

ESSAYS ON COMPARATIVE STATICS ON NON-EXPECTED

UTILITY MODELS

A DISSERTATION

PRESENTED TO THE GRADUATE SCHOOL OF ECONOMICS OF

KYOTO UNIVERSITY

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY HIROYUKI TANAKA

December 2018

Abstract

This dissertation studies comparative statics analysis on risk taking behavior, time preference, and value of information. A trait of our approach is that we derive a comparative statics conclusion so that it can be applied to non-expected utility models, which gains growing interest in response to experimental evidence showing deviation from the expected utility framework. We address a question of when result for expected utility maximizers is extended analogously for non-expected utility maximizers; that is, whether or not a result under expected utility theory is robust to model specification. Motivated by it, we provide conditions for such robust conclusion in each chapter summarized as follows:

Chapter 1 reviews the standard qualitative approach to the comparative statics. we exploit ordinal properties of utility functions and dispense with assumptions such as differentiability to derive conditions for monotone behavior in expected utility models. As a benchmark result, we identify the class of expected utility where a monotone decision occurs to any shift in the sense of monotone likelihood ratio order.

Chapter 2 develops a comparison method of risk aversion for non-expected utility models under exogenous risks. We derive sufficient conditions for the comparison of certainty equivalents to be preserved with random initial wealth. By introducing the notion of Bayesian updating on objectively evaluated events, preservation results for expected utility maximizers are extended for non-expected utility models. As applications, a sufficient condition for the preservation of increasing certainty equivalent is also provided.

Chapter 3 investigates time preference and flexibility. We formulate the notion of impatience for random discounting models where there is uncertainty on discount factor and the decision maker has to commit his future consumption. By monotone

comparative statics approach, we show that the degree of impatience corresponds to the monotone likelihood ratio order on distributions of discount factor.

Chapter 4 examines the value of information for the maximin expected utility models. We formulate a decision problem of a decision maker exhibiting dynamic consistency and characterize comparison of experiments which is formulated so as to accommodate ambiguity on probability distributions. As a main result, we characterize how information is lost by replacing another information structure with Markov matrices.

Acknowledgment

I would thank my advisor Atsushi Kajii for advices regarding all chapters and encouragements. I would also thank Chiaki Hara for comments regarding Chapter 2 and Youichiro Higashi, Kazuya Hyogo, Norio Takeoka for the joint research regarding Chapter 3.

Contents

1	Monotone comparative statics	3
1.1	Single crossing condition and monotonicity of optimal set	3
1.2	Supermodularity and monotone comparative statics	4
1.3	Monotone comparative statics for expected utility	6
1.4	A variation of Athey's result	7
1.5	Log-supermodularity and supermodularity	10
1.6	Applications to portfolio problems	11
2	Comparison of Risk Taking Behavior with Random Initial Wealth	14
2.1	Introduction	14
2.2	Preliminaries	15
2.2.1	Utility	15
2.3	Comparison of certainty equivalents	17
2.4	Preservation of comparative certainty equivalent	18
2.5	Preservation of comparative certainty equivalent in non-expected utility models	21
2.5.1	Counter example to the third agent condition in non-expected utility models	21
2.5.2	Bayesian updating on objectively evaluated events	22
2.6	Main result	23
2.7	Other interpretations and related results	25
2.7.1	Preservation of decreasing certainty equivalent	25
3	Impatience and commitment	28
3.1	Introduction	28
3.2	Comparative impatience	29
3.3	Main result	31
3.4	Fundamental lemma	31
3.5	Proof of sufficiency	33

3.6	Proof of necessity	34
3.6.1	Definitions	34
3.6.2	Proof	35
3.7	Application to Goldman model	41
3.7.1	Model	41
3.7.2	Monotone money demand	42
4	Deficiency in the Maximin Expected Utility Models	45
4.1	Introduction	45
4.2	Preliminaries	47
4.3	Main result	50
4.4	Discussion	53
4.4.1	Three Ellsberg urns	53
4.4.2	Application: CEU model	54

1 Monotone comparative statics

1.1 Single crossing condition and monotonicity of optimal set

In this chapter, we review results on the monotonicity of an optimal solution with respect to the so called single crossing condition. To motivate our argument, recall the standard comparative static analysis based on differentiability, i.e., the sign of the derivative of the first order condition. For an optimal solution x^* satisfying the first order condition

$$U_x(x^*, \alpha) = 0$$

of an objective function $U(x, \alpha)$ on $X \subset \mathbb{R}$ parametrized by α in $A \subset \mathbb{R}$, its change with respect to α is, by the implicit function theorem,

$$\frac{dx^*}{d\alpha} = -\frac{U_{x\alpha}(x^*, \alpha)}{U_{xx}(x^*, \alpha)}.$$

Thus, the optimal solution is increasing in the parameter if and only if the right hand side of the above equation is positive. However, this argument requires the differentiability of U which is assumed only by a technical reason. In fact, for monotonicity of the optimal solution, we can replace the condition of the sign of derivative above by the following one:

Definition 1 (Increasing differences) *A function $f : X \times A \rightarrow \mathbb{R}$ has increasing differences if for any $x'' > x'$ and $\alpha'' > \alpha'$*

$$f(x'', \alpha'') - f(x', \alpha'') \geq f(x'', \alpha') - f(x', \alpha').$$

Note that this condition is not robust to monotone transformation, i.e., $\phi \circ f$ does not necessarily have increasing differences for monotone transformation ϕ even if f has. By a fact that a utility representation is essentially the same with its monotone transformation, it is plausible to introduce an ordinal version of increasing differences. To consider this question, let x^* be the optimal solution when the parameter takes a value α' and x^{**} such that $\alpha'' \geq \alpha'$, respectively. By the definition of optimality, we always have

$$f(x^*, \alpha') \geq f(x^{**}, \alpha').$$

To guarantee that if $x^* \geq x^{**}$, then x^* is also an optimal solution under α'' , we need $f(x^*, \alpha'') \geq f(x^{**}, \alpha'')$ (actually equality holds by definition of x^{**}). By this consideration, we reach the following ordinal formulation of increasing differences.

Definition 2 (Single crossing condition) *A function $f : X \times A \rightarrow \mathbb{R}$ is single crossing if for any $x'' \geq x'$ and $\alpha'' \geq \alpha'$*

$$f(x'', \alpha') \geq (>)f(x', \alpha') \implies f(x'', \alpha'') \geq (>)f(x', \alpha')$$

From discussion so far, we have the following result:

Theorem 1 *Suppose a function $f : X \times A \rightarrow \mathbb{R}$ satisfies the single crossing condition and let*

$$x^* \in \arg \max_x U(x, \alpha') \text{ and } x^{**} \in \arg \max_x U(x, \alpha'').$$

If $x^ \geq x^{**}$, then $x^*(x^{**})$ is also an optimal for $U(\cdot, \alpha'')(U(\cdot, \alpha'))$.*

This result excludes the situation where an optimal choice decreases while the parameter α increases.

1.2 Supermodularity and monotone comparative statics

The result we stated above has no implication to multivariate settings. Then, a natural question arising next is to obtain a condition for monotonicity applicable to n -dimensional Euclidean spaces. Milgrom and Shannon [6] use the lattice theory to order the solution sets in such a way and call the approach they introduce monotone comparative statics. In this section, we review their approach and a result.

One novelty of Milgrom and Shannon's approach is that they consider the models where the domain of the objective function is closed in two operations, meet and join, to order abstract partially ordered sets. Such sets are called lattice:

Definition 3 (Meet, join, and lattice) *Let x and y be elements of a partially ordered*

set X . The meet $x \wedge y$ of x and y in a partially ordered set X is

$$\sup\{z : z \leq x, y\}$$

The join $x \vee y$ of x and y is

$$\inf\{z : z \geq x, y\}$$

If these uniquely exist, X is called a lattice. If a subset $X' \subset X$ is also a lattice, it is called a sublattice.

By these operations, we can order subsets of a lattice by the following way.

Definition 4 (Strong set order) For two partially ordered sets X and Y $X \geq_s Y$ if for any $x \in X$ and $y \in Y$

$$x \vee y \in X \text{ and } x \wedge y \in Y$$

we call this order $\geq_s$ strong set order.

Note that X and Y themselves are not necessarily lattices.

For a function on a lattice, we can formulate an analogue of increasing differences by using meet and join.

Definition 5 (Supermodularity) A function $f : X \rightarrow \mathbb{R}$ is supermodular if

$$f(x \vee y) + f(x \wedge y) \geq f(x) + f(y)$$

When f is a function on the product of a set of choice variables X and a parameter set Θ , we say f supermodular if it is so as a function on $X \times \Theta$ and supermodular in X if it is so by fixing θ . In this sense, the supermodularity is used like increasing differences while we do not need specify which coordinate X_i of $X = X_1 \times \cdots \times X_n$ is used as a parameter set.

Analogous to increasing differences, the supermodularity has an ordinal version as follows:

Definition 6 (Quasisupermodularity) A function $f : X \rightarrow \mathbb{R}$ is quasisupermodular if for any x and y

$$f(x) \geq f(x \wedge y) \implies f(x \vee y) \geq f(y)$$

Note that a monotone transformation of a quasisupermodular function is also quasisupermodular.

A fundamental result by Milgrom and Shannon is that single crossing condition and quasisupermodularity in choice variable yield monotone optimal behavior in the sense of strong set order:

Theorem 2 (Milgrom and Shannon) *Suppose $U : X \times \Theta \rightarrow \mathbb{R}$ is supermodular in X and single crossing. For any $\alpha'' \geq \alpha'$ and a sublattice $A \subset X$*

$$\arg \max_{x \in A} U(x, \alpha'') \geq_s \arg \max_{x \in A} U(x, \alpha').$$

1.3 Monotone comparative statics for expected utility

If the utility function is expected utility, a condition monotonicity of optimal solution can be obtained as a condition for the von Neumann-Morgenstern function and a condition for probability law. This problem is treated by Athey [2]. The key condition is the supermodularity of logarithm:

Definition 7 (Log-supermodular) *Let X be a lattice. $f : X \rightarrow \mathbb{R}$ is log-supermodular if for any x and y*

$$f(x \vee y)f(x \wedge y) \geq f(x)f(y).$$

Let $\theta \subset \mathbb{R}$ and $\{f(\cdot, \theta)\}_{\theta \in \Theta}$ be a set of Radon-Nykodim derivative with respect to some probability law, we call it is ordered by monotone likelihood ratio (Henceforth, MLR) order if the inequality above holds.

A useful property of this condition is log-supermodularity is preserved by integration. This is proved in Ahlswede and Daykin [1].

Lemma 1 *Let f_1, f_2, f_3 , and f_4 be nonnegative functions on a lattice X . If*

$$f_1(x \vee y)f_2(x \wedge y) \geq f_3(x)f_4(y),$$

then,

$$\int f_1 d\nu \int f_2 d\nu \geq \int f_3 d\nu \int f_4(y) d\nu$$

where ν is a probability law on X . Since log-supermodularity implies single crossing and quasisupermodular, this result directly yields the following conclusion:

Theorem 3 (Athey) *Let $f : S \times \Theta \rightarrow \mathbb{R}$ be a log-supermodular function. Then, for $\theta'' \geq \theta'$*

$$\arg \max_{x \in X} \int u(x, s) f(s, \theta'') \geq_s \arg \max_{x \in X} \int u(x, s) f(s, \theta') \text{ for all log-supermodular } u.$$

1.4 A variation of Athey's result

In this section, we derive an analogous result to Athey's one stated as follows:

Proposition 1 *Let u be a log-supermodular function. Then, for $\theta'' \geq \theta'$*

$$\arg \max_{x \in X} \int u(x, s) f(s, \theta'') \geq_s \arg \max_{x \in X} \int u(x, s) f(s, \theta') \text{ for all log-supermodular } f.$$

The difference between this result and Athey's result is the role of utility and that of probability density. Athey's result identifies the environment under which all log-supermodular utilities behave monotonically with respect to parameters. To the contrary, our result identifies the class of utility behaving monotonically under all environments ordered by log-supermodular. Thus, the result meets a demand from an analyst who try to classify preferences but the preference is estimated only through its behavior.

In the rest of this section, we shall prove a generalization of this proposition and obtain the proposition above as a corollary.

To derive such generalization, one candidate for characterization is single crossing and quasisupermodular von Neumann-Morgenstern function as in Milgrom and Shannon's result. However, it is impossible as seen in the following example:

Example 1 (Quasisupermodular is not preserved by MLR) Define u on $\{0, 1\}^2 \times \{0, 1\}$ so that

$$u(1, 0; 0) - u(0, 0; 0) = 1$$

$$u(1, 0; 1) - u(0, 0; 1) = \epsilon$$

$$u(1, 1; 0) - u(0, 1; 0) = -\epsilon$$

$$u(1, 1; 1) - u(0, 1; 1) = -1$$

Then, u is quasi-supermodular in $\{0, 1\}^2$ and single crossing. However, For point masses δ_0 and δ_1 we have δ_1 MLR-dominates δ_0 and

$$\int u(1, 0; s) \delta_0 - \int u(0, 0; s) \delta_0 = 1 - \epsilon > 0$$

and

$$\int u(1, 0; s) \delta_0 - \int u(0, 0; s) \delta_1 = \epsilon - 1 < 0$$

Thus, the MLR ordering does not guarantee the preservation of quasi-supermodular.

While the example above shows quasi-supermodularity is not preserved by the MLR ordered family, supermodularity is preserved no matter what the distribution is.

On the other hand, the single crossing condition is preserved under integration with MLR ordered density, i.e.,

Lemma 2 *u is single crossing and f is log-supermodular. Then,*

$$U(x, \theta) = \int u(x, s) f(s, \theta)$$

is single crossing.

Proof See Karlin and Rubin [4].

Proposition 2 *(i) implies (ii) :*

1. *u is single crossing in s and supermodular in a*

2. *Monotone comparative statics holds for all probability density functions $p : S \times \Theta \rightarrow \mathbb{R}_+$ that are log-supermodular*

Proof By lemma.

The condition of the MLR ordering can not be weakened to first order stochastic dominance as checked by the following example:

Example 2 (Investment problem) Take any real numbers $\bar{s} > \underline{s}$, and $\epsilon \in (0, 1/2)$. Define $u : \{0, 1\} \times \mathbb{R} \rightarrow \mathbb{R}$ by

$$u(1, s) := \begin{cases} -\epsilon & \text{if } s < \underline{s} \\ -1 & \text{if } \underline{s} \leq s < \bar{s} \\ 2\epsilon & \text{if } \bar{s} \leq s \end{cases}$$

$$u(0, s) := 0$$

An interpretation is that the decision maker chooses “invest” (1) or “not invest” (0) and receive the return depending on state s . States are divided into three parts: “bad” $s \leq \underline{s}$ “very bad” $s \in [\underline{s}, \bar{s})$ and “good” $s \geq \bar{s}$. Since if choosing “invest” at s is better than “not invest”, so does for all $s'' \geq s$, u is single crossing. But $\int u(s) dP_\theta(s)$ is not single crossing under the following probability laws: Take a, b, c, d so that $a < \underline{s} < b < \bar{s} < c < d$ and define P_0 by $1/2\delta_a + 1/2\delta_c$ and P_1 by $1/2\delta_b + 1/2\delta_d$ ¹. Note that $P_1 \geq_{FOSD} P_0$ because $a < b < c < d$. Since

$$u(1, s) - u(0, s) := \begin{cases} -\epsilon & \text{if } s < \underline{s} \\ -1 & \text{if } \underline{s} \leq s < \bar{s} \\ 2\epsilon & \text{if } \bar{s} \leq s \end{cases}$$

we have $\int u(1, s) - u(0, s) dP_0 > 0$ and $\int u(1, s) - u(0, s) dP_1 < 0$. Thus, as in the example above, the single crossing property is not preserved through the first order stochastic dominance. Although the single crossing property implies that if “not invest” is better than “invest” at s so does for all $s' \leq s$, it does not put any condition how it is. Then, the

¹ δ_a is the point mass at a

preservation of the single crossing property can not be guaranteed under risks like P_0 and P_1 .

1.5 Log-supermodularity and supermodularity

Note that generally, when a function is increasing, the log-supermodularity implies the supermodularity while the converse is not necessarily true. To see this, consider the following function $f(z_1, z_2) = z_1 + z_2$ where the domain of f is $\mathbb{R}_{++}$. Verify that

$$f\left(\frac{1}{2}, 1\right) + f\left(1, \frac{1}{2}\right) \leq f(1, 1) + f(0, 0)$$

while

$$f\left(\frac{1}{2}, 1\right)f\left(1, \frac{1}{2}\right) > f(1, 1)f(0, 0).$$

The reason the log-supermodularity implies the supermodularity is that the supermodularity is preserved by an increasing convex transformation given in Topkis [1978]. To see this, let f be a function from lattice X into $\mathbb{R}$ and $T : \mathbb{R} \rightarrow \mathbb{R}$ an increasing convex transformation. Then, for any x and y in X

$$\begin{aligned} & T(f(x)) + T(f(y)) - T(f(x \vee y)) - T(f(x \wedge y)) \\ &= \{T(f(x)) - T(f(x \vee y) + f(x \wedge y) - f(y))\} \\ &+ \{T(f(x \vee y) + f(x \wedge y) - f(y)) - T(f(x \vee y)) - T(f(x \wedge y)) + T(f(y))\}. \end{aligned}$$

Note that the first brace is nonpositive since T is increasing and f is supermodular and that the second brace is nonpositive since T is convex and f is supermodular. Thus the whole expression above is nonpositive.

If a positive function is log-supermodular, the log transformation is well-defined and it is supermodular. The exponential transformation is increasing and convex. Hence, the log-supermodularity implies the supermodularity in case functions are positive. Thus proposition 2 is a generalization of proposition 1.

1.6 Applications to portfolio problems

In this section, we will apply results derived in the previous section to portfolio problems. There are two applications. The first one is a saving problem and the second one is a portfolio choice problem.

Application 1: Consider the following portfolio problem as an application.

$$\max_{a \in [0, w]} u(w - a) + \delta \int u(a - \epsilon) p(\epsilon, \theta) d\epsilon$$

A risk averse decision maker u with discount factor $\delta \in (0, 1)$ has initial wealth w . He chooses saving a for the future where there will be negative shocks $\epsilon \in [0, \bar{\epsilon}]$. Densities $p(\cdot, \theta)$ of these distributions are parametrized by θ in Θ where Θ is a poset. In this setting, we will verify that the optimal solution set will increase in response to increasing in w , θ and $\bar{\epsilon}$ as in the usual case that u is twice differentiable.

Assume that distributions are MLR-ordered. Since the decision maker is risk averse, u is a concave function. Pick any $a > a'$ in $[0, w]$ and any $\epsilon > \epsilon'$. Then, for some $\lambda \in [0, 1]$

$$\begin{aligned} & u(a - \epsilon) + u(a' - \epsilon') \\ &= u(\lambda(a - \epsilon') + (1 - \lambda)(a' - \epsilon)) + u(\lambda(a' - \epsilon) + (1 - \lambda)(a - \epsilon')) \\ &\geq \lambda u(a - \epsilon') + (1 - \lambda)u(a' - \epsilon) + \lambda u(a' - \epsilon) + (1 - \lambda)u(a - \epsilon') \\ &= u(a - \epsilon') + u(a' - \epsilon) \end{aligned}$$

To derive the first inequality, the concavity of u is used. Then, u has increasing differences in (a, ϵ) . The supermodularity of u in a is obvious since this is a one-variate function. To sum up, we have the monotone comparative statics conclusion of this portfolio problem by Proposition 2: As θ increases, the optimal solution set will increase. Note that we did not assume the differentiability of u and that u is not necessarily nonnegative as required in Athey's result. An interpretation is that since the decision maker is risk averse, his savings will increase when negative shocks are serious and the wealth in hand is large

so that his consumption will be smoothed. A large θ is a signal that the distribution of negative shocks will be worse, and thus it makes the decision maker save more.

Application 2: Consider the following utility maximization problem.

$$\max_{\alpha \in \{0,1\}} \int u(z + \alpha f(s)) p(s, \theta) ds$$

where there is $\bar{s}$ such that $f(s') < f(\bar{s}) < f(s'')$ for all $s' < \bar{s} < s''$ and $f(\bar{s}) = 0$. An interpretation of this model is that the decision maker chooses whether to buy or not an asset which will pay $f(s)$ which is affected by the characteristics s of the firm. While the relation between the characteristic and the payoff is not necessarily monotonic, the range of the characteristic is separated into the negative range and the positive range by the zero-payoff range. We will verify that if there is uncertainty about characteristic of the firm, the MLR shift of the distribution makes the decision maker take more risk.

von Neumann-Morgenstern function u is supermodular since u is a one-variate function and single crossing in (α, s) since $u(z + f(s')) < u(z + f(\bar{s})) < u(z + f(s''))$. Note that u does not need to have increasing differences since f may be a non-monotonic function. The objective function is an integration of supermodular in α and single crossing in (α, s) . Then, if $\theta > \theta'$, then we have that the decision maker takes a risk under θ' , so does under θ .

References

- [1] Ahlswede, Rudolf, and David E. Daykin. "An inequality for the weights of two families of sets, their unions and intersections." *Probability Theory and Related Fields* 43.3 (1978): 183-185.
- [2] Athey, Susan. "Monotone comparative statics under uncertainty." *quarterly Journal of Economics* (2002): 187-223.
- [3] Gollier, Christian. *The economics of risk and time*. MIT press, (2004).

- [4] Karlin, Samuel, and Herman Rubin. "The theory of decision procedures for distributions with monotone likelihood ratio." *The Annals of Mathematical Statistics* (1956): 272-299.
- [5] Karlin, Samuel, and Yosef Rinott. "Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions." *Journal of Multivariate Analysis* 10.4 (1980): 467-498.
- [6] Milgrom, Paul, and Chris Shannon. "Monotone comparative statics." *Econometrica: Journal of the Econometric Society* (1994): 157–180.
- [7] Topkis, Donald M. "Minimizing a submodular function on a lattice." *Operations research* 26.2 (1978): 305–321.

2 Comparison of Risk Taking Behavior with Random Initial Wealth

2.1 Introduction

Portfolio choice between a risky asset and a safe asset is determined by preferences on risk and uncertainty and these preferences are characterized by their certainty equivalent. As shown in the classical study by Pratt [11], under the expected utility theory, comparison among investors' certainty equivalent can be examined by comparison through absolute risk aversion, risk premium, concave transformation, and so on. However, if we introduce an exogenous risk to the initial wealth to analyze the more realistic situation, these comparison may be reversed as seen in an example by Kihlstrom, Romer, and Williams [6]. Motivated to obtain a robust conclusion to such perturbations, Nachman [10] and Pratt [12] derive sufficient conditions for preservation of comparison of certainty equivalent. Roughly speaking, their results are summarized as follows: If we find the third agent with increasing certainty equivalent between these of two agents, then the comparison of these two agents is preserved under random initial wealth.

Since their results are obtained for expected utility maximizers, recent development of decision theory motivates the extension to models not necessarily expected utility models. For example, if investors are not confident on their probability distribution, their portfolio choice depends not only on risk but also ambiguity of probability distributions. In such situation, their utility functions deviate from the expected utility framework. One celebrated formulation describing this is Gilboa and Schmeidler[3] among others.

Our main results is an extension of preservation of comparative statics on portfolio choice for expected utility to that for non-expected utility so as to accommodate multiple probabilities such as ambiguity models where a decision maker evaluates risky assets by multiple probability measures and does not behave as if he has the unique probability which makes him as the expected utility maximizer.

In ambiguity models, potential probability measures are in most cases interpreted as

subjective probability measures because ambiguity emerges only when there is no objectively plausible probability and some likelihood to occur some event may be disagreed among agents. However, there can be a situation where certain events are objectively evaluated, that is, there can be a collection of events which all potential probabilities are equivalent if restricted on them. To illustrate it, consider the classical experiment known as Ellsberg's experiment in which subjects draw balls from an urn containing 90 balls whose composition is: 30 red balls and 60 black or yellow balls. A speculation on the share of black balls in 60 balls may differ in subjects and may be multiple as suggested by the experiment. However, it seems to be natural to put probability $1/3$ to red balls and to consider drawing red as an objectively evaluated event. Thus, natural description of Ellsberg's experiment assumes that potential probabilities are equivalent on $\{\{\text{red}\}\{\text{black,yellow}\}\}$, that is, they must assign $1/3$ to $\{\text{red}\}$ and $2/3$ to $\{\text{black,yellow}\}$. Motivated by this observation, we classify events into subjective and objective evaluated events and derive their consequences to portfolio choice.

2.2 Preliminaries

2.2.1 Utility

Let $\mathcal{P}_1$ be a set of probability measures on a measurable space $(S_1, \mathcal{S}_1)$ and let P_2 be a probability measure on a measurable space $(S_2, \mathcal{S}_2)$. The state space is the product measurable space $S := S_1 \times S_2$, endowed with the product σ -algebra $\mathcal{S} := \mathcal{S}_1 \otimes \mathcal{S}_2$ and the set $\mathcal{P}$ of probability measures on it. Each probability measure $P \in \mathcal{P}$ is defined as a product probability measure $P_1 \times P_2$ for some $P_1 \in \mathcal{P}_1$.

A decision maker chooses whether to have a risky alternative $\hat{x}$ or not. The risky alternative is a random variable with a finite range and defined on a probability space S . The decision maker thinks its realization obeys some probability measure in $\mathcal{P}$ but may not be confident on which probability measure is true. Since the multiplicity of probabilities is due to $\mathcal{P}_1$, $\mathcal{S}_1$ is interpreted as the set of events whose probability is evaluated subjectively. On the other hand, $\mathcal{S}_2$ is interpreted as the set of events whose probability is evaluated

objectively. Accordingly, a random variable $\hat{x}$ is subjective if for each value x

$$\{s \in S : \hat{x} = x\} = S'_1 \times S_2 \text{ for some } S'_1 \in \mathcal{S}_1$$

and a random variable $\hat{x}$ is objective if for each value x

$$\{s \in S : \hat{x} = x\} = S_1 \times S'_2 \text{ for some } S'_2 \in \mathcal{S}_2.$$

A random variable may be both subjective and objective. However, in this case it must be a constant function.

As another remark, since an inverse image of a subjective (objective) random variable is a union of cylinder set of the form $S'_1 \times S_2$ ($S_1 \times S'_2$), we implicitly assume that a subjective random variable and an objective random variable are automatically independent.

Assumption 1 (Non-atomicity of S_2) *For any positive real number $p \in (0, 1)$ there is a measurable set $S'_2 \subset S_2$ such that*

$$P_2(S'_2) = p.$$

Utility function U_i of agent i assigns a random variable with a finite range on S a real number. We assume:

Assumption 2 (Monotonicity for random variables) *For any risky alternatives $\hat{x}$ and $\hat{y}$, if $\hat{x} \geq \hat{y}$, then $U_i(\hat{x}) \geq U_i(\hat{y})$.*

Furthermore, we also assume :

Assumption 3 (Differentiability) *For each constant c*

$$\lim_{\epsilon \rightarrow 0} \frac{U_i(c + \epsilon) - U_i(c)}{\epsilon}$$

exists.

Example 3 (Expected utility) Let $\mathcal{P} = \{P\}$ and define utility function as:

$$U_i(\hat{x}) = \int u_i(\hat{x}(s)) dP(s).$$

In this case, Assumption 3 is differentiability of the von-Neumann Morgenstern function u .

Example 4 (Maximin expected utility)

$$U_i(\hat{x}) = \min_{P \in \mathcal{P}} \int u_i(\hat{x}(s)) dP(s)$$

In this case, Assumption 3 is differentiability of the von-Neumann Morgenstern function u .

Example 5 (Utility with interim decision making) Let $\mathcal{P} = \{P\}$ and define utility function as:

$$U_i(\hat{x}) = \int \max_{q \cdot c \leq x} u_i(c_1, \dots, c_n) dP(s)$$

where $q \in \mathbb{R}_{++}^n$ and $c \in \mathbb{R}_+^n$. An interpretation is that the decision maker chooses his consumption in the budget constraint generated by monetary prize by the random variable $\hat{x}$. In this case, Assumption 3 is differentiability of the (interim) value function.

2.3 Comparison of certainty equivalents

Comparison of attitude toward uncertainty can be defined in various ways. Our definition is given with certainty equivalent:

Definition 8 (Comparison of certainty equivalents) Consider decision makers U_i ($i=1,2$). For a random variable $\hat{x}$ and initial wealth w define the constant $C_i(\hat{x}, w)$ as

$$U_i(\hat{x} + w) = U_i(C_i(\hat{x}, w) + w)$$

U_2 is ready to pay more than U_1 to take a risk at w if

$$C_1(\hat{x}, w) \leq C_2(\hat{x}, w)$$

If the inequality holds for all w , we say U_2 is ready to pay more than U_1 to take a risk.

An interpretation of $C_i(\hat{x}, w)$ is monetary value of the risky asset $\hat{x}$ for agent i with initial wealth w . Note that the certainty equivalent is affected by ambiguity. Thus, certainty equivalent is not determined only by attitudes toward risks but also by attitudes toward uncertainty.

An equivalent condition of comparison is defined in terms of a behavioral condition:

Definition 9 (Comparative risk taking behavior) U_2 takes more risks than U_1 if for any risky alternative $\hat{x}$ and a constant w ,

$$U_1(\hat{x} + w) - U_1(w) \geq (>)0$$

implies

$$U_2(\hat{x} + w) - U_2(w) \geq (>)0.$$

This condition means that if agent 1 takes a risk, then so does agent 2. Note that this comparison of attitude toward risks is also affected by attitude toward uncertainty as in certainty equivalent and is equivalent to the previous criterion formulated with certainty equivalent.

The key property for our main result is monotonicity of certainty equivalent:

Definition 10 (Increasing certainty equivalent) Decision maker U_i has increasing certainty equivalent if for any random variable $\hat{x}$ $C_i(\hat{x}, w)$ is increasing in w .

This can be formulated with a behavioral condition such that the agent tends to take risks more often as he becomes wealthy.

2.4 Preservation of comparative certainty equivalent

Since even if no action is taken, we must incur exogenous risk in reality, it is plausible to change the problem to choice between taking an exogenous risk or taking an additional risk to the exogenous risk. The comparison of attitude toward uncertainty in such situations is formulated as follows:

Definition 11 (Comparison of certainty equivalent under random initial wealth)

U_2 is ready to pay more than U_1 under random initial wealth if for any subjective random variable $\hat{x}$ and any objective random variable $\hat{w}$

$$C_1(\hat{x}, \hat{w}) \leq C_2(\hat{x}, \hat{w})$$

where $C_i(\hat{x}, \hat{w})$ is a constant such that

$$U_i(\hat{x} + \hat{w}) = U_i(C_i(\hat{x}, \hat{w}) + \hat{w}).$$

As in the previous section, certainty equivalent measures a risky asset $\hat{x}$ by amount of money C_i . The difference is that an risky alternative is a subjective random variable and initial wealth is an objective random variable. Equivalently,

Definition 12 (Comparison of risk taking behavior) U_2 takes more risks than U_1 under random initial wealth if for any subjective random variable $\hat{x}$ and any objective random variable $\hat{w}$

$$U_1(\hat{x} + \hat{w}) - U_1(\hat{w}) \geq (>)0.$$

then

$$U_2(\hat{x} + \hat{w}) - U_2(\hat{w}) \geq (>)0,$$

These two definitions for comparison are equivalent as in the case where wealth is constant.

One reason we assume independence of $\hat{x}$ and $\hat{w}$ is that without this condition, the preservation result is quite demanding to utility functions. In the expected utility model, Ross [14] shows that for accommodating any dependence of risky alternative and random initial wealth von Neumann-Morgenstern function u_1 with smaller certainty equivalent than u_2 must satisfy $u_1 = au_2 + g$ for some positive constant a and decreasing concave function g . Thus, to put relatively mild assumption to utility functions, we assume independence induced by the subjective-objective classification.

The preservation of comparison of certainty equivalent does not hold in full generality even if utility function is expected utility. Actually, Kihlstrom, Romer, and Williams [6] gives the following example:

Example 6 (Kihlstrom-Romer-Williams) Define the von Neumann-Morgenstern functions on $(0, 1)$ of expected utility as

$$u_n(x) = x - \frac{1}{n}x^n$$

Absolute risk aversion of these function is

$$\frac{(n-1)x^{n-2}}{1-x^{n-1}} \quad (1)$$

which is larger when $n = 2$ than when $n = 11$. Thus, u_{11} takes more risks than u_2 . Under random initial wealth $\hat{w}$ they become

$$\frac{(n-1)E[x^{n-2} + \hat{w}]}{1 - E[x^{n-1} + \hat{w}]} \quad (2)$$

For $n = 2, 11$ set $\hat{w} = 0.01$ or 0.99 equally likely and $x = 0$. Then, (1) is larger when $n = 2$ but (2) is larger when $n = 11$. This implies that certainty equivalent of u_{11} is larger when initial wealth is constant but that of u_2 is larger under random initial wealth.

One observation of the example above is that both utility functions do not exhibit decreasing absolute risk aversion, i.e, increasing certainty equivalent. Nachman [10] and Pratt [12] derive a sufficient condition for that comparison is preserved under random initial wealth. This preservation result uses the condition that for each random variable $\hat{x}$ there is an increasing function g such that

$$C_1(\hat{x}, w) \leq g(w) \leq C_2(\hat{x}, w).$$

This condition is interpreted as follows: there is a third agent such that he is more risk averse than agent 1, less risk averse than agent 2, and has increasing certainty equivalent, in our notation:

Theorem 4 (Preservation of comparison of risk taking behavior for expected utility)

Let $\mathcal{P} = \{P\}$ and U_i ($i = 1, 2, 3$) be expected utilities:

$$U_i(\hat{x}) = \int u_i(\hat{x}(s))dP(s).$$

Suppose that U_2 is ready to pay more than U_1 to take a risk. If decision maker U_3 is ready to pay more than U_1 to take a risk, less than U_2 , and its certainty equivalent is increasing, then U_2 is ready to pay more than U_1 to take a risk under random initial wealth.

Note that if either decision maker 1 or 2 has increasing certainty equivalent, then the existence of a third agent is immediate, i.e. one of them can be used as this.

2.5 Preservation of comparative certainty equivalent in non-expected utility models

2.5.1 Counter example to the third agent condition in non-expected utility models

By the result in the previous section, we reviewed a characterization of the robust comparative statics for expected utility models. In this section, we shall seek a preservation result of comparative certainty equivalent analogue in the non-expected utility model. We first verify that the existence of a third agent is not sufficient as shown in the following example.

Example 7 (Mean-variance preference) Let $\mathcal{P} = \{P\}$ and ϕ_i nonnegative increasing function with $\phi_i(0) = 0$. Let

$$U_i(\hat{x}) = E\hat{x} - \phi_i(V\hat{x})$$

where E and V are the expectation and variance respectively with respect to P . Assume that

$$\phi_1 \geq \phi_2$$

and

$$\phi_1(a) - \phi_1(b) \leq \phi_2(a) - \phi_2(b)$$

for some $a > b > 0$. Since for any w

$$U_i(\hat{x} + w) - U_i(w) = E\hat{x} - \phi_i(V\hat{x}) + \phi_i(0) = E\hat{x} - \phi_i(V\hat{x}),$$

U_2 takes more risks than U_1 and certainty equivalent of these two agents is independent of wealth (then, trivially exhibits increasing certainty equivalent). However, under random initial wealth $\hat{w}$

$$U_i(\hat{x} + \hat{w}) - U_i(\hat{w}) = E\hat{x} - \phi_i(V\hat{x} + V\hat{w}) + \phi_i(V\hat{w})$$

Then, U_1 takes risk but U_2 does not for $\hat{x}$ and $\hat{w}$ such that $V\hat{x} + V\hat{w} = a$ and $V\hat{w} = b$.

As seen in the example, the preservation result is impossible to obtain in full generality even if there is a third agent between two compared decision makers. Then, in the subsequent subsection, we shall introduce a condition which makes non-expected utilities partially similar to the expected utilities.

2.5.2 Bayesian updating on objectively evaluated events

As seen in the mean-variance model in the previous subsection, nonlinearity of probability seems to be a major obstacle for the preservation of comparison. Then, we introduce the notions to make the model closer to the one linear in probability.

The next condition is a consistency condition making the utility a little closer to the expected utility.

Assumption 4 (Bayesian updating on objectively evaluated events) *Let $\hat{x}$ be a subjective random variable and $\hat{y}$ be an objective random variable whose range is $\{y_1, \dots, y_m\}$ where all y_k s are different. Then, we have*

$$U_i(\hat{x} + \hat{y}) = \sum_{k=1}^m P(E_k)U(\hat{x} + y_k)$$

where

$$E_k := \{s \in S : \hat{y} = y_k\} = S_1 \times \{s_2 \in S_2 : \hat{y}(s_1, s_2) = y_k\}.$$

This condition means that the agent is able to make a thought experiment with Bayesian updating. Consider a situation where the decision maker knows that he is in some event E . At this event, he evaluates the restriction of $\hat{x}$ to E but, since E is an objective event, the probability distribution of $\hat{x}$ should not be affected by this information. Thus, summing these interim utility with weight $P(E_k)$, he evaluates his ex ante payoff as this weighted sum above.

2.6 Main result

In this section, we shall prove the main result of this chapter. The way to prove the result is analogous to the usual expected utility preservation proof with slight difference such as conditional utility may not have an expected utility functional form.

Theorem 5 *Suppose that decision maker U_2 is ready to pay more than U_1 to take a risk. If there is an agent U_3 such that it is ready to pay more than U_1 to take a risk and less than U_2 , and has increasing certainty equivalent, then U_2 is ready to pay more than U_1 to take a risk under initial random wealth.*

Proof Take any subjective random variable $\hat{x}$ and $\hat{w}$ any objective random variable. Suppose that $g(w)$ is third agent's certainty equivalent:

$$g(w) = C_3(\hat{x}, w).$$

Lemma 3 $C_1(\hat{x}, \hat{w}) \leq C_1(g(\hat{w}), \hat{w})$ and $C_2(g(\hat{w}), \hat{w}) \leq C_2(\hat{x}, \hat{w})$.

Proof Let $\{w_1, \dots, w_n\}$ be the range of $\hat{w}$ with $w_k \neq w_l$ for all $k \neq l$ and

$$E_k = \{s : \hat{w} = w_k\}.$$

we have

$$\begin{aligned} U_1(C_1(\hat{x}, \hat{w}) + \hat{w}) &= U_1(\hat{x} + \hat{w}) \\ &= \sum_k P(E_k) U_1(\hat{x} + w_k) \\ &= \sum_k P(E_k) U_1(C_1(\hat{x}, w_k) + w_k) \\ &\leq \sum_k P(E_k) U_1(g(w_k) + w_k) = U_1(g(\hat{w}) + \hat{w}) \\ &= U_1(C_1(g(\hat{w}), \hat{w}) + \hat{w}) \end{aligned}$$

Thus, by monotonicity, $C_1(\hat{x}, \hat{w}) \leq C_1(g(\hat{w}), \hat{w})$. $C_2(g(\hat{w}), \hat{w}) \leq C_2(\hat{x}, \hat{w})$ can be verified in the similar argument.

Lemma 4

$$C_1(g(\hat{w}), \hat{w}) \leq C_2(g(\hat{w}), \hat{w})$$

Proof Take any w, w' , and w'' so that $w' < w < w''$. Let E be an objectively evaluated event. By strict monotonicity for constants and continuity implied by differentiability, we can construct the following risky asset:

$$\hat{z}(s) := \begin{cases} w'' & \text{if } s \in E \\ w' & \text{if } s \in E^c \end{cases}$$

such that

$$pU_1(w'') + (1 - p)U_1(w') = U_1(w)$$

where $p = P(E)$ for all $P \in \mathcal{P}$. Then,

$$p = \frac{U_1(w) - U_1(w')}{U_1(w'') - U_1(w')}.$$

Since

$$p(U_2(w'') - U_2(w)) + (1 - p)(U_2(w') - U_2(w)) \geq 0,$$

we have

$$\frac{U_1(w) - U_1(w')}{U_1(w'') - U_1(w')} (U_2(w'') - U_2(w)) + \frac{U_1(w'') - U_1(w)}{U_1(w'') - U_1(w')} (U_2(w') - U_2(w)) \geq 0.$$

By rearranging, we have

$$\frac{U_1(w'') - U_1(w)}{U_2(w'') - U_2(w)} \leq \frac{U_1(w) - U_1(w')}{U_2(w) - U_2(w')}$$

By the Cauchy mean value theorem, there are $\hat{w}' \in [w', w]$ and $\hat{w}'' \in [w, w'']$ such that

$$\frac{U_1'(w'')}{U_2'(w'')} \leq \frac{U_1'(\hat{w}'')}{U_2'(\hat{w}'')} = \frac{U_1(w'') - U_1(w)}{U_2(w'') - U_2(w)} \leq \frac{U_1(w) - U_1(w')}{U_2(w) - U_2(w')} = \frac{U_1'(\hat{w}')}{U_2'(\hat{w}')} \leq \frac{U_1'(w')}{U_2'(w')}$$

which implies that

$$\begin{aligned} \frac{U_2(w) - U_2(w')}{U_2'(w'')} &\leq \frac{U_1(w) - U_1(w')}{U_1'(w'')} \\ \frac{U_2(w'') - U_2(w)}{U_2'(w')} &\geq \frac{U_1(w'') - U_1(w)}{U_1'(w')} \end{aligned}$$

Then, we have

$$\frac{U_1(g(w) + w) - U_1(C_1(g(\hat{w}), \hat{w}) + w)}{U_1(b + C(g(\hat{w}), \hat{w}))} \leq \frac{U_2(g(w) + w) - U_2(C_1(g(\hat{w}), \hat{w}) + w)}{U_2'(b + C(g(\hat{w}), \hat{w}))}$$

where $b := \inf\{w : g(w) \geq C_1(g(\hat{w}), \hat{w})\}$. By taking expectation by $P(E_k)$, the left hand side is zero. Hence, $C_2(g(\hat{w}), \hat{w}) \geq C_1(g(\hat{w}), \hat{w})$.

By Lemma 3 and 4, we have

$$C_1(\hat{x}, \hat{w}) \leq C_1(g(\hat{w}), \hat{w}) \leq C_2(g(\hat{w}), \hat{w}) \leq C_2(\hat{x}, \hat{w}).$$

2.7 Other interpretations and related results

2.7.1 Preservation of decreasing certainty equivalent

We can interpret two different agents as the same person with different situations. For example, an agent has initial wealth w and this agent with initial wealth w' may be interpreted as different agents and can be compared by the analysis in the previous section. As a corollary of the main result, we have the following result.

Corollary 1 *Let U be an agent with increasing certainty equivalent. Then, it has increasing certainty equivalent under random initial wealth.*

Proof For $w_1 < w_2$ let

$$U_i(\hat{x}) = U(\hat{x} + w_i)$$

Then, one of U_i can be the third agent between U_1 and U_2 . Then, apply the theorem.

References

- [1] Ellsberg, Daniel. "Risk, ambiguity, and the Savage axioms." The quarterly journal of economics (1961): 643-669.
- [2] Epstein, Larry G., and Jiankang Zhang. "Subjective probabilities on subjectively unambiguous events." Econometrica 69.2 (2001): 265-306.

- [3] Gilboa, I and Schmeidler, D. "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics* 18, 1989: 141-153.
- [4] Gollier, Christian. *The economics of risk and time*. MIT press, 2004.
- [5] Gollier, Christian, and John W. Pratt. "Risk vulnerability and the tempering effect of background risk." *Econometrica: Journal of the Econometric Society* (1996): 1109-1123.
- [6] Kihlstrom, Richard E., David Romer, and Steve Williams. "Risk aversion with random initial wealth." *Econometrica: Journal of the Econometric Society* (1981): 911-920.
- [7] Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. "A smooth model of decision making under ambiguity." *Econometrica* 73.6 (2005): 1849-1892.
- [8] Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. "Definitions of ambiguous events and the smooth ambiguity model." *Economic Theory* 48.2-3 (2011): 399-424.
- [9] Machina, Mark J., and David Schmeidler. "A more robust definition of subjective probability." *Econometrica: Journal of the Econometric Society* (1992): 745-780.
- [10] Nachman, David C. "Preservation of more risk averse under expectations." *Journal of Economic Theory* 28.2 (1982): 361-368.
- [11] Pratt, John W. "Risk aversion in the small and in the large." *Econometrica: Journal of the Econometric Society* (1964): 122-136.
- [12] Pratt, John W. "Aversion to one risk in the presence of others." *Journal of Risk and Uncertainty* 1.4 (1988): 395-413.
- [13] Quiggin, John. "Background risk in generalized expected utility theory." *Economic Theory* 22.3 (2003): 607-611.

- [14] Ross, Stephen A. "Some stronger measures of risk aversion in the small and the large with applications." *Econometrica: Journal of the Econometric Society* (1981): 621-638.
- [15] Yaari, Menahem E. "Some remarks on measures of risk aversion and on their uses." *Journal of Economic theory* 1.3 (1969): 315-329.
- [16] Zhang, Jiankang. "Subjective ambiguity, expected utility and Choquet expected utility." *Economic Theory* 20.1 (2002): 159-181.

3 Impatience and commitment¹

3.1 Introduction

In this chapter, we investigate comparative statics analysis for time trade-off and flexibility. To motivate our discussion, recall a basic model with time trade-off of the following form

$$\sum_t \alpha^t u(c_t).$$

While this model is ubiquitous in economic analysis, it has the assumption behind the formulation that the decision maker understands what he actually consume in the future. That is, he knows how he discount the future and there is no uncertainty about time preference. However, there are situations where the timing of choice and that of actual consumption are different. This difference may generate problems from the descriptive view point if his preference or mood changes as time goes on. For example, if there is uncertainty about how he discount the future, a problem such as whether a security yielding the fixed payoff should be hold is not a simple problem because the decision maker should compare states where the security is desirable for him and states where it is not.

To describe such situations, Higashi, Hyogo, and Takeoka [12] axiomatize preference under such subjective uncertainty and derive utility representation called random discounting model. In the subsequent sections, we consider the following model based on their study:

$$U(M) = \sum_{\alpha} \max_{c \in M} \sum_t \alpha_t u(c_t).$$

An interpretation is that the decision maker commit a menu M which he can choose in the future and choose an actual consumption c from the menu after the discount factor is revealed.

¹This chapter is based on an early draft of a joint work "Comparative Impatience under Random Discounting" with Y. Higashi, K. Hyogo, and N. Takeoka, the final draft of which has been published as Higashi, Hyogo, Takeoka, and Tanaka [14] where the n -period model of this chapter is revised to be a simple two-period model.

We will characterize comparison of time preference for this class of preference in the subsequent discussions.

3.2 Comparative impatience

In this section, we introduce concepts used in the later discussion.

To formulate impatience, we first define order for menus which impatient decision maker tends to prefer. Such menus are defined as menus containing consumption streams which weights more on the present than the future. We define such consumption streams by the following way:

Definition 13 (Comparison of consumption streams) *For two consumption streams c and c' $c \succ c'$ if*

1. *there are s and t*

$$c_s \geq c'_s \text{ and } c'_t \geq c_t$$

2. *for any $t > s$*

$$c'_s \geq (>)c_s \implies c'_t \geq (>)c_t.$$

Geometrically, this condition means that c' crosses c from below at most once as a function of periods.

To define order for menus, we introduce the following auxiliary condition.

Definition 14 *A chain C is divided by A and B for any $a \in A \cap C$ and $b \in B \cap C$ if $b \succ a$, then*

$$b \in A \text{ and } a \in B.$$

This means that on the chain C , A weights more on the present consumption than B . Note that the role of A and that of B are not symmetric, i.e., “divided by A and B ” and “by B and A ” are different conditions.

We can define the relation of “weighting-more-on-the-present” as we can take such chain sufficiently large so that an alternative not contained in the chain is irrelevant for the desirability of the menu.

Definition 15 $M \triangleright M'$ if there is a chain C divided by M and M' such that for any $c \in M (c' \in M')$ if $c(c')$ is not in C , then it is dominated in $M(M')$ ².

To see the meaning of this condition more concretely, let

$$E(M) := \{z : \text{There is no } w \in M \text{ such that } w > z\}.$$

Since $E(M)$ means the set of all efficient elements of M , an element in $E(M)$ must be in C . Since this argument is also true for $E(M')$, $E(M) \cup E(M')$ is totally ordered and $E(M)$ is higher than $E(M')$ with respect to $\triangleright$.

Based on this order, we can define comparative impatience as the following way:

Definition 16 (Comparison of impatience) $\succsim_1$ is more impatience than $\succsim_2$ if for any $M \triangleright M'$

$$M \succsim_2 (\succ_2) M' \implies M \succsim_1 (\succ_1) M'$$

Note that this condition is a single crossing condition and thus has the direct implication to the utility maximization problem.

In components of random discounting model u and μ , time preference should corresponds to μ . Then, we characterize comparative impatience by the condition that for probability laws μ_i whose support is MLR-ordered, μ_1 MLR-dominates μ_2 .

Definition 17 (MLR ordered discount factors) For two discount factors (α_t) and (β_t) $\alpha \geq_{MLR} \beta$ if for any $s < t$

$$\alpha_t \beta_s \geq \alpha_s \beta_t.$$

We denote this relation by $\alpha \geq_{MLR} \beta$ and if $\{\alpha, \beta, \gamma, \dots\}$ satisfies that each pair is ordered by $\geq_{MLR}$, it is a MLR-ordered chain.

² c is dominated in M if there is $z \in M$ such that $z > c$

Definition 18 (MLR ordered probability laws) For two probability laws μ_1 and μ_2 $\mu_1 \geq_{MLR} \mu_2$ if for any $\alpha \geq_{MLR} \beta$

$$\mu_1(\alpha)\mu_2(\beta) \geq \mu_1(\beta)\mu_2(\alpha)$$

3.3 Main result

To prove the main result, we assume the following two conditions.

Assumption 5 (Totally ordered support) The support $\sigma(\mu_1) \cup \sigma(\mu_2)$ of μ_1 and μ_2 is a MLR-ordered chain.

Assumption 6 (Finite support) For each i there is N_i such that $|\sigma(\mu_i)| = N_i$.

Under these assumptions, we shall show the following result:

Theorem 6 For two preferences with random discounting representation the following two conditions are equivalent:

1. $\succsim_2$ is more impatient than $\succsim_1$
2. μ_1 MLR dominates μ_2 .

3.4 Fundamental lemma

In this section, we prove the following result used in the proof of the main theorem:

Lemma 5 for any preferences $\succsim_\alpha$ and $\succsim_\beta$ having representations $\max_{c \in \cdot} \alpha_t u(c_t)$ and $\max_{c \in \cdot} \beta_t u(c_t)$ respectively, the following conditions are equivalent.

1. $\succsim_\beta$ is more impatient than $\succsim_\alpha$.
2. $\alpha \geq_{MLR} \beta$.

Proof (2) $\implies$ (1): Let M and M' be menus with $M \triangleright M'$ and suppose $M \succ_\alpha (\lesssim_\alpha) M'$. Take

$$c^* \in \arg \max_{c \in M'} \sum_t \beta_t u(c_t) \text{ and } c^{**} \in \arg \max_{c \in M} \sum_t \alpha_t u(c_t).$$

By definition of $\triangleright$, $c^{**} \triangleright c^*$ or $c^{**} \triangleright c^*$. For the first case, c^{**} must be in M' and $c^* \in M$ and thus, $M' \lesssim_\alpha M$ and $M \lesssim_\beta M'$. For the second case, let $\bar{t} := \inf\{t : c_t^* < c_t^{**}\}$. Then, we have

$$\begin{aligned} & \sum_t \beta_t (u(c_t^*) - u(c_t^{**})) \\ &= \sum_{t \leq \bar{t}} \beta_t (u(c_t^*) - u(c_t^{**})) + \sum_{t > \bar{t}} \beta_t (u(c_t^*) - u(c_t^{**})) \\ &\geq \frac{\beta_{\bar{t}}}{\alpha_{\bar{t}}} \sum_{t \leq \bar{t}} \alpha_t (u(c_t^*) - u(c_t^{**})) + \frac{\beta_{\bar{t}}}{\alpha_{\bar{t}}} \sum_{t > \bar{t}} \alpha_t (u(c_t^*) - u(c_t^{**})) \\ &= \frac{\beta_{\bar{t}}}{\alpha_{\bar{t}}} \sum_t \alpha_t (u(c_t^*) - u(c_t^{**})) \end{aligned}$$

yielding the conclusion.

(1) $\implies$ (2): Suppose that for some $s < t$, $\alpha_t \beta_s < \alpha_s \beta_t$. Let c and c' be sequences of consumptions defined as

$$u(c_\tau) := \begin{cases} \delta_1 & \text{if } \tau < s \\ \alpha_t + \varepsilon & \text{if } \tau = s \\ 0 & \text{if } s < \tau < t \\ -\alpha_s & \text{if } \tau = t \\ -\delta_2 & \text{if } \tau > t \end{cases} \text{ and } u(c'_\tau) := \begin{cases} -\delta_1 & \text{if } \tau < s \\ -\alpha_t - \varepsilon & \text{if } \tau = s \\ 0 & \text{if } s < \tau < t \\ \alpha_s & \text{if } \tau = t \\ \delta_2 & \text{if } \tau > t \end{cases}$$

where $\delta_1 > 0$, $\delta_2 > 0$, and $\varepsilon > 0$ are specified later. It is easy to see that $c \triangleright c'$. Then, we have

$$\begin{aligned} \sum_\tau \alpha_\tau (u(c_\tau) - u(c'_\tau)) &= \delta_1 \sum_{\tau < s} \alpha_\tau + \alpha_s (\alpha_t + \varepsilon) - \alpha_t \alpha_s - \delta_2 \sum_{\tau > t} \alpha_\tau \\ &\quad - \left(-\delta_1 \sum_{\tau < s} \alpha_\tau - \alpha_s (\alpha_t + \varepsilon) + \alpha_t \alpha_s + \delta_2 \sum_{\tau > t} \alpha_\tau \right) \\ &= 2 \left(\delta_1 \sum_{\tau < s} \alpha_\tau + \alpha_s \varepsilon - \delta_2 \sum_{\tau > t} \alpha_\tau \right) \end{aligned}$$

and

$$\begin{aligned}
\sum_{\tau} \beta_{\tau}(u(c_{\tau}) - u(c'_{\tau})) &= \delta_1 \sum_{\tau < s} \beta_{\tau} + \beta_s(\alpha_t + \varepsilon) - \beta_t \alpha_s - \delta_2 \sum_{\tau > t} \beta_{\tau} \\
&\quad - \left(-\delta_1 \sum_{\tau < s} \beta_{\tau} - \beta_s(\alpha_t + \varepsilon) + \beta_t \alpha_s + \delta_2 \sum_{\tau > t} \beta_{\tau} \right) \\
&= 2\delta_1 \sum_{\tau < s} \beta_{\tau} + 2(\beta_s \alpha_t - \beta_t \alpha_s) + 2\beta_s \varepsilon - 2\delta_2 \sum_{\tau > t} \beta_{\tau} \\
&= 2 \left(\delta_1 \sum_{\tau < s} \beta_{\tau} + \beta_s \varepsilon - \delta_2 \sum_{\tau > t} \beta_{\tau} \right) + 2(\alpha_t \beta_s - \alpha_s \beta_t).
\end{aligned}$$

Since δ_1, δ_2 , and ε can be arbitrarily small, we have $\{c\} \succ_{\alpha} \{c'\}$ and $\{c'\} \succ_{\beta} \{c\}$, which contradicts the assumption that $\succ_{\beta}$ is more impatient than $\succ_{\alpha}$.

3.5 Proof of sufficiency

Suppose $\mu^1 \geq_{MLR} \mu^2$ and let $M \triangleright M'$. By the fundamental lemma, for any k

$$V(M, \alpha^{k+1}) - V(M', \alpha^{k+1}) \geq (>)0 \implies V(M, \alpha^k) - V(M', \alpha^k) \geq (>)0,$$

where $V(M, \alpha) := \max_{c \in M} \sum_{t=1}^T \alpha_t u(c_t)$. Let $\bar{k} := \inf\{k \mid V(M, \alpha^k) - V(M', \alpha^k) < 0\}$. Since if such $\bar{k}$ does not exist, the conclusion is immediate, we may concentrate on the case where such $\bar{k}$ exists.

Suppose $M \succsim_1 M'$. Thus,

$$\begin{aligned}
U^2(M) - U^2(M') &= \sum_k V(M, \alpha^k) \mu^2(\alpha^k) - \sum_k V(M', \alpha^k) \mu^2(\alpha^k) \\
&\geq \sum_{k: \mu^1(\alpha^k) \neq 0} \frac{\mu^2(\alpha^k)}{\mu^1(\alpha^k)} \mu^1(\alpha^k) (V(M, \alpha^k) - V(M', \alpha^k)) \\
&= \sum_{k < \bar{k}: \mu^1(\alpha^k) \neq 0} \frac{\mu^2(\alpha^k)}{\mu^1(\alpha^k)} \mu^1(\alpha^k) (V(M, \alpha^k) - V(M', \alpha^k)) \\
&\quad + \sum_{k \geq \bar{k}: \mu^1(\alpha^k) \neq 0} \frac{\mu^2(\alpha^k)}{\mu^1(\alpha^k)} \mu^1(\alpha^k) (V(M, \alpha^k) - V(M', \alpha^k)) \\
&\geq \frac{\mu^2(\alpha^{\bar{k}})}{\mu^1(\alpha^{\bar{k}})} \sum_{k: \mu^1(k) \neq 0} \mu^1(\alpha^k) (V(M, \alpha^k) - V(M', \alpha^k)) = \frac{\mu^2(\alpha^{\bar{k}})}{\mu^1(\alpha^{\bar{k}})} (U^1(M) - U^1(M')) \geq 0,
\end{aligned}$$

which implies $M \succsim_2 M'$. For the case $M \succ_1 M'$ an analogous argument yields $M \succ_2 M'$.

3.6 Proof of necessity

3.6.1 Definitions

Definition 19 A sequence $(v^k)_{k=1}^n$ of vectors $v^k \in \mathbb{R}^T$ is consistent with a sequence of $(\alpha^k)_{k=1}^n$ of vectors $\alpha^k \in \mathbb{R}^T$ if for each k

$$\alpha^k \cdot v^k > \alpha^k \cdot v^l \quad \text{for all } l \neq k.$$

This definition means that v^k is the unique maximizer of $\alpha^k \cdot v$ among $v^1, \dots, v^n$.

Definition 20 A sequence $(v^k)_{k=1}^n$ of vectors $v^k \in \mathbb{R}^T$ satisfies a single crossing condition if for any k and l with $l > k$ and $t > s$

$$v_s^l - v_s^k \geq (>)0 \implies v_t^l - v_t^k \geq (>)0.$$

In other words, for three sets defined by

$$T_0 := \{t \mid v_t^l - v_t^k = 0\},$$

$$T_- := \{t \mid v_t^l - v_t^k < 0\},$$

$$T_+ := \{t \mid v_t^l - v_t^k > 0\},$$

the single crossing condition requires $T_- \leq T_0 \leq T_+$ for each $l > k$ ³. For ease of notation, we use the same notation $\triangleright$ for vectors as in the consumption stream, $v^l \triangleright v^k$ for example.

Definition 21 A sequence of $(v^k)_{k=1}^n$ of vectors $v^k \in \mathbb{R}^T$ satisfies an upward weak single crossing condition if for any k and l with $l > k$ and $t > s$

$$v_s^l - v_s^k \geq 0 \implies v_t^l - v_t^k \geq 0.$$

The upward weak single crossing condition requires $T_- \leq T_0 \cup T_+$ for each $l > k$. As its name suggests, the upward weak single crossing condition is weaker than the single crossing condition. However, as verified later, some translation makes an upward weak single crossing sequence into a single crossing one.

³ $A \leq B$ means $a \leq b$ for all $a \in A$ and $b \in B$.

3.6.2 Proof

Seeking a contradiction, suppose there are $\alpha^j \geq_{MLR} \alpha^i$, $j > i$, such that

$$\mu^1(\alpha^j)\mu^2(\alpha^i) < \mu^1(\alpha^i)\mu^2(\alpha^j). \quad (3)$$

First consider the following optimization problem: For each k ,

$$\max_u \alpha^k \cdot u \quad \text{s.t.} \quad \|u\| \leq 1.$$

By the first order condition, as the unique candidate for the solution, we have

$$u^k := \frac{1}{\|\alpha^k\|} \alpha^k.$$

Since this satisfies the transversality condition $\lim_{t \rightarrow \infty} \|\alpha\|(1 - \sum_{s \leq t} (u_s^k)^2) = 0$ for all k , u^k is actually the unique optimal solution for k . We shall construct two menus M and M' based on $\{u^k\}$ such that $M \succ^1 M'$, $M' \succ^2 M$, and $M \triangleright M'$, which is a contradiction.

Lemma 6 $(u^k)_{k=1}^n$ is an upward weak single crossing sequence consistent with $(\alpha^k)_{k=1}^n$.

Proof Since for each k and l with $l > k$ we have

$$u_t^l - u_t^k = \begin{cases} \frac{\alpha_t^l}{\|\alpha^l\|} - \frac{\alpha_t^k}{\|\alpha^k\|} & \text{if } t \leq T(k) \\ \frac{\alpha_t^l}{\|\alpha^l\|} & \text{if } T(k) < t \leq T(l) \\ 0 & \text{if } T(l) < t \end{cases}.$$

Since for $s < t \leq T(k)$ α_t^k and α_t^l are both nonzero, the MLR condition

$$\alpha_t^l \alpha_s^k \geq \alpha_t^k \alpha_s^l$$

directly yields the conclusion because if $u_s^l \geq u_s^k$, it implies $u_t^l \geq u_t^k$ and if $u_s^l > u_s^k$, it implies $u_t^l > u_t^k$. For $t > T(k)$ u_t^l is nonnegative and u_t^k is zero. Consistency follows from the definition of $(u^k)_{k=1}^n$.

Lemma 7 *There is a number $\bar{\delta} > 0$ such that for any $\delta_1 \in (0, \bar{\delta})$ and $\delta_2 \in (0, \bar{\delta})$, $(\hat{u}^k)_{k=1}^n$ is an upward weak single crossing sequence consistent with $(\alpha^k)_{k=1}^n$ where $\hat{u}^k$ are defined by*

$$\hat{u}^k := \begin{cases} u^k & \text{if } k \neq i, j \\ u^i + \delta_1 e^1 & \text{if } k = i \\ u^j + \delta_2 e^1 & \text{if } k = j \end{cases}$$

where $e^1 := (1, 0, 0, \dots)$.

Proof We shall first verify $u_1^1 > \dots > u_1^n$. To prove this, it is sufficient to show that $u_1^k > u_1^{k+1}$ for all k . If not, we must have for some k

$$\frac{\alpha_1^{k+1}}{\alpha_1^k} \frac{\|\alpha^k\|}{\|\alpha^{k+1}\|} \geq 1.$$

Since (α^k) is ordered in $\geq_{MLR}$, this condition implies $u_t^{k+1} \geq u_t^k$ for all t . Then, we have $\alpha^k \cdot u^{k+1} \geq \alpha^k \cdot u^k$ which violates consistency. Then, $u_1^1 > \dots > u_1^n$, and thus, we can take $\delta > 0$ so that $\hat{u}_1^1 > \dots > \hat{u}_1^n$ and $(\hat{u}^k)$ is an upward single crossing sequence consistent with (α^k) .

Lemma 8 *There is a positive vector $\bar{\varepsilon}$ such that for any positive vector ε with $\varepsilon \leq \bar{\varepsilon}$, $(\hat{u}^k)_{k=1}^n$ satisfies single crossing condition and is consistent with $(\alpha^k)_{k=1}^n$ where $\hat{u}^k$ are defined by*

$$\hat{u}^k := u^k + k\varepsilon$$

Proof By (3), $n \geq 2$. Then, we can define h by

$$h_t = \begin{cases} \varepsilon_t^1 & \text{if } \hat{u}_t^k - \hat{u}_t^{k'} \text{ for some } k \neq k' \\ 1 & \text{otherwise} \end{cases}$$

where ε_t^1 is taken so that

$$0 < \varepsilon_t^1 < \frac{1}{n-1} \inf\{|\hat{u}_t^k - \hat{u}_t^{k'}| \mid \hat{u}_t^k \neq \hat{u}_t^{k'}, 1 \leq k, k' \leq n\}.$$

Define $\bar{\varepsilon}$ by

$$\bar{\varepsilon} := \varepsilon^2 h$$

where ε^2 is a positive number. Since $\hat{u}^k$ are consistent, we may take ε^2 sufficiently small so that for any $k \neq k'$

$$\alpha^k \cdot (\hat{u}^k - \hat{u}^{k'}) > |(k - k')\alpha \cdot \bar{\varepsilon}|.$$

Take any positive vector $\varepsilon \leq \bar{\varepsilon}$. Since

$$|(k - k')\alpha \cdot \bar{\varepsilon}| \geq |(k - k')\alpha \cdot \varepsilon|,$$

$\hat{u}^k$ are consistent. Suppose that for $l > k$ and s

$$\hat{u}_s^l - \hat{u}_s^k \geq 0.$$

By definition of $\bar{\varepsilon}$, this implies $\hat{u}_s^l - \hat{u}_s^k \geq 0$. Recalling that $(\hat{u}^k)$ is an upward weak single crossing sequence, we have $\hat{u}_t^l - \hat{u}_t^k \geq 0$ for all $t > s$, and thus, we have

$$\hat{u}_t^l - \hat{u}_t^k + (l - k)\varepsilon_t > 0$$

which is equivalent to $\hat{u}_t^l - \hat{u}_t^k > 0$.

Lemma 9 *There is a positive vector $\bar{\gamma}$ such that for any positive vector $\gamma \leq \bar{\gamma}$, $(\tilde{u}^k)_{k=1}^n$ satisfies*

$$k' > k \implies \tilde{u}^k \triangleright \hat{u}^{k'} \tag{4}$$

$$k > k' \implies \hat{u}^{k'} \triangleright \tilde{u}^k$$

where $\tilde{u}^k$ is defined by

$$\tilde{u}^k := \begin{cases} \hat{u}^k + \gamma & \text{if } k \leq i \\ \hat{u}^k & \text{if } i < k < j \\ \hat{u}^k + \gamma & \text{if } k \geq j \end{cases}.$$

Moreover,

$$(\tilde{u}^1, \dots, \tilde{u}^i, \hat{u}^{i+1}, \dots, \hat{u}^n) \text{ and } (\hat{u}^1, \dots, \hat{u}^{j-1}, \tilde{u}^j, \dots, \tilde{u}^n)$$

are consistent with $(\alpha^k)_{k=1}^n$.

Proof By (3), $n \geq 2$. Then, we can define g by

$$g_t := \begin{cases} \gamma_t^1 & \text{if } \hat{u}_t - \hat{u}_t^{k'} \neq 0 \\ 1 & \text{otherwise} \end{cases}$$

where γ_t^1 is taken so that

$$0 < \gamma_t^1 < \inf\{|\hat{u}_t^k - \hat{u}_t^{k'}| \mid \hat{u}_t^k \neq \hat{u}_t^{k'}, 1 \leq k, k' \leq n\}.$$

Define $\bar{\gamma}$ by

$$\bar{\gamma} := \gamma^2 g$$

where γ^2 is a positive number. Since $\hat{u}^k$ are consistent, we may take γ^2 sufficiently small so that for any $k \neq k'$

$$\alpha^k \cdot (\hat{u}^k - \hat{u}^{k'}) > |(k - k')\alpha \cdot \bar{\gamma}|.$$

Take any positive vector $\gamma \leq \bar{\gamma}$. We have for any $l > k$

$$\tilde{u}^l - \hat{u}^k = \begin{cases} \hat{u}^l - \hat{u}^k + \gamma & \text{if } l \geq j > k > i \\ \hat{u}^l - \hat{u}^k - \gamma & \text{if } j > l > i \geq k \\ \hat{u}^l - \hat{u}^k & \text{otherwise} \end{cases} \quad (5)$$

and

$$\hat{u}^l - \tilde{u}^k = \begin{cases} \hat{u}^l - \hat{u}^k - \gamma & \text{if } j > l > i \geq k \\ \hat{u}^l - \hat{u}^k + \gamma & \text{if } l \geq j > i > k \\ \hat{u}^l - \hat{u}^k & \text{otherwise} \end{cases} \quad (6)$$

For any $k' > k$, by (5) and the definition of $\bar{\gamma}$, single crossing property of $(\hat{u}^k)$ implies $\hat{u}^k \triangleright \tilde{u}^{k'}$. For any $k > k'$, by (6) and the definition of $\bar{\gamma}$, single crossing property of $(\hat{u}^k)$ also implies $\tilde{u}^{k'} \triangleright \hat{u}^k$. By the definition of $\bar{\gamma}$, we have

$$\alpha^k \cdot (\hat{u}^k - \hat{u}^{k'}) > |(k - k')\alpha \cdot \bar{\gamma}| \geq |(k - k')\alpha \cdot \gamma|$$

Thus, $(\tilde{u}^1, \dots, \tilde{u}^i, \hat{u}^{i+1}, \dots, \hat{u}^n)$ and $(\hat{u}^1, \dots, \hat{u}^{j-1}, \tilde{u}^j, \dots, \tilde{u}^n)$ are consistent.

Take sequences of consumption streams $\tilde{c}^k$, $k = 1, \dots, i, j, \dots, n$, and $\hat{c}^k$, $k = 1, \dots, n$ such that $(u(\tilde{c}_t^k))_{t=1}^T = \tilde{u}^k$ and $(u(\hat{c}_t^k))_{t=1}^T = \hat{u}^k$. Define menus M and M' by

$$M := \{\tilde{c}^1, \dots, \tilde{c}^i\} \cup \{\hat{c}^1, \dots, \hat{c}^n\}, \quad M' := \{\hat{c}^1, \dots, \hat{c}^n\} \cup \{\tilde{c}^j, \dots, \tilde{c}^n\}.$$

Then,

$$u(E(M)) = \{\tilde{u}^1, \dots, \tilde{u}^i, \hat{u}^{i+1}, \dots, \hat{u}^n\}, \quad u(E(M')) = \{\hat{u}^1, \dots, \hat{u}^{j-1}, \tilde{u}^j, \dots, \tilde{u}^n\}$$

and thus,

$$\begin{aligned} u(E(M) \setminus M') &= u(E(M)) \setminus u(M') = \{\tilde{u}^1, \dots, \tilde{u}^i\} \\ u(E(M') \setminus M) &= u(E(M')) \setminus u(M) = \{\tilde{u}^j, \dots, \tilde{u}^n\} \\ u(E(M') \cap E(M)) &= u(E(M')) \cap u(E(M)) = \{\hat{u}^{i+1}, \dots, \hat{u}^{j-1}\}. \end{aligned}$$

By these equations, we have:

Lemma 10 $M \triangleright M'$.

Proof Take a chain C divided by M and M' as

$$C := \{\tilde{c}^1, \dots, \tilde{c}^i, \hat{c}^{i+1}, \dots, \hat{c}^{j-1}, \tilde{c}^j, \dots, \tilde{c}^n\}$$

Finally, we shall verify $M \succ^1 M'$ and $M' \succ^2 M$ for some δ, ε , and γ . By the later part of Lemma 9, we have

$$\begin{aligned} U^1(M) &= \sum_k \left(\max_{c \in M} \sum_t \alpha_t^k u(c_t) \right) \mu^1(\alpha^k) = \sum_{k \leq i} \alpha^k \cdot \tilde{u}^k \mu^1(\alpha^k) + \sum_{k > i} \alpha^k \cdot \hat{u}^k \mu^1(\alpha^k) \\ &= \sum_{k < i} \alpha^k \cdot (u^k + k\varepsilon + \gamma) \mu^1(\alpha^k) \\ &\quad + \sum_{k > i} \alpha^k \cdot (u^k + k\varepsilon) \mu^1(\alpha^k) + \alpha^i \cdot (u^i + i\varepsilon + \gamma + \delta_1 e^1) \mu^1(\alpha^i) \\ &= \sum_k \alpha^k \cdot u^k \mu^1(\alpha^k) + \alpha_1^i \delta_1 \mu^1(\alpha^i) + R(\varepsilon, \gamma), \end{aligned}$$

where

$$R(\varepsilon, \gamma) := \sum_{k \leq i} \alpha^k \cdot (k\varepsilon + \gamma) \mu^1(\alpha^k) + \sum_{k > i} \alpha^k \cdot (k\varepsilon) \mu^1(\alpha^k).$$

Similarly,

$$\begin{aligned} U^1(M') &= \sum_{k < j} \alpha^k \cdot \hat{u}^k \mu^1(\alpha^k) + \sum_{k \geq j} \alpha^k \cdot \tilde{u}^k \mu^1(\alpha^k) \\ &= \sum_{k < j} \alpha^k \cdot (u^k + k\varepsilon) \mu^1(\alpha^k) \\ &\quad + \sum_{k > j} \alpha^k \cdot (u^k + k\varepsilon + \gamma) \mu^1(\alpha^k) + \alpha^j \cdot (u^j + j\varepsilon + \gamma + \delta_2 e^1) \mu^1(\alpha^j) \\ &= \sum_k \alpha^k \cdot u^k \mu^1(\alpha^k) + \alpha_1^j \delta_2 \mu^1(\alpha^j) + R'(\varepsilon, \gamma), \end{aligned}$$

where

$$R'(\varepsilon, \gamma) := \sum_{k < j} \alpha^k \cdot (k\varepsilon) \mu^1(\alpha^k) + \sum_{k \geq j} \alpha^k \cdot (k\varepsilon + \gamma) \mu^1(\alpha^k).$$

Then,

$$\begin{aligned} U^1(M) - U^1(M') &= \sum_k \alpha^k \cdot u^k \mu^1(\alpha^k) + \alpha_1^i \delta_1 \mu^1(\alpha^i) + R(\varepsilon, \gamma) \\ &\quad - \left(\sum_k \alpha^k \cdot u^k \mu^1(\alpha^k) + \alpha_1^j \delta_2 \mu^1(\alpha^j) + R'(\varepsilon, \gamma) \right) \\ &= \alpha_1^i \delta_1 \mu^1(\alpha^i) - \alpha_1^j \delta_2 \mu^1(\alpha^j) + R(\varepsilon, \gamma) - R'(\varepsilon, \gamma), \end{aligned} \tag{7}$$

and similarly,

$$U^2(M) - U^2(M') = \alpha_1^i \delta_1 \mu^2(\alpha^i) - \alpha_1^j \delta_2 \mu^2(\alpha^j) + R(\varepsilon, \gamma) - R'(\varepsilon, \gamma). \tag{8}$$

Recall assumption (3), we have $\mu^1(\alpha^i) \neq 0$ and $\mu^2(\alpha^j) \neq 0$. Since

$$R(\varepsilon, \gamma) \rightarrow 0 \text{ and } R'(\varepsilon, \gamma) \rightarrow 0$$

as $\sup_t \varepsilon_t$ and $\sup_t \gamma_t$ approach to zero, we can choose ε and γ sufficiently small so that (7) is positive and (8) is negative. Indeed, for all sufficiently small ε and γ , if $\mu^1(\alpha^j) = 0$, (7)

is positive, and if $\mu^2(\alpha^i) = 0$, (8) is negative. If $\mu^1(\alpha^j) \neq 0$ and $\mu^2(\alpha^i) \neq 0$, we can find small δ_1 and δ_2 satisfying

$$\frac{\alpha_1^j \mu^2(\alpha^j)}{\alpha_1^i \mu^2(\alpha^i)} > \frac{\delta_1}{\delta_2} > \frac{\alpha_1^j \mu^1(\alpha^j)}{\alpha_1^i \mu^1(\alpha^i)}.$$

For example, letting $\delta_1 = \frac{\alpha_1^j \mu^1(\alpha^j)}{N} + \frac{1}{N'}$ and $\delta_2 = \frac{\alpha_1^i \mu^1(\alpha^i)}{N}$ for some large $N' > N > 0$ and such ε and γ yields $M \succ^1 M'$ and $M' \succ^2 M$.

3.7 Application to Goldman model

In this section, we apply the result in the previous section to a liquidity constraint model introduced by Goldman [10].

3.7.1 Model

There are three periods $t = 0, 1, 2$.

At period 0, the decision maker chooses his money M and bond B yielding 1 per unit. Since the price of the bond is p , the constraint is $M + pB = W$.

At period 1, he chooses consumption stream (c_1, c_2) from the budget set generated by the decision at period 0. If consumption at the first period satisfies $c_1 \leq M$, $M - c_1$ can be carried to period 2. If consumption at the first period exceeds money $c_1 > M$, the decision maker must sell bonds to compensate $c_1 - M$ with the proportional transaction cost $(1 - t)$. Then, in this case, the consumption stream satisfies

$$c_1 = M + ptB_1$$

$$c_2 = B - B_1$$

where B_1 is the amount of bonds sold at period 1. Thus, the budget set generated by choosing M at period 0 is

$$D(M) = \left\{ (c_1, c_2) : \begin{array}{l} c_1 + c_2 \leq M + B \text{ if } c_1 \leq M \\ c_1 + ptc_2 \leq M + ptB \text{ if } c_1 > M \\ B = \frac{W-M}{p} \end{array} \right\}$$

Then, the decision maker maximizes the following utility function under uncertainty about his discount factor:

$$U(M) = \sum_{\alpha} \max_{c \in D(M)} u(c_1) + \alpha u(c_2) \mu(\alpha)$$

Note that since the decision consists of two stages, it may be complicated to obtain the solution explicitly.

3.7.2 Monotone money demand

Proposition 3 *Let μ_1 MLR-dominates μ_2 . Then,*

$$\arg \max_{M \in [0, W]} U_1(M) \geq_s \arg \max_{M \in [0, W]} U_2(M)$$

Specification of the random discounting model by Goldman's model can not make the condition for monotone comparative statics a weaker condition like first order stochastic dominance order. This is implied by the following result.

Proposition 4 *For CRRA von Neumann-Morgenstern function $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ where $\sigma \neq 1$ let*

$$U(M) = \sum_{\alpha} \left(\max_{c \in D(M)} u(c_1) + u(c_2) \right) \mu(\alpha)$$

If the ratio M^/B^* of optimal money over bond is not 1. then, there is ν FOSD-dominates μ and its ratio M^{**}/B^{**} is larger than M^*/B^* .*

Proof See [14].

References

- [1] Athey, S. Monotone comparative statics under uncertainty. *Quarterly Journal of Economics*, Vol. 117, 187-223, 2002.
- [2] Benoît, J. P., and E. A. Ok. Delay aversion. *Theoretical Economics*, Vol. 2, 71-113, 2007.

- [3] Camargo, B., and B. Lester. Trading dynamics in decentralized markets with adverse selection. *Journal of Economic Theory*, Vol. 153, 534-568, 2014.
- [4] Chatterjee, S., D. Corbae, M. Nakajima, and J.-V. Rios-Rull. A quantitative theory of unsecured consumer credit with risk of default. *Econometrica*, Vol. 75, 1525-1589, 2007.
- [5] Dekel, E., B. Lipman, and A. Rustichini. Representing preferences with a unique subjective state space. *Econometrica*, Vol. 69, 891-934, 2001.
- [6] Dekel, E., B. Lipman, and A. Rustichini. Temptation-Driven Preferences. *Review of Economic Studies*, Vol. 76, 937-971, 2009.
- [7] Diamond, D. W., and Dybvig, P. H. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*. Vol. 91, 401-419, 1983
- [8] Dillenberger, D., J. S. Lleras, P. Sadowski, and N. Takeoka. A theory of subjective learning. *Journal of Economic Theory*, Vol. 153, 287-312, 2014.
- [9] Fishburn, P. C., and R. B. Porter. Optimal portfolios with one safe and one risky asset: Effects of changes in rate of return and risk. *Management Science*, Vol. 22, 1064-1073, 1976.
- [10] Goldman, S. M. Flexibility and the demand for money. *Journal of Economic Theory*, Vol. 9, 203-222, 1974.
- [11] Gul, F., and W. Pesendorfer. Self-control and the theory of consumption. *Econometrica*, Vol. 72, 119-158, 2004.
- [12] Higashi, Y., K. Hyogo, and N. Takeoka. Subjective random discounting and intertemporal choice. *Journal of Economic Theory*, Vol. 144, 1015-1053, 2009.
- [13] Higashi, Y., K. Hyogo, and N. Takeoka. Stochastic endogenous time preference. *Journal of Mathematical Economics*, Vol. 51, 77-92, 2014.

- [14] Higashi, Y., Hyogo, K., Takeoka, N., & Tanaka, H. (2017). Comparative impatience under random discounting. *Economic Theory*, 63(3), 621-651.
- [15] Horowitz, J. K. Comparative impatience. *Economics Letters*, Vol. 38, 25-29, 1992.
- [16] Karlin, S., and H. Rubin. The theory of decision procedures for distributions with monotone likelihood ratio. *Annals of Mathematical Statistics*, Vol. 27, 272-299, 1956.
- [17] Karni, E., and I. Zilcha. Saving behavior in stationary equilibrium with random discounting. *Economic Theory*, Vol. 15, 551-564, 2000.
- [18] Kreps, D. M. A representation theorem for preference for flexibility. *Econometrica*, Vol. 47, 565-578, 1979.
- [19] Kreps, D. M. Static choice and unforeseen contingencies. In Dasgupta, P. and Gale, D. and Hart, O. and Maskin, E., editors, *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn*, 259-281. MIT Press, Cambridge, MA, 1992.
- [20] Krishna, V., and P. Sadowski. Dynamic preference for flexibility. *Econometrica*, Vol. 82, 655-703, 2014.
- [21] Krusell, P., and A. Smith. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, Vol. 106, 867-896, 1998.
- [22] Landsberger, M., and I. Meilijson. Demand for risky financial assets: A portfolio analysis. *Journal of Economic Theory*, Vol. 50, 204-213, 1990.
- [23] Lin, P. Banking, incentive constraints, and demand deposit contracts with nonlinear returns. *Economic Theory*, Vol. 8, 27-39, 1996.
- [24] Rustichini, A. Preference for flexibility in infinite horizon problems. *Economic Theory*, Vol. 20, 677-702, 2002.
- [25] Shaked, M., and J. G. Shanthikumar. *Stochastic Orders*, Springer, 2007.

4 Deficiency in the Maximin Expected Utility Models

4.1 Introduction

It is known as Blackwell's theorem that a signal is better than another signal if the later is reproduced by garbling the former. This result enable the comparison of information structures for expected utility models and is applied in various economic situations. However, recent developments of non-expected utility models raises the question whether Blackwell's theorem is valid in these models. Among a line of research investigating this question, Çelen [3] considers it in the maximin expected utility model whose foundation is given by Gilboa and Schmeidler [5] and derives an analogue of Blackwell's theorem by assuming that the probability of a signal value conditional on a state is uniquely determined. However, since a maximin expected utility model may not be consistent with the Bayes' rule, this assumption is too stringent in the setting where the decision is made with probability of signal conditional on state. For instance, Heyen and Wiesenfarth[7] gives an example showing Çelen's definition of value of information is not formulated as a dynamically consistent way, that is, an ex-ante optimal strategy may not be constructed by backward induction.

In this chapter, we formulate the optimization problem so as to be dynamically consistent. Thus, the decision maker does not turn over his ex ante plan of actions after he receives a signal. Our main result is an extension of the Blackwell theorem for maximin expected utility by more general criterion by Le Cam[8] where he defines the notion called deficiency so as to compare information structures not necessarily comparable in the sense of Blackwell.

To motivate our result, consider the following Ellsberg-like problem: There is an urn containing three small urns, A,B, and C containing balls as follows:

A : 10 red balls and 20 balls with green or blue color

B : 10 green balls and 20 balls with red or blue color

C : 10 blue balls and 20 balls with red or blue color

and the decision maker chooses whether he draws a ball or not. The payoff of the draw is \$ 100 if a red ball is drawn while \$ -20 if a ball with the other colors is drawn after one of the small urns is chosen at random. For example, if the decision maker evaluates probability of color of a chosen ball as follows:

$$\begin{aligned}\mathcal{P}_A &= \left\{ \left(\frac{1}{3}, p, \frac{2}{3} - p \right) : p \in [0, \frac{2}{3}] \right\} \\ \mathcal{P}_B &= \left\{ \left(\frac{2}{3} - p, \frac{1}{3}, p \right) : p \in [0, \frac{2}{3}] \right\} \\ \mathcal{P}_C &= \left\{ \left(p, \frac{1}{3}, \frac{2}{3} - p \right) : p \in [0, \frac{2}{3}] \right\}\end{aligned}$$

where the subscript on a set $\mathcal{P}$ of probability distributions represents the small urn a chosen ball is picked from. Then, in the ex ante stage, the set $\mathcal{P}$ of probability distributions is:

$$\mathcal{P} = \frac{1}{3}\mathcal{P}_A + \frac{1}{3}\mathcal{P}_B + \frac{1}{3}\mathcal{P}_C$$

For example, probability of small urn A conditional on red ball is

$$\frac{\frac{1}{9}}{\frac{1}{9} + \frac{1}{3}p_B + \frac{1}{3}p_C}$$

where $\frac{1}{3}p_B$ is the probability of red ball from small urn B and $\frac{1}{3}p_C$ is from C. Since p_B and p_C can take values from 0 to $\frac{2}{3}$, probability of urn A conditional on red ball is not uniquely determined. Note that this situation cannot be described by Çelen's argument.

In this setting, thinking of A, B, and C as values of a signal and color of balls as states, knowing the small urn a ball is taken from seems to be valuable for the decision maker because he can take a strategy such as “draw” if the small urn turns out to be A and “not draw” otherwise.

The organization of this chapter is as follows: The next section formulates the model and introduce deficiency due to Le Cam [8]. The third section states the main theorem and proves this. The last section applies the main result to the motivating example above and discuss other applications.

4.2 Preliminaries

Let $\Omega := \{1, 2, \dots, m\}$ be a set of states and $S := \{1, 2, \dots, n\}$ a set of signals where generic elements of these are $\omega \in \Omega$ and $s \in S$, respectively. A is a finite set of actions and a strategy is a function from S to the simplex $\Delta(A)$ of A . Then, we denote by $f(s, a)$ the probability for action $a \in A$ chosen by strategy f if signal $s \in S$ received. The decision maker has a state-dependent utility $u : A \times \Omega \rightarrow \mathbb{R}$ and has a set $\mathcal{P}$ of probability distributions P on $\Omega \times S$ which we call information structure. An interpretation is that the realization of states and signals obeys some probability distribution, but there is ambiguity on which probability distribution is true. We assume that $\mathcal{P}$ is closed, convex, and for each $s \in S$

$$P(\Omega \times \{s\}) = Q(\Omega \times \{s\}) \quad (9)$$

for any P and Q in $\mathcal{P}$. This assumption requires that the marginal distribution of signals is not ambiguous. Since Ω and S are finite, we can identify a probability distribution P on $\Omega \times S$ as an $|\Omega| \times |S|$ matrix whose ωs element $P_{\omega s}$ is defined by

$$P_{\omega s} = P(\{\omega\} \times \{s\})$$

For simplicity of notation, for probability distribution P on $\Omega \times S$ let

$$P_s := \sum_{\omega \in \Omega} P_{\omega s},$$

$$P_{\omega|s} := \frac{P_{\omega s}}{P_s},$$

and for a set $\mathcal{P}$ of probability distributions let

$$\mathcal{P}_{post} := \left\{ \left(\frac{P_{\omega s}}{P_s} \right)_{s \in S, \omega \in \Omega} : P \in \mathcal{P} \right\}$$

Then, the utility maximization problem we consider is:

$$\min_{P \in \mathcal{P}} \sum_{s \in S} P_s \max_{f(s, \cdot) \in \Delta(A)} \min_{P_{\omega|s} \in \mathcal{P}_{post}} \sum_{\omega \in \Omega} P_{\omega|s} \sum_{a \in A} f(a, s) u(a, \omega).$$

Note that by the assumption (9) we posed,

$$\begin{aligned}
& \min_{P \in \mathcal{P}} \sum_{s \in S} p_s \max_{f(s, \cdot) \in \Delta(A)} \min_{P_{\omega|s} \in \mathcal{P}_{Post}} \sum_{\omega \in \Omega} P_{\omega|s} \sum_{a \in A} f(a, s) u(a, \omega) \\
&= \sum_{s \in S} P_s \max_{f(s, \cdot) \in \Delta(A)} \min_{P_{\omega|s} \in \mathcal{P}_{Post}} \sum_{\omega \in \Omega} P_{\omega|s} \sum_{a \in A} f(a, s) u(a, \omega) \\
&= \max_{f \in \Delta(A)^{|S|}} \sum_{s \in S} P_s \min_{P_{\omega|s} \in \mathcal{P}_{Post}} \sum_{\omega \in \Omega} P_{\omega|s} \sum_{a \in A} f(a, s) u(a, \omega) \\
&= \max_{f \in \Delta(A)^{|S|}} \min_{P \in \mathcal{P}} \sum_{\omega, s, a} P_{\omega, s} f(a, s) u(a, \omega)
\end{aligned}$$

Thus, an ex-ante optimal strategy $f^* \in \Delta(X)^{|S|}$ can be constructed from interim optimal strategies $f^*(s, a)$, and vice versa. Hence, in our formulation, dynamic inconsistency cannot occur.

For an $n_1 \times n_2$ matrix X , denote $\|X\|_1 := \sum_{1 \leq i \leq n_1, 1 \leq j \leq n_2} |X_{ij}|$ and $\|X\|_\infty := \sup_{1 \leq i \leq n_1, 1 \leq j \leq n_2} |X_{ij}|$. For two $n_1 \times n_2$ matrices X and Y denote $X \cdot Y = \sum_{1 \leq i \leq n_1, 1 \leq j \leq n_2} X_{ij} Y_{ij}$. Let $\mathcal{M}$ be the set of all Markov matrices which are nonnegative matrices $M \in \mathbb{R}^{|S| \times |S|}$ with $\sum_{s' \in S} M_{ss'} = 1$ for each $s \in S$. For two sets of probability measures $\mathcal{P}$ and $\mathcal{Q}$ let $\delta(\mathcal{P}, \mathcal{Q})$ be the smallest number ϵ satisfying the following condition: For each $P \in \mathcal{P}$ there are $Q \in \mathcal{Q}$ and $M \in \mathcal{M}$ such that

$$\|PM - Q\|_1 \leq \epsilon.$$

In other words,

$$\delta(\mathcal{P}, \mathcal{Q}) = \sup_{P \in \mathcal{P}} \inf_{Q \in \mathcal{Q}, M \in \mathcal{M}} \|PM - Q\|_1.$$

Under information structure PM , signal s is interpreted as signal s' with probability $M_{ss'}$. Then, M is a randomization device garbling information structure P . δ measures how close each P approaches to $\mathcal{Q}$ with the randomization device M . However, note that δ is not a metric since it is not symmetric.

Example 8 Consider the following two matrices:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}$$

$$Q = \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

P is an information structure such that the first signal reveals the true state 1 but the second one is uninformative. On the other hand, Q is an information structure such that the first signal tells nothing about the true state but the second one reveals the true state 2. With this interpretation, they can not be comparable with respect to Blackwell's criterion, i.e. $\delta(\{P\}, \{Q\}) \neq 0$. By computation, we actually have $\delta(\{P\}, \{Q\}) = \frac{3}{4}$. To verify this, consider the following minimization problem:

$$\min_{M: \text{Markov}} \|PM - Q\|_1 = \left| \frac{1}{2}M_{11} + \frac{1}{4}M_{21} - \frac{1}{4} \right| + \left| \frac{1}{2}M_{12} + \frac{1}{4}M_{22} \right| + \left| \frac{1}{4}M_{21} - \frac{1}{4} \right| + \left| \frac{1}{4}M_{22} - \frac{1}{2} \right|$$

s.t. $M_{11} + M_{12} = 1$,

$M_{21} + M_{22} = 1$.

By using $M_{12} = 1 - M_{11}$ and $M_{22} = 1 - M_{21}$, we have

$$\left| \frac{1}{2}M_{11} + \frac{1}{4}(M_{21} - 1) \right| + \left| \frac{1}{2}(1 - M_{11}) + \frac{1}{4}(1 - M_{21}) \right| + \left| \frac{1}{4}M_{21} - \frac{1}{4} \right| + \left| \frac{1}{4}M_{21} \right|.$$

For the sum of the first two terms, we have

$$\begin{aligned} & \left| \frac{1}{2}M_{11} + \frac{1}{4}(M_{21} - 1) \right| + \left| \frac{1}{2}(1 - M_{11}) + \frac{1}{4}(1 - M_{21}) \right| \\ & \geq \left| \frac{1}{2}M_{11} + \frac{1}{4}M_{21} - \frac{1}{4} + \frac{1}{2}(1 - M_{11}) + \frac{1}{4}(1 - M_{21}) \right| = \frac{1}{2} \end{aligned} \quad (10)$$

and $M_{11} = M_{21} = 1$ yields this lower bound. Since $0 \leq M_{21} \leq 1$, for the sum of the last two terms, we have

$$\left| \frac{1}{4}M_{21} - \frac{1}{4} \right| + \left| \frac{1}{4}M_{21} \right| = \frac{1}{4}. \quad (11)$$

Thus, by (10) and (11), the value of the minimization problem is $\frac{3}{4}$.

Definition 22 (Deficiency for multiple priors) Let $\mathcal{P}$ and $\mathcal{Q}$ be sets of probabilities on $\Omega \times S$. $\mathcal{P}$ is ϵ -deficient relative to $\mathcal{Q}$ if for any von Neumann-Morgenstern function u on $A \times \Omega$

$$\begin{aligned} & \sum_{s \in S} Q_s \max_{f \in \Delta(A)} \min_{Q_{\cdot|s} \in \mathcal{Q}_{post}} \sum_{\omega \in \Omega} Q_{\omega|s} \sum_{a \in A} f(a, s) u(a, \omega) \\ & \leq \sum_{s \in S} P_s \max_{f \in \Delta(A)} \min_{P_{\cdot|s} \in \mathcal{P}_{post}} \sum_{\omega \in \Omega} P_{\omega|s} \sum_{a \in A} f(a, s) u(a, \omega) + \epsilon \|u\|_{\infty} \end{aligned}$$

The notion of deficiency is introduced by Le Cam [8]. Deficiency measures how much information is lost by replacing $\mathcal{Q}$ by $\mathcal{P}$. It generalizes comparison of informativeness by Blackwell in the sense that 0-deficiency is sufficiency. Le Cam's celebrated results are the equivalence between utility criterion as above and the approximation by Markov matrices⁴.

In literature on informativeness, the prior is defined to be unique, that is, marginals $P_{\omega} := \sum_s P_{\omega s}$ and $Q_{\omega} := \sum_s Q_{\omega s}$ must be same when comparison is possible. For multiple priors, if $\mathcal{P}$ is ϵ -deficient relative to $\mathcal{Q}$, then the marginals must satisfy $\{(\sum_s P_{\omega s}) : P \in \mathcal{P}\} \subset \{(\sum_s Q_{\omega s}) : Q \in \mathcal{Q}\}$, but the converse is not necessarily true. However, since concepts such as deficiency or informativeness are motivated by the question “Which information structure is finer than the other”, it is natural to assume that the set of marginals are same. Then, while the result in the next section can be derived without assuming this, it is normatively plausible to assume the equivalence of the sets of marginals.

4.3 Main result

In this section, we shall show the following result:

Theorem 7 For any two sets of information structures $\mathcal{P}$ and $\mathcal{Q}$ the following two statements are equivalent: For any $\epsilon \geq 0$

1. $\delta(\mathcal{P}, \mathcal{Q}) \leq \epsilon$.

⁴While we assume that S, Ω , and X are finite to avoid technical difficulty, his argument does not assume finiteness of these sets.

2. $\mathcal{P}$ is ϵ -deficient relative to $\mathcal{Q}$

Proof (1 $\implies$ 2) Take any $P \in \mathcal{P}$. Suppose that there is $Q \in \mathcal{Q}$ and a Markov matrix M with $\|PM - Q\|_1 \leq \epsilon$. Let

$$U := (u(a, \omega))_{a \in A, \omega \in \Omega}$$

and let

$$F := (f(s, a))_{s \in S, a \in A}$$

Then,

$$P \cdot FU = \sum_{\omega \in \Omega, s \in S} P_{\omega s} \sum_a f(s, a) u(a, \omega).$$

Thus,

$$\begin{aligned} \epsilon \|U\|_\infty &\geq \|PM - Q\|_1 \|U\|_\infty \\ &\geq (Q - PM) \cdot FU \\ &= Q \cdot FU - PM \cdot FU \\ &= Q \cdot FU - P \cdot M^\top FU \geq Q \cdot FU - \max_F P \cdot FU. \end{aligned}$$

where the second inequality is due to the Holder inequality and the last inequality is due to $M^\top F$ is also a strategy.

(2 $\implies$ 1) Suppose that there is $P \in \mathcal{P}$ such that for any Markov matrix M and any $Q \in \mathcal{Q}$

$$\|PM - Q\|_1 \geq \epsilon$$

In other words,

$$(\mathcal{Q} - p\mathcal{M}) \cap \epsilon B_1 = \emptyset$$

where $\mathcal{M}$ is the set of all Markov matrices and

$$B_1 = \{x \in \mathbb{R}^{|\Omega| \times |S'|} : \sum_{\omega, s'} |x_{\omega s'}| \leq 1\}$$

Since $\mathcal{Q} - p\mathcal{M}$ and B_1 are nonempty convex sets, by the hyperplane separating theorem, there exists a matrix $X \in \mathbb{R}^{|\Omega| \times |S|}$ such that

$$(Q - PM) \cdot X \geq \epsilon b \cdot X$$

for any matrix $p \in \mathcal{P}$, any Markov matrix M , and any matrix $b \in B_1$. Let $\bar{\omega}\bar{s}'$ be an index such that

$$|X_{\bar{\omega}\bar{s}'}| = \sup_{\omega, s'} |X_{\omega, s'}|$$

and take $b \in B_1$ so that

$$b_{\omega s'} = \begin{cases} \text{sign of } X_{\bar{\omega}, \bar{s}'} & \text{if } \omega = \bar{\omega}, s' = \bar{s}' \\ 0 & \text{otherwise} \end{cases}.$$

Then, we have

$$Q \cdot X - PM \cdot X \geq \epsilon \|X\|_\infty \quad \text{for all } Q \in \mathcal{Q} \text{ and Markov matrices } M.$$

Define a set of actions A and a von-Neumann Morgenstern function u so that

$$A = S$$

$$u(s, \omega) = X_{\omega s}$$

Under the information structure $\mathcal{Q}$,

$$f(a, s) = s \text{ for each } s$$

is a feasible strategy which gives the following payoff:

$$Q \cdot X \quad \text{for any } Q \in \mathcal{Q}.$$

On the other hand, under the information structure $\mathcal{P}$, a strategy is to choose a probability distribution on S for each signal s in S , which is equivalent to choose an $S \times S$ Markov matrix. Namely, the feasible payoff is of the form:

$$PM \cdot X \quad \text{for some Markov matrix } M.$$

Hence, we have

$$\min_{Q \in \mathcal{Q}} \sum_{\omega, s} Q_{\omega, s} u(s, \omega) \geq \max_{f \in \Delta(A)^{|S|}} \sum_{\omega, s} P_{\omega, s} f(s, a) u(a, \omega) + \|U\|_\infty$$

which implies that condition 2 cannot hold.

4.4 Discussion

4.4.1 Three Ellsberg urns

As mentioned in the introduction, the main result obtained above has implication to the three Ellsberg urns decision problem.

Example 9 Let $\Omega = \{R, G, B\}$ and $S = \{A, B, C\}$. A risk neutral decision maker is asked whether to pick a ball from an urn or not. The urn consists of three small urns A, B , and C and he cannot check which small urn the ball is picked from. The set of actions consists of a_1 (draw) or a_0 (not draw). If he chooses a_1 , he gets payoff;

$$u(a_1, \omega) := \begin{cases} 100 & \text{if } \omega = R \\ -20 & \text{otherwise} \end{cases}$$

and if a_0 ,

$$u(a_0, \omega) := 0 \text{ for all } \omega.$$

Suppose he suspects that small urn A does not contain a green ball or not contain a blue ball and evaluates the other two small urns symmetrically, namely.

$$\begin{aligned} \mathcal{P}_A &:= \left\{ \left(\frac{1}{3}, p, \frac{2}{3} - p \right) : p \in [0, \frac{2}{3}] \right\} \\ \mathcal{P}_B &:= \left\{ \left(\frac{2}{3} - p, \frac{1}{3}, p \right) : p \in [0, \frac{2}{3}] \right\} \\ \mathcal{P}_C &:= \left\{ \left(p, \frac{1}{3}, \frac{2}{3} - p \right) : p \in [0, \frac{2}{3}] \right\} \end{aligned}$$

and

$$\mathcal{P} := \frac{1}{3}\mathcal{P}_A + \frac{1}{3}\mathcal{P}_B + \frac{1}{3}\mathcal{P}_C.$$

Then, when without knowing the small urn,

$$\begin{aligned} \min_{P \in \mathcal{P}} \sum_{\omega} P_{\omega} u(a_1, \omega) &= \frac{1}{9} u(a_1, R) + \frac{8}{9} u(a_1, B) = \frac{100 - 160}{9} < 0, \\ \min_{P \in \mathcal{P}} \sum_{\omega} P_{\omega} u(a_0, \omega) &= 0. \end{aligned}$$

Thus, he chooses a_0 and his payoff is 0. Next, suppose that he knows the small urn which he will pick a ball from. If the small urn is A

$$\min_{P \in \mathcal{P}_A} \sum_{\omega} P_{\omega} u(a_1, \omega) = 20,$$

$$\min_{P \in \mathcal{P}_A} \sum_{\omega} P_{\omega} u(a_0, \omega) = 0.$$

If the small urn is B or C , he evaluate the probability of red ball is zero. Then, it is optimal to choose a_0 , and thus, the interim value function is:

$$\max_{f \in \Delta(\{a_1, a_0\})} \min_{p \in \mathcal{P}_s} \sum_{\omega} P_{\omega} \sum_a f(s, a) u(a, \omega) = \begin{cases} 20 & \text{if } s = A \\ 0 & \text{otherwise.} \end{cases}$$

Hence, knowing the small urn form which the ball comes from is valuable for the decision maker.

4.4.2 Application: CEU model

Since the Choquet expected utility (CEU) model with convex capacity has MEU representation, our result has implication to the CEU model.

Let ν be a capacity on $2^{\Omega \times S}$ satisfying

$$\nu(X \cup Y) + \nu(X \cap Y) \geq \nu(X) + \nu(Y)$$

for any subsets X and Y of $\Omega \times S$. We assume

$$\nu(\Omega \times S) = \sum_{s \in S} \nu(\Omega \times \{s\})$$

Define the core $C(\nu)$ of capacity ν as

$$C(\nu) := \{P \in \Delta(\Omega \times S) : \sum_{(\omega, s) \in X} P_{\omega s} \geq \nu(X) \quad \forall X \subset \Omega \times S\}$$

By using these notations, we have the following representation result:

$$\begin{aligned} U(f) &= \int_{\Omega \times S} \sum_a f(s, a) u(a, \omega) d\nu = \min_{P \in C(\nu)} \sum_{\omega \in \Omega, s \in S} P_{\omega s} \sum_a f(s, a) u(a, \omega) \\ &= \sum_{s \in S} P_s \min_{P \in C(\nu_s)} \sum_{\omega \in \Omega} P_{\omega|s} \sum_a f(s, a) u(a, \omega) \end{aligned}$$

where $C(\nu_s) := \{P \in \Delta(\Omega) : \sum_{\omega \in X} P_\omega \geq \frac{\nu(X \times \{s\})}{\nu(\Omega \times \{s\})}\}$. The CEU utility with convexity can be expressed by the MEU model. Then, the result derived above is also applicable to this class of utility functions.

Corollary 2 *Let ν and μ be capacities. Then, the following two conditions are equivalent:*

1. $\delta(C(\nu), C(\mu)) \leq \epsilon.$

- 2.

$$\max_f \int_{\Omega \times S} \sum_a f(s, a) u(a, \omega) d\nu + \epsilon \|U\|_\infty \geq \max_f \int_{\Omega \times S} \sum_a f(s, a) u(a, \omega) d\mu.$$

References

- [1] Blackwell, D., (1951): “Comparison of experiments,” In Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, 93-102.
- [2] Blackwell, D., (1953): “Equivalent comparison of experiments,” *Annals of Mathematical Statistics* 24, 265-272.
- [3] Çelen, B(2012): “Informativeness of experiments for meu”, *Journal of Mathematical Economics* 48,404-406
- [4] Ellsberg, D.(1961) Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics* 75, 643-669
- [5] Gilboa, I. and Schmeidler, D., (1989): “Maxmin expected utility with a non-unique prior.” *Journal of Mathematical Economics* 18, 141-153.
- [6] Gilboa, I., and Schmeidler, D (1994): Additive Representations of Non-Additive Measures and the Choquet Integral, *Annals of Operations Research*, 51,43-65
- [7] Heyen, D, and Boris R. Wiesenfarth(2015): ”Informativeness of experiments for MEU —A recursive definition.” , *Journal of Mathematical Economics*, 57 28-30.

- [8] Le Cam, L. (1964):“ Sufficiency and approximate sufficiency” *Ann. Math. Statist*, 35
1419-1455.