Theoretical Economics 6 (2011), 423–472 1555-7561/20110423

Intertemporal substitution and recursive smooth ambiguity preferences

Takashi Hayashi
Department of Economics, University of Texas at Austin

Jianjun Miao
Department of Economics, Boston University, CEMA, Central University of Finance and Economics, and AFR, Zhejiang University

In this paper, we establish an axiomatically founded generalized recursive smooth ambiguity model that allows for a separation among intertemporal substitution, risk aversion, and ambiguity aversion. We axiomatize this model using two approaches: the second-order act approach à la Klibanoff et al. (2005) and the two-stage randomization approach à la Seo (2009). We characterize risk attitude and ambiguity attitude within these two approaches. We then discuss our model's application in asset pricing. Our recursive preference model nests some popular models in the literature as special cases.

Keywords. Ambiguity, ambiguity aversion, risk aversion, intertemporal substitution, model uncertainty, recursive utility, dynamic consistency.

JEL classification. D80, D81, D90.

1. Introduction

The rational expectations hypothesis is a workhorse assumption in macroeconomics and finance. However, it rules out ambiguity-sensitive behavior. In addition, it faces serious difficulties when confronted with experimental evidence (Ellsberg 1961) or asset markets data (Hansen and Singleton 1983 and Mehra and Prescott 1985). Since Gilboa and Schmeidler's (1989) and Schmeidler's (1989) seminal contributions, there is a growing body of literature that develops theoretical models of decision making under ambiguity.[1] In addition, there is also a growing body of literature that applies these utility models to finance and macroeconomics.[2] This literatures demonstrates that these models are useful for explaining many economic phenomena.

In this paper, we establish an axiomatically founded generalized recursive smooth ambiguity model that allows for a separation among intertemporal substitution, risk

Takashi Hayashi: th925@eco.utexas.edu
Jianjun Miao: miaoj@bu.edu
We are grateful to three anonymous referees and a co-editor (Gadi Barlevy) for helpful comments and suggestions. We also thank Bart Lipman, Massimo Marinacci, and Kyoungwon Seo for helpful comments.

[1] See Cerreia-Vioglio et al. (2008) for a comprehensive study and the references cited therein.
[2] See Backus et al. (2005) and Hansen and Sargent (2007b) for surveys.

DOI: 10.3982/TE843

aversion, and ambiguity aversion.[3] An axiomatic foundation is important because the choice-based assumptions on preferences make the model testable in principle. We axiomatize our model using two approaches: the second-order act approach à la Klibanoff et al. (henceforth KMM) (2005) and the two-stage randomization approach à la Seo (2009). We characterize risk attitude and ambiguity attitude within these two approaches. We then apply our model to asset pricing and derive its pricing kernel using a homothetic specification. We show that an ambiguity averse agent attaches more weight on the pricing kernel when his continuation value is low in a recession. This feature generates countercyclical market price of uncertainty and is useful in explaining asset pricing puzzles (Hansen 2007, Hansen and Sargent 2010, and Ju and Miao forthcoming).

Our dynamic model is built on the static smooth ambiguity model developed by KMM (2005). This static model delivers a utility function over the space of random consumption as[4]

$$V(c) = v^{-1}\left(\int_{\mathcal{P}} v \circ u^{-1}\left(\int_S u(c)\, d\pi(s)\right) d\mu(\pi)\right), \quad c : S \to \mathbb{R}_+, \tag{1}$$

where S is the state space, $\mathcal{P}$ is a set of probability measures on S, μ is a probability measure over $\mathcal{P}$, u describes risk attitude, and $v \circ u^{-1}$ describes ambiguity attitude. The set $\mathcal{P}$ reflects model uncertainty or the decision maker's ambiguity about the "true" distribution of consumption. This model permits a separation between ambiguity and ambiguity attitude, and allows smooth, rather than kinked, indifference curves. Both features are conceptually important and empirically useful. In addition, KMM (2005) show that this model includes the multiple-priors model of Gilboa and Schmeidler (1989) as a special case when ambiguity aversion goes to infinity under some technical regularity conditions.

Embedding their static model in a dynamic environment, KMM (2009a) develop a recursive smooth ambiguity preference model. This dynamic model suffers from a limitation that intertemporal substitution and attitudes toward risk or uncertainty are intertwined. This inflexibility limits its empirical applications and makes comparative statics of risk aversion hard to interpret.[5] For example, calibrating this model in a representative–agent consumption-based asset-pricing setting, Ju and Miao (2007) show that somewhat implausible parameter values are needed to explain the equity premium puzzle. By contrast, after separating out intertemporal substitution as in our generalized recursive smooth ambiguity model, Ju and Miao (forthcoming) show that the empirical performance improves significantly.

We summarize our preference model when restricted to the space of adapted consumption processes as follows. Consider an infinite-horizon setting and denote time by

[3] Roughly speaking, risk refers to situations where known probabilities are available to guide choices, while ambiguity refers to situations where probabilities are vague so that multiple probabilities may be available. Ambiguity aversion means that individuals dislike ambiguity.

[4] For alternative axiomatizations of an essentially identical functional form, see Chew and Sagi (2008), Ergin and Gul (2009), Nau (2006), and Seo (2009).

[5] See Epstein and Zin (1989) for an early discussion of the importance of the separation between intertemporal substitution and risk aversion in a pure risk setting.

$t = 0, 1, 2, \ldots$. The state space in each period is S. At time t, the decision maker's information consists of histories $s^t = \{s_1, s_2, \ldots, s_t\}$ with $s_0 \in S$ given and $s_t \in S$. The decision maker ranks adapted consumption plans $c = (c_t)_{t \geq 0}$. That is, c_t is a measurable function of s^t. His preferences are represented by the recursive form

$$V_{s^t}(c) = W\left(c_t, v^{-1}\left(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\left(\int_S u(V_{(s^t, s_{t+1})}(c))\, d\pi(s_{t+1})\right) d\mu_{s^t}(\pi)\right)\right), \tag{2}$$

where $V_{s^t}(c)$ is conditional utility or continuation value at history s^t, $W : \mathbb{R}^2 \to \mathbb{R}$ is a time aggregator, $\mathcal{P}_{s^t}$ is a set of one-step-ahead probability measures on S at history s^t, and μ_{s^t} is a probability measure over $\mathcal{P}_{s^t}$. The measure μ_{s^t} represents second-order beliefs about distributions governing one-step-ahead resolution of uncertainty. Given some assumptions similar to those in KMM (2009a), we show that μ_{s^t} is obtained by Bayesian updating from an initial prior.

When the set $\mathcal{P}_{s^t}$ consists of a set of conditional likelihood distributions $\pi_z(\cdot|s^t)$ indexed by an unknown parameter $z \in Z$, we use (2) to derive a model with learning,

$$V_{s^t}(c) = W\left(c_t, v^{-1}\left(\int_Z v \circ u^{-1}\left(\int_S u(V_{(s^t, s_{t+1})}(c))\, d\pi_z(s_{t+1}|s^t)\right) d\mu_{s^t}(z)\right)\right), \tag{3}$$

where $\mu_{s^t}(z)$ is the posterior distribution of z given s^t. More generally, the learning model in (3) allows z to be a hidden state that follows a Markov process because $\mathcal{P}_{s^t}$ can be history dependent.

Our generalized recursive smooth ambiguity model nests some popular models in the literature as special cases.

- The subjective version of the recursive expected utility model of Kreps and Porteus (1978) and Epstein and Zin (1989) is obtained by setting $v = u$ in (2). In this case, the two distributions μ_{s^t} and π can be reduced to a one-step-ahead predictive distribution:

$$p(s_{t+1}|s^t) = \int_{\mathcal{P}_{s^t}} \pi(s_{t+1})\, d\mu_{s^t}(\pi). \tag{4}$$

  This is the standard Bayesian approach that rules out ambiguity-sensitive behavior. If we further set $v(x) = u(x) = -\exp(-x/\theta)$, we obtain the multiplier preference model or the risk-sensitivity model discussed in Hansen and Sargent (2001).[6] Here θ is a robustness parameter, which enhances risk aversion.

- The generalized recursive multiple-priors model of Hayashi (2005) is obtained as the limit of (2) under some technical regularity conditions when ambiguity aver-

[6] This multiplier model is dynamically consistent according to the standard definition and the definition in this paper. Hansen and Sargent (2001, 2007b) also propose several other models of robustness. Some of them, e.g., "constraint preferences," are dynamically inconsistent according to the standard definition as pointed out by Epstein and Schneider (2003). However, the constraint preferences satisfy a different notion of dynamic consistency defined in Section 19.4 of Hansen and Sargent (2007b, pp. 407–412).

sion goes to infinity:

$$V_{s^t}(c) = W\left(c_t, u^{-1}\left(\min_{\pi \in \mathcal{P}_{s^t}} \int_S u(V_{(s^t, s_{t+1})}(c))\, d\pi(s_{t+1})\right)\right).$$

This model nests the recursive multiple-priors model of Epstein and Wang (1994) and Epstein and Schneider (2003) as a special case, as discussed in Hayashi (2005).

- The recursive smooth ambiguity model of KMM (2009a) has a discounted aggregator and takes the form

$$V_{s^t}(c) = u(c_t) + \beta\phi^{-1}\left(\int_Z \phi\left(\int_S V_{(s^t, s_{t+1})}(c)\, d\pi_z(s_{t+1}|s^t)\right) d\mu_{s^t}(z)\right).$$

The concavity of ϕ characterizes ambiguity aversion. The curvature of u describes both intertemporal substitution and risk aversion. Thus, they are intertwined.

- The multiplier preference model with hidden states of Hansen (2007) and Hansen and Sargent (2007a) is obtained by setting $W(c, y) = h(c) + \beta y$, $u(x) = -\exp(-x/\theta_1)$, and $v(x) = -\exp(-x/\theta_2)$, $\theta_1, \theta_2 > 0$, in (3). In this model, there are two risk-sensitivity adjustments. The first risk-sensitivity adjustment for the distribution $\pi_z(\cdot|s^t)$ reflects the decision maker's concerns about the misspecification in the conditional distribution given the parameter value z. The second risk-sensitivity adjustment for the distribution μ_{s^t} reflects the decision maker's concerns about the misspecification of the posterior distribution.

To provide an axiomatic foundation for the model in (2), we need to choose a suitable domain for preferences. As is well known from Kreps and Porteus (1978) and Epstein and Zin (1989), one needs to define a hierarchical domain of choices so as to separate intertemporal substitution from risk aversion. In our second-order act approach, we take the product space of current consumption and the continuation compound lottery acts as the primary preference domain. Hayashi (2005) first introduces the domain of compound lottery acts to provide an axiomatic foundation for a generalized recursive multiple-priors model. A compound lottery act is a random variable that maps today's state of the world into a joint lottery over current consumption and a compound lottery act for tomorrow. It is the dynamic counterpart of the horse-race roulette-wheel act introduced by Anscombe and Aumann (1963).

Our first axiomatic characterization consists of five standard axioms to deliver recursive expected utility under uncertainty and two additional axioms related to ambiguity. The five standard axioms deliver W and u in (2). The two additional axioms deliver v and μ_{s^t}. To pin down a unique v and a unique μ_{s^t}, we need more choices available to the decision maker. Because μ_{s^t} is a second-order probability measure over the first-order probability measures on S, to elicit this belief, it seems natural and intuitive to assume that choices contingent on the first-order probability measures are observable. These choices are modelled as second-order acts in KMM (2005) in a static setting. Extending their insight to our dynamic Anscombe–Aumann setting, we define a second-order act as a mapping that maps a probability measure on S to a compound consumption lottery. We then define auxiliary preferences over second-order acts and impose an axiom that these preferences are represented by subjective expected utility. This representation can be delivered by imposing additional primitive axioms from the standard subjective expected utility theory.[7]

As in KMM (2005, 2009b), we impose the last axiom that connects preferences over second-order acts and the original preferences over pairs of current consumption and continuation compound lottery acts. In doing so, we introduce the notion of a one-step-ahead act—a compound lottery act in which subjective uncertainty resolves in just one period. We then construct a second-order act associated with a one-step-ahead act that maps a probability measure on S to a compound lottery on the consumption space. This compound lottery is obtained by averaging out states in the one-step-ahead act using the probability measure on S. The last axiom says that the decision maker orders pairs of current consumption and the one-step-ahead act identically to the second-order acts associated with the one-step-ahead acts. The intuition is that the decision maker's ranking of the former choices reflects his uncertainty about the underlying distribution of the choices, which is the domain of the second-order acts.

One critique of the KMM (2005) model raised by Seo (2009) is that second-order acts and preferences over second-order acts are typically unobservable in the financial markets. For example, investors typically bet on the realization of stock prices, but not on the true distribution underlying stock prices. A similar critique applies to the Anscombe–Aumann acts as well because these acts are also unobservable in financial markets: the realizations of stock prices are monetary values, not lotteries. However, both Anscombe–Aumann acts and second-order acts are useful modelling devices and are available from laboratory and thought experiments.[8] More concretely, when measures in $\mathcal{P}_{s^t}$ correspond to conditional distributions indexed by an unknown parameter as in (3), the second-order acts are bets on the value of the parameter. In an asset pricing application studied by Ju and Miao (forthcoming), $\mathcal{P}_{s^t}$ consists of two distributions for consumption growth in a boom and in a recession so that the second-order acts are bets on the economic regime. In a portfolio choice application studied by Chen et al. (2009), $\mathcal{P}_{s^t}$ consists of two distributions for the possibly misspecified stock return models so that the second-order acts are simply bets on the statistical model of stock returns.

It is possible to dispense with the auxiliary domain of second-order acts following Seo's (2009) axioms. Building on his insight, we provide an alternative axiomatization for (2) without second-order acts. Adapting Seo's (2009) static setup, we introduce an

[7] In a recent critique of the static KMM model, Epstein (2010) argues that Ellsbergian choices on S should lead to Ellsbergian choices for second-order acts. In response, KMM (2009a) argue that second-order acts are modelling devices to deliver Ellsbergian choices on the state space S of primary interest. To accommodate Ellsbergian choices for second-order acts, one can simply expand the state space to incorporate measures on S.

[8] To further illustrate this point, we quote Kreps (1988, p. 101): "This procedure of enriching the set of items to which preference must apply is quite standard. It makes perfectly good sense in normative applications, *as long as* the Totrep involved is able to envision the extra objects and agree with the axiom applied to them. This need be no more than a thought experiment for Totrep, as long as he is willing to say that it is a valid (i.e., conceivable) thought experiment."

extra stage of randomization. As a by-product contribution, we construct a set of *two-stage compound lottery acts*, which allows for randomization both before and after the realization of the state of the world. We then define the product space of current consumption and the continuation lotteries over two-stage compound lottery acts as the single domain of preferences. We impose five axioms analogous to the first five axioms in the second-order act approach. We replace the last two axioms in that approach with a first-stage independence axiom and a dominance axiom adapted from Seo (2009). Given these seven axioms, we establish a dynamic version of Seo's static model. To the best of our knowledge, our paper provides the first dynamic extension of Seo's static model.

We should mention that each of our two different adopted axiomatic approaches is debatable. For example, some researchers (e.g., Seo 2009 and Epstein 2010) argue that second-order acts or preferences on these acts are either unobservable or may not be totally plausible. In the two-stage randomization approach, a failure of the reduction of compound lotteries may not be normatively appealing.

After providing axiomatic foundations, we characterize risk attitude and ambiguity attitude. Our characterization in the second-order act approach is similar to that of KMM (2005), suitably adapted to our dynamic setting with Anscombe–Aumann-type acts. In this approach, ambiguity aversion is associated with aversion to the variation of ex ante evaluations of one-step-ahead acts due to model uncertainty. In the two-stage randomization approach, we distinguish between attitudes toward risks in the two stages. We define absolute ambiguity aversion as an aversion to a first-stage mixture of acts before the realization of the state of the world compared to the second-stage mixture of these acts after the realization of the state. We show that this notion of ambiguity aversion is equivalent to risk aversion in the first stage. In particular, ambiguity aversion is associated with the violation of reduction of compound lotteries. We also show that in both approaches, risk attitude and ambiguity attitude are characterized by the shapes of the functions u and v, respectively.

The remainder of the paper proceeds as follows. Section 2 reviews the atemporal models of KMM (2005) and Seo (2009). Section 3 embeds the KMM (2005) model in a dynamic setting and axiomatizes it using the second-order act approach. Section 4 embeds the Seo (2009) model in a dynamic setting and axiomatizes it using the two-stage randomization approach. Section 5 applies our model to asset pricing. Section 6 discusses related literature. Appendices A–E contain proofs.

2. Review of the atemporal models

In this section, we provide a brief review of the atemporal models of ambiguity proposed by KMM (2005) and Seo (2009). We embed these models in a dynamic setting in Sections 3 and 4. Both atemporal models when restricted to the space of random consumption deliver an identical representation in (1). The two models differ in domain and axiomatic foundation. For both models, we take a complete, transitive, and continuous preference relation $\succeq$ as given.

Consider the KMM model first. KMM originally study Savage acts over $S \times [0, 1]$, where the auxiliary state space $[0, 1]$ is used to describe objective lotteries. Here we translate their model into the Anscombe–Aumann domain. Let S be the set of states, which is assumed to be finite for simplicity. Let $\mathcal{C}$ be a compact metric space and let $\Delta(\mathcal{C})$ be the set of lotteries over $\mathcal{C}$.[9] An Anscombe–Aumann act is defined as a mapping $g: S \to \Delta(\mathcal{C})$. Let $\mathcal{G}$ denote the set of all such acts. To pin down second-order beliefs, KMM introduce an auxiliary preference ordering $\succeq^2$ over second-order acts. A *second-order act is* a mapping $\mathfrak{g}: \mathcal{P} \to \Delta(\mathcal{C})$, where $\mathcal{P} \subset \Delta(S)$. Let $\mathcal{G}^2$ denote the set of all second-order acts.

The preference ordering $\succeq$ over $\mathcal{G}$ and the preference ordering $\succeq^2$ over $\mathcal{G}^2$ are represented by

$$U(g) = \int_{\mathcal{P}} \phi\left(\sum_{s \in S} \pi(s)\bar{u}(g(s))\right) d\mu(\pi) \quad \forall g \in \mathcal{G} \tag{5}$$

and

$$U^2(\mathfrak{g}) = \int_{\mathcal{P}} \phi(\bar{u}(\mathfrak{g}(\pi)))\, d\mu(\pi) \quad \forall \mathfrak{g} \in \mathcal{G}^2,$$

where $\phi: \bar{u}(\Delta(\mathcal{C})) \to \mathbb{R}$ is a continuous and strictly increasing function and $\bar{u}: \Delta(\mathcal{C}) \to \mathbb{R}$ is a mixture-linear function.[10]

The previous representation is characterized by the following axioms.[11] (i) The preference $\succeq$ satisfies the mixture-independence axiom over the set of constant acts $\Delta(\mathcal{C})$. (ii) The preference over second-order acts $\succeq^2$ is represented by the subjective expected utility of Savage (1954). (iii) The two preference relations $\succeq$ and $\succeq^2$ are consistent with each other in the sense that $g \succeq h$ if and only if $g^2 \succeq^2 h^2$, where g^2 is the second-order acts associated with g defined by $g^2(\pi) = \sum_{s \in S} g(s)\pi(s)$ for each $\pi \in \mathcal{P}$ and h^2 is defined similarly. The interpretation for the last axiom is the following. If the decision maker prefers f to g, then the average of f across states over all possible beliefs (distributions) should also be preferred to that of g. The reverse is also true.

The last two axioms are controversial as argued by Epstein (2010). To illustrate the plausibility of these axioms, consider the following example. Suppose there is an Ellsberg urn containing 90 balls. A decision maker is told that there are 30 black balls and 60 white or red balls in the urn. But he does not know the composition of white or red balls. There are four bets as in Table 1. The Ellsbergian choice is

$$g_1 \succ g_2 \text{ but } g_4 \succ g_3.$$

One justification is that the decision maker is unsure about the probabilities of white and red balls and is averse to this ambiguity.

Suppose there are two possible distributions over the set of ball color $S = \{\text{b}, \text{w}, \text{r}\}$: $\pi_1 = (1/3, 2/3, 0)$ and $\pi_2 = (1/3, 0, 2/3)$. Consider the second-order acts associated with

[9] We use the following notations and assumptions throughout the paper. Given a compact metric space Y, let $\mathcal{B}(Y)$ be the family of Borel subsets of Y and let $\Delta(Y)$ be the set of Borel probability measures defined over $\mathcal{B}(Y)$. Endow $\Delta(Y)$ with the weak convergence topology. Then $\Delta(Y)$ is a compact metric space.

[10] A function f is mixture linear on some set X if $f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda) f(y)$ for any $x, y \in X$ and any $\lambda \in [0, 1]$.

[11] The proof can be obtained from our proof of Theorem 1 in Appendix A.

	b	w	r
g_1	10	0	0
g_2	0	10	0
g_3	10	0	10
g_4	0	10	10

Table 1. This table shows four acts, g_1, g_2, g_3, and g_4, with payoffs contingent on three events {b}, {w}, and {r}.

	π_1	π_2
g_1^2	10/3	10/3
g_2^2	20/3	0
g_3^2	10/3	10
g_4^2	20/3	20/3

Table 2. This table shows four second-order acts, g_1^2, g_2^2, g_3^2, and g_4^2, associated with four acts g_1, g_2, g_3, and g_4, respectively. Their payoffs are contingent on two distributions π_1 and π_2.

g_i, $i = 1, \dots, 4$, $g_i^2(\pi_j) = \sum_{s \in \{\mathrm{b},\mathrm{w},\mathrm{r}\}} g_i(s)\pi_j(s)$, where $j = 1, 2$. We write their payoffs in Table 2. The previous consistency axiom implies that

$$g_1^2 \succ^2 g_2^2 \text{ but } g_4^2 \succ^2 g_3^2.$$

This behavior can be consistent with expected utility over second-order acts as long as the decision maker is risk averse, because g_1^2 and g_4^2 give sure outcomes, but g_2^2 and g_3^2 are risky bets. The intuition is that second-order acts average out uncertain states (ball color) by definition and such hedging may eliminate ambiguity (see Gilboa and Schmeidler 1989). So it is possible that the decision maker is ambiguity neutral for second-order acts, but ambiguity averse for bets on the Ellsberg urn. Of course, one can design thought experiments to display Ellsbergian choices for second-order acts, which are ruled out by the KMM model.

Seo (2009) provides a different axiomatic foundation for (5) by dispensing with the auxiliary set of second-order acts and the associated preferences over this set. He considers the domain of lotteries over Anscombe–Aumann acts, $\Delta(\mathcal{G})$, and a single preference relation $\succeq$ defined over it. Notice that by restricting attention to lotteries over constant acts, we have the domain of two-stage lotteries $\Delta(\Delta(\mathcal{C}))$ as a subset of $\Delta(\mathcal{G})$, and by further making the first-stage randomization degenerate, we have $\Delta(\mathcal{C})$ as a subset of $\Delta(\Delta(\mathcal{C}))$, hence of $\Delta(\mathcal{G})$ too.

Seo (2009) shows that the representation of preference $\succeq$ takes the form[12]

$$U(p) = \int_{\mathcal{G}} \int_{\mathcal{P}} \phi\left(\sum_{s \in S} \pi(s)\bar{u}(g(s))\right) d\mu(\pi)\, dp(g) \quad \forall p \in \Delta(\mathcal{G}). \tag{6}$$

[12] In Seo's (2009) original representation, he takes $\mathcal{P} = \Delta(S)$. When we adapt his dominance axiom for $\mathcal{P}$, we can allow $\mathcal{P}$ to be an arbitrary subset of $\Delta(S)$. For example, the proof in Seo's appendix gives an example of a finite set $\mathcal{P}$.

When p is degenerate at some $g \in \mathcal{G}$, (6) reduces to (5).

In Seo's approach, the representation is characterized by the following axioms. (i) The preference $\succeq$ satisfies the mixture-independence axiom on the set of one-stage lotteries $\Delta(\mathcal{C})$. (ii) The preference $\succeq$ satisfies the mixture-independence axiom on the set of lotteries over acts $\Delta(\mathcal{G})$. (iii) A dominance condition holds. To state the dominance axiom formally, we define a two-stage lottery $a(p, \pi) \in \Delta(\Delta(\mathcal{C}))$ induced by $p \in \Delta(\mathcal{G})$ and $\pi \in \mathcal{P}$ as $a(p, \pi)(B) = p(\{g \in \mathcal{G} : g^2(\pi) \in B\})$ for any Borel set B on $\Delta(\mathcal{C})$. Dominance says that for any $p, q \in \Delta(\mathcal{G})$, $p \succeq q$ if $a(p, \pi) \succeq a(p, \pi)$ for all $\pi \in \mathcal{P}$. Seo's approach does not deliver uniqueness of the second-order belief μ in general. For example, if ϕ is linear, then any μ with an identical mean $\int_{\mathcal{P}} \pi \, d\mu(\pi)$ yields the same ranking. It is unique in some special cases, for example, if ϕ is some exponential function. We refer to Seo (2009) for a characterization of the uniqueness of μ.

3. Axiomatization with second-order acts

We embed the atemporal KMM model reviewed in Section 2 in a discrete-time infinite-horizon environment. Time is denoted by $t = 0, 1, 2, \ldots$. Let S be a finite set of states at each period. The full state space is S^∞. Let $\mathcal{C}$ be a complete, separable, and compact metric space, which is the set of consumption choices in each period.

3.1 *Primary domain*

To introduce the domain of choices of primary interest, we consider the set of compound lottery acts introduced by Hayashi (2005). A compound lottery act is identified as a random variable that maps a state of the world into a joint lottery over consumption and a compound lottery act for the next period. Hayashi (2005) shows that the set of all such acts $\mathcal{G}$ satisfies the homeomorphic relation[13]

$$\mathcal{G} \simeq (\Delta(\mathcal{C} \times \mathcal{G}))^S.$$

It is a compact metric space with respect to the product metric. By abuse of notation, we may refer S to the set of states as well as its cardinality.

Up to homeomorphic transformations, the domain $\mathcal{G}$ of compound lottery acts includes subdomains $\mathcal{G}^*$, $\mathcal{F}$, $\mathcal{M}$, and $\mathcal{C}^\infty$, which are defined as follows.

- Adapted processes of consumption lotteries are

$$\mathcal{G}^* \simeq (\Delta(\mathcal{C}) \times \mathcal{G}^*)^S.$$

- Adapted processes of consumption levels are

$$\mathcal{F} \simeq (\mathcal{C} \times \mathcal{F})^S.$$

[13]Two topological spaces X and Y are called homeomorphic (denoted $X \simeq Y$) if there is a one-to-one continuous map f from X onto Y such that f^{-1} is continuous too. The map f is called a homeomorphism.

- Compound lotteries are

$$\mathcal{M} \simeq \Delta(\mathcal{C} \times \mathcal{M}).$$

- Deterministic consumption streams are

$$\mathcal{C}^\infty \simeq \mathcal{C} \times \mathcal{C}^\infty.$$

The subdomain $\mathcal{G}^*$ is obtained by randomizing current consumption only, but not over acts. This domain corresponds to the one adopted by Epstein and Schneider (2003). The subdomain $\mathcal{F}$ is obtained when there is no randomization. It is adopted by Wang (2003). The subdomain $\mathcal{M}$ is obtained by taking constant acts. Epstein and Zin (1989) define recursive utility under objective risk over the domain $\mathcal{C} \times \mathcal{M}$, while Chew and Epstein (1991) axiomatize this utility over the domain $\mathcal{M} \simeq \Delta(\mathcal{C} \times \mathcal{M})$. The space of deterministic consumption plans $\mathcal{C}^\infty$ is obtained by taking constant acts with no randomization.

Relations among these subdomains are expressed as

$$\begin{array}{ccccccc} \mathcal{G} & \supset & & \mathcal{G}^* & \supset & \mathcal{F} \\ \cup & & & \cup & & \cup \\ \mathcal{M} & \supset \Delta(\mathcal{C}^\infty) & \supset & (\Delta(\mathcal{C}))^\infty & \supset & \mathcal{C}^\infty. \end{array}$$

For any $c \in \mathcal{C}$, we use $\delta[c]$ to denote the degenerate lottery over c. When no confusion arises, we tend to omit the symbol of degenerate lottery and write down the deterministic component as it is. For example, a deterministic sequence $y = (c_0, c_1, c_2, \dots)$ is used as it is, instead of being denoted like $(c_0, \delta[(c_1, \delta[(c_2, \delta[\cdots])])])$.

3.2 *Preferences*

We consider two preference relations over two domains. Of primary interest are the decision maker's preferences $\succeq_{s^t}$ at each history s^t over pairs of current consumption and continuation compound lottery acts $\mathcal{C} \times \mathcal{G}$. Each pair is called a consumption plan. To recover the decision maker's second-order beliefs, we introduce another preference ordering over second-order acts. Take a set of one-step-ahead probability measures $\mathcal{P}_{s^t} \subset \Delta(S)$ as a primitive for each history s^t. A second-order act on $\mathcal{P}_{s^t}$ is a mapping $\mathfrak{f}: \mathcal{P}_{s^t} \to \mathcal{M}$. Let $\Im(\mathcal{P}_{s^t})$ denote the set of all the second-order acts on $\mathcal{P}_{s^t}$. Let $\succeq^2_{s^t}$ denote the conditional second-order preference defined over $\Im(\mathcal{P}_{s^t})$ at each history s^t.

3.3 *Axioms*

We start by introducing five standard axioms for the preference process $\{\succeq_{s^t}\}$. First, we assume weak order (complete and transitive), continuity, and sensitivity. This ensures the existence of a continuous functional representation of preference (see Debreu 1954).

Axiom A1 (Order). *For all t and s^t, $\succeq_{s^t}$ is a continuous weak order over $\mathcal{C} \times \mathcal{G}$ and there exist $y, y' \in \mathcal{C}^\infty$ such that $y \succ_{s^t} y'$.*

Second, we assume that preference over acts for the future is independent of current consumption. This axiom is adapted from Koopmans (1960) and is essential for the representation to have a form in which current consumption and continuation value are separable.

Axiom A2 (Current Consumption Separability). *For all t and s^t, for all $c, c' \in \mathcal{C}$ and $g, g' \in \mathcal{G}$,*

$$(c, g) \succeq_{s^t} (c, g') \iff (c', g) \succeq_{s^t} (c', g').$$

Third, we assume that preference over risky consumption is independent of history. This axiom ensures that utility is stationary (or time invariant) in the pure risk domain $\mathcal{C} \times \mathcal{M}$. It also implies that preference over deterministic consumption streams is independent of history.

Axiom A3 (History Independence of Risk Preference). *For all t, $\tilde{t}$ and s^t, $\tilde{s}^{\tilde{t}}$, for all $(c, m), (c', m') \in \mathcal{C} \times \mathcal{M}$,*

$$(c, m) \succeq_{s^t} (c', m') \iff (c, m) \succeq_{\tilde{s}^{\tilde{t}}} (c', m').$$

Fourth, we impose an independence axiom à la von Neumann and Morgenstern (vNM) for timeless gambles. This axiom is essential to have an expected utility representation in the pure risk domain.

Axiom A4 (Independence for Timeless Lotteries). *For all t and s^t, for all $m, m', n \in \mathcal{M}$ and $\lambda \in (0, 1)$,*

$$(c, m) \succeq_{s^t} (c, m') \iff (c, \lambda m + (1-\lambda)n) \succeq_{s^t} (c, \lambda m' + (1-\lambda)n).$$

Fifth, we impose dynamic consistency to connect conditional preferences across histories. It is essential to deliver a recursive form of utility representation. The idea is that if two plans give the same consumption today, but may differ in the continuation choices, then the plan that is preferred tomorrow is also preferred today. Because of our large choice domain, we need to define the notion of stochastic dominance so as to formulate our dynamic consistency condition.

Definition 1. Given $p, q \in \Delta(\mathcal{C} \times \mathcal{G})$, say that p stochastically dominates q with regard to $\succeq_{s^t}$ if

$$p\big(\{(c', g') \in \mathcal{C} \times \mathcal{G} \colon (c', g') \succeq_{s^t} (c, g)\}\big) \geq q\big(\{(c', g') \in \mathcal{C} \times \mathcal{G} \colon (c', g') \succeq_{s^t} (c, g)\}\big)$$

for all $(c, g) \in \mathcal{C} \times \mathcal{G}$. If, in addition, there is some $(c, g) \in \mathcal{C} \times \mathcal{G}$ such that $\geq$ is replaced with $>$, then we say p strictly stochastically dominates q.

When $p, q \in \Delta(\mathcal{C} \times \mathcal{G})$ stochastically dominate each other, we say that p and q are *stochastically equivalent* with regard to $\succeq_{s^t}$. Note that in the above definition, we allow p or q to be a measure on $\mathcal{C} \times \mathcal{M}$, say $p \in \Delta(\mathcal{C} \times \mathcal{M})$. In this case, we view $p \in \Delta(\mathcal{C} \times \mathcal{G})$ with the support $\mathcal{C} \times \mathcal{M}$.

Axiom A5 (Dynamic Consistency). *For all t and s^t, for all $c \in \mathcal{C}$ and $g, g' \in \mathcal{G}$, if $g(s)$ (strictly) stochastically dominates $g'(s)$ with regard to $\succeq_{s^t,s}$ for each $s \in S$, then $(c, g) \succeq_{s^t} (\succ_{s^t}) (c, g')$.*

Because we allow lotteries as outcomes of acts whereas preference at each period is defined over pairs of current consumption and continuation acts, our preceding formulation of dynamic consistency is more general than that in the literature (e.g., Epstein and Zin 1989, Epstein and Schneider 2003, Hayashi 2005, and KMM 2009a). When we restrict attention to smaller domains used in the literature, we obtain the standard definition. For example, suppose the choice domain is the adapted consumption processes $\mathcal{C} \times \mathcal{F}$ and the utility representation is given by (2). Our Axiom A5 implies the following: For all $c \in \mathcal{C}$ and $d, d' \in \mathcal{F} \simeq (\mathcal{C} \times \mathcal{F})^S$, if $d(s) \succeq_{s^t} (\succ_{s^t}) d'(s)$ for each $s \in S$, then $(c, d) \succeq_{s^t} (\succ_{s^t}) (c, d')$.

Now, we introduce two axioms on $\{\succeq^2_{s^t}\}$ so as to embed the atemporal KMM model in the dynamic setting. First, we follow KMM (2009a) and assume that the preference over second-order acts falls in the subjective expected utility (SEU) theory of Savage (1954), in which $\mathcal{P}_{s^t}$ is the state space and $\mathcal{M}$ is the set of pure outcomes.

Axiom A6 (SEU Representation of Preference Over Second-Order Acts). *For each s^t, there exists a unique countably additive probability measure $\mu_{s^t} : \mathcal{P}_{s^t} \to [0, 1]$ and a continuous and strictly increasing function $\psi : \mathcal{M} \to \mathbb{R}$ such that for all $\mathfrak{f}, \mathfrak{g} \in \Im(\mathcal{P}_{s^t})$,*

$$\mathfrak{f} \succeq^2_{s^t} \mathfrak{g} \quad \Longleftrightarrow \quad \int_{\mathcal{P}_{s^t}} \psi(\mathfrak{f}(\pi))\, d\mu_{s^t}(\pi) \geq \int_{\mathcal{P}_{s^t}} \psi(\mathfrak{g}(\pi))\, d\mu_{s^t}(\pi).$$

Moreover, ψ is unique up to a positive affine transformation if $\mu_{s^t}(J) \in (0, 1)$ for some $J \subset \mathcal{P}_{s^t}$.[14]

Second, we introduce an axiom that connects preference relations $\{\succeq_{s^t}\}$ and $\{\succeq^2_{s^t}\}$ using one-step-ahead acts and their corresponding second-order acts. A one-step-ahead act $g_{+1} \in \mathcal{G}$ is a compound lottery act in which subjective uncertainty resolves in just one period. Define the set of one-step-ahead acts as

$$\mathcal{G}_{+1} = \{g_{+1} \in \mathcal{G} \colon\ g_{+1}(s) \in \mathcal{M}, \forall s \in S\}.$$

Definition 2. Given a one-step-ahead act $g_{+1} \in \mathcal{G}_{+1}$, its corresponding second-order act on $\mathcal{P}_{s^t}$ is given by $g^2_{+1} : \mathcal{P}_{s^t} \to \mathcal{M}$, where

$$g^2_{+1}(\pi) = \sum_{s \in S} g_{+1}(s)\pi(s)$$

for each $\pi \in \mathcal{P}_{s^t}$.

[14] Because ψ is independent of history s^t in this axiom, we implicitly assume that $\succeq^2_{s^t}$ restricted to constant acts in $\Im(\mathcal{P}_{s^t})$ is independent of s^t.

The axiom below states that the preference $\{\succeq_{s^t}\}$ over the subdomain of one-step-ahead acts and the preference $\{\succeq^2_{s^t}\}$ over the subdomain of the corresponding second-order acts are consistent with each other.

Axiom A7 (Consistency With Preference Over Second-Order Acts). *For each s^t, for every $c \in \mathcal{C}$ and $g_{+1}, h_{+1} \in \mathcal{G}_{+1}$,*

$$(c, g_{+1}) \succeq_{s^t} (c, h_{+1}) \iff g^2_{+1} \succeq^2_{s^t} h^2_{+1}.$$

3.4 *Representation*

Now we state our first representation theorem.

Theorem 1 (Representation). *The preference process $\{\succeq_{s^t}, \succeq^2_{s^t}\}$ satisfies Axioms A1–A7 if and only if there exists representation $(\{V_{s^t}\}, W, u, v, \{\mu_{s^t}\})$ such that the following conditions are valid.*

(i) On $\mathcal{C} \times \mathcal{G}$, each $\succeq_{s^t}$ is represented by

$$V_{s^t}(c, g) = W\biggl(c, v^{-1}\biggl(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\biggl(\sum_{s\in S} \pi(s) \times \int_{\mathcal{C}\times\mathcal{G}} u(V_{s^t,s}(c', g'))\, dg(s)(c', g')\biggr)\, d\mu_{s^t}(\pi)\biggr)\biggr) \tag{8}$$

for each $(c, g) \in \mathcal{C} \times \mathcal{G}$, where $W : \mathcal{C} \times \mathbb{R} \to \mathbb{R}$ is continuous and strictly increasing in the second argument, and $u, v : \mathbb{R} \to \mathbb{R}$ are continuous and strictly increasing functions.[15]

(ii) On $\mathcal{C} \times \mathcal{M}$, each V_{s^t} coincides with

$$V(c, m) = W\biggl(c, u^{-1}\biggl(\int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')\biggr)\biggr) \quad \forall (c, m) \in \mathcal{C} \times \mathcal{M}. \tag{9}$$

(iii) On $\Im(\mathcal{P}_{s^t})$, each $\succeq^2_{s^t}$ is represented by the function

$$V^2_{s^t}(\mathfrak{g}) = \int_{\mathcal{P}_{s^t}} v \circ u^{-1} \circ \bar{u}(\mathfrak{g}(\pi))\, d\mu_{s^t}(\pi) \quad \forall \mathfrak{g} \in \Im(\mathcal{P}_{s^t}),$$

where $v \circ u^{-1} \circ \bar{u} = \psi$ and $\bar{u} : \mathcal{M} \to \mathbb{R}$ is a mixture-linear function

$$\bar{u}(m) = \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m') \quad \forall m \in \mathcal{M}. \tag{10}$$

In addition, we have the following uniqueness result, up to some monotonic transformations.

[15]Note that the domains of W, u, and v may be smaller than those specified in the theorem. We do not make this explicit so as to avoid introducing additional notations.

THEOREM 2 (Uniqueness). *Let* $\{\succeq_{s^t}, \succeq^2_{s^t}\}$ *satisfy Axioms A1–A7. If both* $(\{\tilde{V}_{s^t}\}, \tilde{W}, \tilde{u}, \tilde{v}, \{\tilde{\mu}_{s^t}\})$ *and* $(\{V_{s^t}\}, W, u, v, \{\mu_{s^t}\})$ *represent* $\{\succeq_{s^t}, \succeq^2_{s^t}\}$*, then there exist a strictly increasing function* Φ *and constants* $A > 0$ *and* B *such that for all* s^t*,* $\tilde{\mu}_{s^t} = \mu_{s^t}$ *and*

$$\tilde{V}_{s^t} = \Phi \circ V_{s^t}, \qquad \tilde{W}(\cdot, \cdot) = \Phi(W(\cdot, \Phi^{-1}(\cdot)))$$
$$\tilde{u} \circ \Phi = Au + B, \qquad \tilde{v} \circ \Phi = v.$$

Our representation in the theorem nests several models as special cases.

1. When there is no randomization (i.e., $g(s)$ is a degenerate lottery for all $s \in S$), the representation reduces to (2) on $\mathcal{C} \times \mathcal{F}$. As discussed in Section 1, many popular utility models are special cases of (2).

2. In the pure risk case, (9) is the recursive expected utility model of Kreps and Porteus (1978) and Epstein and Zin (1989).

3. In the deterministic case, the representation reduces to the Koopmans form

$$V(c, y) = W(c, V(y)), \tag{11}$$

where $(c, y) \in \mathcal{C} \times \mathcal{C}^\infty$.

Our preference model incorporates an information structure with hidden states. As an example, suppose z is a parameter taking values in a finite set Z. Let

$$\mathcal{P}_{s^t} = \{\pi_z(\cdot|s^t) \in \Delta(S) \colon z \in Z\}, \tag{12}$$

where $\pi_z(\cdot|s^t)$ is a conditional distribution on S given history s^t and the parameter value z. The distribution $\pi_z(\cdot|s^t)$ may be derived by the Bayes rule from a more primitive family of distributions $\{\pi_z\}_{z\in Z}$ on S^∞. Each π_z represents a statistical model. Then the representation in (8) reduces to (3) when $g(s)$ is a degenerate lottery for all $s \in S$.

Theorem 1 does not say anything about how measures μ_{s^t} at all histories s^t are related. In application, it is natural that μ_{s^t} is obtained by Bayesian updating from the initial prior $\mu = \mu_{s^0}$. To deliver this result, we consider (2) in the special case with (12). We follow KMM (2009a) and assume a full rank condition. Extending KMM's (2009a) Assumption 10, we introduce a marginal rate of substitution assumption for smooth functions W, u, and v:

$$\left.\frac{\partial V_{s^t}(c)/\partial c(s^{t+n})}{\partial V_{s^t}(c)/\partial c(s^t)}\right|_{c=\bar{c}} = \beta(\bar{c})^n \int_Z \pi_z(s^{t+n}|s^t)\, d\mu_{s^t}(z),$$

where $\pi_z(s^{t+n}|s^t)$ is the conditional probability of s^{t+n} given s^t, $\bar{c}$ is a constant consumption plan, and $c(s^{t+n})$ and $c(s^t)$ are consumption levels at histories s^{t+n} and s^t, respectively. In addition, $\beta(\bar{c}) = W_2(\bar{c}, \bar{V})$ is the discount factor, where $\bar{V} = W(\bar{c}, \bar{V})$. As in KMM (2009a), we can show that if the full rank condition holds, then the marginal rate of substitution assumption is equivalent to Bayesian updating of μ_{s^t}.

Theorem 1 does not say anything about the existence and uniqueness of a solution for $\{V_{s^t}\}$ to the recursive equation (8). Following a similar argument to that in the proof of Theorem 2 in KMM (2009a), we can show that $\{V_{s^t}\}$ exists. We need additional conditions for the uniqueness. Epstein and Zin (1989) provide sufficient conditions for recursive expected utility. KMM (2009a) give sufficient conditions for their recursive smooth ambiguity model. Marinacci and Montrucchio (2010) derive sufficient conditions for general recursive equations that may be applied to our model.

3.5 *Ambiguity attitude*

Because our model nests the deterministic case (11) and the pure risk case (9), we immediately deduce that the function W characterizes intertemporal substitution and the function u characterizes risk aversion in the usual way. We turn to the characterization of ambiguity aversion. We adopt the behavioral foundation of ambiguity attitude developed by Ghirardato and Marinacci (2002) and KMM (2005). Epstein (1999) provides a different foundation. The main difference is that the benchmark ambiguity neutral preference is the expected utility preference according to Ghirardato and Marinacci (2002), while Epstein's (1999) benchmark is the probabilistic sophisticated preference.

We first consider absolute ambiguity aversion. According to our first axiomatization, ambiguity comes from the multiplicity of distributions in the set $\mathcal{P}_{s^t}$. The decision maker's ambiguity attitude is toward uncertainty about the possible distributions in $\mathcal{P}_{s^t}$. To characterize this attitude, we define the lottery $m(g_{+1}, \mu_{s^t}) \in \mathcal{M}$ associated with the one-step-ahead act g_{+1} and the second-order belief μ_{s^t} on $\mathcal{P}_{s^t}$ as

$$m(g_{+1}, \mu_{s^t}) = \int_{\mathcal{P}_{s^t}} \sum_{s\in S} g_{+1}(s)\pi(s)\, d\mu_{s^t}(\pi).$$

Since $\sum_{s\in S} g_{+1}(s)\pi(s)$ is the outcome of the second-order act $g^2_{+1}(\pi)$ associated with g_{+1}, $m(g_{+1}, \mu_{s^t})$ is simply the mean value of g^2_{+1} with respect to the second-order belief μ_{s^t}. Alternatively, from the definition of predictive distribution in (4), we observe that the lottery $m(g_{+1}, \mu_{s^t})$ is also the mean value of the act g_{+1} with respect to the predictive distribution induced by μ_{s^t}. The following definition of ambiguity aversion states that the decision maker is ambiguity averse if he prefers a sure lottery obtained as the mean value of a given act to the act itself.

Definition 3. The decision maker with $\{\succeq_{s^t}\}$ exhibits ambiguity aversion if for all s^t, for all $c \in \mathcal{C}$ and $g_{+1} \in \mathcal{G}_{+1}$,

$$(c, m(g_{+1}, \mu_{s^t})) \succeq_{s^t} (c, g_{+1}).$$

Similarly to this definition, we can define ambiguity loving and ambiguity neutrality in the usual way. An immediate consequence of this definition is the following proposition.

Proposition 1. *Suppose* $\{\succeq_{s^t}\}$ *satisfies Axioms A1–A7. Then* $\{\succeq_{s^t}\}$ *exhibits ambiguity aversion if* $\phi \equiv v \circ u^{-1}$ *is concave.*[16]

[16] If $v \circ u^{-1}$ is concave, it is easy to check that $\{\succeq_{s^t}\}$ satisfies the uncertainty aversion axiom of Gilboa and Schmeidler (1989).

The proof of this proposition is straightforward and is omitted. Clearly, when $v \circ u^{-1}$ is linear, $\{\succeq_{s^t}\}$ displays ambiguity neutrality. Thus, the ambiguity neutrality benchmark is the recursive expected utility model. We need additional conditions to establish the converse statement that ambiguity aversion implies concavity of $v \circ u^{-1}$. The reason is that, to prove this statement, one needs to know preferences over binary bets on some $\mathcal{P}_{s^t}$, but our axioms and representation hold only for fixed $\mathcal{P}_{s^t}$. To deal with this issue in the KMM model, KMM (2005) consider a family of preference relations indexed by rich supports of second-order beliefs, and impose an assumption that ambiguity attitude and risk attitude are invariant across these supports (see their Assumption 4). We can adapt their assumption to establish the converse statement. Since the proof is similar to the proof of Proposition 1 in their paper, we omit it here.

We now turn to comparative ambiguity aversion.

Definition 4. Let the representations of the preferences of persons i and j share the same second-order belief μ_{s^t} on the same support $\mathcal{P}_{s^t}$ for all s^t. Say that $\{\succeq^i_{s^t}\}$ is more ambiguity averse than $\{\succeq^j_{s^t}\}$ if for all s^t, for all $c \in \mathcal{C}$, $m \in \mathcal{M}$, and $g_{+1} \in \mathcal{G}_{+1}$,

$$(c, m) \succeq^j_{s^t} (c, g_{+1}) \implies (c, m) \succeq^i_{s^t} (c, g_{+1}),$$

and if this property also holds for strict preference relations $\succ^i_{s^t}$ and $\succ^j_{s^t}$.

The interpretation of this definition is similar to that of Definition 5 in KMM (2005). The idea is that if person i prefers a lottery over an uncertain act whenever person j does so, then this must be due to person i's comparatively higher aversion to uncertainty. This cannot be due to aversion to risk, because the act g_{+1} itself may be a lottery and the conditions in the definition imply that persons i and j rank lotteries in the same way. Because the difference in beliefs is ruled out in the definition, the behavior in the definition must be due to differences in ambiguity attitude. The following proposition is a partial characterization. We omit its straightforward proof.

Proposition 2. *Suppose $\{\succeq^i_{s^t}\}$ and $\{\succeq^j_{s^t}\}$ satisfy Axioms A1–A7 and their representations share the same second-order belief μ_{s^t} on the same support $\mathcal{P}_{s^t}$ for all s^t. Then $\{\succeq^i_{s^t}\}$ is more ambiguity averse than $\{\succeq^j_{s^t}\}$ if there exist corresponding utility representations such that $V^i|_{\mathcal{C}\times\mathcal{M}} = V^j|_{\mathcal{C}\times\mathcal{M}}$, $W^i = W^j$, $u^i = u^j$, and $v^i = \Psi \circ v^j$, where Ψ is a strictly increasing and concave function.*

As in the case of absolute ambiguity aversion, one needs more information to establish the converse statement that comparative ambiguity aversion implies concavity of Ψ. As discussed earlier, we may make an assumption similar to Assumption 4 in KMM (2005) to establish this statement.

4. Axiomatization with two-stage compound lottery acts

To embed Seo's (2009) atemporal model in a dynamic setting, we adapt his atemporal domain—the set of lotteries over Anscombe–Aumann acts—to a dynamic setting. This leads us to consider the set of two-stage compound lottery acts.

4.1 *Domain*

We consider preference relations $\succeq_{s^t}$ at each history s^t defined on the domain $\mathcal{C} \times \Delta(\mathcal{H})$, where $\mathcal{H}$ is a set of two-stage compound lottery acts constructed as follows. Inductively define the family of sets $\{\mathcal{H}_0, \mathcal{H}_1, \dots\}$ by

$$\mathcal{H}_0 = (\Delta(\mathcal{C}))^S$$
$$\mathcal{H}_1 = \big(\Delta(\mathcal{C} \times \Delta(\mathcal{H}_0))\big)^S$$
$$\vdots$$
$$\mathcal{H}_t = \big(\Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))\big)^S$$

and so on. By induction, $\Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$ and $\mathcal{H}_t$ are compact metric spaces, for every $t \geq 1$. Let $\mathcal{H}^* = \prod_{t=0}^{\infty} \mathcal{H}_t$. It is a compact metric space with respect to the product metric.

We consider sequences of acts $(h_0, h_1, h_2, \dots)$ in $\mathcal{H}^*$ that are *coherent*. That is, h_t and h_{t+1} must be consistent for all $t \geq 0$ in the sense that is made precise in Appendix B. The domain of coherent acts, a subset of $\mathcal{H}^*$, is denoted by $\mathcal{H}$. The details of the definition of coherent acts and formal construction of the domain are given in Appendix B. The domain $\mathcal{H}$ satisfies a homeomorphic property analogous to those shown in Epstein and Zin (1989), Chew and Epstein (1991), Wang (2003), Gul and Pesendorfer (2004), and Hayashi (2005).

Theorem 3. *The set $\mathcal{H}$ is homeomorphic to $(\Delta(\mathcal{C} \times \Delta(\mathcal{H})))^S$, denoted as*

$$\mathcal{H} \simeq \big(\Delta(\mathcal{C} \times \Delta(\mathcal{H}))\big)^S.$$

When attention is restricted to constant acts, we obtain the subdomain consisting of two-stage compound lotteries, which satisfies the homeomorphism

$$\mathcal{L} \simeq \Delta(\mathcal{C} \times \Delta(\mathcal{L})).$$

Relations among the domains defined so far are summarized as

$$\begin{array}{ccccccccc}
\mathcal{H} \supset & \mathcal{G} & & \supset & & \mathcal{G}^* & \supset & \mathcal{F} \\
\cup & \cup & & & & \cup & & \cup \\
\mathcal{L} \supset & \mathcal{M} & \supset & \Delta(\mathcal{C}^\infty) & \supset & (\Delta(\mathcal{C}))^\infty & \supset & \mathcal{C}^\infty.
\end{array}$$

In particular, the set of compound lottery acts $\mathcal{G}$ and the set of compound lotteries $\mathcal{M}$ studied in Section 3 are subsets of $\mathcal{H}$ and $\mathcal{L}$, respectively.

We now introduce some useful notations. For any two-stage compound lottery acts, $h, h' \in \mathcal{H}$, and any $\lambda \in (0,1)$, we use $\lambda h + (1-\lambda)h' \in \Delta(\mathcal{H})$ to denote a lottery that gives h with probability λ and h' with probability $1-\lambda$. We use $\lambda h \oplus (1-\lambda)h' \in \mathcal{H}$ to denote a statewise mixture. That is, for each $s \in S$ and each Borel set $B \in \mathcal{B}(\mathcal{C} \times \Delta(\mathcal{H}))$, $\lambda h \oplus (1-\lambda)h'(s)(B) = \lambda h(s)(B) + (1-\lambda)h'(s)(B)$. For any $p, q \in \Delta(\mathcal{H})$, $\lambda p + (1-\lambda)q \in \Delta(\mathcal{H})$ denotes the usual mixture.

4.2 *Axioms*

We impose the following axioms on the preference process $\{\succeq_{s^t}\}$. The first three axioms are analogous to Axioms A1–A3.

Axiom B1 (Order). *For all t and s^t, $\succeq_{s^t}$ is a continuous weak order over $\mathcal{C} \times \Delta(\mathcal{H})$ and there exist $y, y' \in \mathcal{C}^\infty$ such that $y \succ_{s^t} y'$.*

Axiom B2 (Current Consumption Separability). *For all t and s^t, for all $c, c' \in \mathcal{C}$ and $p, q \in \Delta(\mathcal{H})$,*

$$(c, p) \succeq_{s^t} (c, q) \quad \Longleftrightarrow \quad (c', p) \succeq_{s^t} (c', q).$$

Axiom B3 (History Independence of Risk Preference). *For all t, $\tilde{t}$ and s^t, $\tilde{s}^{\tilde{t}}$, for all $(c, a), (c', a') \in \mathcal{C} \times \Delta(\mathcal{L})$,*

$$(c, a) \succeq_{s^t} (c', a') \quad \Longleftrightarrow \quad (c, a) \succeq_{\tilde{s}^{\tilde{t}}} (c', a').$$

Next, we assume independence conditions for "timeless" gambles, similar to Axiom A4. There are two kinds of such timeless gambles here: One is made before the realization of the one-step-ahead subjective uncertainty, and the other is made after that.

Axiom B4 (First-stage Independence). *For all t and s^t, for all $p, q, r \in \Delta(\mathcal{H})$ and $\lambda \in (0, 1)$,*

$$(c, p) \succeq_{s^t} (c, q) \quad \Longleftrightarrow \quad (c, \lambda p + (1-\lambda)r) \succeq_{s^t} (c, \lambda q + (1-\lambda)r).$$

Axiom B5 (Second-Stage Independence). *For all t and s^t, for all $c \in \mathcal{C}$, for all $l, m, n \in \Delta(\mathcal{C} \times \Delta(\mathcal{H}))$ and $\lambda \in (0, 1)$,*

$$(c, \delta[l]) \succeq_{s^t} (c, \delta[m]) \quad \Longleftrightarrow \quad \big(c, \delta[\lambda l \oplus (1-\lambda)n]\big) \succeq_{s^t} \big(c, \delta[\lambda m \oplus (1-\lambda)n]\big).$$

To connect preferences across histories, we impose a dynamic consistency axiom, similar to Axiom A5.

Definition 5. Given $a, b \in \Delta(\mathcal{C} \times \Delta(\mathcal{H}))$, say that a stochastically dominates b with regard to $\succeq_{s^t}$ if

$$a\big(\{(c', p') \in \mathcal{C} \times \Delta(\mathcal{H}) \colon (c', p') \succeq_{s^t} (c, p)\}\big) \ge b\big(\{(c', p') \in \mathcal{C} \times \Delta(\mathcal{H}) \colon (c', p') \succeq_{s^t} (c, p)\}\big)$$

for all $(c, p) \in \mathcal{C} \times \Delta(\mathcal{H})$. If in addition there is some $(c, p) \in \mathcal{C} \times \Delta(\mathcal{H})$ such that $\ge$ is replaced with $>$, then we say a strictly stochastically dominates b. If a and b stochastically dominate each other, we say that a and b are stochastically equivalent with regard to $\succeq_{s^t}$.

Note that in this definition, we allow a or b to be a measure on $\mathcal{C} \times \Delta(\mathcal{L})$, say $a \in \Delta(\mathcal{C} \times \Delta(\mathcal{L}))$. In this case, we view $a \in \Delta(\mathcal{C} \times \Delta(\mathcal{H}))$ with the support $\mathcal{C} \times \Delta(\mathcal{L})$.

Axiom B6 (Dynamic Consistency). *For all t and s^t, for all $c \in \mathcal{C}$ and $h, h' \in \mathcal{H}$, if $h(s)$ (strictly) stochastically dominates $h'(s)$ with regard to $\succeq_{s^t,s}$ for each $s \in S$, then $(c, \delta[h]) \succeq_{s^t} (\succ_{s^t})\ (c, \delta[h'])$.*

Finally, we embed Seo's (2009) dominance axiom to the set of one-step-ahead acts. A one-step-ahead act is an act for which subjective uncertainty resolves in just one period. We define the set of one-step-ahead acts as

$$\mathcal{H}_{+1} = \{h_{+1} \in \mathcal{H}\colon\ h_{+1}(s) \in \mathcal{L}, \forall s \in S\}.$$

Definition 6. Given $h_{+1} \in \mathcal{H}_{+1}$ and $\pi \in \Delta(S)$, define $l(h_{+1}, \pi) \in \mathcal{L}$ by

$$l(h_{+1}, \pi) = \sum_{s \in S} h_{+1}(s)\pi(s).$$

Given $p_{+1} \in \Delta(\mathcal{H}_{+1})$ and $\pi \in \Delta(S)$, define $a(p_{+1}, \pi) \in \Delta(\mathcal{L})$ by

$$a(p_{+1}, \pi)(L) = p_{+1}\bigl(\{h_{+1} \in \mathcal{H}_{+1}\colon\ l(h_{+1}, \pi) \in L\}\bigr)$$

for every Borel subset $L \subset \mathcal{L}$.

We take a set of one-step-ahead probability measures, $\mathcal{P}_{s^t}$, as given for each history s^t and impose the following dominance axiom on this set. We allow this set to be different from $\Delta(S)$ to permit more flexibility in applications as discussed in Section 1.

Axiom B7 (Dominance). *For all t and s^t, for all $c \in \mathcal{C}$ and $p_{+1}, p'_{+1} \in \Delta(\mathcal{H}_{+1})$,*

$$(c, a(p_{+1}, \pi)) \succeq_{s^t} (c, a(p'_{+1}, \pi)) \quad \forall \pi \in \mathcal{P}_{s^t} \quad \Longrightarrow \quad (c, p_{+1}) \succeq_{s^t} (c, p'_{+1}),$$

where $\mathcal{P}_{s^t} \subset \Delta(S)$.

To interpret this axiom, imagine that $\mathcal{P}_{s^t}$ is a set of probability distributions, which contains the "true" distribution unknown to the decision maker. Given the same current consumption c, if the decision maker prefers the continuation two-stage lottery $a(p_{+1}, \pi)$ induced by p_{+1} over another one $a(p'_{+1}, \pi)$ induced by p'_{+1} for each probability distribution $\pi \in \mathcal{P}_{s^t}$, then he must also prefer (c, p_{+1}) over (c, p'_{+1}).

Compared to the axioms in Section 3, First-Stage Independence (Axiom B4) and Dominance (Axiom B7) are the counterparts of the SEU Representation of Preference Over Second-Order Acts (Axiom A6) and Consistency With Preference Over Second-Order Acts (Axiom A7). Thus, we can dispense with second-order acts.

4.3 *Representation*

The following theorem gives our second representation result.

Theorem 4 (Representation). *The preference process* $\{\succeq_{s^t}\}$ *satisfies Axioms B1–B7 if and only if there exists a family of functions* $(\{V_{s^t}\}, W, u, v)$ *and a process of probability measures* $\{\mu_{s^t}\}$ *over* $\mathcal{P}_{s^t}$ *such that for each* s^t, *the function* $V_{s^t} : \mathcal{C} \times \Delta(\mathcal{H}) \to \mathbb{R}$ *represents* $\succeq_{s^t}$ *and has the form*

$$V_{s^t}(c,p) = W\Bigg(c, v^{-1}\Bigg(\int_{\mathcal{H}}\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\Bigg(\sum_{s\in S}\pi(s) \times \int_{\mathcal{C}\times\Delta(\mathcal{H})} u(V_{s^t,s}(c',a'))\,dh(s)(c',a')\Bigg)\,d\mu_{s^t}(\pi)\,dp(h)\Bigg)\Bigg) \tag{13}$$

for $(c,p) \in \mathcal{C}\times\Delta(\mathcal{H})$, *where* W *is continuous and strictly increasing in the second argument, and* u *and* v *are continuous and strictly increasing.*

We also have the following uniqueness result, up to some monotonic affine transformations.

Theorem 5 (Uniqueness). *Let* $\{\succeq_{s^t}\}$ *satisfy Axioms B1–B7. If both* $(\{\tilde{V}_{s^t}\}, \tilde{W}, \tilde{u}, \tilde{v}, \{\tilde{\mu}_{s^t}\})$ *and* $(\{V_{s^t}\}, W, u, v, \{\mu_{s^t}\})$ *represent* $\{\succeq_{s^t}\}$, *then there exist a strictly increasing function* Φ *and constants* A, B, D, E *with* $A, D > 0$, *such that*

$$\tilde{V}_{s^t} = \Phi \circ V_{s^t}, \qquad \tilde{W}(\cdot,\cdot) = \Phi(W(\cdot, \Phi^{-1}(\cdot)))$$
$$\tilde{u}\circ\Phi = Au + B, \qquad \tilde{v}\circ\Phi = Dv + E.$$

As in the static model of Seo (2009), the process of second-order beliefs $\{\mu_{s^t}\}$ is not unique in general. For example, when $\phi = v \circ u^{-1}$ is linear, $\{\mu_{s^t}\}$ is indeterminate. It is unique if ϕ is some exponential function. The existence of a solution for $\{V_{s^t}\}$ to the recursive equation (13) follows a similar argument in the proof of Theorem 2 in KMM (2009a). We may apply sufficient conditions in Marinacci and Montrucchio (2010) to establish uniqueness. The following list shows how the above model nests the existing models.

1. On the subdomain $\mathcal{C}\times\mathcal{G}$, the representation reduces to (8), which further reduces to (2) on $\mathcal{C}\times\mathcal{F}$.

2. On the subdomain $\mathcal{C}\times\Delta(\mathcal{L})$, we obtain a pure risk setting where the two-stage randomization is present. In this case, each V_{s^t} coincides with the common representation

$$V(c,a) = W\Bigg(c, v^{-1}\Bigg(\int_{\mathcal{L}} v\circ u^{-1}\Bigg(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c',a'))\,dl(c',a')\Bigg)\,da(l)\Bigg)\Bigg), \tag{14}$$

where $(c,a)\in\mathcal{C}\times\Delta(\mathcal{L})$.

3. On the subdomain $\mathcal{C}\times\mathcal{M}$, we obtain a pure risk setting where only the second-stage randomization is present. In this case, the model reduces to (9).

4.4 *Risk aversion and ambiguity aversion*

As discussed before, the function W describes intertemporal substitution. Now, we discuss how ambiguity aversion is separated from risk aversion in the two-stage randomization approach. We begin by characterizing risk aversion. In doing so, we restrict attention to the subdomain $\mathcal{C} \times \Delta(\mathcal{L})$ without subjective uncertainty. In this case, the utility representation takes the form in (14). Because there is two-stage randomization, we have two risk attitudes toward the risk in the two stages (or in the first order and the second order).

For the risk in the second stage, we remove the first-stage risk by assuming that the first-stage lottery is degenerate. We then obtain the representation of recursive risk preference given in (9). We can define risk aversion in the second stage in a standard way and show that it is completely characterized by the concavity of u.

Turn to risk aversion in the first stage. We define absolute risk aversion in the first stage as follows.

Definition 7. The decision maker with preference $\{\succeq_{s^t}\}$ exhibits risk aversion in the first stage if for all s^t, $c \in \mathcal{C}$ and $l, l' \in \mathcal{L}$, $\lambda \in [0, 1]$,

$$\big(c, \delta[\lambda l \oplus (1-\lambda) l']\big) \succeq_{s^t} (c, \lambda\delta[l] + (1-\lambda)\delta[l']). \tag{15}$$

We can similarly define risk loving and risk neutrality in the first stage. In Definition 7, $\lambda\delta[l] + (1-\lambda)\delta[l'] \in \Delta(\mathcal{L})$ represents a lottery in the first stage and $\delta[\lambda l \oplus (1-\lambda) l']$ represents a degenerate lottery over the mixture $\lambda l \oplus (1-\lambda) l'$ in the second stage. According to this definition, the decision maker may not be indifferent between these two lotteries, even though they give the same final outcome distribution. In particular, if the decision maker believes that the degenerate lottery is like a sure outcome and must be preferred, then he displays risk aversion in the first stage.

Note that if we replace $\succeq_{s^t}$ with $\sim_{s^t}$ in (15), we obtain a dynamic counterpart of Seo's (2009) Reduction of Compound Lotteries axiom. Thus, according to our Definition 7, violation of the Reduction of Compound Lotteries reflects the decision maker's attitude toward the risk in the first stage. The following proposition characterizes this risk attitude.

Proposition 3. *Suppose $\{\succeq_{s^t}\}$ satisfies Axioms B1–B7. Then $\{\succeq_{s^t}\}$ exhibits risk aversion in the first-stage if and only if $v \circ u^{-1}$ is concave.*

An immediate corollary of this proposition is that, given Axioms B1–B7, the Reduction of Compound Lotteries axiom is satisfied if and only if $v \circ u^{-1}$ is a strictly increasing affine function. In this case, the two lotteries l and a in (14) can be reduced to a compound lottery and hence (14) reduces to a model belonging to the class of recursive expected utility under objective risk.

Next, we consider comparative risk aversion.

Definition 8. Say that $\{\succeq^i_{s^t}\}$ is more risk averse than $\{\succeq^j_{s^t}\}$ in the first stage if for all s^t, $c \in \mathcal{C}$, $l \in \mathcal{L}$, and $a \in \Delta(\mathcal{L})$,

$$(c, \delta[l]) \succeq^j_{s^t} (c, a) \quad \Longrightarrow \quad (c, \delta[l]) \succeq^i_{s^t} (c, a),$$

and if this property also holds true for strict preference relations $\succ^j_{s^t}$ and $\succ^i_{s^t}$.

Take current consumption c as given. Suppose person j prefers a "sure" outcome (with the outcome being a lottery) to an arbitrary lottery for tomorrow. This must be due to j's aversion to risk. Facing the same choices, if person i is more risk averse than person j in the first stage, then person i should dislike what person j dislikes.

Proposition 4. *Suppose $\{\succeq^i_{s^t}\}$ and $\{\succeq^j_{s^t}\}$ satisfy Axioms B1–B7. Then $\{\succeq^i_{s^t}\}$ is more risk averse than $\{\succeq^j_{s^t}\}$ in the first stage if and only if there exist corresponding utility representations such that $V^i|_{\mathcal{C}\times\Delta(\mathcal{L})} = V^j|_{\mathcal{C}\times\Delta(\mathcal{L})}$, $W^i = W^j$, $u^i = u^j$, and $v^i = \Psi \circ v^j$, where Ψ is a strictly increasing and concave transformation.*

By Definition 8, persons i and j rank deterministic consumption plans in the same way and rank lotteries in the second stage in the same way. Thus, (W^i, u^i) and (W^j, u^j) are ordinally equivalent. Proposition 4 shows that person i is more risk averse than person j in the first stage if and only if v^i is a monotone concave transformation of v^j.

Now, we consider ambiguity attitude. Because ambiguity attitude deals with subjective uncertainty, we focus on the subdomain $\mathcal{C} \times \Delta(\mathcal{H}_{+1})$ in which uncertainty resolves in just one period. We define absolute ambiguity aversion as follows.

Definition 9. The decision maker with $\{\succeq_{s^t}\}$ exhibits ambiguity aversion if for all s^t, $c \in \mathcal{C}$, $h_{+1}, h'_{+1} \in \mathcal{H}_{+1}$, and $\lambda \in [0, 1]$,

$$\big(c, \delta[\lambda h_{+1} \oplus (1-\lambda)h'_{+1}]\big) \succeq_{s^t} (c, \lambda\delta[h_{+1}] + (1-\lambda)\delta[h'_{+1}]). \tag{16}$$

We can similarly define ambiguity loving and ambiguity neutrality. Definition 9 says that if a first-stage mixture of acts is preferred to their second-stage mixture, then the decision maker is ambiguity averse. The intuition for this definition is that hedging across ambiguous states is valuable compared to randomization of acts before the realization of the states. It is related to Gilboa and Schmeidler's (1989) definition of ambiguity aversion, which states that hedging across states for two indifferent acts is valuable to an ambiguity averse decision maker.[17] When $\succeq_{s^t}$ is replaced with $\sim_{s^t}$ in (16), then it becomes the dynamic counterpart of Seo's Reversal of Order axiom. Thus, ambiguity attitude is associated with the violation of the Reversal of Order axiom.

An example taken from Seo (2009) illustrates Definition 9. Restrict attention to a static setting. Consider an Ellsberg urn that contains 100 black or white balls, but the exact composition is unknown. The state of the world is the color of the ball. Let f be the

[17] Given Axiom B4 (First-Stage Independence), our definition implies the following Gilboa and Schmeidler definition of ambiguity aversion: $(c, \delta[h_{+1}]) \sim_{s^t} (c, \delta[h'_{+1}]) \Longrightarrow (c, \delta[\lambda h_{+1} \oplus (1-\lambda)h'_{+1}]) \succeq_{s^t} (c, \delta[h_{+1}])$ for all s^t, $c \in \mathcal{C}$, $h_{+1}, h'_{+1} \in \mathcal{H}_{+1}$, and $\lambda \in [0, 1]$.

act that gives \$100 if the chosen ball is black and nothing otherwise. Let g be the act that gives \$100 if the chosen ball is white and nothing otherwise. Let p be a lottery with 50% chance of winning \$100. Experimental evidence reveals that most people are indifferent between f and g, but prefer p to f and p to g. The first-stage mixture $\frac{1}{2}f + \frac{1}{2}g$ is still an ambiguous act. But the second-stage mixture $\frac{1}{2}f \oplus \frac{1}{2}g$ gives an identical lottery p no matter whether the chosen ball is black or white. Thus, it is intuitive that an ambiguity averse decision maker prefers $\frac{1}{2}f \oplus \frac{1}{2}g$ to $\frac{1}{2}f + \frac{1}{2}g$.

As Seo (2009) and Segal (1987, 1990) point out, ambiguity attitude is associated with violation of the Reduction of Compound Lotteries.[18] We now characterize this relationship. In his atemporal model, Seo (2009) shows that Reduction of Compound Lotteries and Reversal of Order are equivalent under Dominance. Adapting his argument to our dynamic two-stage compound lottery acts framework, we show below that ambiguity aversion is identical to risk aversion in the first stage.

Proposition 5. *Suppose $\{\succeq_{s^t}\}$ satisfies Axioms B1–B7. Then $\{\succeq_{s^t}\}$ exhibits ambiguity aversion if and only if $\{\succeq_{s^t}\}$ exhibits risk aversion in the first stage.*

An immediate implication of this proposition is that ambiguity aversion is equivalent to concavity of $v \circ u^{-1}$. In addition, the decision maker is ambiguity neutral if and only if $v \circ u^{-1}$ is a strictly increasing affine function. As a result, the four distributions h, π, μ_{s^t}, and p can be reduced to a compound distribution and the model reduces to recursive expected utility under uncertainty.

Finally, we study comparative ambiguity aversion.

Definition 10. Let the utility representations of $\{\succeq^i_{s^t}\}$ and $\{\succeq^j_{s^t}\}$ share the same second-order belief μ_{s^t} on the same support $\mathcal{P}_{s^t}$. Say that $\{\succeq^i_{s^t}\}$ is more ambiguity averse than $\{\succeq^j_{s^t}\}$ if for all s^t, all $c \in \mathcal{C}$, $l \in \mathcal{L}$, and $h_{+1} \in \mathcal{H}_{+1}$,

$$(c, \delta[l]) \succeq^j_{s^t} (c, \delta[h_{+1}]) \quad \Longrightarrow \quad (c, \delta[l]) \succeq^i_{s^t} (c, \delta[h_{+1}]),$$

and if this property also holds true for strict preference relations $\succ^j_{s^t}$ and $\succ^i_{s^t}$.

To interpret this definition, fix current consumption at c and consider two sure outcomes for tomorrow, with one outcome being a lottery and the other outcome being a one-step-ahead act. Suppose person j prefers the sure lottery outcome to the sure one-step-ahead act. This must be due to person j's aversion to subjective uncertainty or ambiguity. Facing the same choices, if person i dislikes what person j dislikes, then person i must be more ambiguity averse than person j because differences in beliefs are ruled out.

The following proposition states that in the framework of two-stage randomization, comparative ambiguity aversion is identical to comparative risk aversion in the first stage.

[18] Halevy (2007) finds experimental evidence to support this view. This view is controversial because nonreduction of compound lotteries is arguably a "mistake."

Proposition 6. *Suppose that* $\{\succeq^i_{s^t}\}$ *and* $\{\succeq^j_{s^t}\}$ *satisfy Axioms B1–B7 and that their representations share the same second-order belief* μ_{s^t} *on the same support* $\mathcal{P}_{s^t}$ *for all* s^t. *Then* $\{\succeq^i_{s^t}\}$ *is more ambiguity averse than* $\{\succeq^j_{s^t}\}$ *if and only if* $\{\succeq^i_{s^t}\}$ *is more risk averse than* $\{\succeq^j_{s^t}\}$ *in the first stage.*

Given Axioms B1–B7, an immediate corollary of this proposition is that a decision maker's preferences have a representation with a concave function $v \circ u^{-1}$ if and only if he is more ambiguity averse than a decision whose preferences are represented by recursive expected utility. This result connects our definition of ambiguity aversion in Definition 5 to our definition of comparative ambiguity aversion in Definition 6. It shows that recursive expected utility is the dividing line between ambiguity loving and ambiguity aversion.

What is the relationship between the notion of ambiguity aversion defined in this section and that in Section 3? Because the preference domain of choices is different under the two approaches in these two sections, ambiguity aversion reflects different natures. But the utility representations under these two approaches give identical functionals in the domain of adapted consumption processes. In addition, these two approaches give identical characterizations of ambiguity attitude in terms of the function v for fixed u or $v \circ u^{-1}$.

Unlike the second-order act approach in Section 3 or KMM (2005), the two-stage randomization approach does not need to have a rich support of μ_{s^t} to establish that absolute or comparative ambiguity aversion implies concavity or comparative concavity of $v \circ u^{-1}$. The reason is that the presence of two-stage randomization provides rich choices of lotteries, which allow us to use the standard analysis for objective risk.

5. Application

We use the representation in (3) to illustrate the application of our general model in finance. In that model, the decision maker does not observe a finite parameter $z \in Z$ and has ambiguous beliefs about the possible consumption distributions π_z indexed by z ($\mathcal{P}_{s^t}$ in (2) is a set indexed by z). We first derive the utility gradient (Duffie and Skiadas 1994) for the utility function defined in (3). The utility gradient is useful for solving an individual's optimal consumption and investment problem. It is also useful for equilibrium asset pricing. We define the gradient of a utility function V_0 at c given z as the adapted process (ξ^z_t) such that

$$\lim_{\alpha \downarrow 0} \frac{V_0(c+\alpha\delta) - V_0(c)}{\alpha} = \mathbb{E}\left[\sum_{t=0}^{\infty} \xi^z_t \delta_t\right]. \tag{17}$$

Let V_t denote $V_{s^t}(c)$ in (3) and define

$$\mathcal{R}_t(V_{t+1}) = v^{-1}\big(\mathbb{E}_{\mu_t} v \circ u^{-1}(\mathbb{E}_{\pi_{z,t}} u(V_{t+1}))\big),$$

where we use μ_t and $\pi_{z,t}$ to denote the posterior distribution μ_{s^t} and the conditional distribution $\pi_z(\cdot|s^t)$, respectively.

Proposition 7. *Suppose W, u, and v are differentiable. Then the utility gradient (ξ_t^z) at c for the generalized smooth ambiguity model is given by $\xi_t^z = \lambda_t \mathcal{E}_t^z$ for all t, where*

$$\lambda_t = W_1(c_t, \mathcal{R}_t(V_{t+1})) \tag{18}$$

$$\mathcal{E}_t^z = \prod_{s=0}^{t-1} \frac{W_2(c_s, \mathcal{R}_s(V_{s+1}))}{v'(\mathcal{R}_s(V_{s+1}))} \frac{v' \circ u^{-1}(\mathbb{E}_{\pi_{z,s}}[u(V_{s+1})])}{u'(u^{-1}(\mathbb{E}_{\pi_{z,s}}[u(V_{s+1})]))} u'(V_{s+1}), \qquad \mathcal{E}_0^z = 1. \tag{19}$$

This proposition demonstrates that under some regularity conditions, our generalized recursive smooth ambiguity model delivers a unique utility gradient, which is tractable for applications. By contrast, the widely adopted recursive multiple-priors model implies a set of utility supergradients due to its kinked indifference curves (see Epstein and Wang 1994). After we obtain the utility gradient, we can easily derive the pricing kernel. The pricing kernel M_{t+1}^z between date t and $t+1$ is defined as $M_{t+1}^z = \xi_{t+1}^z / \xi_t^z$. The pricing kernel is often referred to in the literature as the intertemporal marginal rate of substitution or the stochastic discount factor.

In applications, it proves important to work with tractable parametric utility functionals. Our model permits flexible parametric specifications. Inspired by Epstein and Zin (1989), we consider the following homothetic functional forms in (3):

$$W(c, y) = [(1-\beta)c^{1-\rho} + \beta y^{1-\rho}]^{\frac{1}{1-\rho}}, \quad \rho > 0 \tag{20}$$

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}, \quad \gamma > 0, \neq 1 \tag{21}$$

$$v(x) = \frac{x^{1-\eta}}{1-\eta}, \quad \eta > 0, \neq 1, \tag{22}$$

where $\beta \in (0,1)$ is the subjective discount factor, $1/\rho$ represents the elasticity of intertemporal substitution (EIS), γ is the risk aversion parameter, and η is the ambiguity aversion parameter. If $\eta = \gamma$, the decision maker is ambiguity neutral and our model reduces to the recursive utility model of Epstein and Zin (1989) and Weil (1989). The decision maker displays ambiguity aversion if and only if $\eta > \gamma$. By the property of certainty equivalent, a more ambiguity averse agent with a higher value of η has a lower utility level. The preceding interpretations are justified by our axiomatic foundations in previous sections. We refer the reader to Ju and Miao (forthcoming) for more discussions on the specification in (20)–(22).

The key to understanding asset pricing puzzles in a representative–agent consumption-based framework is to understand the pricing kernel. We now derive the pricing kernel for the homothetic generalized recursive ambiguity model. As is well known in the literature of recursive utility, we can write the pricing kernel in two ways.

Proposition 8. *The pricing kernel in terms of continuation values satisfies*

$$M_{t+1}^z = \beta \left(\frac{c_{t+1}}{c_t}\right)^{-\rho} \left(\frac{V_{t+1}}{\mathcal{R}_t(V_{t+1})}\right)^{\rho-\gamma} \left(\frac{(\mathbb{E}_{\pi_{z,t}}[V_{t+1}^{1-\gamma}])^{\frac{1}{1-\gamma}}}{\mathcal{R}_t(V_{t+1})}\right)^{-(\eta-\gamma)}. \tag{23}$$

Alternatively, the pricing kernel in terms of the market return under complete markets satisfies

$$M_{t+1}^{z}=\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho}\right)^{\frac{1-\gamma}{1-\rho}}\left(\frac{1}{R_{t+1}}\right)^{1-\frac{1-\gamma}{1-\rho}}\left(\mathbb{E}_{\pi_{z,t}}\left[\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho}R_{t+1}\right)^{\frac{1-\gamma}{1-\rho}}\right]\right)^{-\frac{\eta-\gamma}{1-\gamma}}, \tag{24}$$

where R_{t+1} is the market return from periods t to $t+1$ that satisfies

$$X_{t+1}=R_{t+1}(X_t-c_t), \tag{25}$$

where (X_t) is the wealth process.

When $\eta=\gamma$, the homothetic recursive ambiguity model reduces to the Epstein–Zin–Weil model. In this case, the pricing kernel in (23) or (24) reduces to that in Epstein and Zin (1989) and Hansen et al. (2008). Why is our generalized recursive smooth ambiguity model useful in explaining asset pricing puzzles? Equation (23) reveals that there are two adjustments to the standard pricing kernel $\beta(c_{t+1}/c_t)^{-\rho}$. The first adjustment is present for recursive expected utility of Epstein and Zin (1989). This adjustment is the second term on the right-hand side of (23). The second adjustment is due to ambiguity aversion, which is given by the last term on the right-hand side of (23). This adjustment has the feature that an ambiguity averse agent with $\eta>\gamma$ puts a higher weight on the pricing kernel when his continuation value is low in recessions. This pessimistic behavior helps explain the equity premium puzzle and the risk-free rate puzzle, and also generates the time-varying equity premium.

Ju and Miao (forthcoming) study the quantitative implications of the above homothetic specification using the pricing kernel in (23), when z is governed by a regime switching process. They show that our model proves successful in explaining many asset pricing puzzles quantitatively.

6. Related literature

Our paper is related to a small literature on axiomatically founded dynamic models of ambiguity. Our second-order act approach is closely related to KMM (2009a).[19] Unlike that paper, we adopt a hierarchical Anscombe–Aumann-type domain. This domain allows us to impose simple and intuitive axioms. More importantly, it permits a separation of intertemporal substitution from attitudes toward risk or uncertainty. Our utility representation allows for flexible parametric specifications, and nests the KMM (2009a) model and some other popular models in the literature as special cases such as the recursive expected utility model (Kreps and Porteus 1978 and Epstein and Zin 1989) and the multiplier preference model with hidden states (Hansen 2007 and Hansen and Sargent 2007a). In addition, this representation permits an information structure with hidden states that could be unknown parameters as in KMM (2009a) or Markov processes.

[19] Hanany and Klibanoff (2009) also provide a dynamic extension of the KMM (2005) model. Their approach is nonrecursive in that they first define preference over consumption plans and then determine conditional preferences by updating beliefs.

Our preference domain in the second-order act approach is built on Hayashi (2005), who first constructs the domain of compound lottery acts $\mathcal{G}$. He chooses $\Delta(\mathcal{C} \times \mathcal{G})$ as the preference domain, while we adopt $\mathcal{C} \times \mathcal{G}$ as the primary domain of preference $\{\succeq_{s^t}\}$. Our domain choice proves to be more convenient in our setting. Hayashi (2005) embeds Gilboa and Schmeidler's (1989) static multiple-priors model in a dynamic environment and establishes a generalized recursive multiple-priors model. His model permits a separation of intertemporal substitution from attitudes toward risk or uncertainty. It generalizes the recursive multiple-priors model of Epstein and Wang (1994) and Epstein and Schneider (2003). Epstein and Schneider (2003) axiomatize the recursive multiple-priors model and prove that dynamic consistency leads to rectangular sets of priors and to prior-by-prior Bayesian updating as the updating rule for such sets of priors. Wang (2003) also axiomatizes this model and some updating rules for preferences that are not necessarily in the expected utility class. Hanany and Klibanoff (2007) follow a nonrecursive approach and extend the Epstein and Schneider model by allowing updating of the set of priors to violate conseqentialism. As KMM (2005) point out, one limitation of the multiple-priors model is that there is no separation between ambiguity and ambiguity attitude. The set of priors may reflect the decision maker's perceived ambiguity or his attitude toward ambiguity. This confounding makes comparative static analysis hard to interpret.

Our second axiomatization using the two-stage randomization approach extends Seo's (2009) static model to a dynamic setting. To the best of our knowledge, our paper provides the first dynamic extension of Seo's static model. As a by-product contribution, we construct a domain of two-stage compound lottery acts $\mathcal{H}$, which contains $\mathcal{G}$ and allows for randomization both before and after the realization of the state of the world. We then dispense with second-order acts and the associated preferences over these acts. We define a single preference relation $\{\succeq_{s^t}\}$ over $\mathcal{C} \times \Delta(\mathcal{H})$.

Our characterization of ambiguity attitude in the two axiomatic approaches is based on the foundation of Ghirardato and Marinacci (2002). Ambiguity aversion reflects somewhat different natures in the two approaches because of different choice domains, though the characterization in terms of concavity of $v \circ u^{-1}$ is identical. In the second-order act approach, ambiguity aversion is an aversion to the subjective uncertainty about ex ante evaluations of one-step-ahead acts. In the two-stage randomization approach, ambiguity aversion is associated with the violation of reduction of compound lotteries, as also pointed out by Segal (1987, 1990) and Seo (2009). Segal uses the anticipated utility model and considers objective lotteries, while the Ellsberg paradox is often viewed as a phenomenon associated with subjective uncertainty. Seo does not provide a formal definition of ambiguity aversion and characterizations of ambiguity attitude. We provide such an analysis and characterize the link between ambiguity aversion and reduction of compound lotteries. Our result that ambiguity aversion is identical to risk aversion in the first stage is similar to Theorem 5 in Ergin and Gul (2009), who refer to risk aversion in the first stage as second-order risk aversion.

Maccheroni et al. (2006b) provide a dynamic extension of the static variational model of ambiguity developed by Maccheroni et al. (2006a). The static variational model includes the multiple-priors model and the multiplier preference model of Hansen and

Sargent (2001) as special cases. The dynamic extension does not separate intertemporal substitution from attitude toward risk or ambiguity. Variational preferences are also subject to the limitation concerning the separation of ambiguity from ambiguity attitude.

Our model is also related to the literature on recursive utility under risk or uncertainty (Kreps and Porteus 1978, Epstein and Zin 1989, Chew and Epstein 1991, Skiadas 1998, and Klibanoff and Ozdenoren 2007). This literature does not deal with ambiguity. In the framework of Klibanoff and Ozdenoren (2007) or Skiadas (1998), preferences depend on the filtration. Unlike their framework, we take the filtration as given and, thus, cannot make comparisons of representations across filtrations.

Recursive utility models allow for preferences for the timing of the temporal resolution of uncertainty. As is well known in the literature on recursive expected utility preferences, a nonlinear time aggregator is needed to permit nonindifference to the timing of the temporal resolution of uncertainty. Strzalecki (2009) shows that even without a nonlinear aggregator, or with standard discounting, most dynamic models of ambiguity aversion (including the models discussed above) result in timing nonindifference. In particular, decision makers with such preferences prefer earlier resolution of uncertainty. The only model of ambiguity aversion that exhibits indifference to timing is the multiple-priors utility model. Our paper does not study this issue. Presumably, Strzalecki's analysis can be applied to our setting.

Finally, like most papers in the literature on dynamic models, we follow a recursive approach and maintain dynamic consistency. This approach is normatively appealing and computationally simple in applications because the usual dynamic programming method can be applied. This approach typically shares the drawback of lacking a "reduction" or "closure" property as discussed in KMM (2009a).[20] This means that our recursive model (2) over adapted consumption processes does not have the reduced static KMM smooth ambiguity functional form. Siniscalchi (2011) follows a different approach to formulating dynamic models of ambiguity. He takes an individual's preferences over decision trees, rather than acts, as primitive. His approach allows for dynamic inconsistency. He formalizes sophistication as an assumption about the way individuals resolve conflicts between preferences at different decision points. It remains to see whether dynamic smooth ambiguity preferences can be formulated in his framework.

Appendix A: Proofs of Theorems 1 and 2

We prove the sufficiency of the axioms. The proof of necessity is routine.

A1 *Representation of risk preference*

When $\{\succeq_{s^t}\}$ is restricted to the domain $\mathcal{C} \times \mathcal{M}$, Axiom A3 (History Independence of Risk Preference) implies that $\{\succeq_{s^t}\}$ induces a single preference relation $\succeq$ defined on $\mathcal{C} \times \mathcal{M}$. By Axiom A1 (Order) and Debreu's (1954) theorem, there is a continuous representation $V : \mathcal{C} \times \mathcal{M} \to \mathbb{R}$ of $\succeq$. We fix such a representation.

[20]Epstein and Schneider (2003) and Maccheroni et al. (2006b) are exceptions.

Fix some arbitrary $\widehat{c} \in \mathcal{C}$ throughout the proof. By Axiom A2 (Current Consumption Separability), $V(c, \cdot)$ and $V(\widehat{c}, \cdot)$ represent the same ranking over $\mathcal{M}$; hence V has the form

$$V(c, m) = \widehat{W}(c, V(\widehat{c}, m)) \tag{26}$$

for some function $\widehat{W}$ that is strictly increasing in the second argument. By Axiom A4 (Independence for Timeless Lotteries), $V(\widehat{c}, m)$ is ordinally equivalent to an expected utility representation on $\mathcal{C} \times \mathcal{M}$. Thus, we have the form

$$V(\widehat{c}, m) = \zeta\bigg(\int_{\mathcal{C}\times\mathcal{M}} \widehat{u}(c', m')\, dm(c', m')\bigg), \tag{27}$$

where $\widehat{u}$ is a vNM index and ζ is a monotone transformation.

Lemma 1. *Given Axioms A1 and A3, Axiom A5 implies that*

$$(c, \delta[(c', m')]) \succeq (c, \delta[(c'', m'')]) \quad \Longleftrightarrow \quad (c', m') \succeq (c'', m'')$$

for any $c \in \mathcal{C}$ and $m, m' \in \mathcal{M}$.

Proof. We restrict attention to the subdomain $\mathcal{C} \times \mathcal{M}$. By Axiom A3 (History Independence of Risk Preference), we can replace $\{\succeq_{s^t}\}$ with $\succeq$ in Axiom A5 (Dynamic Consistency). Suppose $(c', m') \succeq (c'', m'')$. Then $(c'', m'') \succeq (c_0, m_0) \Longrightarrow (c', m') \succeq (c_0, m_0)$ for any $(c_0, m_0) \in \mathcal{C} \times \mathcal{M}$. Thus, $\delta[(c', m')]$ stochastically dominates $\delta[(c'', m'')]$. By Axiom A5, $(c, \delta[(c', m')]) \succeq (c, \delta[(c'', m'')])$. Suppose $(c, \delta[(c', m')]) \succeq (c, \delta[(c'', m'')])$, but $(c', m') \prec (c'', m'')$. By continuity of $\succeq$ from Axiom A1, there exists some $(c_0, m_0) \in \mathcal{C} \times \mathcal{M}$ such that $(c', m') \prec (c_0, m_0) \preceq (c'', m'')$. Thus, $\delta[(c'', m'')]$ strictly stochastically dominates $\delta[(c', m')]$. By Axiom A5, $(c, \delta[(c', m')]) \prec (c, \delta[(c'', m'')])$, which is a contradiction. $\square$

Now, we deduce

$$\begin{aligned}
\widehat{u}(c', m') \geq \widehat{u}(c'', m'') \quad &\Longleftrightarrow \quad (\widehat{c}, \delta[(c', m')]) \succeq (\widehat{c}, \delta[(c'', m'')]) \quad \text{(by (27))}\\
&\Longleftrightarrow \quad (c', m') \succeq (c'', m'') \quad \text{(by Lemma 1)}\\
&\Longleftrightarrow \quad V(c', m') \geq V(c'', m'').
\end{aligned}$$

Hence, $\widehat{u}$ and V are ordinally equivalent representations of $\succeq$ on $\mathcal{C} \times \mathcal{M}$, implying that there is a monotone transformation u such that $\widehat{u} = u \circ V$. Plugging this equation into (27) yields

$$V(\widehat{c}, m) = \zeta\bigg(\int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')\bigg). \tag{28}$$

Define W by $W(c, x) = \widehat{W}(c, \zeta(u(x)))$, which is strictly increasing in the second argument. Then

$$V(c, m) = \widehat{W}\bigg(c, \zeta\bigg(\int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')\bigg)\bigg)$$

$$= \widehat{W}\left(c, \zeta\left(u \circ u^{-1} \circ \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')\right)\right)$$
$$= W\left(c, u^{-1}\left(\int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')\right)\right).$$

A2 *Extension to the whole domain*

By an argument similar to the proof of Lemmas 8 and 9 in Hayashi (2005), we can use continuity of $\succeq_{s^t}$ from Axiom A1, Dynamic Consistency (Axiom A5), and compactness of $\mathcal{C}$ to show that for each $(c, g) \in \mathcal{C} \times \mathcal{G}$, there exists a *risk equivalent* $(c, m) \in \mathcal{C} \times \mathcal{M}$ such that $(c, g) \sim_{s^t} (c, m)$ for each s^t. Thus, for each s^t, define $V_{s^t} : \mathcal{C} \times \mathcal{G} \to \mathbb{R}$ by

$$V_{s^t}(c, g) = V(c, m), \tag{29}$$

where m is such that $(c, g) \sim_{s^t} (c, m)$. Using this definition and (26), we obtain

$$V_{s^t}(c, g) = V(c, m) = \widehat{W}(c, V(\widehat{c}, m)) = \widehat{W}(c, V_{s^t}(\widehat{c}, g)). \tag{30}$$

From Axioms A6 (SEU Representation of Preference Over Second-Order Acts) and A7 (Consistency With the Preference Over Second-Order Acts), we obtain

$$V_{s^t}(\widehat{c}, g_{+1}) = \xi_{s^t}\left(\int_{\mathcal{P}_{s^t}} \psi\left(\sum_{s\in S} g_{+1}(s)\pi(s)\right) d\mu_{s^t}(\pi)\right), \tag{31}$$

where ξ_{s^t} is a monotone transformation. By restricting attention to $\mathcal{M}$, we can use Axiom A3 (History Independence of Risk Preference) to set $\xi_{s^t} = \xi$ for all s^t.

Define $v = \psi \circ \bar{u}^{-1} \circ u$, where $\bar{u}$ is defined in (10).[21] Using Axiom A6, we immediately obtain part (iii) of the theorem. Plugging this definition of v into (31) yields

$$V_{s^t}(\widehat{c}, g_{+1}) = \xi\left(\int_{\mathcal{P}_{s^t}} v \circ u^{-1} \circ \bar{u}\left(\sum_{s\in S} g_{+1}(s)\pi(s)\right) d\mu_{s^t}(\pi)\right)$$
$$= \xi\left(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\left(\sum_{s\in S} \pi(s) \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dg_{+1}(s)(c', m')\right) d\mu_{s^t}(\pi)\right),$$

where the second equality follows from (10).

When restricting V_{s^t} to the domain $\mathcal{M}$, we obtain

$$V_{s^t}(\widehat{c}, m) = \xi \circ v \circ u^{-1}\left(\int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')\right)$$
$$= V(\widehat{c}, m) = \zeta\left(\int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')\right)$$

for all $m \in \mathcal{M}$, where the last equality follows from (28). Therefore, we have $\xi \circ v \circ u^{-1} = \zeta$, implying $u^{-1} \circ \zeta^{-1} \circ \xi = v^{-1}$.

[21] Note that $\bar{u}$ is increasing on $\mathcal{M}$ when $\mathcal{M}$ is ordered by first-order stochastic dominance; therefore, its inverse exists.

Define

$$y \equiv V_{s^t}(\widehat{c}, g_{+1}) = \xi\biggl(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\biggl(\sum_{s\in S} \pi(s) \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dg_{+1}(s)(c', m')\biggr)\, d\mu_{s^t}(\pi)\biggr).$$

Using (30), we obtain

$$\begin{aligned} V_{s^t}(c, g_{+1}) &= \widehat{W}(c, y) = \widehat{W}\bigl(c, \zeta \circ u(u^{-1} \circ \zeta^{-1}(y))\bigr) = W(c, u^{-1} \circ \zeta^{-1}(y)) \\ &= W\biggl(c, u^{-1} \circ \zeta^{-1} \circ \xi\biggl(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\biggl(\sum_{s\in S} \pi(s) \\ &\qquad\qquad \times \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dg_{+1}(s)(c', m')\biggr)\, d\mu_{s^t}(\pi)\biggr)\biggr) \\ &= W\biggl(c, v^{-1}\biggl(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\biggl(\sum_{s\in S} \pi(s) \\ &\qquad\qquad \times \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dg_{+1}(s)(c', m')\biggr)\, d\mu_{s^t}(\pi)\biggr)\biggr), \end{aligned}$$

where the third equality follows from the definition of W in Appendix A1.

For any $g \in \mathcal{G}$, for each $s \in S$ and each (c', g') in the support of $g(s) \in \Delta(\mathcal{C} \times \mathcal{G})$, there exists a risk equivalent $(c', m') \in \mathcal{C} \times \mathcal{M}$ such that $(c', m') \sim_{s^t, s} (c', g')$. Let g_{+1} be a one-step-ahead act such that $g_{+1}(s)(L') = g(s)(L)$ holds for all pairs $L \subset \mathcal{C} \times \mathcal{G}$ and $L' \subset \mathcal{C} \times \mathcal{M}$, where L' consists of all risk equivalents (c', m') of corresponding elements (c', g') in L. By construction, $g_{+1}(s)$ and $g(s)$ are stochastically equivalent. By Axiom A5 (Dynamic Consistency), $(c, g) \sim_{s^t} (c, g_{+1})$. Therefore,

$$\begin{aligned} V_{s^t}(c, g) &= V_{s^t}(c, g_{+1}) \\ &= W\biggl(c, v^{-1}\biggl(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\biggl(\sum_{s\in S} \pi(s) \\ &\qquad\qquad \times \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dg_{+1}(s)(c', m')\biggr)\, d\mu_{s^t}(\pi)\biggr)\biggr) \\ &= W\biggl(c, v^{-1} \circ \biggl(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\biggl(\sum_{s\in S} \pi(s) \\ &\qquad\qquad \times \int_{\mathcal{C}\times\mathcal{G}} u(V_{s^t,s}(c', g'))\, dg(s)(c', g')\biggr)\, d\mu_{s^t}(\pi)\biggr)\biggr), \end{aligned}$$

where we have used the fact that $g(s)$ and $g_{+1}(s)$ are stochastically equivalent to derive the second equality.

A3 *Proof of uniqueness*

Suppose $(\{\tilde{V}_{s^t}\}, \tilde{W}, \tilde{u}, \tilde{v}, \{\tilde{\mu}_{s^t}\})$ and $(\{V_{s^t}\}, W, u, v, \{\mu_{s^t}\})$ represent the same preference. On the domain of deterministic consumption streams $\mathcal{C}^\infty$, each $\tilde{V}_{s^t}$ coincides with the common function $\tilde{V}$ and each V_{s^t} coincides with the common function V. Since $\tilde{V}$ and V are ordinally equivalent over $\mathcal{C}^\infty$, there is a monotone transformation Φ such that

$$\tilde{V}(y) = \Phi \circ V(y) \quad \text{for all } y \in \mathcal{C}^\infty.$$

By (29), we have $\tilde{V}_{s^t} = \Phi \circ V_{s^t}$.

Since

$$\tilde{W}(c, \tilde{V}(y)) = \tilde{V}(c, y) = \Phi(V(c, y)) = \Phi(W(c, V(y))) = \Phi(W(c, \Phi^{-1}(\tilde{V}(y)))),$$

we deduce that $\tilde{W}(c, \cdot) = \Phi(W(c, \Phi^{-1}(\cdot)))$.

On $\mathcal{M}$, $\int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m')$ and $\int_{\mathcal{C}\times\mathcal{M}} \tilde{u}(\tilde{V}(c', m'))\, dm(c', m')$ are equivalent mixture-linear representations of the risk preference conditional on the fixed current consumption $\widehat{c}$. Therefore, there exist constants A, B with $A > 0$ such that

$$\tilde{u}(\tilde{V}(c', m')) = Au(V(c', m')) + B \quad \text{for all } (c', m') \in \mathcal{M}.$$

Since $\tilde{V} = \Phi \circ V$, we obtain $\tilde{u} \circ \Phi = Au + B$.

By construction from Appendix A2, $\tilde{v} \circ \tilde{u}^{-1} \circ \bar{\tilde{u}} = \psi = v \circ u^{-1} \circ \bar{u}$. By (10), we compute

$$\begin{aligned}
\bar{\tilde{u}}(m) &= \int_{\mathcal{C}\times\mathcal{M}} \tilde{u}(\tilde{V}(c', m'))\, dm(c', m') \\
&= \int_{\mathcal{C}\times\mathcal{M}} \tilde{u} \circ \Phi(V(c', m'))\, dm(c', m') \\
&= \int_{\mathcal{C}\times\mathcal{M}} Au(V(c', m'))\, dm(c', m') + B \\
&= A \int_{\mathcal{C}\times\mathcal{M}} u(V(c', m'))\, dm(c', m') + B = A\bar{u}(m) + B.
\end{aligned}$$

Let $\bar{u}(m) = w$. Then we have

$$\tilde{v} \circ \tilde{u}^{-1}(Aw + B) = v \circ u^{-1}(w).$$

Since $\tilde{u} \circ \Phi(w) = Au(w) + B$, it follows that

$$\tilde{v} \circ \tilde{u}^{-1}(Aw + B) = \tilde{v} \circ \tilde{u}^{-1}(Au \circ u^{-1}(w) + B) = \tilde{v} \circ \Phi(u^{-1}(w)).$$

Thus, we obtain

$$\tilde{v} \circ \Phi(u^{-1}(w)) = v \circ u^{-1}(w).$$

By replacing $u^{-1}(w)$ with x, we obtain $\tilde{v} \circ \Phi(x) = v(x)$. Finally, uniqueness of μ_{s^t} follows from Axiom A5.

Appendix B: Proof of Theorem 3

Given a compact metric space Y, let $\mathcal{B}(Y)$ be the family of Borel subsets of Y and let $\Delta(Y)$ be the set of Borel probability measures defined over $\mathcal{B}(Y)$, which is again a compact metric space with respect to the weak convergence topology. Inductively define the family of domains $\{\mathcal{H}_0, \mathcal{H}_1, \dots\}$ by

$$\begin{aligned}
\mathcal{H}_0 &= (\Delta(\mathcal{C}))^S \\
\mathcal{H}_1 &= \big(\Delta(\mathcal{C} \times \Delta(\mathcal{H}_0))\big)^S \\
&\vdots \\
\mathcal{H}_t &= \big(\Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))\big)^S,
\end{aligned}$$

and so on. By induction, $\Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$ is a compact metric space and so is $\mathcal{H}_t$, for every $t \geq 0$. Let d_t be the metric over $\mathcal{H}_t$. Let $\mathcal{H}^* = \prod_{t=0}^{\infty} \mathcal{H}_t$. This is a compact metric space with respect to the product metric $d(h, h') = \sum_{t=0}^{\infty} (1/2^t) d_t(h_t, h'_t)/(1 + d_t(h_t, h'_t))$.

The domain to be constructed is a *subset* of $\mathcal{H}^*$, which consists of coherent acts. Define a mapping $\pi_0 : \mathcal{C} \times \Delta(\mathcal{H}_0) \to \mathcal{C}$ by

$$\pi_0(c, p_0) = c$$

for each $(c, p_0) \in \mathcal{C} \times \Delta(\mathcal{H}_0)$. Define a mapping $\rho_0 : \mathcal{H}_1 \to \mathcal{H}_0$ by

$$\rho_0(h_1)(s)[B_0] = h_1(s)[\pi_0^{-1}(B_0)]$$

for each $h_1 \in \mathcal{H}_1$, $s \in S$, and $B_0 \in \mathcal{B}(\mathcal{C})$. Define a mapping $\tilde{\rho}_0 : \Delta(\mathcal{H}_1) \to \Delta(\mathcal{H}_0)$ by

$$\tilde{\rho}_0(p_1)[H_0] = p_1[\rho_0^{-1}(H_0)]$$

for each $p_1 \in \Delta(\mathcal{H}_1)$ and $H_0 \in \mathcal{B}(\mathcal{H}_0)$.

Similarly, define $\pi_1 : \mathcal{C} \times \Delta(\mathcal{H}_1) \to \mathcal{C} \times \Delta(\mathcal{H}_0)$ by

$$\pi_1(c, p_1) = (c, \tilde{\rho}_0(p_1))$$

for each $(c, p_1) \in \mathcal{C} \times \Delta(\mathcal{H}_1)$, define $\rho_1 : \mathcal{H}_2 \to \mathcal{H}_1$ by

$$\rho_1(h_2)(s)[B_1] = h_2(s)[\pi_1^{-1}(B_1)]$$

for each $h_2 \in \mathcal{H}_2$, $s \in \Omega$, and $B_1 \in \mathcal{B}(\mathcal{C} \times \Delta(\mathcal{H}_0))$, and define $\tilde{\rho}_1 : \Delta(\mathcal{H}_2) \to \Delta(\mathcal{H}_1)$ by

$$\tilde{\rho}_1(p_2)[H_1] = p_2[\rho_1^{-1}(H_1)]$$

for each $p_2 \in \Delta(\mathcal{H}_2)$ and $H_1 \in \mathcal{B}(\mathcal{H}_1)$.

Inductively, given $\pi_{t-1} : \mathcal{C} \times \Delta(\mathcal{H}_{t-1}) \to \mathcal{C} \times \Delta(\mathcal{H}_{t-2})$, $\rho_{t-1} : \mathcal{H}_t \to \mathcal{H}_{t-1}$, and $\tilde{\rho}_{t-1} : \Delta(\mathcal{H}_t) \to \Delta(\mathcal{H}_{t-1})$, define $\pi_t : \mathcal{C} \times \Delta(\mathcal{H}_t) \to \mathcal{C} \times \Delta(\mathcal{H}_{t-1})$ by

$$\pi_t(c, p_t) = (c, \tilde{\rho}_{t-1}(p_t))$$

for each $(c, p_t) \in \mathcal{C} \times \Delta(\mathcal{H}_t)$, define $\rho_t : \mathcal{H}_{t+1} \to \mathcal{H}_t$ by

$$\rho_t(h_{t+1})(s)[B_t] = h_{t+1}(s)[\pi_t^{-1}(B_t)]$$

for each $h_{t+1} \in \mathcal{H}_{t+1}$, $s \in S$, and $B_t \in \mathcal{B}(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$, and define $\tilde{\rho}_t : \Delta(\mathcal{H}_{t+1}) \to \Delta(\mathcal{H}_t)$ by

$$\tilde{\rho}_t(p_{t+1})[H_t] = p_{t+1}[\rho_t^{-1}(H_t)]$$

for each $p_{t+1} \in \Delta(\mathcal{H}_{t+1})$ and $H_t \in \mathcal{B}(\mathcal{H}_t)$.

Define

$$\mathcal{H} = \{h = (h_0, h_1, h_2, \dots) \in \mathcal{H}^*\colon\ h_t = \rho_t(h_{t+1}),\ t \geq 0\}.$$

For each $s \in S$, the sequence $(h_0(s), h_1(s), h_2(s), \dots) \in \prod_{t=0}^{\infty} \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$ is viewed as a sequence of constant acts since

$$\begin{aligned} & h_0(s) \in \Delta(\mathcal{C}) \subset \mathcal{H}_0 \\ & h_1(s) \in \Delta(\mathcal{C} \times \Delta(\mathcal{H}_0)) \subset \mathcal{H}_1 \\ & \qquad \vdots \\ & h_t(s) \in \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1})) \subset \mathcal{H}_t, \end{aligned}$$

and so on.

The lemmas below verify that such constant acts are also coherent. They are immediate from the definition of $\mathcal{H}$.

Lemma 2. *For every $h \in \mathcal{H}$ and $s \in S$, the sequence $(h_0(s), h_1(s), h_2(s), \dots) \in \prod_{t=0}^{\infty} \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$ satisfies $h_t(s) = \rho_t(h_{t+1}(s))$.*

Lemma 3. *For every $t \geq 0$, $h_t \in \mathcal{H}_t$, and $h_{t+1} \in \mathcal{H}_{t+1}$, if $h_t(s) = \rho_t(h_{t+1}(s))$ for every $s \in S$, then $h_t = \rho_t(h_{t+1})$.*

Let

$$\begin{aligned} Q &= \left\{(q_t) \in \prod_{t=0}^{\infty} \Delta(\mathcal{H}_t)\colon\ q_t = \tilde{\rho}_t(q_{t+1}), \forall t \geq 0\right\} \\ A &= \left\{(a_t) \in \prod_{t=0}^{\infty} \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))\colon\ a_t = \rho_t(a_{t+1}), \forall t \geq 0\right\}. \end{aligned}$$

Lemma 4. *We have the homeomorphic relation*

$$A \simeq \Delta(\mathcal{C} \times Q).$$

Proof. Given $(a_t) \in A \subset \prod_{t=0}^{\infty} \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$, by the Kolmogorov extension theorem there exists a unique $a \in (\mathcal{C} \times \prod_{t=0}^{\infty} \Delta(\mathcal{H}_{t-1}))$ such that

$$\mathrm{mrg}_{\mathcal{C} \times \Delta(\mathcal{H}_{t-1})} a = a_t$$

for each $t \geq 0$, where mrg denotes marginal. Define a mapping $\xi\colon A \to \Delta(\mathcal{C} \times Q)$ by $\xi((a_t)) = a$.

We need to show $a \in \Delta(\mathcal{C} \times Q)$. For each $t \geq 0$, let

$$Q_t = \{(q_t, q_{t+1}) \in \Delta(\mathcal{H}_t) \times \Delta(\mathcal{H}_{t+1})\colon\ q_t = \tilde{\rho}_t(q_{t+1})\} \times \prod_{\tau \neq t, t+1} \Delta(\mathcal{H}_\tau).$$

We derive that

$$\begin{aligned} a(\mathcal{C} \times Q_t) &= \mathrm{mrg}_{\mathcal{C} \times \Delta(\mathcal{H}_t) \times \Delta(\mathcal{H}_{t+1})} a\big(\mathcal{C} \times \{(q_t, q_{t+1}) \in \Delta(\mathcal{H}_t) \times \Delta(\mathcal{H}_{t+1})\colon\ q_t = \tilde{\rho}_t(q_{t+1})\}\big) \\ &= \mathrm{mrg}_{\mathcal{C} \times \Delta(\mathcal{H}_t)} a\big(\mathcal{C} \times \tilde{\rho}_t(\Delta(\mathcal{H}_{t+1}))\big) \\ &= a_{t+1}\big(\mathcal{C} \times \tilde{\rho}_t(\Delta(\mathcal{H}_{t+1}))\big) \\ &= a_{t+1}\big(\pi_{t+1}(\mathcal{C} \times \Delta(\mathcal{H}_{t+1}))\big) \\ &= \rho_{t+1}(a_{t+2})\big(\pi_{t+1}(\mathcal{C} \times \Delta(\mathcal{H}_{t+1}))\big) \\ &= a_{t+2}\big(\pi_{t+1}^{-1}\big(\pi_{t+1}(\mathcal{C} \times \Delta(\mathcal{H}_{t+1}))\big)\big) \\ &= a_{t+2}(\mathcal{C} \times \Delta(\mathcal{H}_{t+1})) = 1. \end{aligned}$$

Therefore,

$$a(\mathcal{C} \times Q) = a\left(\bigcap_{t=0}^{\infty} (\mathcal{C} \times Q_t)\right) = \lim_{T \to \infty} a\left(\bigcap_{t=0}^{T} (\mathcal{C} \times Q_t)\right) = 1.$$

- *Mapping ξ is one-to-one*: This follows from the uniqueness of Kolmogorov extension theorem.
- *Mapping ξ is onto*: For every $a \in \Delta(\mathcal{C} \times Q)$, the inverse is given by $(a_t) \in \prod_{t=0}^{\infty} \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$ such that

  $$a_t = \mathrm{mrg}_{\mathcal{C} \times \Delta(\mathcal{H}_{t-1})} a$$

  for each $t \geq 0$. To show $(a_t) \in \mathcal{A}$, take any $B_t \in \mathcal{B}(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$. We deduce that

  $$\begin{aligned} a_t(B_t) &= a\left(B_t \times \prod_{\tau \neq t-1} \Delta(\mathcal{H}_t)\right) \\ &\geq a\big(\big\{(c, (q_\tau)) \in \mathcal{C} \times Q\colon\ (c, q_t) \in B_t\big\}\big) \\ &= a\big(\big\{(c, (q_\tau)) \in \mathcal{C} \times Q\colon\ (c, \tilde{\rho}_t(q_{t+1})) \in B_t\big\}\big) \\ &= a\big(\big\{(c, (q_\tau)) \in \mathcal{C} \times Q\colon\ (c, q_{t+1}) \in \pi_t^{-1}(B_t)\big\}\big) \\ &= a_{t+1}(\pi_t^{-1}(B_t)) \\ &= \rho_t(a_{t+1})(B_t). \end{aligned}$$

  Since

  $$\begin{aligned} 1 &= a_t(\mathcal{C} \times \Delta(\mathcal{H}_{t-1})) = a_t(B_t) + a_t(B_t^c) \\ &\geq \rho_t(a_{t+1})(B_t) + \rho_t(a_{t+1})(B_t^c) = \rho_t(a_{t+1})(\mathcal{C} \times \Delta(\mathcal{H}_{t-1})) = 1, \end{aligned}$$

we obtain

$$a_t(B_t) = \rho_t(a_{t+1})(B_t).$$

- *Mappings ξ and ξ^{-1} are continuous*: This is immediate from the nature of the product topology. □

LEMMA 5. *We have the homeomorphic relation*

$$\mathcal{H} \simeq A^S.$$

PROOF. Define $\xi : \mathcal{H} \to A^S$ by

$$\xi(h)(s) = (h_0(s), h_1(s), h_2(s), \ldots).$$

It follows from Lemma 2 that $\xi(h) \in A$.

- *Mapping ξ is one-to-one*: Suppose $\xi(h) = \xi(h')$. By definition of ξ, we have $(h_0(s), h_1(s), h_2(s), \ldots) = (h'_0(s), h'_1(s), h'_2(s), \ldots)$ for all $s \in S$, which implies $h = h'$.
- *Mapping ξ is onto*: Take any $\tilde{h} \in A^S$. By definition,

$$\tilde{h}(s) = (\tilde{h}_0(s), \tilde{h}_1(s), \tilde{h}_2(s), \ldots) \in \prod_{t=0}^{\infty} \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{t-1}))$$

for each $s \in S$. Then $\xi^{-1}(\tilde{h}) = (h_0, h_1, h_2, \ldots) \in \mathcal{H}^*$ satisfies $h_t(s) = \tilde{h}_t(s)$ for each t and s. By Lemmas 2 and 3, the sequence $(h_0, h_1, h_2, \ldots)$ is coherent and hence $\xi^{-1}(\tilde{h}) \in \mathcal{H}$.
- *Mappings ξ and ξ^{-1} are continuous*: This is immediate from the nature of the product topology. □

Let

$$P^* = \left\{(p_t) \in \prod_{t=0}^{\infty} \Delta\left(\prod_{\tau=0}^{t} \mathcal{H}_\tau\right) : \operatorname{mrg}_{\prod_{\tau=0}^{t} \mathcal{H}_\tau} p_{t+1} = p_t\right\}.$$

LEMMA 6. *For any $(p_t) \in P^*$, there exists a unique $p \in \Delta(\mathcal{H}^*)$ such that*

$$\operatorname{mrg}_{\prod_{\tau=0}^{t} \mathcal{H}_\tau} p = p_t.$$

Moreover, there exists a homeomorphism $\chi : P^* \to \Delta(\mathcal{H}^*)$.

The proof follows from Lemma 1 in Brandenberger and Dekel (1993).

Let

$$\mathcal{H}^t = \left\{(h_0, \ldots, h_t) \in \prod_{\tau=0}^{t} \mathcal{H}_\tau : h_\tau = \rho_\tau(h_{\tau+1}), \tau = 0, \ldots, t-1\right\}$$

for each $t \geq 0$ and let

$$P = \{(p_t) \in P^* : p_t(\mathcal{H}^t) = 1, t \geq 0\}.$$

Lemma 7. *The equality $\chi(P) = \Delta(\mathcal{H})$ holds. As a result, $P \simeq \Delta(\mathcal{H})$ holds through χ.*

Proof. The $\subset$ *part*: Let $p = \chi((p_t))$ for some $(p_t) \in P$. Let

$$\Gamma_t = \mathcal{H}^t \times \prod_{\tau=t+1}^{\infty} \mathcal{H}_\tau$$

for each $t \geq 0$. Then we have $\mathcal{H} \subset \Gamma_t \subset \mathcal{H}^*$ for each $t \geq 0$, (Γ_t) is decreasing, and $\bigcap_{t\geq 0} \Gamma_t = \mathcal{H}$.

Since p is the Kolmogorov extension of (p_t), we have

$$p(\Gamma_t) = p_t(\mathcal{H}^t) = 1$$

for every $t \geq 0$. Thus, $p(\mathcal{H}) = p(\bigcap_{t\geq 0} \Gamma_t) = \lim p(\Gamma_t) = 1$.

The $\supset$ *part*: Pick any $p \in \Delta(\mathcal{H})$ that satisfies $p(\mathcal{H}) = 1$. Let (p_t) be the sequence of marginals defined by $p_t = \mathrm{mrg}_{\prod_{\tau=0}^{t} \mathcal{H}_\tau} p$ for each $t \geq 0$. Then $p_t(\mathcal{H}^t) = p(\Gamma_t) \geq 1$, where the second inequality follows from $\Gamma_t \supset \mathcal{H}$. Since p_t is a probability measure, we have $p_t(\mathcal{H}^t) = 1$. Since p is the Kolmogorov extension of (p_t), we have $p = \chi((p_t))$. □

Lemma 8. *For every $(q_t) \in Q$, there exists a unique $(p_t) \in P$ such that*

$$\mathrm{mrg}_{\mathcal{H}_t} p_t = q_t.$$

Moreover, Q and P are homeomorphic.

Proof. Define a sequence of mappings (ξ_t), $\xi_t : \mathcal{H}_t \to \prod_{\tau=0}^{t} \mathcal{H}_\tau$ for each $t \geq 0$, by

$$\xi_t(h_t) = (\widehat{h}_0, \dots, \widehat{h}_t),$$

where $\widehat{h}_t = h_t$ and $\widehat{h}_\tau = \rho_\tau(\widehat{h}_{\tau+1})$ for $\tau = 0, 1, \dots, t-1$.

By construction, each (ξ_t) is a one-to-one mapping and $\xi_t(\mathcal{H}_t) = \mathcal{H}^t$. Therefore, we can define the sequence of inverse mappings (ξ_t^{-1}), $\xi_t^{-1} : \mathcal{H}^t \to \mathcal{H}_t$ given by

$$\xi_t^{-1}(h_0, \dots, h_t) = h_t,$$

which is a projection mapping that is continuous.

For $(q_t) \in Q$, define the corresponding sequence $(p_t) \in P$ by

$$p_t(E_t) = q_t(\xi_t^{-1}(E_t))$$

for each $E_t \in \mathcal{B}(\prod_{\tau=0}^{t} \mathcal{H}_\tau)$ and $t \geq 0$. We can see that $(p_t) \in P$ since $p_t(\mathcal{H}^t) = q_t(\xi_t^{-1}(\mathcal{H}^t)) = q_t(\mathcal{H}_t) = 1$. By construction, $\mathrm{mrg}_{\mathcal{H}_t} p_t = q_t$ for each $t \geq 0$. □

Now, Theorem 3 follows from the fact that $\mathcal{H} \simeq A^S$, $A \simeq \Delta(\mathcal{C} \times Q)$, $Q \simeq P$, and $P \simeq \Delta(\mathcal{H})$.

Finite-step-ahead acts and denseness

Finally, we define finite-step-ahead acts and show that the union of all the sets of finite-step-ahead acts is dense. Let

$$\mathcal{H}_{+1} = \left\{h_{+1} \in \left(\Delta(\mathcal{C} \times \Delta(\mathcal{H}))\right)^S \colon \forall s \in S,\, h_{+1}(s) \in \Delta(\mathcal{C} \times \Delta(\mathcal{L}))\right\}.$$

Since $\mathcal{H} \simeq (\Delta(\mathcal{C} \times \Delta(\mathcal{H})))^S$, we can embed $\mathcal{H}_{+1}$ into $\mathcal{H}$, where the range of $\mathcal{H}_{+1}$ is embedded into $\mathcal{L}$ since $\mathcal{L} \simeq \Delta(\mathcal{C} \times \Delta(\mathcal{L}))$. Inductively, define

$$\mathcal{H}_{+\tau} = \left\{h_{+\tau} \in \left(\Delta(\mathcal{C} \times \Delta(\mathcal{H}))\right)^S \colon \forall s \in S,\, h_{+\tau}(s) \in \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{+(\tau-1)}))\right\}.$$

Similarly, we can embed $\mathcal{H}_{+\tau}$ into $\mathcal{H}$. We call $\bigcup_{\tau\geq 1} \mathcal{H}_{+\tau}$ the domain of *finite-step-ahead acts*.

Lemma 9. *The domain of finite-step-ahead acts $\bigcup_{\tau\geq 1} \mathcal{H}_{+\tau}$ is a dense subset of $\mathcal{H}$. Also, $\bigcup_{\tau\geq 1} \Delta(\mathcal{C} \times \Delta(\mathcal{H}_{+\tau}))$ is a dense subset of $\Delta(\mathcal{C} \times \Delta(\mathcal{H}))$.*

This result is analogous to Proposition 1 in Hayashi (2005) and its proof is omitted. It is useful to establish the existence of a risk equivalent as in Lemma 9 of Hayashi (2005). We implicitly applied a similar result in Appendix A.

Appendix C: Proofs of Theorems 4 and 5

We prove the sufficiency of the axioms. The proof of necessity is routine.

C1 *Representation of risk preference*

When $\{\succeq_{s^t}\}$ is restricted to the domain $\mathcal{C} \times \Delta(\mathcal{L})$, Axiom B3 (History Independence of Risk Preference) implies that $\{\succeq_{s^t}\}$ induces a single preference relation $\succeq$ defined on $\mathcal{C} \times \Delta(\mathcal{L})$. By Axiom B1 (Order) and Debreu's (1954) theorem, there is a continuous representation $V : \mathcal{C} \times \Delta(\mathcal{L}) \to \mathbb{R}$ of $\succeq$. We fix such a representation.

By Axiom B2 (Current Consumption Separability), $V(c, \cdot)$ and $V(\widehat{c}, \cdot)$ represent the same ranking over $\Delta(\mathcal{L})$, hence V has the form

$$V(c, a) = \widehat{W}(c, V(\widehat{c}, a)) \quad \forall (c, a) \in \mathcal{C} \times \Delta(\mathcal{L}) \tag{32}$$

for some function $\widehat{W}$ that is strictly increasing in the second argument. Because of Axiom B4 (First-Stage Independence), $V(\widehat{c}, a)$ has the form

$$V(\widehat{c}, a) = \zeta\left(\int_{\mathcal{L}} U(l)\, da(l)\right), \tag{33}$$

where ζ is a strictly increasing function and U is a vNM index.

Because of Axiom B5 (Second-Stage Independence), U has the form

$$U(l) = \phi\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} \widehat{u}(c', a')\, dl(c', a')\right), \tag{34}$$

where ϕ is a strictly increasing function and $\widehat{u}$ is a vNM index.

By Axiom B6 (Dynamic Consistency) and a similar argument as in Appendix A1, $\widehat{u}$ and V are ordinally equivalent. Hence, we deduce that

$$\widehat{u}(c', a') = u(V(c', a')), \tag{35}$$

where u is a strictly increasing function.

Plugging (33), (34), and (35) into (32) yields

$$V(c, a) = \widehat{W}\left(c, \zeta\left(\int_{\mathcal{L}} \phi\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a')\right) da(l)\right)\right).$$

Now define W by

$$W(c, x) = \widehat{W}(c, \zeta \circ \phi \circ u(x)),$$

which is strictly increasing in the second argument. Then we have

$$\widehat{W}(c, \zeta(z)) = W(c, u^{-1} \circ \phi^{-1}(z)) \tag{36}$$

and hence,

$$V(c, a) = W\left(c, u^{-1} \circ \phi^{-1}\left(\int_{\mathcal{L}} \phi\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a')\right) da(l)\right)\right).$$

Let $v = \phi \circ u$. We obtain representation (14).

C2 *Extension to the whole domain*

Define $V_{s^t} : \mathcal{C} \times \Delta(\mathcal{H})$ by

$$V_{s^t}(c, p) = V(c, a) \tag{37}$$

for each $(c, p) \in \mathcal{C} \times \Delta(\mathcal{H})$, where $a \in \Delta(\mathcal{L})$ is such that $(c, p) \sim_{s^t} (c, a)$. The existence of such a risk equivalent a follows from Lemma 9, Dynamic Consistency (Axiom B6), compactness of $\mathcal{C}$, and continuity of $\succeq_{s^t}$ (see Lemma 9 in Hayashi 2005). Using definition (37) and (32), we derive

$$V_{s^t}(c, p) = V(c, a) = \widehat{W}(c, V(\widehat{c}, m)) = \widehat{W}(c, V_{s^t}(\widehat{c}, p)). \tag{38}$$

When our Axioms B1, B4, B5, and B7 are restricted to $\Delta(\mathcal{H}_{+1})$, they satisfy the conditions in Theorem 4.2 in Seo (2009). By this theorem, $V_{s^t}(\widehat{c}, \cdot)$ restricted to $\Delta(\mathcal{H}_{+1})$ is ordinally equivalent to a second-order subjective expected utility representation, and hence has the form

$$V_{s^t}(\widehat{c}, p_{+1}) = \zeta_{s^t}\left(\int_{\mathcal{H}_{+1}} U_{s^t}(h_{+1})\, dp_{+1}(h_{+1})\right) \tag{39}$$

and

$$U_{s^t}(h_{+1}) = \int_{\mathcal{P}_{s^t}} \phi_{s^t}\left(\sum_{s\in S} \pi(s) \int_{\mathcal{C}\times\Delta(\mathcal{L})} \widehat{u}_{s^t}(c', a')\, dh_{+1}(s)(c', a')\right) d\mu_{s^t}(\pi), \tag{40}$$

where ζ_{s^t} and ϕ_{s^t} are strictly increasing functions and $\widehat{u}_{s^t}$ is a vNM index. By Axiom B6 (Dynamic Consistency) and a similar argument in Appendix A1, $\widehat{u}_{s^t}$ and V are ordinally equivalent over $\mathcal{C} \times \Delta(\mathcal{L})$. Thus, there is a monotone transformation u_{s^t} such that

$$\widehat{u}_{s^t}(c', a') = u_{s^t}(V(c', a')) \tag{41}$$

for every $(c', a') \in \mathcal{C} \times \Delta(\mathcal{L})$.

Equations (33) and (37) imply that on $\Delta(\mathcal{L})$,

$$\zeta\left(\int_{\mathcal{L}} U(l)\, da(l)\right) = V(\widehat{c}, a) = V_{s^t}(\widehat{c}, a) = \zeta_{s^t}\left(\int_{\mathcal{L}} U_{s^t}(l)\, da(l)\right),$$

which in turn implies that on $\mathcal{L}$,

$$\zeta(U(l)) = V(\widehat{c}, \delta[l]) = V_{s^t}(\widehat{c}, \delta[l]) = \zeta_{s^t}(U_{s^t}(l)).$$

Hence, we deduce that

$$U_{s^t} = \zeta_{s^t}^{-1} \circ \zeta \circ U,$$

which implies that

$$\zeta\left(\int_{\mathcal{L}} U(l)\, da(l)\right) = \zeta_{s^t}\left(\int_{\mathcal{L}} \zeta_{s^t}^{-1} \circ \zeta \circ U(l)\, da(l)\right).$$

By the additivity of integral formula, we have

$$\zeta_{s^t}^{-1} \circ \zeta(\alpha x + (1-\alpha) y) = \alpha \zeta_{s^t}^{-1} \circ \zeta(x) + (1-\alpha)\zeta_{s^t}^{-1} \circ \zeta(y)$$

for all x, y in the range of U and all $\alpha \in [0, 1]$. Therefore, ζ_{s^t} and ζ are identical up to positive affine transformations. Thus, without loss of generality, we can take $\zeta_{s^t} = \zeta$ and $U_{s^t} = U$ for all s^t.

Equations (34), (35), (40), and (41) imply that on $\mathcal{L}$,

$$\begin{aligned}\phi\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u \circ V(c', a')\, dl(c', a')\right) &= U(l) = U_{s^t}(l)\\ &= \phi_{s^t}\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u_{s^t} \circ V(c', a')\, dl(c', a')\right),\end{aligned}$$

which in turn implies that

$$\phi \circ u \circ V(c', a') = U(\delta[c', a']) = U_{s^t}(\delta[c', a']) = \phi_{s^t} \circ u_{s^t} \circ V(c', a').$$

Hence, we have $\phi \circ u = \phi_{s^t} \circ u_{s^t}$, which implies that

$$\phi\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u \circ V(c', a')\, dl(c', a')\right) = \phi_{s^t}\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} \phi_{s^t}^{-1} \circ \phi \circ u \circ V(c', a')\, dl(c', a')\right).$$

By the same reasoning as above, ϕ_{s^t} and ϕ are identical up to positive affine transformations. Therefore, without loss of generality, we can set $\phi_{s^t} = \phi$ and $u_{s^t} = u$ for all s^t.

Now, plugging (39), (40), and (41) into (38), we obtain that, on $\mathcal{C} \times \Delta(\mathcal{H}_{+1})$,

$$\begin{aligned} V_{s^t}(c, p_{+1}) &= \widehat{W}\bigg(c, \zeta\bigg(\int_{\mathcal{H}_{+1}} \int_{\mathcal{P}_{s^t}} \phi\bigg(\sum_{s\in S} \pi(s) \\ &\qquad \times \int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dh_{+1}(s)(c', a')\bigg)\, d\mu_{s^t}(\pi)\, dp_{+1}(h_{+1})\bigg)\bigg) \\ &= W\bigg(c, u^{-1}\circ\phi^{-1}\bigg(\int_{\mathcal{H}_{+1}} \int_{\mathcal{P}_{s^t}} \phi\bigg(\sum_{s\in S} \pi(s) \\ &\qquad \times \int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dh_{+1}(s)(c', a')\bigg)\, d\mu_{s^t}(\pi)\, dp_{+1}(h_{+1})\bigg)\bigg), \end{aligned}$$

where the second equality follows from (36).

Finally, we extend the above representation to the whole domain $\mathcal{C} \times \Delta(\mathcal{H})$. A risk equivalent always exists as discussed before. By a similar argument as in Appendix A2, for every s^t and every $h \in \mathcal{H}$, there exists a one-step-ahead act $h_{+1} \in \mathcal{H}_{+1}$ such that $h(s)$ and $h_{+1}(s)$ are stochastically equivalent. We call h_{+1} the equivalent one-step-ahead act of h. By Axiom B6 (Dynamic Consistency), we deduce

$$(c, \delta[h]) \sim_{s^t} (c, \delta[h_{+1}]). \tag{42}$$

Suppose that $p \in \Delta(\mathcal{H})$ has a finite support $\{h^1, h^2, \dots, h^m\}$, with $p = \sum_i \alpha_i \delta[h^i]$, $\alpha_i \in (0, 1)$, and $\sum_i \alpha_i = 1$. For each h^i, $i = 1, \dots, m$, let $h^i_{+1} \in \mathcal{H}_{+1}$ be its equivalent one-step-ahead act. Let $p_{+1} \in \Delta(\mathcal{H}_{+1})$ be a probability measure with a finite support such that the support is $\{h^1_{+1}, h^2_{+1}, \dots, h^m_{+1}\}$, and for each $i = 1, \dots, m$, $p_{+1}(\{h^i_{+1}\}) = p(\{h^i\})$. By repeated applications of Axiom B4 (First-Stage Independence) and (42), we obtain $(c, p) \sim_{s^t} (c, p_{+1})$. This relation is also true for arbitrary $c \in \mathcal{C}$ because of Axiom B2 (Current Consumption Separability). By continuity of $\succeq_{s^t}$, the claim extends to arbitrary p. Hence, we have

$$\begin{aligned} V_{s^t}(c, p) &= V_{s^t}(c, p_{+1}) \\ &= W\bigg(c, u^{-1}\circ\phi^{-1}\bigg(\int_{\mathcal{H}_{+1}} \int_{\mathcal{P}_{s^t}} \phi\bigg(\sum_{s\in S} \pi(s) \\ &\qquad \times \int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dh_{+1}(s)(c', a')\bigg)\, d\mu_{s^t}(\pi)\, dp_{+1}(h_{+1})\bigg)\bigg) \\ &= W\bigg(c, u^{-1}\circ\phi^{-1}\bigg(\int_{\mathcal{H}} \int_{\mathcal{P}_{s^t}} \phi\bigg(\sum_{s\in S} \pi(s) \\ &\qquad \times \int_{\mathcal{C}\times\Delta(\mathcal{H})} u(V_{s^t,s}(c', a'))\, dh(s)(c', a')\bigg)\, d\mu_{s^t}(\pi)\, dp(h)\bigg)\bigg), \end{aligned}$$

where we use the fact that $h(s)$ and $h_{+1}(s)$ are stochastically equivalent to derive the last equality. Defining $v = \phi \circ u$, we obtain the representation as in the theorem.

C3 *Proof of uniqueness*

Suppose $(\{\tilde{V}_{s^t}\}, \tilde{W}, \tilde{u}, \tilde{v}, \{\tilde{\mu}_{s^t}\})$ and $(\{V_{s^t}\}, W, u, v, \{\mu_{s^t}\})$ represent the same preference. On the domain of deterministic consumption streams $\mathcal{C}^\infty$, all $\tilde{V}_{s^t}$ coincide with the common function $\tilde{V}$ and all V_{s^t} coincide with the common function V. Since $\tilde{V}$ and V are ordinally equivalent over $\mathcal{C}^\infty$, there is a monotone transformation Φ such that

$$\tilde{V}(y) = \Phi \circ V(y)$$

for all $y \in \mathcal{C}^\infty$. By (37), we have $\tilde{V}_{s^t} = \Phi \circ V_{s^t}$.

Since

$$\tilde{W}(c, \tilde{V}(y)) = \tilde{V}(c, y) = \Phi(V(c, y)) = \Phi(W(c, V(y))) = \Phi(W(c, \Phi^{-1}(\tilde{V}(y)))),$$

we have $\tilde{W}(c, z) = \Phi(W(c, \Phi^{-1}(z)))$.

On $\mathcal{L}$, $\int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a')$ and $\int_{\mathcal{C}\times\Delta(\mathcal{L})} \tilde{u}(\tilde{V}(c', a'))\, dl(c', a')$ are equivalent mixture-linear representations of the second-stage risk preference conditional on the fixed current consumption $\widehat{c}$ defined in Appendix C1. Therefore, there exist constants A, B with $A > 0$ such that

$$\tilde{u}(\tilde{V}(c', a')) = Au(V(c', a')) + B$$

for all $(c', a') \in \mathcal{C} \times \Delta(\mathcal{L})$. Since $\tilde{V} = \Phi \circ V$, we obtain $\tilde{u} \circ \Phi = Au + B$.

On $\Delta(\mathcal{L})$,

$$\int_{\mathcal{L}} \tilde{v} \circ \tilde{u}^{-1}\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} \tilde{u}(\tilde{V}(c', a'))\, dl(c', a')\right) da(l)$$

and

$$\int_{\mathcal{L}} v \circ u^{-1}\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a')\right) da(l)$$

are equivalent mixture-linear representations of the first-stage risk preference conditional on the fixed current consumption $\widehat{c}$. Hence, there exist constants D, E with $D > 0$ such that

$$\tilde{v} \circ \tilde{u}^{-1}\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} \tilde{u}(\tilde{V}(c', a'))\, dl(c', a')\right) = Dv \circ u^{-1}\left(\int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a')\right) + E.$$

From the previous result, we have

$$\begin{aligned}\int_{\mathcal{C}\times\Delta(\mathcal{L})} \tilde{u}(\tilde{V}(c', a'))\, dl(c', a') &= \int_{\mathcal{C}\times\Delta(\mathcal{L})} \tilde{u} \circ \Phi(V(c', a'))\, dl(c', a') \\ &= A \int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a') + B.\end{aligned}$$

Let $\int_{\mathcal{C}\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a') = x$. Then we have

$$\tilde{v} \circ \tilde{u}^{-1}(Ax + B) = Dv \circ u^{-1}(x) + E.$$

From $\tilde{u} \circ \Phi(x) = Au(x) + B$, it follows that

$$\tilde{v} \circ \tilde{u}^{-1}(Ax + B) = \tilde{v} \circ \tilde{u}^{-1}(Au \circ u^{-1}(x) + B) = \tilde{v} \circ \Phi(u^{-1}(x)).$$

Thus, combining the above two equations, we obtain

$$\tilde{v} \circ \Phi(u^{-1}(x)) = Dv \circ u^{-1}(x) + E.$$

So $\tilde{v} \circ \Phi = Dv + E$.

Appendix D: Proofs for Section 4.4

Proof of Proposition 3. If $v \circ u^{-1}$ is concave, it is straightforward to check that $\{\succeq_{s^t}\}$ is risk averse in the first stage. We now prove the reverse direction. Pick any $l_1, l_2 \in \mathcal{L}$ and $\lambda \in [0, 1]$. By Theorem 3, we have

$$V(c, \lambda l_1 + (1 - \lambda) l_2) = W\big(c, v^{-1}\big(\lambda v \circ u^{-1}(V^*(l_1)) + (1 - \lambda) v \circ u^{-1}(V^*(l_2))\big)\big),$$

where

$$V^*(l) = \int_{\mathcal{C} \times \Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a')$$

is a mixture-linear function on $\mathcal{L}$. Also we have

$$V\big(c, \delta[\lambda l_1 \oplus (1 - \lambda) l_2]\big) = W\left(c, v^{-1}\left(\int_{\mathcal{L}} v \circ u^{-1}(\lambda V^*(l_1) + (1 - \lambda) V^*(l_2))\right)\right).$$

From the definition of risk aversion in the first stage, we have $(c, \delta[\lambda l_1 \oplus (1-\lambda) l_2]) \succeq_{s^t} (c, \lambda l_1 + (1 - \lambda) l_2)$. Thus,

$$\lambda v \circ u^{-1}(V^*(l_1)) + (1 - \lambda) v \circ u^{-1}(V^*(l_2)) \le v \circ u^{-1}(\lambda V^*(l_1) + (1 - \lambda) V^*(l_2)).$$

We may vary l_1 and l_2 to cover the whole domain of $v \circ u^{-1}$. The above inequality implies that $v \circ u^{-1}$ is concave. □

Proof of Proposition 4. Suppose $\{\succeq^i_{s^t}\}$ is more risk averse than $\{\succeq^j_{s^t}\}$ in the first stage. By definition they rank deterministic consumption streams in the same way and rank lotteries in the second stage in the same way. So there exist representations such that $V^i = V^j$, $W^i = W^j$, and $u^i = u^j$.

Since $v^i(V^i(\cdot))$ and $v^j(V^j(\cdot))$ are ordinally equivalent over C^∞, there is a monotone transformation Ψ such that $v^i = \Psi \circ v^j$. It remains to show that Ψ is concave. Let $y, y', y'' \in C^\infty$ be such that $(c, \delta[\delta[y]]) \sim^j_{s^t} (c, \lambda\delta[\delta[y']] + (1 - \lambda)\delta[\delta[y'']])$. This is possible due to continuity of preference ordering. Thus, we have $v^j(V^j(y)) = \lambda v^j(V^j(y')) + (1 - \lambda) v^j(V^j(y''))$. Since $\succeq^i_{s^t}$ is more risk averse than $\succeq^j_{s^t}$ in the first stage, we have $(c, \delta[\delta[y]]) \succeq^i_{s^t} (c, \lambda\delta[\delta[y']] + (1 - \lambda)\delta[\delta[y'']])$, which implies $v^i(V^i(y)) \ge \lambda v^i(V^i(y')) + (1 - \lambda) v^i(V^i(y''))$.

Since $v^i(V^i(\cdot)) = v^i((V^j(\cdot))) = \Psi(v^j(V^j(\cdot)))$, we obtain

$$\Psi(u^j(V^j(y))) \ge \lambda\Psi(u^j(V^j(y'))) + (1 - \lambda)\Psi(u^j(V^j(y''))).$$

One can choose y, y', and y'' so as to cover the whole range of $u^j \circ V^j$. Thus, Ψ is concave. The proof of the other direction of the proposition is routine. □

Proof of Proposition 5. When we restrict to the subdomain $\mathcal{C} \times \Delta(\mathcal{L})$, we immediately deduce that ambiguity aversion implies risk aversion in the first stage. Now, we consider the reverse statement. Given $h_{+1}, h'_{+1} \in \mathcal{H}_{+1}$, we define $a \in \Delta(\mathcal{L})$ as in Definition 6. We then have

$$\begin{aligned}
&\bigl(c, a(\lambda\delta[h_{+1}] \oplus (1-\lambda)\delta[h'_{+1}], \pi)\bigr)\\
&\quad= \bigl(c, \lambda a(\delta[h_{+1}], \pi) \oplus (1-\lambda) a(\delta[h'_{+1}], \pi)\bigr)\\
&\quad\succeq_{s^t} \bigl(c, \lambda a(\delta[h_{+1}], \pi) + (1-\lambda) a(\delta[h'_{+1}], \pi)\bigr)\\
&\quad= \bigl(c, a(\lambda\delta[h_{+1}] + (1-\lambda)\delta[h'_{+1}], \pi)\bigr),
\end{aligned}$$

for all $\pi \in \mathcal{P}_{s^t}$, where the relation $\succeq_{s^t}$ follows from the definition of risk aversion in the first stage. By Axiom B7 (Dominance), we obtain

$$(c, \lambda\delta[h_{+1}] \oplus (1-\lambda)\delta[h'_{+1}]) \succeq_{s^t} (c, \lambda\delta[h_{+1}] + (1-\lambda)\delta[h'_{+1}]).$$

It follows from the definition that the decision maker is ambiguity averse. □

Proof of Proposition 6. First, we show that comparative risk aversion in the first stage implies comparative ambiguity aversion. To show that i is more ambiguous averse than j, we need to show

$$(c, \delta[l]) \succeq^j_{s^t} (c, \delta[h_{+1}]) \quad\Longrightarrow\quad (c, \delta[l]) \succeq^i_{s^t} (c, \delta[h_{+1}]). \tag{43}$$

Given $h_{+1} \in \mathcal{H}_{+1}$ and $\mu_{s^t} \in \Delta(\mathcal{P}_{s^t})$, define $b(h_{+1}, \mu_{s^t}) \in \Delta(\mathcal{L})$ as

$$b(h_{+1}, \mu_{s^t})(L) = \mu_{s^t}\bigl(\{\pi \in \mathcal{P}_{s^t} \colon l(h_{+1}, \pi) \in L\}\bigr)$$

for every Borel set $L \subset \mathcal{L}$, where $l(h_{+1}, \pi) = \sum_s h_{+1}(s)\pi(s)$.

For any preferences $\{\succeq_{s^t}\}$ satisfying Axioms B1–B7, we use Theorem 3 to compute

$$\begin{aligned}
&V(c, b(h_{+1}, \mu_{s^t}))\\
&\quad= W\biggl(c, v^{-1}\biggl(\int_{\mathcal{L}} v \circ u^{-1}\biggl(\int_{C\times\Delta(\mathcal{L})} u(V(c', a'))\, dl(c', a')\biggr)\, db(h_{+1}, \mu_{s^t})(l)\biggr)\biggr)\\
&\quad= W\biggl(c, v^{-1}\biggl(\int_{\mathcal{P}_{s^t}} v \circ u^{-1}\biggl(\sum_{s\in S} \pi(s) \int_{C\times\Delta(\mathcal{H}_{+1})} u(V(c', a'))\, dh_{+1}(s)(c', a')\biggr)\, d\mu_{s^t}(\pi)\biggr)\biggr)\\
&\quad= V_{s^t}(c, \delta[h_{+1}]),
\end{aligned}$$

where we use the change of variables theorem (Aliprantis and Border 1999, p. 452) to derive the second equality. This implies that $(c, b(h_{+1}, \mu_{s^t})) \sim_{s^t} (c, \delta[h_{+1}])$. Likewise, we have $(c, b(h'_{+1}, \mu_{s^t})) \sim_{s^t} (c, \delta[h'_{+1}])$. Thus, (43) is equivalent to

$$(c, \delta[l]) \succeq^j_{s^t} (c, b(h_{+1}, \mu)) \quad\Longrightarrow\quad (c, \delta[l]) \succeq^i_{s^t} (c, b(h_{+1}, \mu)).$$

This relation holds true because i is more risk averse than j in the first stage.

Turn to the proof of the converse statement. Fix a set $E \subset S$ such that $\lambda = \int_{\mathcal{P}_{s^t}} \pi(E)\, d\mu_{s^t}(\pi) \in (0,1)$. Suppose $(c, \delta[l]) \succeq^j_{s^t} (c, \lambda\delta[l'] + (1-\lambda)\delta[l''])$ for $l, l', l'' \in \mathcal{L}$. Let h_{+1} be the one-step-ahead act that gives l' if event E happens and gives l'', otherwise. Then by definition, we can show that $b(h_{+1}, \mu_{s^t}) = \lambda\delta[l'] + (1-\lambda)\delta[l'']$. Using the representation in Theorem 3, we can verify that $(c, b(h_{+1}, \mu_{s^t})) \sim^j_{s^t} (c, \delta[h_{+1}])$ or $(c, \lambda\delta[l'] + (1-\lambda)\delta[l'']) \sim^j_{s^t} (c, \delta[h_{+1}])$, which implies $(c, \delta[l]) \succeq^j_{s^t} (c, \delta[h_{+1}])$. By comparative ambiguity aversion, we have $(c, \delta[l]) \succeq^i_{s^t} (c, \delta[h_{+1}])$. Since $(c, \lambda\delta[l'] + (1-\lambda)\delta[l'']) \sim^i_{s^t} (c, \delta[h_{+1}])$ holds as well, we obtain $(c, \delta[l]) \succeq^i_{s^t} (c, \lambda\delta[l'] + (1-\lambda)\delta[l''])$. Hence, we have

$$(c, \delta[l]) \succeq^j_{s^t} (c, \lambda\delta[l'] + (1-\lambda)\delta[l'']) \quad \Longrightarrow \quad (c, \delta[l]) \succeq^i_{s^t} (c, \lambda\delta[l'] + (1-\lambda)\delta[l'']).$$

We can extend this result to all $\lambda \in (0,1)$ by continuity (Axiom B1) and Axiom B4 (First Stage Independence). We can also extend this result to all finite lotteries over $\mathcal{L}$ by repeatedly applying the above argument. We finally extend it to all lotteries over $\mathcal{L}$ by continuity of preferences (Axiom B1). □

Appendix E: Proofs for Section 5

Proof of Proposition 7. Define

$$\phi_t(\alpha) = V_t(c + \alpha\delta)$$

for an adapted process (δ_t). Using (20)–(22), we have

$$\phi_t(\alpha) = W\big(c_t + \alpha\delta_t, \mathcal{R}_t(V_{t+1}(c + \alpha\delta))\big).$$

Taking derivatives in the preceding equation yields

$$\begin{aligned}\phi_t'(0) &= W_1(c_t, \mathcal{R}_t(V_{t+1}))\delta_t \\ &\quad + \frac{W_2(c_t, \mathcal{R}_t(V_{t+1}))}{v'(\mathcal{R}_t(V_{t+1}))}\mathbb{E}_{\mu_t}\left\{\frac{v' \circ u^{-1}(\mathbb{E}_{\pi_{z,t}}[u(V_{t+1})])}{u'(u^{-1}(\mathbb{E}_{\pi_{z,t}}[u'(V_{t+1})]))}\mathbb{E}_{\pi_{z,t}}[u'(V_{t+1})\phi_{t+1}'(0)]\right\}.\end{aligned}$$

Define λ_t as in (18) and $\mathcal{E}_t^z$ as in (19). We obtain

$$\phi_t'(0) = \lambda_t\delta_t + \mathbb{E}_t\left[\frac{\mathcal{E}_{t+1}^z}{\mathcal{E}_t^z}\phi_{t+1}'(0)\right],$$

where $\mathbb{E}_t$ is the conditional expectation operator with respect to the predictive distribution $\sum_z \mu_t(z)\pi_z(\cdot|s^t)$. From this equation and the definition in (17), we can derive that $\xi_t^z = \mathcal{E}_t^z\lambda_t$. □

Proof of Proposition 8. When the utility function takes the homothetic form, we use Proposition 7 and the definition of the pricing kernel to derive (23). Alternatively, we may write the pricing kernel in terms of the market return as in Epstein and Zin (1989).

In a complete market, wealth X_t satisfies

$$X_t = \mathbb{E}_t\left[\sum_{s=t}^{\infty} \frac{\xi_s^z}{\xi_t^z} c_s\right].$$

That is, time t wealth is equal to the present value of the consumption stream. By Lemma 6.25 in Skiadas (2009), we have

$$V_t = \lambda_t X_t. \tag{44}$$

By (18), we have the relation

$$\frac{c_t}{V_t} = \left(\frac{\lambda_t}{1-\beta}\right)^{-1/\rho}.$$

Thus, the consumption–wealth ratio satisfies

$$\frac{c_t}{X_t} = \frac{c_t \lambda_t}{V_t} = \left(\frac{\lambda_t}{1-\beta}\right)^{-1/\rho} \lambda_t. \tag{45}$$

Eliminating λ_t from (44) and (45) yields

$$\begin{aligned} V_t = \lambda_t X_t &= (1-\beta)^{\frac{1}{1-\rho}} c_t^{\frac{-\rho}{1-\rho}} X_t^{\frac{1}{1-\rho}} \\ &= (1-\beta)^{\frac{1}{1-\rho}} c_t^{\frac{-\rho}{1-\rho}} R_t^{\frac{1}{1-\rho}} (X_{t-1} - c_{t-1})^{\frac{1}{1-\rho}}, \end{aligned} \tag{46}$$

where we use (25) to derive the last equality. Note that the second equality implies that

$$\frac{X_t}{c_t} = \frac{1}{1-\beta}\left(\frac{V_t}{c_t}\right)^{1-\rho}.$$

As a result, for unitary EIS ($\rho = 1$), the consumption–wealth ratio is equal to $1-\beta$.

Now, substituting (46) into (23) and manipulating, we derive

$$M_{t+1}^z = \beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho} \left(\frac{c_{t+1}^{\frac{-\rho}{1-\rho}} R_{t+1}^{\frac{1}{1-\rho}}}{\mathcal{R}_t\big(c_{t+1}^{\frac{-\rho}{1-\rho}} R_{t+1}^{\frac{1}{1-\rho}}\big)}\right)^{\rho-\gamma} \left(\frac{\big(\mathbb{E}_{\pi_{z,t}}\big[c_{t+1}^{\frac{-\rho(1-\gamma)}{1-\rho}} R_{t+1}^{\frac{1-\gamma}{1-\rho}}\big]\big)^{\frac{1}{1-\gamma}}}{\mathcal{R}_t\big(c_{t+1}^{\frac{-\rho}{1-\rho}} R_{t+1}^{\frac{1}{1-\rho}}\big)}\right)^{-(\eta-\gamma)}.$$

Writing in terms of consumption growth and manipulating, we obtain

$$\begin{aligned} M_{t+1}^z = \left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho}\right)^{\frac{1-\gamma}{1-\rho}} R_{t+1}^{\frac{1-\gamma}{1-\rho}-1} \left(\mathbb{E}_{\pi_{z,t}}\left[\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho} R_{t+1}\right)^{\frac{1-\gamma}{1-\rho}}\right]\right)^{\frac{-(\eta-\gamma)}{1-\gamma}} \\ \times \left[\mathcal{R}_t\left(\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho} R_{t+1}\right)^{\frac{1}{1-\rho}}\right)\right]^{\eta-\rho}. \end{aligned}$$

In complete markets, the following Euler equation holds:

$$\mathbb{E}_t[M_{t+1}^z R_{t+1}] = 1.$$

By substituting the preceding pricing kernel into this Euler equation, we obtain

$$1 = \mathbb{E}_t \left\{ \left(\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\rho} R_{t+1} \right)^{\frac{1-\gamma}{1-\rho}} \left(\mathbb{E}_{\pi_{z,t}} \left[\left(\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\rho} R_{t+1} \right)^{\frac{1-\gamma}{1-\rho}} \right] \right)^{\frac{-(\eta-\gamma)}{1-\gamma}} \right\} \times \left[\mathcal{R}_t \left(\left(\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\rho} R_{t+1} \right)^{\frac{1}{1-\rho}} \right) \right]^{\eta-\rho}.$$

Noting that $\mathbb{E}_t = \mathbb{E}_{\mu_t} \mathbb{E}_{\pi_{z,t}}$ and using the definition of $\mathcal{R}_t$, we obtain

$$\left[\mathcal{R}_t \left(\left(\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\rho} R_{t+1} \right)^{\frac{1}{1-\rho}} \right) \right]^{1-\rho} = 1.$$

Thus,

$$\mathcal{R}_t \left(\left(\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\rho} R_{t+1} \right)^{\frac{1}{1-\rho}} \right) = 1,$$

so that we can write the pricing kernel as (24). □

References

Aliprantis, Charalambos D. and Kim C. Border (1999), *Infinite Dimensional Analysis*. Springer, Berlin. [466]

Anscombe, Francis J. and Robert J. Aumann (1963), "A definition of subjective probability." *Annals of Mathematical Statistics*, 34, 199–205. [426]

Backus, David E., Bryan R. Routledge, and Stanley E. Zin (2005), "Exotic preferences for macroeconomists." In *NBER Macroeconomics Annual 2004* (Mark Gertler and Kenneth Rogoff, eds.), 319–390, MIT Press, Cambridge. [423]

Brandenberger, Adam and Eddie Dekel (1993), "Hierarchies of beliefs and common knowledge." *Journal of Economic Theory*, 59, 189–198. [458]

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio (2008), "Uncertainty averse preferences." Working paper, Carlo Alberto Notebooks 77, Collegio Carlo Alberto. [423]

Chen, Hui, Nengjiu Ju, and Jianjun Miao (2009), "Dynamic asset allocation with ambiguous return predictability." Unpublished paper, Department of Economics, Boston University. [427]

Chew, Soo Hong and Larry G. Epstein (1991), "Recursive utility under uncertainty." In *Equilibrium Theory in Infinite Dimensional Spaces* (M. Ali Khan and Nicholas C. Yannelis, eds.), 352–369, Springer, Berlin. [432, 439, 450]

Chew, Soo Hong and Jacob S. Sagi (2008), "Small worlds: Modeling attitudes towards sources of uncertainty." *Journal of Economic Theory*, 139, 1–24. [424]

Debreu, Gerard (1954), "Representation of a preference ordering by a numerical function." In *Decision Processes* (Robert M. Thrall, Clyde H. Coombs, and Robert L. Davis, eds.), 159–165, Wiley, New York. [432, 450, 460]

Duffie, Darrell and Costis Skiadas (1994), "Continuous-time security pricing: A utility gradient approach." *Journal of Mathematical Economics*, 23, 107–131. [446]

Ellsberg, Daniel (1961), "Risk, ambiguity, and the Savage axiom." *Quarterly Journal of Economics*, 75, 643–669. [423]

Epstein, Larry G. (1999), "A definition of uncertainty aversion." *Review of Economic Studies*, 66, 579–608. [437]

Epstein, Larry G. (2010), "A paradox for the 'smooth ambiguity' model of preference." *Econometrica*, 78, 2085–2099. [427, 428, 429]

Epstein, Larry G. and Martin Schneider (2003), "Recursive multiple-priors." *Journal of Economic Theory*, 113, 1–31. [425, 426, 432, 434, 449, 450]

Epstein, Larry G. and Tan Wang (1994), "Intertemporal asset pricing under Knightian uncertainty." *Econometrica*, 62, 283–322. [426, 447, 449]

Epstein, Larry G. and Stanley Zin (1989), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework." *Econometrica*, 57, 937–969. [424, 425, 426, 432, 434, 436, 437, 439, 447, 448, 450, 467]

Ergin, Haluk and Faruk Gul (2009), "A theory of subjective compound lotteries." *Journal of Economic Theory*, 144, 899–929. [424, 449]

Ghirardato, Paolo and Massimo Marinacci (2002), "Ambiguity made precise: A comparative foundation." *Journal of Economic Theory*, 102, 251–289. [437, 449]

Gilboa, Itzhak and David Schmeidler (1989), "Maxmin expected utility with a non-unique prior." *Journal of Mathematical Economics*, 18, 141–153. [423, 424, 430, 437, 444, 449]

Gul, Faruk and Wolfgang Pesendorfer (2004), "Self-control and the theory of consumption." *Econometrica*, 72, 119–158. [439]

Halevy, Yoram (2007), "Ellsberg revisited: An experimental study." *Econometrica*, 75, 503–536. [445]

Hanany, Eran and Peter Klibanoff (2007), "Updating preferences with multiple priors." *Theoretical Economics*, 2, 261–298. [449]

Hanany, Eran and Peter Klibanoff (2009), "Updating ambiguity averse preferences." *B.E. Journal of Theoretical Economics (advances)*, 9(1), 1–51. [448]

Hansen, Lars Peter (2007), "Beliefs, doubts and learning: Valuing macroeconomic risk." *American Economic Review Papers and Proceedings*, 97, 1–30. [424, 426, 448]

Hansen, Lars Peter, John C. Heaton, and Nan Li (2008), "Consumption strikes back? Measuring long-run risk." *Journal of Political Economy*, 116, 260–302. [448]

Hansen, Lars P. and Thomas J. Sargent (2001), "Robust control and model uncertainty." *American Economic Review*, 91, 60–66. [425, 449, 450]

Hansen, Lars P. and Thomas J. Sargent (2007a), "Recursive robust estimation and control without commitment." *Journal of Economic Theory*, 136, 1–27. [426, 448]

Hansen, Lars Peter and Thomas J. Sargent (2007b), *Robustness*. Princeton University Press, Princeton. [423, 425]

Hansen, Lars Peter and Thomas J. Sargent (2010), "Fragile beliefs and the price of uncertainty." *Quantitative Economics*, 1, 129–162. [424]

Hansen, Lars Peter and Kenneth J. Singleton (1983), "Stochastic consumption, risk aversion, and the temporal behavior of asset returns." *Journal of Political Economy*, 91, 249–265. [423]

Hayashi, Takashi (2005), "Intertemporal substitution, risk aversion and ambiguity aversion." *Economic Theory*, 25, 933–956. [425, 426, 431, 434, 439, 449, 452, 460, 461]

Ju, Nengjiu and Jianjun Miao (2007), "Ambiguity, learning, and asset returns." Unpublished paper, Department of Economics, Boston University. [424]

Ju, Nengjiu and Jianjun Miao (forthcoming), "Ambiguity, learning, and asset returns." *Econometrica*. [424, 427, 447, 448]

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji (2005), "A smooth model of decision making under ambiguity." *Econometrica*, 73, 1849–1892. [423, 424, 426, 427, 428, 437, 438, 446, 448, 449]

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji (2009a), "On the smooth ambiguity model: A reply." Unpublished paper, Northwestern University. [424, 425, 426, 427, 434, 436, 437, 442, 448, 450]

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji (2009b), "Recursive smooth ambiguity preferences." *Journal of Economic Theory*, 144, 930–976. [427]

Klibanoff, Peter and Emre Ozdenoren (2007), "Subjective recursive expected utility." *Economic Theory*, 30, 49–87. [450]

Koopmans, Tjalling C. (1960), "Stationary ordinal utility and impatience." *Econometrica*, 28, 287–309. [433]

Kreps, David M. (1988), *Notes on the Theory of Choice*. Westview Press, Boulder. [427]

Kreps, David M. and Evan L. Porteus (1978), "Temporal resolution of uncertainty and dynamic choice theory." *Econometrica*, 46, 185–200. [425, 426, 436, 448, 450]

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini (2006a), "Ambiguity aversion, robustness, and the variational representation of preferences." *Econometrica*, 74, 1447–1498. [449]

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini (2006b), "Dynamic variational preferences." *Journal of Economic Theory*, 128, 4–44. [449, 450]

Marinacci, Massimo and Luigi Montrucchio (2010), "Unique solutions for stochastic recursive utilities." *Journal of Economic Theory*, 145, 1776–1804. [437, 442]

Mehra, Rajnish and Edward C. Prescott (1985), "The equity premium: A puzzle." *Journal of Monetary Economics*, 15, 145–161. [423]

Nau, Robert F. (2006), "Uncertainty aversion with second-order utilities and probabilities." *Management Science*, 52, 136–145. [424]

Savage, Leonard J. (1954), *The Foundations of Statistics*. Wiley, New York. [429, 434]

Schmeidler, David (1989), "Subjective probability and expected utility without additivity." *Econometrica*, 57, 571–587. [423]

Segal, Uzi (1987), "The Ellsberg paradox and risk aversion: An anticipated utility approach." *International Economic Review*, 28, 175–202. [445, 449]

Segal, Uzi (1990), "Two-stage lotteries without the reduction axiom." *Econometrica*, 58, 349–377. [445, 449]

Seo, Kyoungwon (2009), "Ambiguity and second-order belief." *Econometrica*, 77, 1575–1605. [423, 424, 427, 428, 430, 431, 438, 441, 442, 443, 444, 445, 449, 461]

Siniscalchi, Marciano (2011), "Dynamic choice under ambiguity." *Theoretical Economics*, 6, 379–421. [450]

Skiadas, Costis (1998), "Recursive utility and preferences for information." *Economic Theory*, 12, 293–312. [450]

Skiadas, Costis (2009), *Asset Pricing Theory*. Princeton University Press, Princeton. [468]

Strzalecki, Tomasz (2009), "Temporal resolution of uncertainty and recursive models of ambiguity aversion." Unpublished paper, Harvard University. [450]

Wang, Tan (2003), "Conditional preferences and updating." *Journal of Economic Theory*, 108, 286–321. [432, 439, 449]

Weil, Philippe (1989), "The equity premium puzzle and the risk-free rate puzzle." *Journal of Monetary Economics*, 24, 401–421. [447]

Submitted 2010-8-6. Final version accepted 2010-11-7. Available online 2010-11-7.