha_j} = \theta^*$.

In this case, $\frac{\alpha_i}{\alpha_i+\alpha_j}\delta_{(x_i,y_i)} + \frac{\alpha_j}{\alpha_i+\alpha_j}\delta_{(x_j,y_j)} \in \sim^*$.
Let $Q' \equiv (\alpha_i + \alpha_j)\delta_{(z,z)} + (1-\alpha_i-\alpha_j)\sum_{k=i,j}\alpha_k\delta_{(x_k,y_k)}$. Then, from $\frac{1}{2}Q + \frac{1}{2}Q'^- \in \sim^*$, I can show that $Q \in \succ^*$ if and only if $Q' \in \succ^*$, $Q^- \in \succ^*$ if and only if $Q'^- \in \succ^*$ and $Q \in \sim^*$ if and only if $Q' \in \sim^*$. As I already proved, $\theta^* = \frac{1}{1-\frac{\Phi(x_i,y_i)}{\Phi(x_j,y_j)}}$. Since $\theta^* = \frac{\alpha_i}{\alpha_i+\alpha_j}$, a trivial computation leads to $\frac{\alpha_i}{\alpha_i+\alpha_j}\Phi(x_i,y_i) + \frac{\alpha_j}{\alpha_i+\alpha_j}\Phi(x_j,y_j) = 0$. Hence, $Q\Phi = Q'\Phi$.

In any case, I proved that I can reduce the cardinality of N_+ and/or N_- by one without affecting choice and value. I can continue this process until either $N_+ = \emptyset$ or $N_- = \emptyset$ happens. I already proved that, in these cases, the statement of the proposition holds. This completes the existence part of the proof.

I still need to prove the uniqueness part of the statement. Suppose the uniqueness statement is false. Then there are two representation functions, Φ and Φ', and two outcome pairs, (x,y) and (v,w), such that:

1. $\delta_{(x,y)} \in \succ^*$ and $\delta_{(v,w)} \in \succ^*$;

2. If $\alpha \equiv \frac{\Phi(x,y)}{\Phi'(x,y)}$ and $\beta \equiv \frac{\Phi(v,w)}{\Phi'(v,w)}$, then $\alpha < \beta$.

We can find a $\theta \in (0,1)$ such that $\theta\Phi'(x,y) + (1-\theta)\Phi'(w,v) = 0$. By the representation result, it must be the case that $\theta\Phi(x,y) + (1-\theta)\Phi(w,v) = 0$. But the choice of (x,y) and (v,w) implies that $\theta\Phi(x,y) + (1-\theta)\Phi(w,v) < \theta\alpha\Phi'(x,y) + (1-\theta)\alpha\Phi'(w,v) = 0$. This is a contradiction. Hence, I get the uniqueness of the representation function up to scale.

This completes the proof.

♠.

5 Applications of the Theory — Solving Various Paradoxes in Expected Utility Theory —

5.1 Introduction

In this section,

1. I introduce three restrictions on the characteristics of a choice maker that causes the paradoxes,

2. I identify the origin of each well-known empirical paradox in expected utility theory by using the characterization of a choice maker in the model and the representation result, and

3. I criticize properties imposed on prospect functions in prospect theory by using the representation result.

This subsection is devoted to the task 1.

Finally, I explain three restrictions imposed on $(\Gamma, \succ)$ in the analysis of paradoxes. The first one is for Γ and the rests are for $\succ$, or, more specifically, for its representation function Φ.

The first restriction is that Γ associates, with each choice situation between lotteries, choice situations in which the two random variables are negatively correlated. This is a form of regret. Suppose that a subject is facing a binary choice situation between lotteries and that she focuses on one of the lotteries. The lottery to which the subject pays more attention gets a priority in choice. The other lottery represents a missed opportunity. Hence, it is plausible for a subject to have a form of regret such that she pays more attention to bad prizes than good prizes in the lottery on which she is focusing and vice versa for the other lottery. The symmetric consideration apply if she focuses on the other lottery. As a result, the subject perceives that events in which bad prizes from one lottery and good prizes from the other occur have relatively high probability.

The second restriction is that $\succ$ exhibit increasing risk aversion.

To define "increasing risk aversion," I need an order on lotteries induced by $\succ$ given a money-prize x. The construction of this order can be carried out for any outcome space. Hence I use the term "outcome" instead of "money-prize" in this paragraph. Fix an outcome x. Suppose that a choice between simple lotteries, (p, q), is given. Let P be an atomless probability measure on the state space and let E be an event such that $P(E) = \frac{1}{2}$. (E exists because P is atomless.) Let f to be a random variable defined on E such that $P_E \circ f^{-1} = p$, where P_E is the conditional probability of P given E. Similarly, let g to be a random variable defined on $\Omega - E$ such that $P_{\Omega - E} \circ g^{-1} = q$, where $P_{\Omega-E}$ is a conditional probability of P given $\Omega - E$. (Since P is atomless, such f and g exist.) Next I define two random variables $\bar{f}$ and $\bar{g}$ as follows.

$$\bar{f}(\omega) = \begin{cases} f(\omega), & \text{if } \omega \in E, \\ x, & \text{if } \omega \in \Omega - E. \end{cases}$$

$$\bar{g}(\omega) = \begin{cases} x, & \text{if } \omega \in E, \\ g(\omega), & \text{if } \omega \in \Omega - E. \end{cases}$$

I define an ordering on simple lotteries, $\succ_x^*$, by the following rule: $(p, q) \in \succ_x^*$ if and only if, for any atomless probability on the state space, P, and random variables, f and g as described, $(\bar{f}, \bar{g}; P) \in \succ$. $\succ_x^*$ is well defined, since $\succ$ is outcome-oriented. Moreover, if $\succ$ satisfies all the axioms of sections 3, then $\succ_x^*$ is asymmetric and negative transitive. This can be verified easily by using the representation of $\succ^*$. It is also clear that $\succ_x^*$ is represented by an expected utility with von-Neumann Morgenstern utility function $\Phi(., x)$. I call $\succ_x^*$ choice order on lotteries with base x.[12]

I define the increasing risk aversion as a property exhibited by a collection of orderings on lotteries, $\{\succ_x^*\}$. Roughly speaking, $\{\succ_x^*\}$ exhibit increasing risk aversion if $\succ_x^*$ becomes more risk-averse with increase in the money-prize x. With the representation result for $\succ^*$, I can formalize this idea. Before giving the definition, I recall that, in expected utility theory, the absolute risk aversion coefficient of a von-Neumann Morgenstern utility function, u, is defined by the formula $-\frac{u'}{u''}$.

[12]It is trivial to see that (1) $\bar{f}$ and $\bar{g}$ are uncorrelated, and (2) f and g are compared indirectly as prospects from x. These are two important characteristics in the definition of $\succ_x^*$. Since we can regard $\Phi(., x)$ as a prospect function with base x defined in [Kahneman & Tversky, 79], I could say that these characteristics give a foundation of the prospect theory in my theory.

Definition 5.1 *Assume that $\succ$ satisfies all the axioms of sections 3. Also assume that each $\succ_x^*$ is risk-averse, i.e., $\Phi(.,x)$ is increasing and concave. Let $A_\Phi(y,x)$ to be the absolute risk aversion coefficient of $\Phi(.,x)$ evaluated at y.*[13] *Then, $\succ$ exhibit increasing risk aversion if, for any money-prize y, $A_\Phi(y,x)$ is increasing with respect to x.*

It is natural to assume the increasing risk aversion for $\succ$. Recall that, in expected utility theory, we can regard the decreasing absolute risk aversion property of a von-Neuman Morgenstern utility function, u, as a comparison of absolute risk aversion for two von-Neuman Morgenstern utility functions, $\bar{u}$ and $\underline{u}$. Both $\bar{u}$ and $\underline{u}$ are constructed from u to measure prospects from some base money-prizes. The expected utility theory explains the base money-prize as choice-maker's wealth. The base money-prize of $\bar{u}$ is set higher than that of $\underline{u}$, and u exhibits decreasing absolute risk aversion if and only if, for any $\bar{u}$ and $\underline{u}$ constructed from u, the absolute risk aversion coefficient of $\bar{u}$ is less than that of $\underline{u}$ in a neighborhood of 0. The increasing risk aversion property of $\{\succ_x^*\}_{x\in R}$ is a counterpart of this decreasing absolute risk aversion property in expected utility theory. The risk aversion increases because, in my theory, a base money-prize x represents an opportunity that a choice maker misses, while a base money-prize in expected utility theory represents wealth that the choice maker already has. The definition of increasing risk aversion is made global in order to analyze paradoxes in expected utility theory.

Though the idea of increasing risk aversion comes from the expected utility theory, it is clear that the expected utility theory cannot exhibit this property. This is because my theory internalize the expected utility theory as a case in which $\Phi(x,y) = u(x) - u(y)$ where u is a von-Neumann-Morgenstern utility function.

The third restriction is similar to the second one, but applied when the outcome space is the set of all probabilities on money-prizes, $\Lambda(R)$. I consider this case in analyzing one of parodoxes, called the common consequence effect. The restriction is that $\succ$ exhibits a change of attitude toward risk.

To explain this restriction, I define the notion of "risk-averter" and "risk-taker" when the outcome space is $\Lambda(R)$. For each $\eta \in \Lambda(R)$, let $\succ_\eta^*$ be as in the previous paragraph. It is clear that

[13] This coefficient is well defined almost everywhere because a concave function defined on real numbers is twice differentiable almost everywhere. For simplicity, I assume that it is well-described everywhere.

$\succ^*_\eta$ induces an order on the outcome space $\Lambda(R)$ that is represented by a utility function $\Phi(.,\eta)$. Let $\succ^*_\eta$ refer also to this order on $\Lambda(R)$, and let (δ_x, ξ) to be a pair of probabilities on R such that the mean of ξ is x. I say that $\succ^*_\eta$ is risk-averse (risk-taking) with respect to (δ_x, ξ) if $(\delta_x, \xi) \in \succ^*_\eta$ $((\xi, \delta_x) \in \succ^*_\eta)$.

The change of attitude toward risk mentioned in the restriction is of the following type; for any (δ_x, ξ) such that x is the mean of ξ, $\succ^*_\eta$ is risk-taking when η is concentrated at bad prizes relative to the support of ξ and risk-averse when η is concentrated at good prizes relative to the support of ξ. It is plausible that a fair bet would be more attractive to a choice maker who pays attention to the possibility of bad prizes from missed alternatives. Hence, it is not unrealistic to impose this restriction. Note that all notions involving risk are defined entirely on the outcome space. The risk that I am describing in the current paragraph has nothing to do with lotteries, i.e., probability measueres on the outcome space. In this sense, a peculiarity of the outcome space, $\Lambda(R)$, causes the change of attitude toward risk.[14]

In the following examples, we show how these restrictions explain paradoxes in expected utility theory. For descriptions of paradoxes, I follow [Machina, 87].

5.2 The Allais Paradox

Consider a choice between the following two lotteries.

Lottery 1 One million dollar for sure.

Lottery 2 Five million dollars with probability 10%, nothing with probability 1%, one million dollar with probability 89%.

Experiments show that subjects tend to choose lottery 1 over lottery 2. Next consider a choice between the following two lotteries.

[14] It seems to me that another property is also natural. To explain this, note that $\Lambda(R)$ can be partially ordered by the first-order stochastic dominance. Then, $\succ$ could be said to exhibit increasing risk aversion with respect to (δ_x, ξ) if $\Phi(\delta_x, \eta) - \Phi(\xi, \eta)$ is increasing with respect to η, where the outcome space, $\Lambda(R)$, is ordered by first-order stochastic dominance. It is plausible that, for any (δ_x, ξ) such that x is the mean of ξ, $\succ$ exhibits increasing risk aversion.

Lottery 3 Five million dollars with probability 10%, nothing with probability 90%.

Lottery 4 One million dollars with probability 11%, nothing with probability 89%.

Again experiments show that subjects tend to choose lottery 3 over lottery 4.

These experimental results contradict the expected utility theory. Let u to be the von-Neumann-Morgenstern utility function of a subject, normalized so that $u(0) = 0$. Then the first choice indicates that

$$u(1) > 0.1u(5) + 0.89u(1),$$

which is equivalent to

$$0.11u(1) > 0.1u(5).$$

Hence, a subject whose choice satisfies the axioms of expected utility theory must choose lottery 4 over lottery 3.

This empirical paradox in expected utility theory was found by M.Allais in the 1950's. His original paper, written in French and published in Econometrica, was translated in English in [Allais, 79].

This paradox is considered as an evidence for the violation of the *independence axiom* of expected utility theory. For any prize x, let δ_x to be the lottery giving x million dollars for sure. Also, let ξ to be the lottery giving five million dollars with probability $\frac{10}{11}$ and nothing with probability $\frac{1}{11}$. Then, we have

$$\text{Lottery } 1 = \delta_1 = \frac{11}{100}\delta_1 + \frac{89}{100}\delta_1,$$

$$\text{Lottery } 2 = \frac{11}{100}\xi + \frac{89}{100}\delta_1,$$

and

$$\text{Lottery } 3 = \frac{11}{100}\xi + \frac{89}{100}\delta_0,$$

$$\text{Lottery } 4 = \frac{11}{100}\delta_1 + \frac{89}{100}\delta_0,$$

where the mixture operations on the right hand sides are defined in an obvious way. The *independence axiom* says that, given a choice situation (α, β), any mixture of this situation with a diagonal choice situation (γ, γ) does not change a subject's choice. It is clear that the independence axiom is violated by the experimental result.

What causes this violation of the independence axiom in the experiment, then? Suppose that all lotteries are induced by random variables on a state space on which a probability is specified. I observe that random variables that induce lottery 1 and lottery 2 must be uncorrelated, but, random variables that induce lottery 3 and lottery 4 can be correlated. The expected utility theory cannot analyze this difference since, in it, correlation between objects of choice is out of scope and, hence, does not have any influence on choice. On the other hand, in my theory, correlation between objects of choice has a significant impact on choice.[15]

In my theory, the surmising process, Γ, incorporates correlation. For illustrative purposes, assume that Γ selects a choice situation in which [5 million dollars for lottery 3 and nothing for lottery 4] happens with probability 10%, [nothing for lottery 3 and 1 million dollars for lottery 4] happens with probability 11%, [nothing for the both lottery] happens with probability 79%. This is the case of maximal negative correlation. The representation result says that a subject chooses the lottery 3 over the lottery 4 if and only if the following inequality holds.

$$\frac{10}{100}\Phi(5,0) > \frac{11}{100}\Phi(1,0).$$

[15]Here, I want to point out that the expected utility theory does not allow a natural mixture operation that can be considered in my theory, because it neglects correlation among random variables. Recall that, given a choice situation between lotteries, (α, β), a surmising process, Γ, selects a set of well-described choice situations between random variables, $(f, g; P)$, such that $\alpha = P \circ f^{-1}$ and $\beta = P \circ g^{-1}$. Assume that P is atomless and f, g are independent. That is, Γ selects a choice situation that is consistent with the expected utility theory. For any given weight $w \in (0,1)$, we can choose an event E in the state space such that $P(E) = w$, and independent random variables f' and g' defined on E such that $P_E \circ f'^{-1} = \alpha$ and $P_E \circ g'^{-1} = \beta$. Also, we can take a random variable h' defined on $\Omega - E$ such that $P_{\Omega - E} \circ h'^{-1} = \gamma$. Let $\bar{f}$ be the obvious concatenation of f' and h'. Similarly, let $\bar{g}$ be the obvious concatenation of g' and h'. Then, $(\bar{f}, \bar{g}; P)$ induces a distribution over pairs of prizes whose marginal distributions are specified as ones in the independence axiom. This is a very intuitive mixture operation for random variables. In my theory, the surmising process, Γ, may select it for a pair of lotteries, $(w\alpha + (1-w)\gamma, w\beta + (1-w)\gamma)$. On the other hand, in expected utility theory, the induced distribution on pairs of prizes by $(\bar{f}, \bar{g}; P)$ cannot be considered because $\bar{f}$ and $\bar{g}$ are clearly not independent. This illustrates that the expected utility theory is more restrictive about the mixture operation than my theory.

Next, consider choice between lottery 1 and lottery 2. According to the theory, a choice maker will choose the lottery 1 over the lottery 2 if and only if the following inequality holds.

$$\frac{10}{100}\Phi(5,1) < \frac{1}{100}\Phi(1,0).$$

Note that $\Phi(5,0)$ appears in the first inequality whereas the $\Phi(5,1)$ appears in the second inequality. Hence, these two inequalities can be treated separately. If $\Phi(5,1)$ is significantly smaller than $\Phi(5,0)$, these inequalities are compatible. For any subject who feels that earning one million dollars is like a dream, this will be the case. This simple result suggests that the Allais paradox happens when Γ selects choice situations in which the two random variables show significant negative correlation.

To see the importance of correlation, it helps to consider the case in which, given the choice between lottery 3 and lottery 4, the surmising process selects a choice situation in which the two random variables show maximal positive correlation. This choice situation is described by; (5 million, 1 million) with probability 10%, (0, 1 million) with probability 1% and (0, 0) with probability 89%. (The first coordinate is the prize from lottery 3 and the second coordinate is that from lottery 4.) By comparing this description to the one for the choice between lottery 1 and lottery 2, it is easy to see that the only difference is that the latter gives (0, 0) a probability 89% while the former gives (1, 1) a probability 89%. Note that, since Φ is skew-symmetric, $\Phi(x, x) = 0$ for all x. Hence, this difference cannot affect choices of subjects in my theory. This shows that, in my theory, the Allais paradox cannot happen if the surmising process associates, with the choice between lottery 3 and lottery 4, a choice situation in which the two random variable show maximal positive correlation.

To analyze the necessity of negative correlation more systematically, I introduce a simple index of negative correlation, $p \in \left[0, \frac{10}{89}\right]$. It is the conditional probability of getting five million dollars from lottery 3 given that lottery 4 gives nothing. It is trivial to see that p measures the degree of negative correlation conveyed by choice situations in Γ(lottery3, lottery 4). Given the choice between lotteries 3 and 4, each well-described choice situation selected by the surmising process can be fully described in the following way with an adequate p,

- Given the information that lottery 4 gives nothing, the conditional probability of getting five million dollars from lottery 3 is p and that of getting nothing from lottery 3 is $1-p$.
- The unconditional distributions of prizes from lottery 3 and lottery 4 are those specified in the description of lottery 3 and lottery 4.

The distribution induced by the choice situation described as above is denoted by $\pi[p]$. It is given by

$$\begin{aligned}
\pi[p](\{(5,0)\}) &= \frac{89}{100}p, \\
\pi[p](\{(0,0)\}) &= \frac{89}{100}(1-p), \\
\pi[p](\{(0,1)\}) &= \frac{11}{100} - \left(\frac{1}{10} - \frac{89}{100}p\right), \\
\pi[p](\{(5,1)\}) &= \frac{1}{10} - \frac{89}{100}p.
\end{aligned}$$

According to the theory, the Allais paradox happens if

$$\frac{10}{11}\Phi(1,5) + \frac{1}{11}\Phi(1,0) > 0,$$

and, for each choice situation in $\Gamma(\text{lottery 3}, \text{lottery 4})$ that is described by p,

$$\frac{89}{11}p\left[(\Phi(5,0) - \Phi(0,0)) - (\Phi(5,1) - \Phi(0,1))\right] > \frac{10}{11}\Phi(1,5) + \frac{1}{11}\Phi(1,0).$$

The first inequality comes from the choice of lottery 1 over lottery 2. It is equivalent to choosing the chance of getting 1 million dollars for sure over the risky opportunity of getting five million dollars with probability $\frac{10}{11}$ and nothing with probability $\frac{1}{11}$. The second inequality comes from the the choice of lottery 3 over lottery 4. I present it in this form to see the implication on two prospect functions, $\Phi(.,0)$ and $\Phi(.,1)$. At first, note that the right-hand side of this inequality is positive if a subject chooses lottery 1 over lottery 2. Therefore, this inequality implies that the increase of prospect given by the prize increase, from nothing to five millions, based on nothing must be bigger than that based on one million dollars. This condition and the first inequality determine necessary conditions on Φ for a subject to exhibit the Allais paradox. Secondly, if Φ satisfies these conditions, the second inequality gives a lower bound, $c(\succ)$, of indexes associated with choice situations selected by the surmising process. By rewriting the second inequality, we get

$$p > \frac{10}{89}\frac{10\Phi(1,5)+\Phi(1,0)}{10\Phi(1,5)+10\left[\Phi(5.0)-\Phi(1.0)\right]}.$$

The implication is that, in order to choose lottery 3 over lottery 4, the surmising process cannot select any choice situations in which the two random variables show significant positive correlation.[16] According to the uniqueness of representation up to scale that I have proved in section 4, $c(\succ)$ is uniquely determined by $\succ$. Since p cannot take a value higher than $\frac{10}{89}$, the following condition must be satisfied.

$$c'(\succ) \equiv \frac{10\Phi(1,5)+\Phi(1,0)}{10\Phi(1,5)+10\left[\Phi(5,0)-\Phi(1,0)\right]} < 1.$$

This gives another necessary condition on Φ for the Allais paradox, that is[17]

$$10\Phi(5,0) > 11\Phi(1,0).$$

The larger the difference $10\Phi(5,0) - 11\Phi(1,0)$ is, the lower is $c(\succ)$.

What the analysis suggests is that, in my theory, the choice maker exhibits the Allais paradox if and only if three strict inequalities for the function Φ are satisfied and the parameter p is above a lower bound which depends also on Φ. Three inequalities are verifiable by asking subjects. The first inequality is verified in the experiment 1. To verify the second inequality, I prepare a box and 30 balls. 10 balls are colored blue, other 10 balls are red, and the rest are white. The box is carefully selected so that nobody can see its content. I offer the following procedure to subjects, in which they have a chance to earn some money. I say that I will put all balls in the box and pick up one ball from it. Before I pick a ball, I allow for each subject to choose either "red" or "blue." I

[16] If $c(\succ)$ is less than $\frac{1}{10}$, then the subject exhibits the Allais paradox even when random variables that generates the lottery 3 and the lottery 4 are believed to be independent.

[17] We already encountered this condition in the illustration with maximal negative correlation.

declare that, if the subject chooses "red," then the consolation color is "white," while, if he chooses "blue," then the consolation color is "red." I guarantee prizes as follows. If the color which the subject has specified matches the color of ball I pick up, then I give five million dollars. If the color of the ball I pick up is the consolation color for subject's choice, then I give one million dollars. Otherwise, I give nothing. The second inequality is equivalent to say that subjects chooses "blue" over "red" in the experiment. A similar experiment, without consolation colors, can be arranged to verify the third inequality. I summarize these experiments (including an obvious one for the first inequality) in the following tables.

	B	R
B	1	1
∨		
R	5	0
	$\frac{10}{11}$	$\frac{1}{11}$

	B	R	W
B	5	1	0
∨			
R	0	5	1
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

	B	R
B	5	0
∨		
R	0	1
	$\frac{10}{21}$	$\frac{11}{21}$

Tables 1

Note that choice according to the expected utility theory is inconsistent with the experimental result in the second table, since both random variables in the table induce the same lottery.

Proposition 5.1 *Assume that $\succ$ satisfies the axioms in the section 3, so that it is represented by an additive formula with skew-symmteric function $\Phi : R \times R \to R$. Then, $\succ$ exhibits the Allais paradox if and only if:*

1. *$\succ$ is consistent with the experimental results in tables;*

2. *For any $(f, g; P) \in \Gamma$(lottery 3, lottery 4) for which the correlation between f and g is indexed by p,*

$$p > \frac{10}{89} c'(\succ),$$

where $c'(\succ) \equiv \frac{10\Phi(1,5)+\Phi(1,0)}{10\Phi(1,5)+10[\Phi(5,0)-\Phi(1,0)]}$.

The trick of the Allais paradox is that the prizes, five million and one million dollars, are carefully configured so that $\Phi(5,1)$ is very small relative to probabilities of the event in which $(5,1)$ happens, while the difference between $\Phi(5,0)$ and $\Phi(1,0)$ is not negligible. When $\Phi(5,1)$ is very small, the experimental results in the first two tables are likely to follow. When $\Phi(5,0)$ is recognizably larger than $\Phi(1,0)$, the experimental result in the third table is likely to follow and $c'(\succ)$ is likely to be low. It seems to me that the choice of unit as million dollars helps to accomplish both of them.

To test my explanation, note that my theory says that subjects, who are inconsistent with the experimental results in tables, does not exhibit the Allais paradox. Hence, if an experimenter could collect many subjects who are inconsistent with at least one of experimental results in tables and exhibits the Allais paradox, then my explanation should be rejected.

5.3 Common Consequence Effect

I specify two lotteries $\bar{\ell}$ and $\underline{\ell}$ as follows. I describe $\bar{\ell}$ by $(z_1, z_2; p, 1-p)$, meaning that the money-prize z_1 is obtained with probability p and z_2 is obtained with probability $1-p$. Similarly, I describe $\underline{\ell}$ by $(z_1', z_2'; p, 1-p)$, meaning that z_1' is obtained with probability p and z_2' is obtained with probability $1-p$. I assume that $z_1 > z_2$, $z_1' > z_2'$, $z_1 > z_1'$ and $z_2 > z_2'$. In other words, I assume that $\bar{\ell}$ stochastically dominates $\underline{\ell}$ by the first order. Let x be a money-prize and δ_x be a "lottery" that gives x for sure. Also let ξ be a lottery with finite support such that its support contains prizes both greater and less than x and a subject is "indifferent" between δ_x and ξ or chooses δ_x over ξ. Finally let $w \in (0,1)$ be a weight to construct choice situations in experiments.

I define lottery 1 and lottery 2 as follows. Lottery 1 is constructed by $w\delta_x + (1-w)\bar{\ell}$. Lottery 2 is constructed by $w\xi + (1-w)\bar{\ell}$. If outcomes are money-prizes, these are two probabilities on R. If outcomes are probabilities on R, then lottery 1 is described by $(\delta_x, \bar{\ell}; w, 1-w)$ and lottery 2 is described by $(\xi, \bar{\ell}; w, 1-w)$. These are probabilities on $\Lambda(R)$.

Next, I define lottery 3 and lottery 4 as follows. Lottery 3 is constructed by $w\delta_x + (1-w)\underline{\ell}$. Lottery 4 is constructed by $w\xi + (1-w)\underline{\ell}$. If R is the outcome space, these are two distributions on R. If $\Lambda(R)$ is the outcome space, lottery 3 is described by $(\delta_x, \underline{\ell}; w, 1-w)$ and lottery 4 is described by $(\xi, \underline{\ell}; w, 1-w)$. These are probabilities on $\Lambda(R)$.

Results in several experiments suggest that there are configurations of x and ξ such that subjects show tendency to choose lottery 1 over lottery 2 and lottery 4 over lottery 3 when the weight w is sufficiently bounded away from 0 and 1. This tendency is called *common consequence effect*. An argument used in 5.2 suggests that we may regard the Allais paradox as a particular case of the common consequence effect.

It is clear that the common consequence effect violates the independence axiom in expected utility theory if outcomes are money-prizes. Traditional explanation about this paradox is formed in terms of change of attitude toward risk. According to this explanation, a subject is risk-averse when she is asked the choice between lottery 1 and lottery 2 and risk-taking when she is asked the choice between lottery 3 and lottery 4. Hence a mixture with a sure outcome is chosen in the former choice situation while a mixture with risky lottery is chosen in the latter choice situation. Note that the analysis of the Allais paradox in 5.2 did not address such change of attitude toward risk. Hence, if we regard the Allais paradox as a particular case of the common consequence effect, we cannot justify the traditional explanation in my theory (or I have to analyze the Allais paradox in an entirely different way).

This puzzle is solved by examining reduction of compound lotteries in experiments. In experiments dedicated to the common consequence effect, an experimenter usually gives subjects compound structure of lotteries. On the other hand, in experiments to examine the Allais paradox, the experimenter does not mention anything about compound structure of lotteries. If subjects do not reduce compound lotteries in experiments, then we can analyze these two paradoxes in totally different way. Hence, the question is whether or not subjects reduce compound structure of lotteries in experiments dedicated to the common consequence effect. I claim that, except in very simple configurations, subjects do not reduce lotteries. The reason is as follows. It is natural to assume that resources available to subjects (memory, computation device, experience in choice, etc.) are bounded. I believe that most of subjects prefer a simple presentation of choice situations than a complicated one because they do not want to consume much resources just to figure out problems to which they are facing. I believe that they want to allocate more resources to solve problems.[18] Reducing compound lotteries may consume too much resources. Note that computation of reduced lotteries is not trivial even when compound structure is fairly simple. Hence, if the experimenter offers choice with simple compound structure of lotteries, subjects would not have any incentive to reduce lotteries.[19]

I show that, if subjects do not reduce compound lotteries, then my theory justifies the traditional explanation. Without reduction of compound lotteries, the outcome space must be $\Lambda(R)$. If $\succ$ exhibits change of attitude toward risk as explained in the introduction, then the common consequence effect happens when $\bar{\ell}$ is concentrated on good prizes relative to the support of ξ and $\underline{\ell}$ is concentrated on bad prizes relative to the support of ξ.

[18]Here I am concerned about model selection. A complete theory of choice requires characterization of a choice maker to reveal a criterion for model selection. What I suggest in this paragraph is that such criterion would maximize the possibility of solving a proposed problem. This is a form of outcome-orientation. There are other important aspects that a criterion for model selection should reveal. For example, it should restrict natures of acceptable solutions. Unfortunately, model selection is beyond the scope of my theory. At the current stage, I have not found a good characterization of a choice maker including model selection.

[19]In general, it is possible for subjects to reduce a compound lottery to a probability on money-prizes. But, the converse operation is impossible because there are multiple candidates as target. Hence, once we agree that subjects do not reduce compound lotteries in experiments for the common consequence effect, we cannot regard the Allais paradox as a special case of the common consequence effect. Some researchers (for example, [Krantz, Luce, Suppes & Tversky, 90]) still regard the Allais paradox as a special case of the common consequence effect. I believe that is wrong.

To see the point formally, I describe each choice situation selected by the surmising process by using a parameter r and r'. It turns out that r and r' measures the degree of negative correlation. Consider the choice between lottery 1 and lottery 2. Let r to be the conditional probability of getting $\bar{\ell}$ from lottery 2 given that lottery 1 yields δ_x. Each choice situation in Γ(lottery 1, lottery 2) is described in the following way:

- Given the outcome from lottery 1 is δ_x, the probability to get the outcome $\bar{\ell}$ from lottery 2 is r and the probability to get the outcome ξ from lottery 2 is $1 - r$;
- The unconditional distribution of outcomes from lottery 1 is $(\delta_x, \bar{\ell}; w, 1 - w)$ and that from lottery 2 is $(\xi, \bar{\ell}; w, 1 - w)$.

It is easy to verify that, with outcome-orientation, this description gives all relevant information for choice. The range of r is $\left[0, \frac{1-w}{w}\right]$. Next, consider the choice between lottery 3 and lottery 4. Let r' to be the conditional probability of getting δ_x from lottery 3 given that lottery 4 yields $\underline{\ell}$. Each choice situation in Γ(lottery 3, lottery 4) is described in the following way:

- Given the outcome from lottery 4 is $\underline{\ell}$, the probability to get the outcome δ_x from the lottery 3 is r' and the probability to get the outcome $\underline{\ell}$ from lottery 3 is $1 - r'$;
- The unconditional distribution of outcomes from lottery 3 is $(\delta_x, \underline{\ell}; w, 1 - w)$ and that from the lottery 4 is $(\xi, \underline{\ell}; w, 1 - w)$.

The range of r' is $\left[0, \frac{w}{1-w}\right]$. For simplicity, I assume that a subject is "indifferent" between δ_x and ξ.

According to my theory, a subject chooses lottery 1 over lottery 2 if and only if, for any choice situation in Γ(lottery 1, lottery 2) that is described by r, $w(1-r)\Phi(\delta_x, \xi) + wr\Phi(\delta_x, \bar{\ell}) + wr\Phi(\bar{\ell}, \xi) > 0$. Since $\Phi(\delta_x, \xi) = 0$ is assumed, this is equivalent to $\Phi(\delta_x, \bar{\ell}) > \Phi(\xi, \bar{\ell})$. Similarly, a subject chooses lottery 4 over lottery 3 if and only if, for any choice situation in Γ(lottery 3, lottery 4) that is described by r', $[w - (1 - w)r']\Phi(\xi, \delta_x) + (1 - w)r'\Phi(\underline{\ell}, \delta_x) + (1 - w)r'\Phi(\xi, \underline{\ell}) > 0$. Again this is equivalent to $\Phi(\delta_x, \underline{\ell}) < \Phi(\xi, \underline{\ell})$. Note that these conditions has nothing to do with the parameter

r and r'. In the setup with the outcome space $\Lambda(R)$, the common consequence effect has nothing to do with the surmising process Γ if ξ is a fair bet with respect to x.[20]

When subjects reduce compound lotteries, it is not clear whether the traditional explanation can be supported in my theory. I analyze only a very simple case by assuming that the surmising process selects a choice situation that shows the maximal negative correlation between the two random variables. This restriction on the surmising process is effectively expressed by the following simple statement;

> Given the lowest prize from lottery to which a subject is paying attention, the probability to get the lowest prize from the other lottery is 0.

I show that, if configuration of choice situations satisfies regularity conditions (just as in the Allais paradox), the common consequence effect happens. This indicates that, if the surmising process selects choice situations in which the two random variables are significantly negatively correlated, then the common consequence effect is likely to happen. This explanation is consistent with the one for the Allais paradox, as it must be.

At first I describe a simple case that I analyze. I believe that the only simple case like the one here could be reducible. I assume that only three prizes $\underline{x} < x < \bar{x}$ are available. ξ is described by $(\bar{x}, \underline{x}; q, 1-q)$. $\bar{\ell}$ is described by $(\bar{x}, x; p, 1-p)$. $\underline{\ell}$ is described by $(x, \underline{x}; p, 1-p)$.[21] By reducing compound structure, lottery 1 is described by $(\bar{x}, x; (1-w)p, w+(1-w)(1-p))$ and lottery 2 is described by $(\bar{x}, x, \underline{x}; (1-w)p+wq, (1-w)(1-p), w(1-q))$. The surmising process (obeying the simple rule stated in the previous paragraph) determines a distribution π over pairs of money-prizes as follows.

[20] If we take ξ to be a favorable bet, i.e., $\Phi(\xi, \delta_x) > 0$, or an unfavorable bet, i.e., $\Phi(\xi, \delta_x) < 0$, then results are not independent of the surmising process. The analysis for these cases are similar to the one for the case with fair bet. Since results are not so illustrative, I omit them here. In any case, it is sufficient for the common consequence effect to happen that subjects exhibits change of attitude toward risk with adequate strength.

[21] Note that the analysis is more complicated than that for the Allais Paradox because both $\bar{\ell}$ and $\underline{\ell}$ are not sure-outcome lotteries. In particular, no simple index for negative correlation is available. This is the reason why I restrict my analysis to the case in which the surmising process selects a choice situation showing maximal negative correlation.

$$\begin{aligned}\pi(\bar{x}, x) &= (1-w)p - w(1-q), \\ \pi(\bar{x}, \underline{x}) &= w(1-q), \\ \pi(x, \bar{x}) &= wq + (1-w)p, \\ \pi(x, x) &= w(1-q) + (1-w)(1-2p).\end{aligned}$$

The restriction for the weight w is

$$\frac{2p-1}{1-q} \leq \frac{w}{1-w} \leq \frac{p}{1-q}.$$

This restriction says that w must be sufficiently bounded away from 0 and 1. According to my theory, a subject chooses lottery 1 over lottery 2 if and only if $\frac{\Phi(\bar{x},x)}{\Phi(\bar{x},\underline{x})} < 1-q$. Similarly, lottery 3 is described by $(x, \underline{x}; w + (1-w)p, (1-w)(1-p))$ and lottery 4 is described by $(\bar{x}, x, \underline{x}; wq, (1-w)p, w(1-q)$ The surmising process determines a distribution π' over pairs of money-prizes as follows.

$$\begin{aligned}\pi'(\underline{x}, x) &= (1-w)(1-p) - wq, \\ \pi'(\bar{x}, \underline{x}) &= wq, \\ \pi'(x, \underline{x}) &= w(1-q) + (1-w)(1-p), \\ \pi'(x, x) &= wq + (1-w)(2p-1).\end{aligned}$$

The restriction on the weight w is

$$\frac{1-2p}{q} \leq \frac{w}{1-w} \leq \frac{1-p}{q}.$$

Again this says that w must be sufficiently bounded away from 0 and 1. According to my theory, a subject chooses lottery 4 over the lottery 3 if and only if $\frac{\Phi(x,\underline{x})}{\Phi(\bar{x},\underline{x})} < q$.

Now we get the following inequalities,

$$\frac{\Phi(x,\bar{x})}{\Phi(\underline{x},\bar{x})} < 1-q,$$
$$\frac{\Phi(x,\underline{x})}{\Phi(\bar{x},\underline{x})} < q.$$

Assuming that prospect functions are all concave, the first inequality gives the restriction on the minimum curvature of the prospect function $\Phi(.,\bar{x})$ and the second inequality gives the restriction on the maximum curvature of the prospect function $\Phi(.,\underline{x})$.

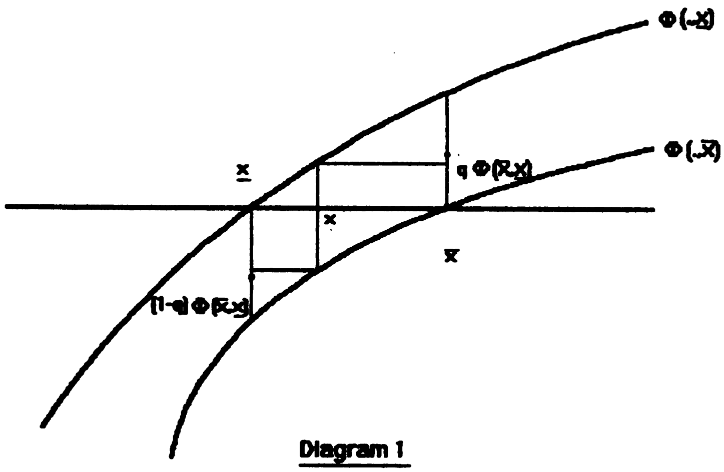

Diagram 1

Both of these inequalities can be checked empirically by asking subjects, since the first inequality is equivalent to $\left((x,\bar{x}),(\bar{x},\underline{x});\frac{1}{2-q},\frac{1-q}{2-q}\right)\in\succ^*$ while the second inequality is equivalent to $\left((\underline{x},x),(\bar{x},\underline{x});\frac{1}{1+q},\frac{q}{1+q}\right)\in\succ^*$.

Finally, I note the following important fact. If the outcome space is $\Lambda(R)$, then we cannot regard the common consequence effect as a violation of the independence axiom. This is also pointed out by [Krantz, Luce, Suppes & Tversky, 90].

5.4 Common Ratio Effect

Let $y > x > 0$ and $p > q$. The lottery1 is specified by $(x, 0; p, 1-p)$ and the lottery 2 is specified by $(y, 0; q, 1-q)$. An experimenter asks subjects to choose between lottery 1 and lottery 2. The values of p and q are adjusted so that a subject chooses lottery 1 over lottery 2. Then the experimenter asks them to choose between lottery 3 and lottery 4, where lottery 3 is specified by $(x, 0; rp, 1-rp)$ and lottery 4 is specified by $(y, 0; rq, 1-rq)$. Experiments shows that, for sufficiently small r, majority of subjects choose lottery 4 over lottery 3. This result is called common ratio effect. The common ratio effect violates the independence axiom because $(x, 0; rp, 1-rp) = r(x, 0; p, 1-p) + (1-r)\delta_0$ and $(y, 0; rq, 1-rq) = r(y, 0; q, 1-p) + (1-r)\delta_0$. Hence, the common ratio effect gives another paradox to the expected utility theory.

The popular explanation of this paradox is that subjects shift their attention, from probabilities to prizes. More concretely, this explanation is described in the following way. When the choice between lottery 1 and lottery 2 is asked, the experimenter fixes p and q carefully so that subjects take probabilities of obtaining positive prizes seriously. When the choice between lottery 3 and the lottery 4 is asked, probabilities of obtaining positive prizes are scaled down, and so is the difference of these probabilities. When the downscaling exceeds some degree, subjects no longer cares the difference of probabilities and takes only the difference of prizes seriously, since the latter is far more easily noticeable than the former.

Intuitively, this popular explanation should be related to the nature of surmising process. Consider the choice between lottery 3 and lottery 4. Suppose that the surmising process selects a choice situation in which the two random variables are significantly positively correlated. In this case, a choice maker believes that pairs (x, y) or $(0, 0)$ happens with very high probability. Hence, the difference of prizes would be relatively less noticeable (especially when x is close to y). Then, a choice maker would be relatively more careful about probabilities of obtaining positive prizes. According to the popular explanation, this should not be the case. Next, consider the choice between lottery 1 and lottery 2. Choosing lottery 1 over lottery 2 is a relatively easy task if the choice maker does not pay much attention to the event in which $(0, y)$ happens. This consideration leads to the following guess; the strength of the tendency that the surmising process selects choice situations showing negative correlation is increasing with respect to probabilities of obtaining the worst prize from either of lotteries. It seems that this is a strong restriction.

Fortunately, I do not need this strong restriction on the surmising process if the configuration of experiment satisfies some regularity conditions. What I require for characteristic of a choice maker is a weak property that the surmising process does not select choice situations showing significant positive correlation, and the increasing risk aversion as introduced in 5.1. In the rest of this subsection, I assume these properties.

To illustrate my explanation formally, I describe choice situations selected in the surmising process by using indexes of negative correlation. Consider the choice between lottery 1 and lottery 2. Let t be the conditional probability of obtaining y from lottery 2 given nothing from lottery 1. The range of t is $\left[0, \frac{q}{1-p}\right]$. It is easy to see that distributions over pairs of money-prizes induced by choice situations in Γ(lottery 1, lottery 2) are uniquely described by t. Similarly, consider the choice between lottery 3 and lottery 4. Let t' be the conditional probability of obtaining y from lottery 4 given nothing from lottery 3. The range of t' is $\left[0, \frac{rq}{1-rp}\right]$. Distributions over pairs of money-prizes induced by choice situations in Γ(lottery 3, lottery 4) are uniquely described by t'.

In my theory, the common ratio effect happens if, for all choice situations in
Γ(lottery 1, lottery 2) and Γ(lottery 3, lottery 4) described with t and t', the following inequalities are satisfied.

$$
\begin{aligned}
t3(1-p)\left[\frac{1}{3}\Phi(y,x)+\frac{1}{3}\Phi(x,0)+\frac{1}{3}\Phi(0,y)\right] &> p\left[\frac{q}{p}\Phi(y,x)+\frac{p-q}{q}\Phi(0,x)\right].\\
t'3(1-rp)\left[\frac{1}{3}\Phi(y,x)+\frac{1}{3}\Phi(x,0)+\frac{1}{3}\Phi(0,y)\right] &< rp\cdot\left[\frac{q}{p}\Phi(y,x)+\frac{p-q}{p}\Phi(0,x)\right].
\end{aligned}
$$

The first inequality corresponds to the choice of lottery 1 over lottery 2 and the second inequality corresponds to the choice of lottery 4 over lottery 3. Suppose that the following conditions are satisfied.

$$\tfrac{q}{p}\Phi(y,x)+\tfrac{p-q}{p}\Phi(0,x)<0 \text{ or } \left((x,y),(x,0);\tfrac{q}{p},\tfrac{p-q}{p}\right)\in\succ^{*}.$$

It is not difficult to see that the increasing risk aversion implies that

$$\frac{1}{3}\Phi(y,x)+\frac{1}{3}\Phi(x,0)+\frac{1}{3}\Phi(0,y)>0,$$

i.e., $\left((x,y),(0,x),(y,0);\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) \in \succ^*$.[22] By manipulating previous two inequalities, we get the following implication for parameters t and t'.[23]

$$\begin{aligned} t &< \frac{p}{3(1-p)}\frac{\frac{q}{p}\Phi(y,x)+\frac{p-q}{p}\Phi(0,x)}{\frac{1}{3}\Phi(y,x)+\frac{1}{3}\Phi(x,0)+\frac{1}{3}\Phi(0,y)}, \\ t' &> \frac{r(1-p)}{1-rp}\frac{p}{3(1-p)}\frac{\frac{q}{p}\Phi(y,x)+\frac{p-q}{p}\Phi(0,x)}{\frac{1}{3}\Phi(y,x)+\frac{1}{3}\Phi(x,0)+\frac{1}{3}\Phi(0,y)}. \end{aligned}$$

The implication of these inequalities is as follows. In the first choice situation, a subject believes small negative correlation relative to the set-up of prizes (x and y) and probabilities (p and q). In the second choice situation, the subject believes relatively large negative correlation with respect to r. The right hand side of the second inequality goes to 0 as r goes to 0. Hence, for very small r, the restriction on $\Gamma(\text{lottery } 3, \text{lottery } 4)$ imposed by the second inequality is not strong. On the other hand, the right hand side of the first inequality exceeds 1 if p is set close to 1. In this case, the first inequality is satisfied trivially. Hence, my theory predicts for the common consequence effect to happen.

More formal analysis goes as follows. The second inequality implies, in particular, that $\frac{q\Phi(y,x)+(p-q)\Phi(0,x)}{\Phi(y,x)+\Phi(x,0)+\Phi(0,y)} < q$. It is easy to see that this is equivalent to

$$p\Phi(x,0) < q\Phi(y,0),$$

i.e., $\left((0,x),(y,0);\frac{q}{p+q},\frac{p}{p+q}\right) \in \succ^*$. Combining this condition and the previous condition, we get

$$q\Phi(y,0) > p\Phi(x,0) > q[\Phi(y,x)+\Phi(x,0)].$$

[22] A comprehensive consideration of the increasing risk aversion is presented in the next subsection. So we omit details here.

[23] Note that the distribution over pairs of money-prizes in the condition corresponds to a choice situation between lottery1 and the lottery 2 with maximal positive correlation.

I summarize this condition into the following assumption.

Assumption 5.1 *I assume*

1. $\left((x,y),(x,0);\frac{q}{p},\frac{p-q}{p},\right)\in\succ^*$, *and*

2. $\left((0,x),(y,0);\frac{q}{p+q},\frac{p}{p+q}\right)\in\succ^*$.

These assumptions determine a regularity required for experiments. They can be easily checked by asking subjects. Next, I define a function F of the probabilities $(\bar{p},\bar{q})$ of winning top prizes from both lotteries by

$$F(\bar{p},\bar{q})\equiv\frac{1}{1-\bar{p}}\frac{\bar{q}\Phi(y,x)+(\bar{p}-\bar{q})\Phi(0,x)}{\Phi(y,x)+\Phi(x,0)+\Phi(0,y)}.$$

Similarly, I define a function G by

$$G(\bar{p},\bar{q})\equiv\frac{\bar{q}}{1-\bar{p}}.$$

Finally, I regard the conditional probabilities of getting y from lottery with the highest prize y given nothing from lottery with the highest prize x determined by Γ as a correspondence of the probabilities $(\bar{p},\bar{q})$ of winning the top prizes in both lotteries. I denote it by $t(\bar{p},\bar{q})$. The restriction on the range of t' is rewritten as; $t(rp,rq)\leq G(rp,rq)$ for all $r\in(0,1]$. The two inequalities derived in the previous paragraph are expressed as; (1) $t(p,q)<F(p,q)$, and (2) $t(rp,rq)>F(rp,rq)$ for all sufficiently small r. It is clear that $f(r)\equiv F(rp,rq)$ is monotone increasing and $\lim_{r\to 0}f(r)=0$.[24] Hence, if the surmising processes of subjects do not select choice situations showing significant positive correlation, my theory predicts that most of subjects choose lottery 4 over lottery 3 if r is set close to 0. This means that the whole trick of the common ratio effect comes from the careful specification of prizes (x and y) and probabilities (p and q) for which subjects choose lottery 1 over lottery 2. Since $F(\bar{p},\bar{q})>1$ when p is close to 1, (1) is trivial in that case. Hence, if the configuration of an experiment satisfies the assumption and p is set close to 1, my theory predicts for the common ratio effect to happen.

[24] Note that our assumption implies that $G(rp,rq)>F(rp,rq)$ for all r.

5.5 Utility Evaluation Effect

Let prizes $\bar{x}$ and $\underline{x}$ be such that $\bar{x} > \underline{x}$. An experimenter selects a prize x for a subject so that she cannot choose between lottery $(\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2})$ and δ_x. Next, the experimenter selects a prize y for the same subject so that she cannot choose between $\left(x, \underline{x}; \frac{1}{2}, \frac{1}{2}\right)$ and δ_y. Finally, the experimenter selects a prize y' for the same subject so that she cannot choose between $\left(\bar{x}, \underline{x}; \frac{1}{4}, \frac{3}{4}\right)$ and $\delta_{y'}$. In expected utility theory, it must be the case that $y = y'$. But, empirical literatures suggest that many subjects reveal $y' > y$.

Similarly, let $\bar{x}$, $\underline{x}$ and x be as specified in the previous paragraph. The experimenter selects a prize z for the same subject so that she cannot choose between $\left(\bar{x}, x; \frac{1}{2}, \frac{1}{2}\right)$ and δ_z. Finally, the experimenter selects a prize z' for the same subject so that she cannot choose between $\left(\bar{x}, \underline{x}; \frac{3}{4}, \frac{1}{4}\right)$ and $\delta_{z'}$. Again, the expected utility theory predict that $z = z'$. But, empirical literatures suggest that many subjects reveal $z > z'$.

These phenomena are called utility evaluation effect. The name comes from de Neufville who interprets the phenomena as follows. Suppose the expected utility theory holds. Normalize the von-Neumann-Morgenstern utility function u so that $u(\bar{x}) = 1$ and $u(\underline{x}) = 0$. By plotting all certainty equivalents as in the diagram, de Neufville argues that the higher probability on the best prize $\bar{x}$ inflates the value of the von-Neumann-Morgenstern utility function.

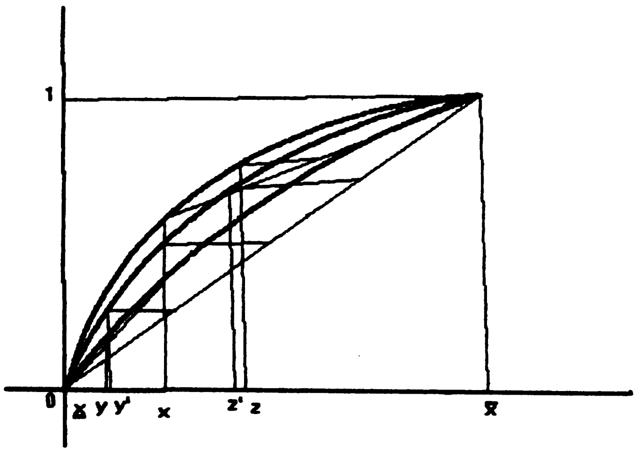

Diagram 2

It has been a convention that the utility evaluation effect is understood as a violation of the independence axiom. On the other hand, the involvement of certainty equivalents in the experiment suggests that this effect can be seen as a violation of transitivity (cf. 5.6). My explanation for the utility evaluation effect regards it as a violation of the transitivity, not as a violation of the independence axiom.

Before presenting the explanation in my theory, I should mention the traditional explanation. It relies on the shift of attention by subjects, from probabilities to prizes. According to this explanation, subjects pay their attention to probabilities when the experimenter tries to get certainty equivalents while they compare risky lotteries with more attention on prizes. To illustrate this, consider the experimental stage in which the experimenter tries to determine the prize x by asking subjects. At this stage, subjects focus on the consideration of distribution over money-prizes. This makes risk-averse subjects select low x. Next, consider the difference between $\left(x, \underline{x}; \frac{1}{2}, \frac{1}{2}\right)$ and $\left(\bar{x}, \underline{x}; \frac{1}{4}, \frac{3}{4}\right)$. The second lottery has a chance to win the highest prize $\bar{x}$ and the first lottery does not. This makes subjects evaluate the second lottery higher than the first one.

There is a way to justify the traditional explanation in my theory. Note that, in the experiment, each subject knows that x is the certainty equivalent of lottery $\left(\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2}\right)$. Hence, it is not impossible for a careful subject to recognize lottery $\left(\bar{x}, \underline{x}; \frac{1}{4}, \frac{3}{4}\right)$ as a lottery on $\Lambda(R)$. $\left((\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2}), \delta_{\underline{x}}; \frac{1}{2}, \frac{1}{2}\right)$.[25] Similarly, it is not impossible for a subject to recognize the lottery $\left(\bar{x}, \underline{x}; \frac{3}{4}, \frac{1}{4}\right)$ as a lottery on $\Lambda(R)$, $\left(\delta_{\bar{x}}, \left(\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2}\right); \frac{1}{2}, \frac{1}{2}\right)$. Consider a choice situation between two lotteries on $\Lambda(R)$, $\left((\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2}), \delta_{\underline{x}}; \frac{1}{2}, \frac{1}{2}\right)$ and $\left(\delta_x, \delta_{\underline{x}}; \frac{1}{2}, \frac{1}{2}\right)$. Assuming that preference reversal (see 5.6) does not happen in the experiment, the utility evaluation effect suggests that the subject chooses the former lottery than the latter one. Similarly, consider a choice situation between two lotteries on $\Lambda(R)$, $\left((\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2}), \delta_{\bar{x}}; \frac{1}{2}, \frac{1}{2}\right)$ and $\left(\delta_x, \delta_{\bar{x}}; \frac{1}{2}, \frac{1}{2}\right)$. The utility evaluation effect suggests that the subject chooses the latter lottery than the former one. We can clearly see that these suggestions correspond to the change of attitude toward risk. The change of attitude toward risk is explained by the shift of attention in the following way. Suppose ξ is a fair bet with respect to a money-prize x. Suppose that $\eta \in \Lambda(R)$ is concentrated on bad prizes relative to the support of ξ. Consider the situation in which a choice maker compares δ_x and ξ indirectly as prospects from η. Since η is a very bad bet, it is plausible that a choice maker pays attention to a chance to win good prizes. In other words, she pays her attention to prizes. Hence she tends to be risk-taking. Next, suppose that η is concentrated on good prizes relative to the support of ξ. In this case, good prizes from ξ is not so attractive because a choice maker obtains good prizes from the alternative bet, η. Hence, it is plausible that a choice maker is more careful about probabilities of obtaining good prizes than prizes themselves. This makes a choice maker risk-averse.

Though this is clearly one way to interpret the utility evaluation effect, I find it unsatisfactory. At first, this explanation assumes that preference reversal does not happen in experiments. As we will see in 5.6, this may not be the case. For some configuration of lotteries, choice between two risky lotteries and choice between their certainty equivalents may not be consistent. In 5.6, I show that, if a configuration of the experiment satisfies a regularity condition, then the preference reversal is likely to happen. It is not difficult to see that experiments considered in this subsections

[25] I admit that this is a very difficult argument to justify. Usually, identifying a compound structure from a reduced lottery is empirically impossible. But, in the course of experiment, the lottery $(\bar{x}, \underline{x}; \frac{1}{4}, \frac{3}{4})$ appears in the last stage. With the knowledge that x is the certainty equivalent of the lottery $(\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2})$, a subject may realize, in the second stage, that $(\bar{x}, \underline{x}; \frac{1}{4}, \frac{3}{4})$ can be obtained by substituting $(\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2})$ into the place of x in lottery $(x, \underline{x}; \frac{1}{2}, \frac{1}{2})$ and reducing it. Knowing this relationship, it is not impossible for the subject to identify $(\bar{x}, \underline{x}; \frac{1}{4}, \frac{3}{4})$ as $\left((\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2}), \delta_{\underline{x}}; \frac{1}{2}, \frac{1}{2}\right)$.

are very likely to satisfy the regularity condition. Secondly, subjects may not be able to realize a compound structure in a reduced lottery. It is possible to reduce a lottery on $\Lambda(R)$ to a lottery on R, but the operation to the other direction is impossible without any additional information. Though this information is given by x being the certainty equivalent of $\left(\bar{x}, \underline{x}; \frac{1}{2}, \frac{1}{2}\right)$ in the experiments, the use of such information for modeling requires a sophistication for subjects. I also warn that model selection is beyond the scope of my theory. My theory does not explain the reason why such a sophistication leads to recognize lottries on money-prizes as lotteries on $\Lambda(R)$.

Hence, I look for the explanation that does not assume either lotteries on $\Lambda(R)$ or the non-existence of preference reversal. My explanation relies on the increasing risk aversion. More concretely, I analyze restrictions on the curvature of three prospect functions, $\Phi(.,\underline{x})$, $\Phi(.,x)$ and $\Phi(.,\bar{x})$, that the increasing risk aversion and the skew-symmetricity impose.[26] I illustrate it in the following diagram.

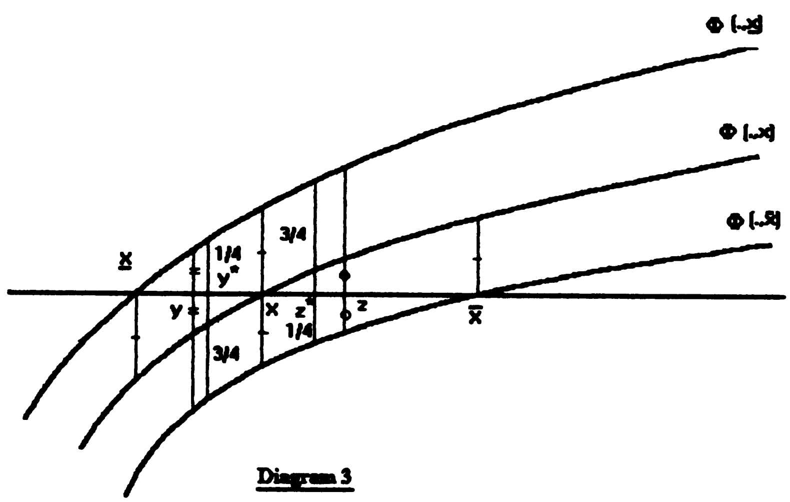

Diagram 3

In the diagram, we have graphs of three prospect functions. These graphs are drawn as follows. At first, draw a graph of the prospect function $\Phi(.,\underline{x})$ freely. By skew-symmetricity, the graph of

[26]It is also closely related to the explanation of de Neufville since the inflation of prospect index (after some normalization) is the issue.

the prospect function $\Phi(.,\bar{x})$ must go through $(\underline{x}, -\Phi(\bar{x},\underline{x}))$ and $(\bar{x},0)$. By moving the graph of $\Phi(.,\underline{x})$ downward so that $(\underline{x},0)$ is identified with $(\underline{x}, -\Phi(\bar{x},\underline{x}))$, we know that the moved graph of $\Phi(.,\underline{x})$ must go through the same points as the grapgh of $\Phi(.,\bar{x})$. Since both $\Phi(.,\underline{x})$ and $\Phi(.,\bar{x})$ are concave, the graph of $\Phi(.,\bar{x})$ on the domain $(\underline{x},\bar{x})$ is either entirely above or below the moved graph of $\Phi(.,\underline{x})$, if they are not equivalent. By the increasing risk aversion, it cannot be below the moved graph of $\Phi(.,\underline{x})$. The reason is as follows. Suppose that is the case. Let c to be a prize in $(\underline{x},\bar{x})$ such that the slope of the moved graph of $\Phi(.,\underline{x})$ is the same as that of the graph of $\Phi(.,\bar{x})$.[27] Then the coefficient of absolute risk aversion for $\Phi(.,\underline{x})$ evaluated at c must be higher than that for $\Phi(.,\bar{x})$ evaluated at c. This contradicts increasing risk aversion. Hence,

The graph of $\Phi(.,\bar{x})$ must be entirely above the moved graph of $\Phi(.,\underline{x})$ on $(\underline{x},\bar{x})$.

Now, draw the graph of $\Phi(.,\bar{x})$ so that it is entirely above the moved graph of $\Phi(.,\underline{x})$ on $(\underline{x},\bar{x})$. The certainty equivalent x is determined as the unique prize x such that $\Phi(x,\underline{x}) = -\Phi(x,\bar{x})$. A consideration analogous to the one I presented gives the following information about the graph of $\Phi(.,x)$.

1. On $(x,\bar{x})$, the graph of $\Phi(.,x)$ must be entirely below the moved graph of $\Phi(.,\bar{x})$ which goes through $(x,0)$ and $(\bar{x}, -\Phi(x,\bar{x}))$.

2. On $(\underline{x},x)$, the graph of $\Phi(.,x)$ must be entirely above the moved graph of $\Phi(.,\underline{x})$ which goes through $(\underline{x}, -\Phi(x,\underline{x}))$ and $(x,0)$.

Finally, draw the graph of $\Phi(.,x)$ so that these conditions are satisfied.

The certainty equivalent y is plotted in the diagram so that $\Phi(y,\underline{x}) = -\Phi(y,x)$. From 2 of the information for $\Phi(.,x)$, it is clear that $\Phi(y,\underline{x}) - \Phi(y,x) = 2\Phi(y,\underline{x}) < \Phi(x,\underline{x})$. By an argument similar to the one given in the previous paragraph, we can show that $\Phi(y,x) - \Phi(y,\bar{x}) > \Phi(x,\underline{x})$. Hence, $-\Phi(y,\bar{x}) > 3\Phi(y,\underline{x})$. In fact, for any $s \leq y$, $\Phi(x,\underline{x}) \geq \Phi(s,\underline{x}) - \Phi(s,x) \geq 2\Phi(s,\underline{x})$ and $\Phi(s,\underline{x}) - \Phi(s,\bar{x}) > 2\Phi(x,\underline{x})$. Hence, $-\Phi(s,\bar{x}) > 3\Phi(s,\underline{x})$ for all $s \in [\underline{x}, y]$. This implies that $y' > y$.

A similar argument can be applied to the second experiment. The certainty equivalent z is plotted in the diagram so that $\Phi(z,\underline{x}) = -\Phi(z,\bar{x})$. For any $s \geq z$, 1 of the information for $\Phi(.,x)$

[27] Such c must exists by the intermediate value theorem when both concave functions are twice differentiable.

says that $\Phi(\bar{x}, x) \geq \Phi(s, x) - \Phi(s, \bar{x}) \geq 2\Phi(\bar{x}, s)$. Also, an argument using the increasing risk aversion shows that $\Phi(s, \underline{x}) - \Phi(s, x) > \Phi(\bar{x}, x)$. This implies that $\Phi(s, \underline{x}) > 3\Phi(\bar{x}, s)$ for all $s \geq z$. Hence, $z' < z$.

I note that the utility evaluation effect is natural phenomenon in my theory. It does not require any regularity of experiments. If a subject can be simulated by my theory and also exhibits the increasing risk aversion, then she must exhibit the utility evaluation effect.

5.6 Preference Reversal Phenomena

An experimenter selects prizes X, x, Y, y so that $X > x$, $Y > y$ and $y < X < Y$. Also he takes probability weights p and q so that $p > q$. Using these prizes and probabilities, the experimenter constructs two risky lotteries. One of them is $(X, x; p, 1-p)$. I call it *probability-bet*. The other is $(Y, y; q, 1-q)$. I call it *prize-bet*. Prizes (X, x, Y and y) and probabilities (p and q) are carefully configured so that subjects become aware of that the probability to win the higher prize is the main attraction of the probability-bet while the relatively high best prize is the main attraction of the prize-bet. The experimenter asks subjects to choose between the probability-bet and the prize-bet. (To make choice situation serious, x and y are often set to small negative values.) Next the experimenter asks them to "value" these lotteries. Typically the experimenter gives subjects each lottery and asks them a price at which they want to sell it. By using a simple version of auction scheme, the experimenter can make subjects reveal their certainty equivalents. It is reported in several literatures that majority of subjects choose the probability-bet when they are offered a choice between the probability-bet and the prize-bet, while they value the prize-bet higher than the probability-bet. There are many different variations of this experiment. Surprisingly, results from these experiments suggest that this phenomenon is robust.

It is clear that the expected utility theory does not allow this to happen. More generally, any *transitive* choice relation on lotteries does not allow this to happen. Hence, I regard this phenomenon as a violation of transitivity.

The explanation by researchers who performed this experiment is as follows. Subjects pay their attention to probabilities in the choice situation between probability-bet and prize-bet, while they pay their attention to prizes when they are asked to value lotteries.

In my theory, this explanation must be related to the nature of surmising process. In the choice situation between probability-bet and prize-bet, the surmising process may take correlation of lotteries seriously. On the other hand, the surmising process has no role to determine certainty equivalents. Unfortunately, this direction of approach is not successful.

The analysis that I give in this subsection relies on the increasing risk aversion. The idea is that the use of certainty equivalents is directly related to properties of prospect functions.

For simplicity, I assume that $x = y$, and denote it by z. I introduce an index r. It is the conditional probability of getting Y from the prize-bet given the prize z from the probability-bet. The range of r is $\left[0, \min(\frac{q}{1-p}, 1)\right]$. It is clear that any distribution over pairs of prizes consistent with the descriptions of probability-bet and prize-bet is determined uniquely by r. Consider the choice between the probability-bet and the prize-bet. A subject chooses probability-bet over prize-bet if, for any choice situation in Γ(probability-bet, prize-bet) whose distribution over prizes is determined by r,

$$r(1-p)[\Phi(Y,X) + \Phi(X,z) + \Phi(z,Y)] > q\Phi(Y,X) + (p-q)\Phi(z,X).$$

To see the implication of this inequality, I draw the graph of three prospect functions $\Phi(.,z)$, $\Phi(.,X)$ and $\Phi(.,Y)$. The procedure is exactly the same as that in the explanation of the utility evaluation effect.

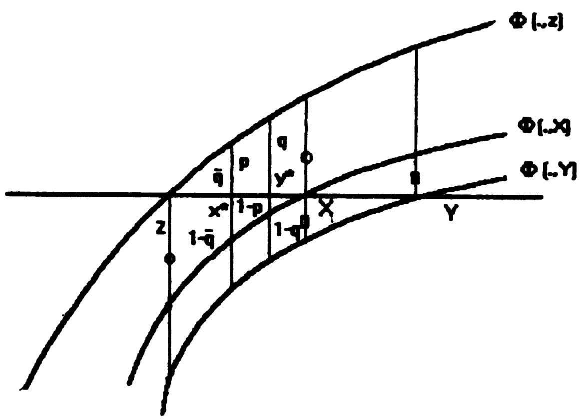

Diagram 4

By using the argument I presented in the analysis of the utility evaluation effect, I can easily show that the increasing risk aversion implies that $\Phi(Y,X)+\Phi(X,z)+\Phi(z,Y)<0$. Hence, if a subject chooses the probability-bet over the prize-bet, it must be the case that $q\Phi(Y,X)+(p-q)\Phi(z,X)<0$. Then I can rewrite the inequality as follows.

$$r<\frac{q}{1-p}\frac{\Phi(Y,X)+\left(\frac{p}{q}-1\right)\Phi(z,X)}{\Phi(Y,X)+\Phi(X,z)+\Phi(Y,z)}.$$

Note that the right-hand side of this inequality exceeds 1 if p is close to 1. Hence, this ineqality does not impose any restriction on the surmising process if p is set close to 1. Also note that, if $p\Phi(X,z)\geq q\Phi(Y,z)$, then this inequality is trivially satisfied because $r\leq\frac{q}{1-p}$. Hence, if this is the case, the inequality does not impose any restriction on the surmising process. These are two different regularity conditions on experiments that I adopt.

The certainty equivalent x^* of the probability-bet is determined by $p\Phi(x^*,X)+(1-p)\Phi(x^*,z)=0$. Let $\bar{q}$ to be such that $\bar{q}\Phi(x^*,Y)+(1-\bar{q})\Phi(x^*,z)=0$. From the diagram, it is clear that the certainty equivalent of prize-bet, y^*, is higher than x^* if and only if

$q > \bar{q}$. Since $p > \bar{q}$, it is possible that the experimenter can take such q, especially when p is set close to 1.

I summarize the analysis as follows. I introduce the following assumption.[28]

Assumption 5.2 $\left((X,z),(z,Y);\frac{p}{p+q},\frac{q}{p+q}\right) \in \succ^*$.

Proposition 5.2 *Assume that assumption 5.2 holds. Also assume that $\succ$ satisfies all axioms in sections 3, and that it exhibits increasing risk aversion. Then the preference reversal happens if and only if $q > \bar{q}$, where $\bar{q}$ is determined by $((x^*,Y),(x^*,z);\bar{q},1-\bar{q}) \in \succ^*$.*

Finally, I want to add a comment about non-transitivity of choice. There is no doubt that preference reversal has been received coldly by researchers in economics, since it violates transitivity. But, the original explanation of this paradox is at least very natural. It claims that the reason why a subject chooses the probability-bet over the prize-bet is completely different from the reason why a subject values the prize-bet higher than the probability-bet. There is no reason to justify transitivity of choice when a subject reveals this type of non-monotonic reasoning. For this non-monotonicity to disappear, a subject must be able to characterize all alternatives involved in choice by an extensive list of characteristics and memorize it. Even a computer has a trouble to do so because of memory restriction. In the real world, non-transitive choice prevails. (To apply the transitivity, we need to restrict the domain very carefully.) My theory is consistent with non-transitivity. A drawback of my theory is that it hardly gives any insight about the mechanism to create non-transitivity. To analyze such a mechanism, we clearly need a deep understanding of non-monotonic nature in information processing, or, more fundamentally, in logic. My theory does not have any contribution in this respect.

5.7 Kahneman-Tversky's Prospect Function

Kahneman and Tversky proposed prospect theory in [Kahneman & Tversky, 79]. The idea of this theory is that choice depends on a base outcome and prospects of alternatives measured from it.

[28] Remember that, in the analysis of the common consequence effect, I assumed (basically) that $\left((z,X),(Y,z);\frac{q}{p+q},\frac{p}{p+q}\right) \in \succ^*$. These assumptions say that configuration of experiments does matter.

They argued that the "prospect function" from a base outcome, say z, is concave on $[z, \infty)$ and convex on $(-\infty, z]$. In other words, they assumed that a subject is risk-averse on "gains" and risk-taking on "losses".[29] In the diagram, a typical graph of their "prospect function" is drawn.

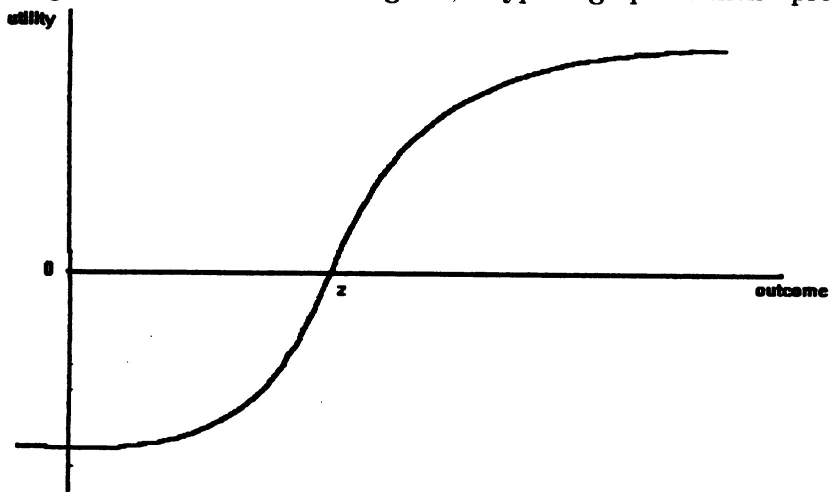

Diagram 5

It seems to me that Kahneman and Tversky regard a base outcome as either current wealth or promised income. In modeling of a choice maker, [Kahneman & Tversky, 79] take such formation of a base outcome as granted. I argue that the use of such fixed base outcome is misleading because of the following reason. Consider the case in which a base outcome is specified by the current wealth of a subject. In this case, the formation of the base outcome is exogenous to experiments. It is independent of any particular choice that an experimenter asks. Then, is it plausible that subjects take their wealth level into account when they are asked a choice? If that is the case, subjects are paying attention to information that is not directly related to the description of alternatives given by the experimenter. I believe that it is not likely. Since objective of a subject in a particular choice situation is to decide an alternative that she chooses, information about alternatives should have a priority over any other information. Assuming that resources available for a subject is bounded and

[29] In this subsection, I use double-quotation mark when I refer to a prospect function defined by Kahneman and Tversky. Without double-quotation mark, prospect function means the one defined in my theory.

that processing information about alternatives is not a trivial task, it is likely that a subject never realizes information that is not related to alternatives. In other words, I do not think that subjects in experiments are aware of their wealth levels. Next, suppose that there is a mechanism, as a part of subject's characteristics, that forms a base outcome depending on a particular experiment. I describe an experiment by a sequence of finitely many choice situations between lotteries. In this case, the formation of a base outcome is endogenous to experiments. In the course of a particular experiment, this mechanism creates a sequence of base outcomes. However, such base outcomes would not be observable in experiments, since it is unlikely that subjects are aware of them clearly. This complicates the task to estimate the mechanism because the experimenter would not get outputs of the mechanism in data. This means that it is possible but very hard to estimate the mechanism from the data collected in experiments.

My theory gives a way to explain how properties on "prospect functions" is derived. The argument relies on misspecification that "shifting" base outcomes are identified as a "fixed" base outcome. See the following diagram.

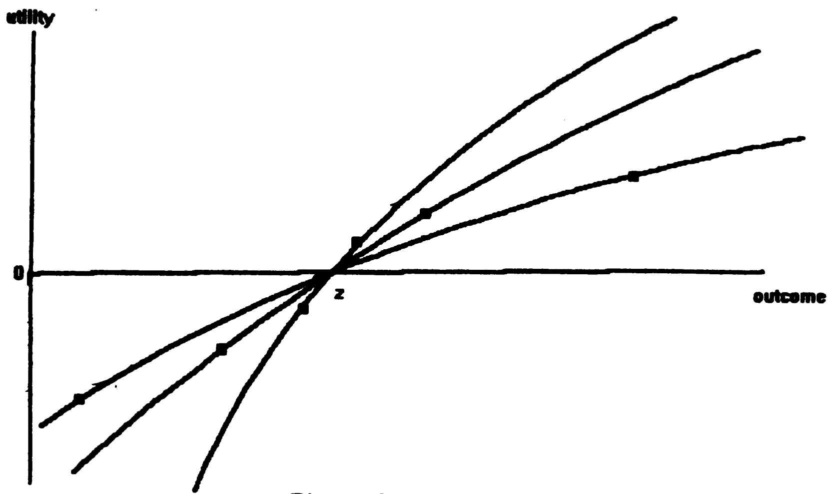

Diagram 6

I get this diagram by gluing up graphs of prospect functions in my theory at zero-prospect level.

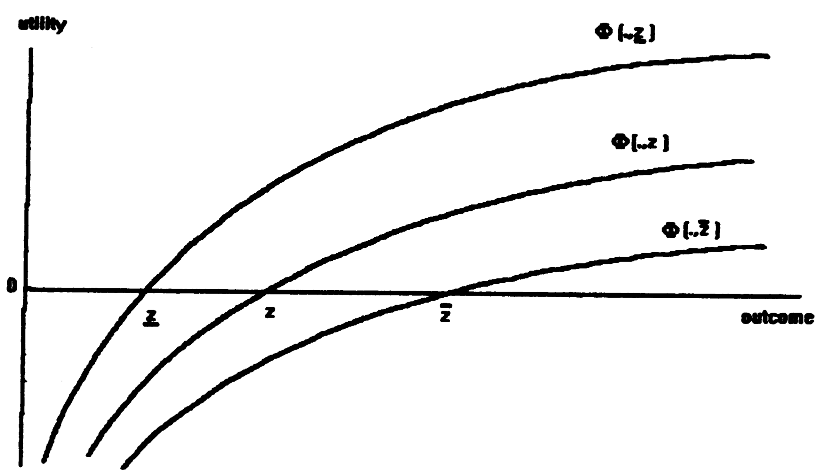

Diagram 7

I interpret the procedure in the following way. Suppose that the experimenter believes that z is the fixed base outcome from which prospects of all outcomes are measured. Suppose also that he is trying to determine the value of "prospect function" on $z \pm x$ ($x > 0$). For simplicity, I assume that each subject is "risk neutral", i.e., prospect functions with different base outcomes are constructed by vertical parallel shift of a concave function. The procedure to find a value of "prospect function" is as follows. Let $\bar{z}$ and $\underline{z}$ to be two prizes such that $[\underline{z}, \bar{z}]$ contains all $z \pm x$ that the experimenter concerns. At first, the experimenter normalizes the values of "prospect function" on $\bar{z}$ and $\underline{z}$. Assume that they are 1 and -1. Next, the experimenter picks a particualr value x and show prizes, $z \pm x$, to a subject. Finally, the experimenter constructs lotteries with support $\{\bar{z}, \underline{z}\}$ whose certainty equivalents for the subject are $z + x$ and $z - x$. I represent these lotteries by $(\bar{z}, \underline{z}; p_{+}(x), 1 - p_{+}(x))$ and $(\bar{z}, \underline{z}; p_{-}(x), 1 - p_{-}(x))$. According to the prospect theory (without a probability modification function), the value of the "prospect function" on $z + x$ ($z - x$) is determined by $1.p_{+}(x) + (-1).(1 - p_{+}(x))$ $(1.p_{-}(x) + (-1).(1 - p_{-}(x)))$. This experiment is repeated by increasing x gradually.

Now, I assume that the experimenter realizes, in the course of experiment, that the valuation of $\bar{z}$ and $\underline{z}$ by the subject is decreasing. Since the subject observes $z + x$ with increasing x, this may happen. Then the normalization should be adjusted adequately. In other words, the experimenter would use $1 - \epsilon$ and $-1 - \epsilon$, instead of 1 and -1, to compute values of "prospect function". If the experimenter plot the values computed with successive adjustment of normalization, then he would get the graph similar to the first diagram in this subsection. My theory explains this by a shift of base outcome to higher prizes. (See the diagram 7.) If this happens in the course of experiment, then I can describe the procedure of experiment by the diagram 6. The successive adjustment of normalization corresponds to the misidentification of shifting base outcomes as the fixed base outcome z.

I note that my theory does not oppose the use of "prospect" to measure desirability of outcomes. What I claim is that such prospect should not be measured based on a fixed base outcome. In the case of binary choice situation, a choice maker knows that, in each state, there are two possible outcomes, one from the first alternative and the other from the second alternative. Hence, it is natural to assume that the prospect of the outcome from the first alternative is measured based on the outcome from the second alternative, and vice versa. This is the essence of SSA theory developed by Fishburn, that is incorporated in my theory.

Finally, I point out that, when the outcome space is $\Lambda(R)$, change of attitude toward risk introduced in 5.1 is similar to properties on "prospect function" imposed by Kahneman and Tversky. The difference is explained as follows. In my theory, change of attitude toward risk is a concept with shifting base outcomes. In the prospect theory, properties on "prospect functions" are based on the presumption of a fixed base outcome.

6 Extension for General Acts

Though analysis of paradoxes in expected utility theory is covered by the theorem 4.1, it is also important to generalize the result to all choice situations in order to be a true generalization of expected utility theory. The argument relies on the approximation of a choice situation by those with finite supports, i.e., choice situations in B^F.[30]

I introduce two additional axioms. The first of these additional axioms, called the *topological continuity axiom*, guarantees that any consequential class of choice situations can be approximated by those with finite supports.

Axiom 6.1 (Topological Continuity) $\{Q \,|\, Q = P \circ (f,g)^{-1} \in \succ$ *for some* $(f,g;P) \in B.\} \subset \Lambda(Z \times Z, \Xi \otimes \Xi)$ *is open with respect to the topology of weak convergence.*

The next axiom, called the *Pareto principle axiom*, is a simple localization principle. It says that, if a choice situation $(f,g;P)$ has a property that a choice maker chooses f over g no matter how she restrict her attention on finite set of pairs of outcomes, then she should choose f over g when she takes into account the entire support of the choice situation.

Axiom 6.2 (Pareto Principle) *If* $(f,g;P) \in B$ *is such that* $Supp(P \circ (f,g)^{-1}) \subseteq \left\{(x,y) \in Z \times Z | \delta_{(x,y)} \in \succ^*\right\}$, *then* $(f,g;P) \in \succ$.

As I did in the section 4, if $\succ$ satisfies the outcome-orientation axiom, I can rewrite these axioms to equivalent statements in $\Lambda(Z \times Z, \Xi \otimes \Xi)$.

Axiom 6.3 (Topological Continuity) $\succ^* \subset \Lambda(Z \times Z, \Xi \otimes \Xi)$ *is open with respect to the topology of weak convergence.*

Axiom 6.4 (Pareto Principle) *If* $\left\{\delta_{(x,y)} | (x,y) \in Supp(Q)\right\} \subset \succ^*$, *then* $Q \in \succ^*$.

At first I prove a lemma that guarantees the boundedness of Φ. I introduce an additional notation. I denote $\{Q \in \Lambda(Z \times Z, \Xi \otimes \Xi) | Supp(Q)$ is countable.$\}$ by $\Lambda^C(Z \times Z, \Xi \otimes \Xi)$.

[30]Compared with the straightforward intuition in our axioms in the section 3, the argument for the extension is not sharp due to introduction of several structural assumptions. Intuitively, the extension is not essential in the sense that it is very difficult for the experimenter to make subjects realize a probability measure with continuum support. The argument adopted here is almost equivalent to those of Fishburn in [Fishburn, 70] and [Fishburn, 89]. I describe the extension just for completeness.

Lemma 6.1 (Boundedness) *Assume that axiom 3.1, 3.2, 3.3, 3.4. 3.5, 6.1 and 6.2 hold on* $B^C \equiv \left\{(f,g;P) \in B | P \circ (f,g)^{-1} \in \Lambda^C(Z \times Z, \Xi \otimes \Xi)\right\}$. *Then the representation function* Φ *given in the theorem 4.1 is bounded.*

Proof:

The proof is almost completely the same as the one given in [Fishburn, 70]. I reproduce it here just for completeness.

Suppose not. Then I can take a sequence of outcome pairs $\{(x_n, y_n)\}$ such that $\Phi(x_n, y_n) \geq 2^n$. Let $Q \in \Lambda^C(Z \times Z, \Xi \otimes \Xi)$ to be such that $Q((x_n, y_n)) > \frac{1}{2^n}$ for all n. By the Pareto principle axiom, $Q \in \succ^*$. Take any $\delta_{(x,y)} \in \succ^*$. Then, by the Archimedian axiom, there exists an $\alpha \in (0,1)$ such that $\alpha\delta_{(x,y)} + (1-\alpha)Q^- \in \succ^*$. I define Q_N by $Q_N \equiv \frac{2^N}{2^N-1} \sum_{n=1}^{N} \frac{1}{2^n}\delta_{(x_n,y_n)}$. Then $\left\{\alpha\delta_{(x,y)} + (1-\alpha)Q_N^-\right\}_{N=1}^{\infty}$ weakly converges to $\alpha\delta_{(x,y)} + (1-\alpha)Q^-$. By the topological continuity axiom, $\alpha\delta_{(x,y)} + (1-\alpha)Q_N^- \in \succ^*$ for all sufficiently large N. But, for sufficiently large N, $[\alpha\delta_{(x,y)} + (1-\alpha)Q_N^-]\Phi = \alpha\Phi(x,y) + (1-\alpha)\frac{2^N}{2^N-1}\sum_{n=1}^{N}\frac{1}{2^n}\Phi(y_n, x_n) \leq \alpha\Phi(x,y) - (1-\alpha)\frac{2^N}{2^N-1}N < 0$. This contradicts the theorem 4.1.

♠

Next I prove the continuity of the representation function Φ. To define the continuity, I need to assume that Z is a topological space and Ξ is a Borel σ-field. I further assume that Z is a metric space.

Lemma 6.2 (Continuity) *Assume that axiom 3.1, 3.2, 3.3, 3.4, 3.5, 6.1 and 6.2 holds. Assume also that* Z *is a metric space and* Ξ *is the Borel* σ*-field. Then the representation function* Φ *in the theorem 4.1 is continuous.*

Proof:

In the proof of the theorem 4.1, I constructed a candidate for the representation function explicitly. By the uniqueness up to the scale factor, it is sufficient to prove the continuity for the one I constructed in the proof of the theorem 4.1. Fix (x_0, y_0) such that $\delta_{(x_0,y_0)} \in \succ^*$ and define $\theta : (Z \times Z, \Xi \otimes \Xi) \to (0,2)$ by the following rule. If (x,y) is such that $\delta_{(x,y)} \in \succ^*$, then I define $\theta(x,y)$ by

$$\theta(x,y)\delta_{(x,y)} + (1-\theta(x,y))\delta_{(y_0,x_0)} \in \sim^* .$$

If (x,y) is such that $\delta_{(y,x)} \in \succ^*$, then I define $\theta(x,y)$ by $2-\theta(y,x)$. If (x,y) is such that $\delta_{(x,y)} \in \sim^*$, then I define $\theta(x,y)$ by 1. I remind that I constructed Φ by the following formula.

$$\Phi(x,y) \equiv \begin{cases} \frac{1-\theta(x,y)}{\theta(x,y)}, & \text{if } \delta_{x,y} \in \succ^*, \\ 0, & \text{if } \delta_{x,y} \in \sim^*, \\ \frac{\theta(x,y)-1}{2-\theta(x,y)}, & \text{if } \delta_{y,x} \in \succ^*. \end{cases}$$

Hence it is sufficient to show that $\theta(.)$ is continuous.

Suppose that $\theta(.)$ is not continuous. Then there is a sequence $(x_n, y_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty}(x_n, y_n) = (x,y)$ and $\lim_{n\to\infty}\theta(x_n,y_n) \neq \theta(x,y)$. By the topological continuity axiom, I can assume, without loss of generality, that either (1) $\delta_{(x,y)} \in \succ^*$ and $\delta_{(x_n,y_n)} \in \succ^*$ for all n, or (2) $\delta_{(x,y)} \in \sim^*$ and $\delta_{(x_n,y_n)} \in \succ^*$ for all n. In the case (1), $Q_n \equiv \theta(x_n,y_n)\delta_{(x_n,y_n)} + (1-\theta(x_n,y_n))\delta_{(y_0,x_0)} \in \sim^*$ for all n and $\lim_{n\to\infty} Q_n = Q \equiv [\lim_{n\to\infty}\theta(x_n,y_n)]\,\delta_{(x,y)} + (1-[\lim_{n\to\infty}\theta(x_n,y_n)])\,\delta_{(y_0,x_0)}$ where the limit is taken with respect to the weak convergence. By the topological continuity axiom, $\sim^*$ is closed with respect to the topology of weak convergence. Hence $Q \in \sim^*$. The replication corollary suggests that $\theta \in (0,1)$ such that $\theta\delta_{(x,y)} + (1-\theta)\delta_{(y_0,x_0)} \in \sim^*$ is uniquely determined. Hence $\lim_{n\to\infty}\theta(x_n,y_n) = \theta(x,y)$. This is a contradiction. In the case (2), $\theta(x,y) = 1$. Since $\theta(x_n,y_n) < 1$ for all n, it must be the case that $\lim_{n\to\infty}\theta(x_n,y_n) < 1$. Since $Q \in \sim^*$ by the topological continuity axiom, this implies that $\delta_{(x,y)} \in \succ^*$. This contradicts the assumption that $\delta_{(x,y)} \in \sim^*$.

♠

Finally I prove the representation theorem on the set of all choice situations.

Theorem 6.1 (Skew-Symmetric Additive Representation: General Case) *Assume that the outcome space Z is a metric space and Ξ is the Borel σ-field. Then axiom 3.1, 3.2, 3.3, 3.4, 3.5, 6.1 and 6.2 hold if and only if there is a bounded continuous function $\Phi : Z \times Z \to R$ such that:*

1. $\Phi(x, y) + \Phi(y, x) = 0$ *for all* $x, y \in Z$;

2. $(f, g; P) \in \succ$ *if and only if* $P\Phi(f, g) > 0$.

If Φ and Φ' are both such representation functions, then there is an $\alpha > 0$ such that $\Phi' = \alpha\Phi$.

Proof :

By the theorem 4.1, the lemma 6.1 and 6.2, the only task left is to prove that the representation property for $(f, g; P)$ such that $Supp(P \circ (f, g)^{-1})$ is infinite. To do this, it is sufficient to show that $Q \in \Lambda(Z \times Z, \Xi \otimes \Xi)$ such that $Supp(Q)$ is infinite belongs to $\succ^*$ if and only if $Q\Phi > 0$.

Suppose that $Supp(Q)$ is infinite and $Q \in \succ^*$. By the topological continuity axiom and the fact that $\Lambda^F(Z \times Z, \Xi \otimes \Xi)$ is dense in $\Lambda(Z \times Z, \Xi \otimes \Xi)$ with respect to the topology of weak convergence, I can take a sequence $\{Q_n\}_{n=1}^{\infty}$ in $\Lambda^F(Z \times Z, \Xi \otimes \Xi) \cap \succ^*$ such that $\lim_{n\to\infty} Q_n = Q$. Since $Q_n\Phi > 0$ and Φ is bounded continuous, $Q\Phi \geq 0$ by the definition of weak convergence. To prove that $Q\Phi > 0$, Take any $Q' \in \succ^*$ such that $Q'\Phi < Q\Phi$. Such Q' exists by the topological continuity axiom. By applying our argument to Q', we know that $Q'\Phi \geq 0$. Hence $Q\Phi > 0$.

Conversely, suppose that $Q\Phi > 0$ and $Supp(Q)$ is infinite. By the topological continuity axiom and the fact that Φ is bounded continuous, we can take a neighborhood of Q, say V_Q, such that $Q'\Phi > 0$ for all $Q' \in V_Q \cap \Lambda^F(Z \times Z, \Xi \otimes \Xi)$. If $Q \notin \succ^*$, then there must exist a $Q' \in V_Q \cap \Lambda^F(Z \times Z, \Xi \otimes \Xi)$ such that $Q'\Phi \leq 0$. This is a contradiction. Hence $Q \in \succ^*$.

This completes the proof.

♠

It is very annoying that the result depends heavily on structural assumptions, axioms 6.1 and 6.2. Though it is interesting to find a better approach, this generalization has nothing to do with the previous analysis of paradoxes. So, it will not be very productive for me to concern about the improvement of the theorem 6.1 in this paper.

7 Conclusion and Remarks

In this paper, I suggested the following viewpoints.

1. I characterized a choice maker by (1) a surmising process that associates with each pair of lotteries a set of well-described choice situations between random variables, and (2) a choice set that is a subset of well-described choice situations between random variables.

2. Given a choice situation between lotteries (p, q) where p and q are probability measures on outcomes, a choice maker chooses the first lottery over the second lottery if and only if, for any choice situation $(f, g; P)$ selected by the surmising process, she chooses f over g.

3. The origin of Allais paradox and the common ratio effect is that the surmising process does not select any choice situation in which the two random variables are significantly positively correlated.

4. If probabilities on money-prizes are common knowledge to an experimenter and subjects as an outcome space, the origin of the common consequence effect is change of attitude toward risk.

5. I regard the utility evaluation effect and the preference reversal as violations of transitivity. The origin of these paradoxes is increasing risk aversion exhibited by a choice set.

Both the selection property by a surmising process and increasing risk aversion by a choice set satisfying allow natural explanation as behavior of a choice maker. I explained change of attitude toward risk in 4 as a peculiarity of the outcome space $\Lambda(R)$. With $\Lambda(R)$ as outcome space, the notion of risk-aversion (and risk-taking) that I adopted is a comparison of outcomes, not a comparison of probabilities on outcomes. I admit that whether probabilities on money-prizes become common knowledge to an experimenter and subjects is a matter of debate. I believe that it depends on the way a choice situation between lotteries is presented to subjects. I adopted a "bounded-resource" argument, that difficulty to reduce compounded lotteries encourages subjects to adopt probabilities on money-prizes as outcome space. Clearly this trade-off should be also a part of subjects' characteristics regarding model selection. Model selection is beyond the scope of this paper. Characterization of a choice maker including model selection is left to a future research.

References

[Allais & Hagen, 79] Allais, M. and O. Hagen, eds., *Expected Utility Hypothesis and the Allais Paradox*, Dordrecht, Holland: D.Reidel, 1979.

[Allais, 79] Allais, M., "The Foundation of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School," in [Allais & Hagen, 79], 1979.

[Beja & Gilboa, 92] Beja, A. and I. Gilboa, "Numerical Representation of Imperfectly Ordered Preferences (A Unified Geometric Approach)," *Journal of Mathematical Psychology*, 1992, *36*, 426-449.

[Chew, 83] Chew, S., "A Generalization of the Quasi-Linear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, *51*, 1983, 1065-1092.

[Dekel, 86] Dekel, E., "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, 1986, *40*, 304-318.

[Fishburn, 70] Fishburn, P.C., *Utility Theory for Decision Making*, 1970, Willey, New York.

[Fishburn, 82] Fishburn, P.C., "Nontransitive Measurable Utility," *Journal of Mathematical Psychology*, 1982, *26*, 31-67.

[Fishburn, 89] Fishburn, P.C., "Non-Transitive Measurable Utility for Decision Under Uncertainty," *Journal of Mathematical Economics*, 1989, *18*, 187-207.

[Krantz, Luce, Suppes & Tversky, 71] Krantz, D.H., Luce, R.D., Suppes, P, and A. Tversky, *Foundation of Measurement, Vol.1, Additive and Polynomial Representation*, 1971, Academic Press, New York and London.

[Krantz, Luce, Suppes & Tversky, 90] Krantz, D.H., Luce, R.D., Suppes, P, and A. Tversky, *Foundation of Measurement, Vol.3, Representation, Axiomatization, and Invariance*, 1971, Academic Press, New York and London.

[Kreps & Porteus, 78] Kreps, D.M. and E.L. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 1978, *46*, 185-200.

[Karni & Schmeidler, 91] Karni, E and D. Schmeidler, "Utility Theory with Uncertainty," Chapter 33, *Handbook of Mathematical Economics*, volume IV, 1763-1831.

[Kahneman & Tversky, 79] Kahneman, D. and A. Tversky, "Prospect Theory: an Analysis of Decision under Risk," *Econometrica*, 1979, *47*, 263-291.

[Loomes & Sugden, 82] Loomes, G. and R. Sugden, "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty," *Economic Journal*, December 1982, 92, 805-24.

[Machina, 87] Machina, M.J., "Choice under Uncertainty: Problems Solved and Unsolved," *Economic Perspectives*, Summer 1987, *vol. 1*, *No. 1*, 121-154

[Quiggin, 82] Quiggin, J., "A Theory of Anticipated Utility Theory," *Journal of Economic Behavior and Organization*, *3*, 323-343.