AN AXIOMATIZATION OF SUBJECTIVE MEAN VARIANCE UTILITY UNDER AMBIGUITY

XIANGYU QU

ABSTRACT. Classical mean variance utility of Markowitz [1952] and Tobin [1958] have relied on objective expected utility hypothesis. This paper presents a choice-based axiomatization of mean variance utility, in a Savage-type setting of ambiguity, which neither assumes nor implies that individual's preferences over portfolios conform to the expected utility hypothesis.

Keywords: subjective mean variance utility, reflexibility, translation invariance, ambiguity aversion.

JEL Codes: D80, D81

1. INTRODUCTION

Mean variance (MV) utility, which is only a function of portfolios' mean and variance, has been widely acknowledged for its simplicity and elegance. It is thus not surprising that for many years, the classical mean variance (CMV) utility of Markowitz [1952] and Tobin [1958] has been one of the leading theories of choices under uncertainty in economics and finance. It has led to the capital asset pricing model of Sharpe [1964], which forms a foundation of modern financial theory. In addition, many current theoretical research in both economics and finance, as well as many applied work in the field are still undertaken in the CMV utility framework.

I am deeply indebted to David Schmeidler for his guidance and generous advices. I also wish to thank Chew Soo Hong, Ani Guerdjikova, Jacob Sagi for helpful comments and discussion. I am especially grateful for Mark Machina's thoughtful suggestion.

Nevertheless, the CMV hypothesis has received considerable scrutiny. First, the CMV hypothesis has a particular assumption that objects of choices consist of probability distributions over monetary outcome or wealth. This representation of uncertainty is known as objective uncertainty or risk. However, in practice uncertainty seldom presents itself in terms of exogenously specified probability. In real world, instead of objective probability distribution, the objects of choices are typically acts, which assign outcomes to the possible events or states. This is known as subjective uncertainty or ambiguity. The importance of the distinction between risk and ambiguity has been realized by economists at least since Knight [1921] and Keynes [1921]. The impacts on choice behavior have been repeatedly demonstrated in the experimental literature, for instance Ellsberg [1961] paradox. Second, the CMV is a special case of expected utility (EU) theory. However, the evidence concerning the descriptive validity of EU is not quite favorable. The best known examples of its systematic violation are Allais [1953] and Kahneman and Tversky [1979]. In the experiments, researchers have found that the choices of majority of subjects violate the independence axiom, and hence are not consistent with EU hypothesis. Third, the CMV is characterized by EU theory where either utility index is quadratic (see Samuelson [1970]), or probability distributions are multivariate normal (see Tobin [1958]) or elliptical (see Chamberlain [1983]). The quadratic utility index implies increasing absolute risk aversion according to Arrow [1971] and Pratt [1964], which is not plausible in most situations. A large amount of studies on the validity of CMV has appeared in the empirical finance literature. For instance, researchers have discovered that stock returns generally are neither normally (see Fama [1976]) nor elliptically (see Zhou [1993]) distributed. Those evidence put doubts on the results that rely heavily on the CMV hypothesis.

Although these findings have enlightened some economists to suggest alternative theories of behavior under ambiguity. MV utility seems to continue to be one dominant framework in economics and fiance literature. It would be crucial that we should have some idea of the realism of the theory to avoid the apparent invalidity of key assumptions of CMV. In other words, do departures from the CMV invalidate our ability to characterize the MV utility?

The purpose of this paper is to investigate the behavioral conditions on preferences under ambiguity which imply that the individual: (i) possesses a unique probability distribution over events,[1] (ii) ranks each act according to the MV utility of her induced probability distribution over monetary outcome, (iii) satisfies monotonicity property, and (iv) does not conform to the expected utility in general. I will call such an individual a *subjective mean variance* (SMV) utility maximizer.

The desirability of my approach comes from three authorities. First, I develop a behavioral method of deriving MV from choice under ambiguity. In fact, the previous axiomatic work have not yet derived MV in the exactly same manner. The choice based approach to MV utility is started from Epstein [1985], and further explored by Maccheroni et al. [2006] and others. However, the MV utility of Epstein [1985] applies to preferences over objective probability distribution, so that it is subject to the criticism that real world uncertainty is mostly ambiguous. Maccheroni et al. [2006][2] use horse race/ roulette lotteries of Anscombe and Aumann [1963], which involves both subjective and objective uncertainty. In the field of, especially, finance, outcome space seldom presents itself in terms of objective prospects, but rather, as alternative "money" or "wealth". Therefore, we stand in need of a behavior foundation in terms of preference over ambiguous acts, which assign different monetary outcome to possible events.

The second is the experimental work of Allais [1953] and others, in which they have displayed systematic violations of EU hypothesis. Since the subjects do not conform to EU hypothesis in the setting of risk, it is predictable that they do so in the setting of ambiguity. Without assumption of independence axiom or sure-thing principle of Savage [1972], SMV thus does not in general coincide with EU and could be consistent with Allais-type behavior.

[1]In other words, the individual is probabilistically sophisticated in the sense of Machina and Schmeidler [1992].

[2]Maccheroni et al. [2006] characterize variational preferences by a set of axioms. They then explore the mathematical conditions when variational preferences can be represented by MV.

The third justification for my work stems from the fact that SMV satisfies monotonicity property. The lack of monotonicity of MV is pointed out by Dybvig and Ingersoll [1982] and Jarrow and Madan [1997]. Without this property, investors sometimes strictly prefer a state-wise dominated portfolio to a dominant one, which is not acceptable in most situations. Not surprisingly, this theoretical drawback has led to *monotone mean variance* of Maccheroni et al. [2006] and Maccheroni et al. [2009]. However this theory coincides with MV utility restricted on the monotone domain, so that individual preferences do not conform to MV utility on the whole set of acts.

The following section provides the framework of my model and gives a formal description of SMV. I suggest a set of axioms over individual preferences in Section 3. Section 4 provides an axiomatization result of SMV which neither assumes nor implies the EU hypothesis. Section 5 provides a discussion of related work. All the proofs appear in an Appendix.

2. Preliminaries

2.1. **Setup.** Let Ω be a finite *state space*, where $|\Omega| \geq 2$. The algebra of Ω is denoted by 2^{Ω}. Elements in 2^{Ω} are *events*. Let $X = [m, M] \subset \mathbb{R}$ with $m < 0 < M$ denote the set of *money*. An *act* f is a function from Ω to X. The set of all acts is denoted by $\mathscr{F}$. A binary act f such that $f(\omega) = x$ if $\omega \in E$ and $f(\omega) = y$ otherwise is denoted by xEy. A constant act f such that $f(\omega) = x$ for all $\omega \in \Omega$ is denoted by $\bar{x}$. My setting of ambiguity is identical to that of Gul [1992]'s version of Savage.

Since X is a convex and compact subset of real space $\mathbb{R}$, it is possible to view $\mathscr{F}$ as a convex and compact subset of Euclidean space $\mathbb{R}^n$, where $n = |\Omega|$. Hence I can define *linear combination* and *translation* operations as usual: (i) for every $f, g \in \mathscr{F}$ and $\alpha \in [0, 1]$, I can linearly combine two acts statewise to obtain a new act $\alpha f + (1-\alpha)g \in \mathscr{F}$, which yields $\alpha f(\omega) + (1-\alpha)g(\omega) \in X$ for every $\omega \in \Omega$; (ii) for every $f \in \mathscr{F}$ and $\delta \in \mathbb{R}$ such that $f(\omega) + \delta \in X$ for every $\omega \in \Omega$, I can add them statewise to obtain a new act $f + \bar{\delta} \in \mathscr{F}$, which yields $f(\omega) + \delta \in X$ for every $\omega \in \Omega$. For every $f, g \in \mathscr{F}$, I write $f \geqslant g$ if $f(\omega) \geq g(\omega)$ for each $\omega \in \Omega$ and $f(\omega) > g(\omega)$ for some $\omega \in \Omega$.

I model the decision makers preferences on $\mathscr{F}$ by a binary relation $\succsim$. As usual, $\succ$ and $\sim$ denote, respectively, the asymmetric and symmetric parts of $\succsim$. A functional V *represents* $\succsim$ on $\mathscr{F}$ *if* V is a function from $\mathscr{F}$ to $\mathbb{R}$ such that $f \succsim g \Leftrightarrow V(f) \geq V(g)$.

2.2. **Subjective Mean Variance Utility.** The primary interest of this paper is the SMV utility function.

Definition 1. An individual is said to be a *subjective mean variance* utility maximizer if there exists a probability measure p on 2^{Ω} with full support and a representation functional V on $\mathscr{F}$ such that

(i): individual preference $\succsim$ can be represented by

$$V(f) = \mathrm{E}(f) - \theta \mathrm{Var}(f) \tag{1}$$

where E and Var are mean and variance with respect to p, and $\theta > 0$;

(ii): for every f and g in $\mathscr{F}$, if $f \geqslant g$, then $V(f) > V(g)$.

3. Axioms

Axiom 1 (Weak Order)**.** $\succsim$ is complete and transitive.

Weak order is a standard axiom and does not need further elaboration.

As I mentioned, $\mathscr{F}$ is a subset of $\mathbb{R}^n$ where $n = |\Omega|$. Hence a set $G \subset \mathscr{F}$ is said to be *closed* if G is a closed subset of $\mathbb{R}^n$.

Axiom 2 (Continuity)**.** For every f in $\mathscr{F}$, the sets $\{g \in \mathscr{F} : g \succsim f\}$ and $\{g \in \mathscr{F} : f \succsim g\}$ are closed.

I assume throughout that if one act assigns to each state a money no less than the money assigned to the same state by a second act, and if for some state the money is strictly higher, then the first act is strictly preferred to the second act. This implies monotonicity over constant act, i.e. if $x, y \in X$ and $x > y$, then $\bar{x} \succ \bar{y}$.

Axiom 3 (Strict Monotonicity)**.** For every f and g in $\mathscr{F}$, if $f \geqslant g$, then $f \succ g$.

If $\succsim$ satisfies weak order, continuity and strict monotonicity, then for each act $f \in \mathscr{F}$, there exists a unique constant act $\bar{x}$ which is *certainty equivalent* for it, i.e. $f \sim \bar{x}$.

An event $E \in 2^{\Omega}$ is *null* if for every $x, y, z \in X$, $yEx \sim zEx$. An event E is *nonnull* if it is not null. Strict monotonicity axiom implies that every event is nonnull.

Ambiguity aversion, due to Schmeidler [1989], states that a linear combination of two indifferent acts is always (weakly) preferred to either acts[3]. Here I strengthen it by requiring that it holds in the strict sense.

Axiom 4 (Strict Ambiguity Averion)**.** For every f and g in $\mathscr{F}$ and $\alpha \in (0, 1)$,

$$f \sim g \Rightarrow \alpha f + (1 - \alpha)g \succ f.$$

The next axiom in spirit is first initiated by Chew et al. [1991] in the risk environment. It requires that given two indifferent acts, a linear combination of the two acts is also indifferent to its conjugate combination.

Axiom 5 (Reflexibility)**.** For every f and g in $\mathscr{F}$ and $\alpha \in (0, 1)$,

$$f \sim g \Rightarrow \alpha f + (1 - \alpha)g \sim (1 - \alpha)f + \alpha g.$$

Clearly reflexibility axiom is implied by the standard independence axiom, but not vice verse. In reflexibility axiom, a linear combination of two indifferent acts, $\alpha f + (1 - \alpha)g$, may sufficiently distinct from acts f and g. This is contrary to the description of standard independence axiom in which $\alpha f + (1 - \alpha)g$ is indifferent to f. On the other hand, if I combine two indifferent acts differently, say $\alpha f + (1 - \alpha)g$ and $\beta f + (1 - \beta)g$, then for some

[3] The discussion of this axiom in length can be found at Schmeidler [1989], and Gilboa and Schmeidler [1989].

plausible α and β the two combinations may have the same effects, and two combinations are therefore indifferent. In this axiom, it is the case where α and β are conjugate, i.e. $\alpha+\beta=1$.

The next axiom also plays a critical role in my representation theorem. It requires that given $f \sim g$, any equitranlation, which adds the same amounts of prizes on f and g state-wisely, will make them maintain the indifference relation.

Axiom 6 (Translation Invariance)**.** For every f and g in $\mathscr{F}$ and $\delta \in \mathbb{R}$ such that $f+\bar{\delta}$ and $g+\bar{\delta}$ are in $\mathscr{F}$,

$$f \sim g \Rightarrow f+\bar{\delta} \sim g+\bar{\delta}.$$

My objective is to formulate a theory of preference which can be represented by a SMV function. The normative argument of translation is worth discussion. If two acts f and g are indifferent, then an income increase (or decrease) by the same amounts will not affect the indifference relation. This reflects that equitranslation does not affect the decision maker's ambiguity evaluation. This axiom provides a comparison with constant independence axiom by Gilboa and Schmeidler [1989], which states that if two acts are indifferent, then the same mixtures with any constant acts will not change the indifference relation. In constant independence, the indifference relation is maintained through a mixture with a constant act, but not through an equitranlation.

4. Representation Results

In this section the preceding axioms and SMV are related. A characterization of SMV in the absence of EU hypothesis involves talking the Axioms 1-6. The main result of this paper demonstrate that if preferences satisfy this set of axioms, there exists a subjective probability over states, and individual behaves as if she evaluates acts based on SMV with respect to the subjective probability. This characterization is stated precisely in the following theorem.

Theorem 1. *Let $\succsim$ be a binary relation on $\mathscr{F}$. Then $\succsim$ satisfies Axioms 1-6 if and only if it can be represented by a subjective mean variance utility as in (1). Furthermore, if V_1 and V_2*

are two subjective mean variance utilities representing $\succsim$ on $\mathscr{F}$, then there exist α and β in $\mathbb{R}$ with $\alpha > 0$ such that $V_1(f) = \alpha V_2(f) + \beta$ for every $f \in \mathscr{F}$.

5. Discussion

It is not surprising that SMV violates Ellsberg Paradox because it is probabilistically sophisticated in the sense of Machina and Schmeidler [1992]. In order to maintain the simplicity of two parameter model as well as conform to Ellsberg-type behavior, Grant and Polak [2011] develop a mean-dispersion utility model in an Anscombe and Aumann [1963] framework. Their results can be applied to obtain a characterization of non-Bayesian models, which yields the class of Choquet expected utility model of Schmeidler [1989], maxmin expected utility model of Gilboa and Schmeidler [1989] and others. In another line of work, their results can be applied to obtain a characterization of probabilistical sophistication, which yields the class of non-expected utility models developed by Gilboa [1987]. However it is not clear how to characterize SMV in the AA framework since the translation invariance axiom is not well-defined.

6. Appendix: Proof of Theorem 1

6.1. **Sufficiency Proof of Theorem 1 where $|\Omega| = 2$.** This appendix deals exclusively with state space where $\Omega = \{E, E^c\}$. In this subsection we will interpret X^2 to mean the set of acts $\mathscr{F}$, and we will also understand that for each $(x, y) \in X^2$, x and y are the prizes associated with events E and E^c respectively. By Debreu (1964), there is no loss of generality in assuming that $\succsim$ can be represented by a utility functional $V : X^2 \to \mathbb{R}$. We also write $V(x, y)$ as the numerically representation of act xEy. The continuity and monotonicity axioms imply that V is continuous and increasing on X^2.

We will show that the preference ordering can be represented by a monotone mean variance utility function. The (income) *expansion paths* (EP) of the indifference curves may play a critical role in the following argument. Recall that we say two points a and b in X^2 lie on the same EP *if* the subgradient to the indifference curves at a and b are the same.

The proof presented below are sketched as follows: we first show that all EPs are linear and parellel to each other. Mixture symmetry and Translation invariance along with other axioms imply that the loci of midpoints of parallel chords of every indifference curves are parallel lines. Then we can use a theorem by Coxeter (1974) and CES (1991) to show that every indifference curve is a conic. As a consequence, indifference curve can be represented by a quadratic function. We further show some additional properties of this quadratic function implied by translation invariance and other axioms. Those properties finally lead it to be a mean variance utility function. Finally, since EPs are linear, we can normalize the coordinate system such that a mean variance utility function can represent $\succsim$ on X^2.

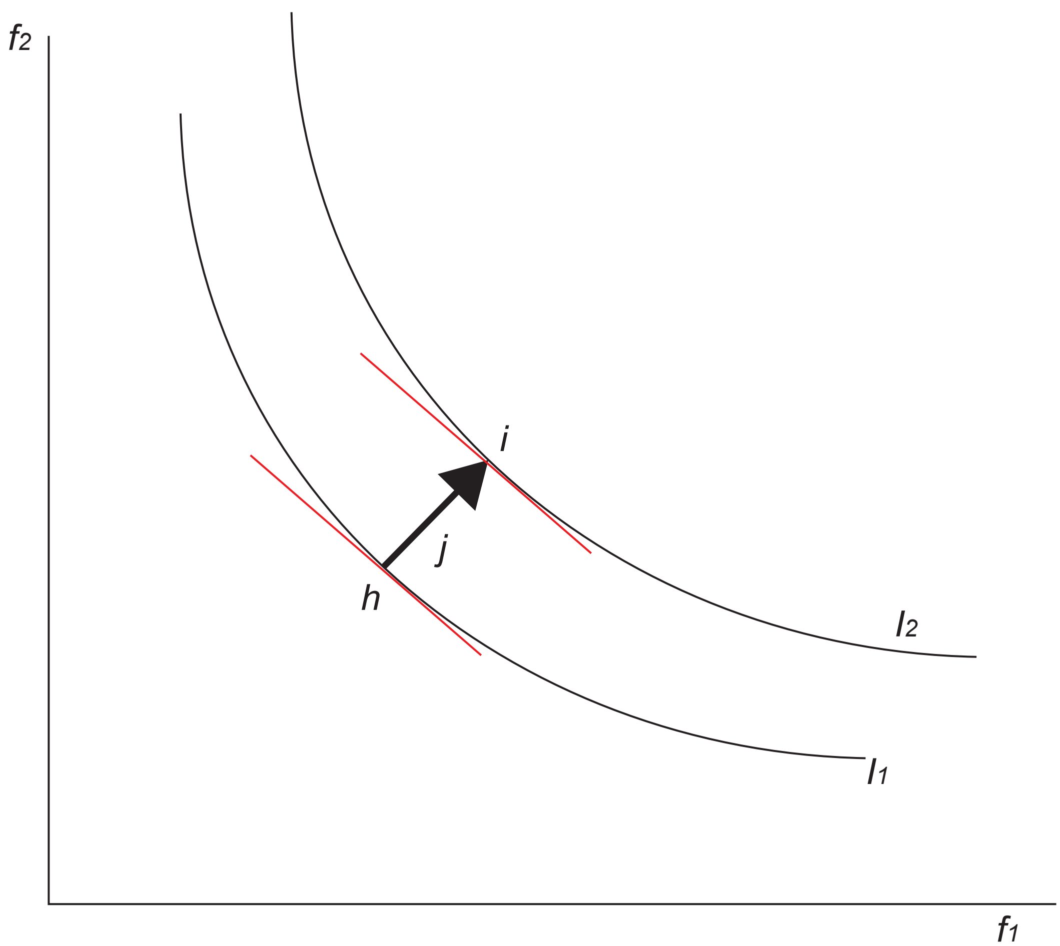

FIGURE 1.

Lemma 1. *All expansion paths are straight parallel lines.*

Proof. Refer to Figure (1). Start with an indifference curve I_1. Draw the curve I_2 by shifting the coordinates of each point of I_1 sufficiently small $\delta > 0$ unit to the upper right. By Translation invariance axiom, I_2 is also an indifference curve. Pick an point h on I_1 and a corresponding point i on I_2 such that $h + \bar{\delta} = i$. By plane geometry, the price ration underlying h is parallel to its underlying i. Therefore h and i are on the same EP. Locate point j such that $j = h + \bar{\delta}'$ where $0 < \delta' < \delta$. Repeating the same argument, the point j is on the EP of h and i. That is j is on the line segment hi. Therefore the EP through h and i must be linear. Since the selection of h is arbitrary, every EP must be linear. Also it is obvious to see that the slopes of every EP are one. Thus, we conclude that EPs are parallel lines. □

Lemma 2. *The loci of midpoints of parallel chords parallel to each other.*

Proof. Refer to Figure(2). Start with two parallel chords ad and bc. Draw the points h and i such that h and i are located at the middle of chords bc and ad respectively. We claim that segment hi is an EP.

To see this, it is sufficient to show that the slopes of chords ad and bc are the same as the price rations at points h and i. Since h is the middle point of chord bc, h is strictly preferred to any other points on bc. This is implied by mixture symmetry axiom. Therefore the indifference curve through h has to be tangent to chord bc at h. Similar argument applies to point i. Thus h and i are on the same EP. According to plane geometry, the locus of midpoints of chords paralleled to ad must be located on the line connecting h and i. By lemma 1, the straight line through h and i is an EP whose slope is one. Thus every locus of midpoints of parallel chords of an indifference curve must be an EP. □

Now we can apply a theorem developed by Coxeter(1974) and Chew, Epstein and Segal(1991) to claim that every indifference curve is a conic. For completeness, we first recall

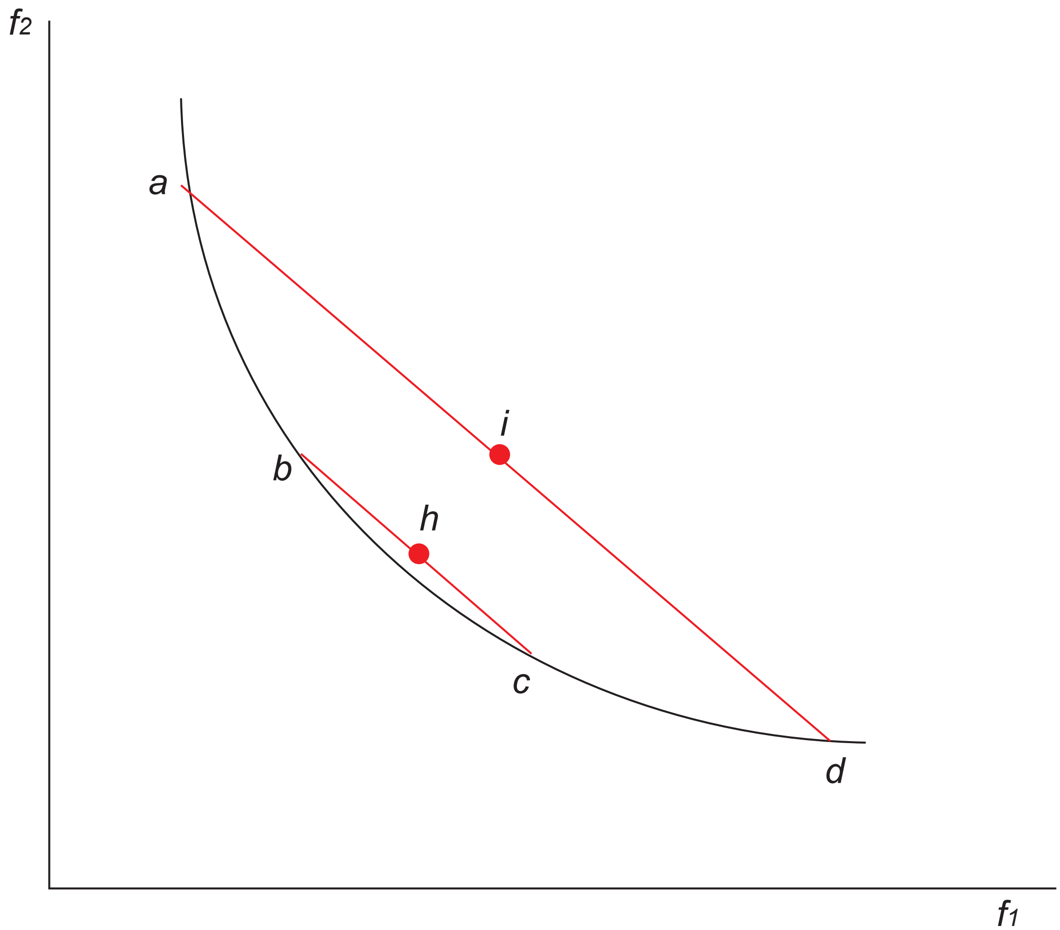

FIGURE 2.

some notation and then restate their theorem below. We say that an indifference curve in X^2 possess the *projection property* if the loci of midpoints of parallel chords either have a common point of intersection or are parallel lines.

Theorem 2 (Coxeter(1974) and CES(1991))**.** *A curve in $\mathbb{R}^2$ is a conic if and only if it has the projection property.*

We have proved that the loci of midpoints of parallel chords are parallel lines, thereby satisfies the projection property. Therefore each indifference curve is a conic. A conic can be represented by a quadratic function, i.e. $ax^2 + by^2 + cxy + dx + ey$. Since EPs are straight parallel lines, we can translate the coordinate system in a way such that the representation function V is homothetic in the new system. It results that the same quadratic function applies to all indifference curves. Next we will display some additional properties that this quadratic function has to have.

Lemma 3. *The representation V on X^2 has the following properties:*

- **(i):** $d > 0$ *and* $e > 0$*;*
- **(ii):** $a = b < 0$ *and* $c = -2a$*.*

Proof. (i) Let $x > 0$ and $-x$ be in X. We have $V(x,0) = ax^2 + dx$ and $V(-x,0) = ax^2 - dx$. By strict monotonicity axiom, $V(x,0) > V(-x,0)$. Hence $d > 0$. Similarly, we can see $e > 0$.

(ii) Let two non-constant acts be such that $(x_1, y_1) \sim (x_2, y_2)$. Let $\delta \in \mathbb{R}\backslash\{0\}$ be such that $(x_1+\delta, y_1+\delta)$ and $(x_2+\delta, y_2+\delta)$ are in X^2. By translation invariance axiom $(x_1+\delta, y_1+\delta) \sim (x_2+\delta, y_2+\delta)$. Then we have

$$ax_1^2 + by_1^2 + cx_1y_1 + dx_1 + ey_1 = ax_2^2 + by_2^2 + cx_2y_2 + dx_2 + ey_1 \tag{2}$$

$$\begin{aligned} a(x_1+\delta)^2 + b((y_1+\delta)^2 + c(x_1+\delta)(y_1+\delta) + d(x_1+\delta) + e(y_1+\delta) = \\ a(x_1+\delta)^2 + b((y_1+\delta)^2 + c(x_1+\delta)(y_1+\delta) + d(x_1+\delta) + e(y_1+\delta) \end{aligned} \tag{3}$$

Subtracting (3) by (2), we get

$$2ax_1\delta + 2by_1\delta + c(x_1\delta + y_1\delta) = 2ax_2\delta + 2by_2\delta + c(x_2\delta + y_2\delta)$$

This implies $(2a+c)(x_1 - x_2) + (2b+c)(y_1 - y_2) = 0$.

First observe that if $c = -2a$, then $c = -2b$. Suppose that $c \neq -2b$. We must have

$$\frac{y_1 - y_2}{x_1 - x_2} = -\frac{2a + c}{2b + c}$$

Weak order, Monotonicity and Continuity guarantee that there exists a constant act (z, z) such that $(x_1, y_1) \sim (x_2, y_2) \sim (z, z)$. By the similar argument as above, we get

$$\frac{y_1 - y_2}{x_1 - x_2} = \frac{y_2 - z}{x_2 - z} = \frac{z - y_1}{z - x_1}$$

This requires that points $(x_1, y_1), (x_2, y_2)$ and (z, z) must lie on the same line since they share the same slope. Thus one point must be a mixture of the other two points. However this contradicts the strict ambiguity aversion axiom. Hence $a = b$ and $c = -2a$.

Now we can write the quadratic function as $V(x, y) = ax^2 + ay^2 - 2axy + dx + ey$, where $d, e > 0$. We show that $a < 0$. By strict ambiguity aversion, the function V is strictly quasiconcave. Since $V : X^2 \to \mathbb{R}$ is differentiable, its twice differentiation matrix is negative definite. Let $z \in \mathbb{R}^2$. Then

$$\begin{aligned} zD^2(V)z' &= (z_1, z_2) \begin{pmatrix} 2a & -2a \\ -2a & 2a \end{pmatrix} (z_1, z_2)' \\ &= 2a(z_1 - z_2)^2 < 0 \end{aligned}$$

Hence $a < 0$.

□

Lemma 4. *A quadratic function V satisfying the properties above is a mean variance utility function.*

Proof. Suppose that V satisfies the above properties. Then we can write it as the following form.

$$V(x, y) = -ax^2 - ay^2 + 2axy + bx + cy$$

where $a, b, c > 0$. Rewrite it, we can get

$$V(x,y) = (b+c)\big[\frac{b}{b+c}x + \frac{c}{b+c}y - \frac{a}{b+c}(x-y)^2\big]$$

Let $p = \frac{b}{b+c}$, $q = \frac{c}{b+c}$ and $\theta = \frac{a(b+c)}{bc}$. So $p, q > 0$ with $p + q = 1$, and $\theta > 0$.

$$\begin{aligned} V(x,y) &= (b+c)\big[px + qy - \theta pq(x-y)^2\big] \\ &= (b+c)\big[px + qy - \theta(p^2q + pq^2)(x-y)^2\big] \\ &= (b+c)\big[px + qy - \theta[p(x-(px+qy))^2 + q(y-(px+qy))^2\big] \end{aligned}$$

Hence, this quadratic function is a mean variance utility function as in (1).

□

The above lemmas imply that each indifference curve is a graph of a monotone mean variance utility function. (The condition (ii) of the definition is guaranteed by the monotonicity axiom.) Since EPs are linear and parallel to each other, we can normalize the coordinate system such that the utility function is homothetic in the new system. Therefore each indifference curve can be represented in the following way:

$$I(c) = \big\{(x,y) \in X^2 : \phi(x,y) = c \text{ and } \phi(c,c) = c\big\}$$

where ϕ is the normalized mean variance utility function and $c \in X$. Thus the above normalized function represents $\succsim$ on X^2. Hence we have established the desired representation on X^2.

6.2. **Sufficiency Proof of Theorem 1 where $|\Omega|$ is finite.** In this section, we extend the representation result to the class of acts with finite state space where $|\Omega| = n$. Similarly as we did before, the domain of representation function can be taken to be X^n. For every $(x_1, \ldots, x_n) \in X^n$, x_i is the prize associated with the ith state in Ω.

Our strategy to show the desired representation on X^n is by induction. The result of induction implies that every indifference set have a local mean variance utility representation.

We therefore can write $X^n = \cup_{i=1}^{\infty} O_i$ where O_i is an open ball having a mean variance utility representation V_i on it and $O_i \cap O_{i+1} \neq \emptyset$ for every $i \geq i$. Therefore V_i and V_{i+1} are ordinally equivalent on $O_i \cap O_{i+1}$ and hence cardinally equivalent by the lemma (5) shown as below. Thus we can redefine $V_{i+1} = V_i$ restricted on $O_i \cap O_{i+1}$. By starting construction from $i = 1$ and similarly normalizing process as last section, we can acquire a mean variance utility function V on X^n. Thus it is straightforward to see that this V is constant along the indifference sets and represents $\succsim$ on X^n.

We first state and prove the following lemma, which says that a pair of ordinally equivalent mean variance utility functions are also cardinally equivalent. As a consequence, the uniqueness of representation follows.

Lemma 5. *Let ϕ and ψ be two mean variance utility functions defined on X^n. If there exists an increasing function ξ on $\psi(X^n)$, the range of ψ, then ξ must be linear on $\psi(X^n)$.*

Proof. Let ϕ and ψ define as follows: for every $h = (h_1, \dots, h_n) \in X^n$,

$$\phi(h) = \sum_{i=1}^{n} h_i p_i - \theta \sum_{i=1}^{n} (h_i - \sum_{j=1}^{n} h_j p_j)^2 p_i$$

$$\psi(h) = \sum_{i=1}^{n} h_i p'_i - \theta' \sum_{i=1}^{n} (h_i - \sum_{j=1}^{n} h_j p'_j)^2 p'_i$$

From the above expression, ϕ is twice differentiable. So is $\xi \circ \psi$ by assumption. Taking twice differentiation on both ϕ and $\xi \circ \psi$, we have for $i \neq j$,

$$\frac{\partial^2 \phi}{\partial h_i \partial h_j} = 2\theta p_i p_j \tag{4}$$

$$\frac{\partial^2 \xi}{\partial h_i \partial h_j} = (2\theta' p'_i p'_j) \cdot \frac{\partial \xi}{\partial \psi} + \frac{\partial^2 \xi}{\partial \psi^2} \cdot \frac{\partial \psi}{\partial h_i} \tag{5}$$

Since $\xi \circ \psi$ is only twice differentialbe, $\frac{\partial^2 \xi}{\partial \psi^2} = 0$. Therefore $2\theta p_i p_j = (2\theta' p'_i p'_j) \cdot \frac{\partial \xi}{\partial \psi}$, which implies $\frac{\partial \xi}{\partial \psi} = \frac{\theta p_i p_j}{\theta' p'_i p'_j}$. Thus ξ must be linear on the range of ψ. $\square$

To formally state the induction, we need some notation first. We say that $L \subset X^n$ is a *diagonal plane of dimension* k if $L = H \cap X^n$, where H is a $k-$dimensional subspace of $\mathbb{R}^n$ such that the diagonal line $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 = \cdots = x_n\}$ lies in H.

$S(k)$**:** Let L be a $k-$dimensional diagonal plane in X^n and let $\hat{h} \in int(L)$ be such that the indifference set containing $\hat{h}$ is not a singleton. Then there exists an open neighborhood N of $\hat{h}$ in L and a mean variance utility function $\phi : N \to \mathbb{R}$ satisfying that for every indifference set I, there is a $c \in \mathbb{R}$ such that

$$I \cap N = \{h \in N : \phi(h) = c\}$$

In other words, we want to show by induction that for every interior point $\hat{h}$ in X^n, we can find an open neighborhood N of $\hat{h}$ such that every indifference set restricted on N can be represented by a mean variance utility function.

Step 1: $k = 2$. Let L be a two dimensional diagonal plane in X^n. Now we show that the desired property holds for L.

Lemma 6. *Let $\hat{h}$ be in $int(L)$. Then $\hat{h}$ is not a minimum point of $\succsim$ on L.*

Proof. Assume that $\hat{h}$ is a minimum point. Since $\hat{h}$ is an interior point of L, we can always find two points h, h' in L such that $\hat{h} \in (h, h')$ the open line segment between h and h'. By strict ambiguity aversion axiom, $\hat{h}$ is at least strictly preferred to one of the two points, which contradicts to our assumption. □

Lemma 7. *Suppose $\hat{h}$ is a maximum point of $\succsim$ in L. Then it is a unique maximum point. Furthermore, $\hat{h}$ is a constant act.*

Proof. Let $\hat{h}$ be a maximum point. Then that it is unique follows directly from strict ambiguity aversion. By continuity and monotonicity, there exists a unique constant act that is indifference to $\hat{h}$. since each constant act is an element of L, $\hat{h}$ must be a constant act. □

We proceed to assume that $\hat{h}$ is neither a maximum nor a minimum point in L. Let $h^* = (\delta, \ldots, \delta)$ be the maximum point in L. By continuity, there are two neighborhoods $B(h^*, \epsilon)$ and $B(\hat{h}, \epsilon)$, which are centered at h^* and $\hat{h}$ respectively with radius $\epsilon > 0$, such that for each $a \in B(h^*, \epsilon)$ and each $b \in B(\hat{h}, \epsilon)$, $a \succsim b$.

Locate the points a, c on $B(h^*, \epsilon)$ and b, d on $B(\hat{h}, \epsilon)$ such that the lines ab and cd are parallel to line $h^*\hat{h}$, and tangent to both $B(h^*, \epsilon)$ and $B(\hat{h}, \epsilon)$. Let a', b' be arbitrary points in $B(h^*, \epsilon)$ and $B(\hat{h}, \epsilon)$ respectively. Draw the lines $a'b'$. We can see that the preference $\succsim$ is monotonic in the direction from b' to a', which follows directly from strict ambiguity aversion.

Since L is a subset of two dimensional subspace of $\mathbb{R}^n$, according to some algebra L can be coordinatized and thereby identified with X^2. Now we choose basis to give a coordinatization. Let $T : L \to \mathbb{R}^2$ be a linear transformation such that $T(\hat{h}) = (0,0)$, $T(a) = (M, 0)$ and $T(c) = (0, M)$. For $x, y \in X^2$, we define a preference $\succsim_T$ on $T(L)$ by $x \succsim_T y \Leftrightarrow T^{-1}(x) \succsim T^{-1}(y)$. This is well-defined since T is a bijection. Next we will show that this new preference $\succsim_T$ around point $T^{-1}(\hat{h}) = (0,0)$ satisfies axioms 1-8. Therefore there should exist a mean variance utility representation of $\succsim_T$ around $(0,0)$. Hence $\succsim$ can be represented locally around $\hat{h}$ by a mean variance utility.

Lemma 8. $\succsim_T$ *around point* $(0,0)$ *satisfies Axioms 1-7.*

Proof. According to the properties of linear transformation, it is straightforward to see that weak order, continuity, essentiality, strict ambiguity aversion and mixture symmetry are satisfied. We are only left to show the translation invariance and monotonicity.

Notice that $h^* = \frac{1}{2}a + \frac{1}{2}c$. Then $T(h^*) = T(\frac{1}{2}a + \frac{1}{2}c) = \frac{1}{2}T(a) + \frac{1}{2}T(c) = (\frac{M}{2}, \frac{M}{2})$. Suppose that $x \sim_T y$. Then $T^{-1}(x) \sim T^{-1}(y)$. Let $\hat{\delta} > 0$ be small enough.

$$\begin{aligned} T^{-1}(x + (\hat{\delta}, \hat{\delta})) &= T^{-1}(x) + T^{-1}((\hat{\delta}, \hat{\delta})) \\ &= T^{-1}(x) \oplus \frac{2\delta\hat{\delta}}{M} \\ &\sim T^{-1}(y) \oplus \frac{2\delta\hat{\delta}}{M} = T^{-1}(y + (\hat{\delta}, \hat{\delta})) \end{aligned}$$

Thus, $x + (\hat{\delta}, \hat{\delta}) \sim_T y + (\hat{\delta}, \hat{\delta})$, which proves the axiom of translation invariance.

Suppose $x \geqslant y$ and x, y are close enough to $(0,0)$ such that $T^{-1}(x), T^{-1}(y) \in B(\hat{h}, \epsilon)$. By the nature of linear transformation, the slope of line $\overline{T^{-1}(x)T^{-1}(y)}$ has to be between the slopes of lines $\overline{T^{-1}(\hat{h})T^{-1}(a)}$ and $\overline{T^{-1}(\hat{h})T^{-1}(c)}$. Therefore $\overline{T^{-1}(x)T^{-1}(y)} \cap B(h^*, \epsilon) \neq \emptyset$. Hence, based on the above argument, $T^{-1}(x) \succ T^{-1}(y)$, which means $x \succ_T y$.

Step 2: Assume $S(k-1)$ and prove $S(k)$.

Since the argument for general k is similar to n but notationally much more complex, we let $k = n$. Let $\hat{h}$ lie in the interior of X^n. We can choose a sequence $\{h^1, \ldots, h^n\}$ of points such that each point is close to $\hat{h}$ and $\{h^i - \hat{h} : i = 1, \ldots, n\}$ is linear independent. The existence of $h^1, \ldots, h^n$ is guaranteed by the strictly ambiguity aversion. (For example, all points can be selected from the indifference set containing $\hat{h}$.) Therefore, $\mathbb{R}^n = span\{h^i - \hat{h} : i = 1, \ldots, n\}$.

Since we can recoordinarize the system through a proper linear transformation, without loss of generality we assume that $h^i - \hat{h} = e^i$ is a unit vector where the ith coordinate is one. Let $h^* = \sum_{i=1}^n e^i$. Hence for each $i = 1, \ldots, n$, $L^i = span\{h^*, j^j - \hat{h} : j \neq i, i+1\}$ is a diagonal hyperplane. Actually $L^i = \{h \in \mathbb{R}^n : h_i = h_{i+1}\}$ is a hyperplane where the ith and $i+1$th coordinates of every vectors are the same. Therefore, by assumption there exists a sequence $\{N^1, \ldots, N^n\}$ of open sets containing $\hat{h}$, and a sequence $\{V^1, \ldots, V^n\}$ of mean variance utility functions such that V^i represents $\succsim$ on N^i for each i. Define V so that it coincides with V^1 on N^1, thereby on $N^1 \cap N^2$. Hence V and V^2 are ordinally equivalent on $N^1 \cap N^2$. By lemma (), the transformation between V and V^2 are linear, and hence V represents $succsim$ on N^2. Repeat this argument, V represents $\succsim$ on each N^i.

Let $\widetilde{L}^i = \{h \in \mathbb{R}^n : h_i = h_{i+1} + (\hat{h}_i - \hat{h}_{i+1})\}$. So $\widetilde{L}^i$ is obtained through parallel translation of L^i such that $\widetilde{L}^i$ goes through $\hat{h}$. Now we claim that there exists an open ball B of $\hat{h}$ such that the V defined above represents $\succsim$ on $\widetilde{N^i} = \widetilde{L}^i \cap B$ for each i. Let N^i be the image of $\widetilde{N^i}$ on L^i through parallel translation. Then there exists a mean variance utility function $\widetilde{V^i}$ that represents $\succsim$ on $\widetilde{N^i}$. Hence $\widetilde{V^i}$ represents $\succsim$ on $\widetilde{N^i} \cap N^j$, which means that V can also represent $\succsim$ on $\widetilde{N^i}$. This argument is applies to each i.

To finish, we need to show that V represents $\succsim$ on B. Take an arbitrary point h on B. Consider the diagonal hyperplane L^1 of h where the nature of mean variance utility function implies locally: (Here we normalize $V(\hat{h}) = 0$.)

$$\begin{aligned} V(h) &= \alpha(h_1 - \hat{h}_1, h_2 - \hat{h}_2)p(h_1, h_2) + \sum_{i=3}^{n}(h_i - \hat{h}_i)p_i(h_1, h_2) - \theta(h_1, h_2)\{[\alpha(h_1 - \hat{h}_1, h_2 - \hat{h}_2)p(h_1, h_2) \\ &- (\alpha(h_1 - \hat{h}_1, h_2 - \hat{h}_2)p(h_1, h_2) + \sum_{i=3}^{n}(h_i - \hat{h}_i)p_i(h_1, h_2)]^2 p(h_1, h_2) + \sum_{i=3}^{n}[(h_i - \hat{h}_i) \\ &- (\alpha(h_1 - \hat{h}_1, h_2 - \hat{h}_2)p(h_1, h_2) + \sum_{i=3}^{n}(h_i - \hat{h}_i)p_i(h_1, h_2)]^2 p_i(h_1, h_2)\} \end{aligned}$$

From the above expression, V has to be twice differentiable. V_{h_1,h_2} is independent of $h_1, \ldots, h_n$. Hence $\alpha(\cdot, \cdot)$ must be linear and $\theta h_1, h_2$ must be a constant. If we repeat this expression by restricted on hyperplanes $L^2, \ldots, L^n$, similar results follows. Hence V represents $\succsim$ on B. V is mean variance utility function on X^n since the restriction on every diagonal hyperplane is. $\square$

6.3. **Necessity Proof of Theorem 1.** The necessity are straightforward. We omit them.

REFERENCES

1. M. Allais. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica*, 21(4):pp. 503–546, 1953.

2. F. Anscombe and R. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.
3. K. Arrow. *Essays in the Theory of Risk Bearing*. Chicago: Markham Publishing Co., 1971.
4. Gary Chamberlain. A characterization of the distributions that imply mean variance utility functions. *Journal of Economic Theory*, 29(1):185 – 201, 1983.
5. S. H. Chew, L. G. Epstein, and U. Segal. Mixture Symmetry and Quadratic Utility. *Econometrica*, 59(1), 1991.
6. Philip H. Dybvig and Jr. Ingersoll, Jonathan E. Mean-Variance Theory in Complete Markets. *The Journal of Business*, 55(2):pp. 233–251, 1982.
7. Daniel Ellsberg. Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.
8. Larry G. Epstein. Decreasing Risk Aversion and Mean-Variance Analysis. *Econometrica*, 53(4):945–961, 1985.
9. Eugene Fama. *Foundations of Finance*. New York: Basic, 1976.
10. Itzhak Gilboa. Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, 16(1):65 – 88, 1987.
11. Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141 – 153, 1989.
12. Simon Grant and Ben Polak. Mean-dispersion preferences and constant absolute uncertainty aversion. working paper, 2011.
13. Faruk Gul. Savag's theorem with a finite number of states. *Journal of Economic Theory*, 57 (1):99 – 110, 1992.
14. Robert A. Jarrow and Dilip B. Madan. Is Mean-Variance Analysis Vacuous: Or was Beta Still Born? *European Finance Review*, 1(1):15–30, 1997.
15. Daniel Kahneman and Amos Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):pp. 263–292, 1979.
16. J. M. Keynes. *A Treatise on Probability*. Macmillan, London, 1921.

17. Frank H. Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston, 1921.

18. Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini. Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica*, 74(6):1447–1498, 2006.

19. Fabio Maccheroni, Massimo Marinacci, Aldo Rustichini, and Marco Taboga. Portfolio Selection with Monotone Mean-Variance Preferences. *Mathematical Finance*, 19(3):487–521, 2009.

20. Mark J. Machina and David Schmeidler. A More Robust Definition of Subjective Probability. *Econometrica*, 60(4):745–780, 1992.

21. H.M. Markowitz. Portfolio selection. *Journal of Fiance*, 7:77–91, 1952.

22. John W. Pratt. Risk Aversion in the Small and in the Large. *Econometrica*, 32(1/2):pp. 122–136, 1964.

23. Paul A. Samuelson. The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments. *The Review of Economic Studies*, 37(4): pp. 537–542, 1970.

24. Leonard J. Savage. *The foundations of statistics*. Second Edition, new York: Dover, 1972.

25. David Schmeidler. Subjective Probability and Expected Utility without Additivity. *Econometrica*, 57(3):571–587, 1989.

26. William F. Sharpe. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3):pp. 425–442, 1964.

27. J. Tobin. Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies*, 25(2):pp. 65–86, 1958.

28. Guofu Zhou. Asset-Pricing Tests Under Alternative Distributions. *The Journal of Finance*, 48(5):pp. 1927–1942, 1993.

DEPARTMENT OF ECONOMICS, OHIO STATE UNIVERSITY

E-mail address: xiangyuqu@gmail.com