A Model of Consumption Smoothing*

Katsutoshi Wakai†

April, 2004

Abstract

This paper develops and axiomatizes a new behavioral notion, *time-variability aversion*, meaning an agent is averse to variations in atemporal *risky* prospects over time. This idea is captured by the agent selecting a sequence of discount factors (from a given set) that minimizes the present discounted value of a given *risk-adjusted utility* stream. One of the results that strongly justifies our postulates is that discount factors necessarily show gain/loss asymmetry once an agent's preferences become *consistent* over time. Thus, the previously noted anomaly on the discounted utility model becomes a norm. This representation is also applied under uncertainty, where an agent first considers time-variability aversion then considers subjective risk over states. Thus, the ranking among uncertain future prospects depends on the level of today's consumption, and discount factors exhibit gain/loss asymmetry over states by taking today's consumption as an endogenous reference point.

Keywords: time-variability aversion, gain/loss asymmetry, intertemporal substitution, risk aversion, recursive utility

---

*This paper is a revised version of Chapter II of my dissertation at Yale University (Wakai (2002)). I would like to express my appreciation to my advisor, Stephen Morris, and dissertation committee members, John Geanakoplos and Benjamin Polak, for their invaluable advice and support. I would also like to thank Larry Epstein and Itzhak Gilboa for valuable discussions and suggestions. I have benefited from comments from Richard Braun, Takao Kobayashi, Giuseppe Moscarini, Robert Shiller, Leeat Yariv, and seminar participants at the State University of New York at Buffalo, University of Iowa, University of Rochester, University of Texas Austin, University of Tokyo, and Washington University. All remaining errors are my own responsibility.

†Department of Economics, SUNY at Buffalo, 427 Fronczak Hall, Buffalo, NY, 14260; kwakai@buffalo.edu; http://pluto.fss.buffalo.edu/classes/eco/kwakai/

1 Introduction

1.1 Motivation and Outline

Consumption smoothing is multi-dimensional. For instance, risk aversion is regarded as the dislike of variations in payoffs of random variables within a period. By contrast, *time-variability* is variation in payoffs over time. We often use the discounted utility model under which atemporal concave utility functions represent a motive for both consumption smoothing over time and consumption smoothing under risk. However, consider the following thought experiment.

Assume that an agent is an expected utility maximizer for atemporal risk. Suppose that an agent needs to select one of the following two-period consumption plans, where payment is not certain: (a) receiving a lottery x on the first day and receiving a lottery y on the second day, (b) receiving a lottery y on the first day and receiving a lottery x on the second day, or (c) receiving a lottery z on both dates.[1] Suppose that lottery z is selected so that atemporal utility of z is an average of the atemporal utilities of x and y. If the agent's preferences follows the discounted utility model, the agent never strictly prefers (c), regardless of the value of the discount factor.[2] However, if the agent does not prefer fluctuations in the "expected utility" of each lottery (atemporal risky prospects themselves instead of realizations from each lottery) over time, it is highly likely that the agent strictly prefers (c).

This paper develops the model under which an agent might want to choose (c). We introduce a new behavioral notion, *time-variability aversion*, meaning an agent is averse to variations in atemporal (objectively) risky prospects over time, where each objective risk is evaluated by atemporal risk-preferences. Formally, in the representation, risk aversion is captured by the concavity of a von Neumann-Morgenstern utility function. Time-variability aversion is captured by the agent selecting a sequence of (normalized) discount factors from a given set that minimizes the present discounted value of a given *utility* stream (i.e., a *multiple-discount-factors* model). We provide an axiomatization for this representation by adapting a method developed in a different context by Gilboa and Schmeidler (1989).

One of the results that strongly justifies our postulates is that discount factors necessarily show gain/loss asymmetry (in fact, *recursive* gain/loss asymmetry) once an agent's preferences become *consistent* over time.[3] Thus, under our representation, the previously noted anomaly on

---

[1] This choice does not concern timing of resolution of uncertainty.

[2] The agent can perform one of the following calculations: (i) compute a reduced lottery by multiplying probabilities from the first lottery and from the second lottery, then take an expected value of the discounted utility or (ii) compute the expected utility of each lottery and apply the discounted utility over them. Both computations result in the same utility.

[3] Gain/loss asymmetry means that gains are discounted more than losses.

the discounted utility model becomes a norm.[4] More formally, the assignment of discount factors is determined recursively. At each time t, the agent compares the utility of present consumption with the present discounted value of utility of future consumption from $t+1$ onward, then selects the time-t discount factor to minimize the weighted sum of these two values. Intuitively, this representation exhibits time-variability aversion by allocating a high discount factor at $t+1$ when future consumption is low, and vice versa.

We also extend our model to the domain that describes uncertainty. Given non-time separability introduced by time-variability aversion, dynamic consistency is maintained only under a recursive representation. In this regard, we propose a new recursive aggregation, where an agent first considers time-variability aversion on a state-by-state basis and then aggregates discounted utility indices on each state with probability weights. Under this operation (applied recursively), discount factors depend on *tomorrow's* states. Then the agent discounts utility in good states more than in bad states so that discount factors show gain/loss asymmetry over *states*. This intertemporal substitution mechanism effectively boosts risk aversion over tomorrow's consumption.[5] This result is also analogous to that of loss aversion by Kahneman and Tversky (1979); in our model, today's utility implicitly becomes an endogenous reference point.

The paper proceeds as follows: Section 1 provides an overview of the paper. Section 2 axiomatizes the notion of time-variability aversion under certainty and derives the utility representation with multiple discount factors. Section 3 extends the representation with time-variability aversion under uncertainty. Section 4 compares our model with other intertemporal utility functions. Section 5 concludes the paper.

1.2 Related Literature

Our approach follows the direction of Gilboa (1989) and Shalev (1997), where they apply the non-additive prior model of Schmeidler (1989) over time and derive a utility representation that depends on the difference between adjacent consumptions. In our model, we apply the multiple-priors model of Gilboa and Schmeidler (1989) under which we can find direct analogy between uncertainty aversion and time-variability aversion. In addition, the multiple-priors model can be applied consistently over time.[6]

In terms of a representation, our model involves time-varying discount factors, where the movement of discount factors demonstrates the aversion to time-variability. In this direction, Uzawa (1968), Epstein (1983), and Shi and Epstein (1993) develop representations with time-varying discount factors that can also show time-variability aversion. In addition, our model involves minimization on a set of discount factors. The variational utility approach proposed by Geoffard (1996)

[4] See Frederick, Loewenstein, and O'Donoghue (2002) for a detailed survey.

[5] This idea is applied to explain asset pricing; see Wakai (2002).

[6] See Section 2 and Section 4.

is based on minimization on a set of discount rates. The main departure of our representation is that it introduces gain/loss asymmetry, which is absent from or not applicable to the other constructions mentioned above.

On the grounds for separating the attitude toward risk from the attitude toward time-variability, the approach suggested by Epstein and Zin (1989) is to consider intertemporal substitution by a recursive aggregator function that has present utility and a continuation value as arguments. This aggregator function implies that an agent first considers risk aversion and then considers intertemporal substitution. By contrast, in our representation, an agent first considers intertemporal substitution and then considers uncertainty. Since both models are based on different preference domains and different objectives, each has its own relative advantage. In Section 4, we will compare the two models in detail.

1.3 Representation

Before proceeding to a formal derivation, we provide an overview of the representation we are going to develop. Suppose that there are two periods, time 0 and time 1, and there are S states of nature at time 1. Our representation takes the following form:

(1) $\mathrm{E}[U(x_0,\, x_{1,s})] \equiv \sum_{s=1}^{S} \pi_s W(u(x_0),\, u(x_{1,s}))$.

(2) $W(u(x_0),\, u(x_{1,s})) \equiv \mathrm{Min}_{\delta\in\Delta}[(1-\delta)u(x_0) + \delta u(x_{1,s})]$,

where Δ is a non-empty, closed, and convex set in (0,1).

An agent first considers atemporal risk attitude expressed by characteristics of $u(.)$, then considers intertemporal attitude toward *time-variability* (by which we mean the fluctuation of $u(.)$ over time) expressed by a non-time-separable aggregator function W, and finally considers uncertainty by aggregating W with a subjective prior. By the nature of the aggregator function W, the discount factor for gains ($u(x_0) < u(x_{1,s})$) and the discount factor for losses ($u(x_0) > u(x_{1,s})$) are different. In addition, this operation is related to but different from the model suggested by Epstein and Zin (1989) and shares characteristics with a neoclassical representation under the discounted utility model.

For the reminder of the paper, *time-preferences* refer to the structure of W (movement of discount factors) that incorporates time-variability aversion. The attitude toward atemporal risk will be called *risk-preferences*. We use the term *intertemporal preferences* to denote overall preference relations under either certainty or uncertainty. Intertemporal preferences consist of risk-preferences, time-preferences, and a subjective prior.

2 Consumption Smoothing under Certainty

2.1 A Two-Period Example

In this subsection, we provide a simple example that motivates our particular representation. Suppose that an agent faces an intertemporal decision problem of a two-period economy under certainty. The agent has three choices: a sequence that yields 4 consumption payoffs in each period, a sequence that yields 1 consumption payoff followed by 9 consumption payoffs, and a sequence that yields 9 consumption payoffs followed by 1 consumption payoff.

x: $(x_0,\, x_1) = (4,\, 4)$
y: $(y_0,\, y_1) = (1,\, 9)$
z: $(z_0,\, z_1) = (9,\, 1)$

Now we follow the steps defined in the previous section. First we convert consumption payoffs into atemporal utility indices. For example, we assume $u(a) = \sqrt{a}$:

x: $(u(x_0),\, u(x_1)) = (2,\, 2)$
y: $(u(y_0),\, u(y_1)) = (1,\, 3)$
z: $(u(z_0),\, u(z_1)) = (3,\, 1)$

Next we aggregate utility indices by an aggregator function W. For any agent with preferences of the form of $u(x_0) + \delta u(x_1)$, the agent will strongly prefer y or z to x (unless $\delta = 1$ in which case the agent is indifferent between all three.). However, an agent who is averse to time-variability might prefer x to y or z because, in terms of *utility* sequences, y hedges the movement of z, and x is a mixture (in fact, an average) of y and z. One way to express the above preferences is to assume the following representation of W:

(3) $$V(x_0,\, x_1) = W(u(x_0),\, u(x_1)) = \mathrm{Min}_{\delta\in\Delta}[(1-\delta)u(x_0) + \delta u(x_1)] \quad \text{with } \Delta = [0.3, 0.7]$$

The following three observations are crucial to our discussion. First, an aggregator function calculates a weighted sum of utility indices with non-fixed weights, and it is not differentiable at $u(x_0) = u(x_1)$ even if $u(.)$ is differentiable (We call it the *first-order effect*; see Section 2.2.). These weights can be regarded as discount factors. Second, each weight is positive and the sum of two weights is one. Hence, we only need to decide a one of two discount factors. Third, effective discount factors are the most pessimistic ones because they minimize the present discounted value.

An application of this aggregator function yields the following:

x: $V(x_0,\, x_1) = 2.$
y: $V(y_0,\, y_1) = 1{\cdot}(1 - 0.3) + 3 \cdot 0.3 = 1.6.$
z: $V(z_0,\, z_1) = 3{\cdot}(1 - 0.7) + 1 \cdot 0.7 = 1.6.$

For y and z (uneven sequences), the fluctuation of atemporal *utility* indices over time decreases the overall value. By assigning a higher discount factor for $u(y_0) = 1$ and a lower discount factor for $u(y_1) = 3$, an agent shifts relative time-preferences from time 1 to time 0, which gives the agent a strong incentive to move consumption from $u(y_1) = 3$ to $u(y_0) = 1$. By achieving complete smoothing, the agent can improve overall utility. Since the representation of (3) involves a set of discount factors, we define (3) as a *multiple-discount-factors* model.

To understand the significance of this representation, we compare (3) with the discounted utility model. Suppose $y \simeq z$. If an agent follows the discounted utility model, a preference relation must satisfy the following condition:

(4) For all $\alpha \in (0, 1)$, $y \simeq z$ if and only if $\alpha y \oplus (1-\alpha)z \simeq \alpha z \oplus (1-\alpha)z \simeq z$,[7]

where $\oplus$ represents a period-by-period addition of utility indices.[8] Then, this condition leads to the following:

(1, 3) $\sim$ (3, 1) implies (2, 2) $\sim$ (1, 3) $\sim$ (3, 1),[9]
so $x \simeq y \simeq z$.

Clearly, the above condition is too strong to admit time-variability aversion. On the other hand, under our representation, the intuition behind the aggregator function is summarized by the following notion of consumption smoothing:

(5) $y \simeq z$ but $x \simeq \frac{1}{2}y \oplus \frac{1}{2}z \succeq z.$

In other words, the mixture of equally preferred *utility* sequences is weakly preferred to the original sequence because the mixture weakly decreases fluctuations in utility indices. Hedging the movement of atemporal utility indices over time increases overall utility.[10]

On the other hand, the multiple-discount-factors model also exhibits gain/loss asymmetry in discount factors. For example, for a gain (i.e., y), an agent assigns a lower discount factor to $u(y_1)$, and for a loss (i.e., z), an agent assigns a higher discount factor to $u(z_1)$. The assignment of discount factors is homogeneous among gains or losses. This result is due to the following notion of consumption smoothing:

---

[7] In our setting, the discounted utility model is based on the independence axiom in Anscombe and Aumann (1963): For all $f, g, h \in \mathcal{H}$ and for all $\alpha \in (0, 1)$, $f \succ g$ if and only if $\alpha f \oplus (1-\alpha)h \succ \alpha g \oplus (1-\alpha)h$.

[8] $\alpha y \oplus (1-\alpha)z$ is a sequence that pays consumption lotteries in each period: The first lottery pays 1 with a probability of α and 9 with a probability of $(1-\alpha)$; the second lottery pays 9 with a probability of $(1-\alpha)$ and 1 with a probability of α.

[9] $0.5(1, 3) \oplus 0.5(3, 1) = (0.5 \cdot 1 + 0.5 \cdot 3, 0.5 \cdot 3 + 0.5 \cdot 1) = (2, 2)$. All numbers are considered to be utils.

[10] This idea is essentially identical to the *preference hedge* in Schmeidler (1989) and Gilboa and Schmeidler (1989).

(6) $\quad y \simeq z$ implies $\frac{1}{2}x\oplus \frac{1}{2}y \simeq \frac{1}{2}x\oplus \frac{1}{2}z$, i.e., (1, 3) $\sim$ (3, 1) implies (1.5, 2.5) $\sim$ (2.5, 1.5).

Note that (1.5, 2.5) and (2.5, 1.5) share the nature of time-variability with (1, 3) and (3, 1), respectively. Then the nature of time-variability determines the preference ordering, and the shift of an overall utility level does not change the preference ordering. In fact, (6) is a weaker version of (4), and it allows (4) to hold among gains (or losses). Then (6) permits an additive representation (i.e., a weighted sum of utility indices) among gains (or losses); in other words, we must have identical assignment of discount factors for y and $\frac{1}{2}x\oplus \frac{1}{2}y$. However, for a case with more than two periods, (6) only delivers the additive representation among subgroups of sequences. For gain/loss asymmetry, we need to incorporate dynamic consistency, which is satisfied under a two-period case (under the assumption of history-independence; see Section 2.2).

Our representation expresses time-variability aversion and gain/loss asymmetry. Note that another concave function W can represent the preference relation in this example. However, our formula has two more advantages: First, it is based on simple axioms, so we can easily understand *why* an agent follows our model. Second, interpretation of time-preferences is direct; we model discount factors themselves. Since our formula becomes a weighted summation of atemporal utility indices at an effective selection of discount factors, the departure from the discounted utility model is minimal. Thus, our model shares the tractability of the discounted utility model.

2.2 Representation of Intertemporal Preferences

In this subsection, we derive a representation of intertemporal preferences of consumption smoothing under certainty. To separate time-variability aversion from risk aversion, we define a preference relation over sequences of *consumption lotteries* by adapting the Anscombe-Aumann (1963) framework with a temporal interpretation. Let X be a set of outcomes and Y be a set of probability distributions over X that satisfies the following:

$Y = \{y|\ y\colon\ X \to [0,1]$ where y has a finite support.$\}$

For convenience, we call $y \in Y$ a lottery and Y a lottery space. Let $\mathcal{T} = \{0, 1, ..., T\}$ be a finite set of periods from time 0 to time T.[11] Let f be an act, where $f : \mathcal{T} \to Y$, and h be a constant act that assigns identical $y \in Y$ for all $t \in \mathcal{T}$. Define $\mathcal{H}$ as a collection of all f and $\mathcal{L}$ as a collection of all constant acts. Denote $f(t) \in Y$ as a lottery assigned at t under f. We also define the following operation: $[\alpha f \oplus (1-\alpha)g](t) = \alpha f(t) + (1-\alpha)g(t)$. We now assume that the following axioms hold for a preference relation on $\mathcal{H}$ at $t = 0$ (i.e., $\succsim_0$):

[11] The result may be extended to an infinite horizon by using the extension theorem in Gilboa and Schmeidler (1989).

A1-WO Weak Order: *For all* $f, g, h \in \mathcal{H}$, (i) $f \succeq_0 g$ *or* $g \succeq_0 f$ *and* (ii) $f \succeq_0 g$ *and* $g \succeq_0 h$ *imply* $f \succeq_0 h$.

A2-C Continuity: *For all* $f, g, h \in \mathcal{H}$ *with* $f \succ_0 g \succ_0 h$, *there exist* α, $\beta \in (0, 1)$ *such that* $\alpha f \oplus (1-\alpha)h \succ_0 g$ *and* $g \succ_0 \beta f \oplus (1-\beta)h$.

A3-ND Non-Degeneracy: *There exist* $f, g \in \mathcal{H}$ *such that* $f \succ_0 g$.

A4-CI Constant-Independence:[12] *For all* $f, g \in \mathcal{H}$, $l \in \mathcal{L}$, *and for all* $\alpha \in (0, 1)$, $f \succ_0 g$ *if and only if* $\alpha f \oplus (1-\alpha)l \succ_0 \alpha g \oplus (1-\alpha)l$.

A1 to A4 imply that for all $l, l' \in \mathcal{L}$, there exists a function $U : Y \to R$ such that $l \succeq_0 l'$ if and only if $U(y) \geq U(z)$, where $l(t) = y \in Y$ and $l'(t) = z \in Y$ for all $t \in \mathcal{T}$. Note that in this model, an agent consumes a lottery itself, not a realization from a lottery at each time. Thus, it is not appropriate to reduce a sequence of lotteries into a single lottery. In addition, constant acts do not involve time-variability, and a preference relation on $\mathcal{L}$ is represented as if it is a preference relation on lotteries assigned for each constant act. For this reason, we consider that a preference relation on $\mathcal{L}$ expresses atemporal risk-preferences on $y \in Y$. Hence, for all $y, z \in Y$, $y \succeq_0 z$ means that $l \succeq_0 l'$, where $l, l' \in \mathcal{L}$ such that for all $t \in \mathcal{T}$, $l(t) = y$ and $l'(t) = z$. We also denote a constant act that assigns identical $y \in Y$ for all $t \in \mathcal{T}$ as y. Under this interpretation, A4-CI states that a preference relation on $\mathcal{H}$ mainly depends on fluctuations in atemporal risk levels. Moreover, A4-CI permits an additive representation among subclasses of acts that share the "same" nature of time-variability (for example, among $\alpha f \oplus (1-\alpha)l$ for all $\alpha \in [0, 1]$, where $f \simeq_0 l$ and $l \in \mathcal{L}$).

Given atemporal risk-preferences, we introduce the following axiom:

A5-SM Strict Monotonicity: *For all* $f, g \in \mathcal{H}$ *such that* $f = (f(0), ..., f(T))$ *and* $g = (g(0), ..., g(T))$, *if* $f(t) \succeq_0 g(t)$ *for all* $t \in \mathcal{T}$, *then* $f \succeq_0 g$. *In addition, if for some* t, $f(t) \succ_0 g(t)$, *then* $f \succ_0 g$.

The rationale for this axiom is the result of the existence of U on Y. If $U(y) \geq U(z)$, monotonicity on *lotteries* should imply that $f \succeq_0 z$ where $f(t) = y$ for $t \in \mathcal{T}' \subset \mathcal{T}$ and $f(t) = z$ for $t \in \mathcal{T} \backslash \mathcal{T}'$. By extending this logic for $\mathcal{H}$, we deduce A5-SM. In fact, A5-SM separates risk from time-variability. An agent first considers risk embedded in a lottery separately at each time, and derives cardinal utility for each lottery. Then, consumption smoothing is defined over these utility indices by the following axiom:

[12] This axiom is called certainty-independence in Gilboa and Schmeidler (1989).

A6-TVA Time-Variability Aversion:[13,14] *For all* $f, g \in \mathcal{H}$ *and for all* $\alpha \in (0,1)$, $f \simeq_0 g$ *implies* $\alpha f \oplus (1-\alpha)g \succeq_0 f$.

Gilboa and Schmeidler (1989) have proven that the above axioms imply the following representation of a preference relation on $\mathcal{H}$:

Theorem 1 Temporal Version of Gilboa and Schmeidler (1989):[15,16] *A preference relation on* $\mathcal{H}$ *satisfies A1 to A6 if and only if there exists a non-empty, closed, and convex set of discount factors,* Δ_0, *with* $\sum_{t=0}^{T} \delta_t = 1$ *and* $\delta_t > 0$ *for all* $t \in \mathcal{T}$ *for* $\delta = (\delta_0, ..., \delta_T) \in \Delta_0$ *such that*

(7) *For all* $f, g \in \mathcal{H}$, $f \succeq_0 g$ *if and only if* $V_0(f) \geq V_0(g)$,

where $V_0(f) \equiv \min_{\delta \in \Delta_0} \sum_{t=0}^{T} \delta_t U(f(t))$ *and* $U(f(t)) \equiv \mathrm{E}_{f(t)}[u(x)]$ *with* $x \in X$.

Moreover, under these conditions, Δ_0 *is unique and* u: $Y \to R$ *is unique up to a positive affine transformation.*

Under A1 to A5, the representation becomes $W(U(f(0)), ..., U(f(T)))$, and A4-CI and A6-TVA determine the structure of W. Under the representation of (7), attitude toward risk is expressed by a von Neumann-Morgenstern utility function $u(.)$. Time-variability aversion is captured by the agent selecting discount factors to minimize the weighted sum of atemporal expected utility indices. We derive W for an entire stream of consumption lotteries, and (7) becomes concave and non-time-separable. In fact, time-variability aversion is independent of the structure of $u(.)$, which can be either concave or convex.[17]

We also confirm that the representation of (7) satisfies "local" (or "first-order") time-variability aversion at a constant act. By the previous result in Epstein and Wang (1994) applied to uncertainty aversion, we achieve the following:

Theorem 2 Local (or First-Order) Time-Variability Aversion (Temporal Version of Epstein and Wang, 1994): *Suppose that a preference relation on* $\mathcal{H}$ *can be represented by* (7) *and* $u(.)$ *is differentiable. Consider sequences of degenerated consumption lotteries denoted by* x. *Let* c *be constant consumption* (i.e., *a sequence of identical consumption*), *and let* $\delta^* \in \Delta_0$. *Then for a consumption stream* x *such that* $\sum_{t=0}^{T} \delta_t^* x(t) = 0$,

[13] This axiom is called uncertainty aversion in Gilboa and Schmeidler (1989).

[14] We can change the attitude toward time-variability as we do for the attitude toward risk. For example, *time-variability seeking* is defined as follows: For all $f, g \in \mathcal{H}$ and for all $\alpha \in (0,1)$, $f \simeq g$ implies $\alpha f \oplus (1-\alpha)g \preceq f$. Then, the representation replaces *min* with *max*. In this sense, we consider the discounted utility representation to be *time-variability neutral.*

[15] (7) includes the discounted utility model as a subset.

[16] We refer to propositions proven by other authors as theorems.

[17] See Wakai (2002) for an application of (7) under a concave or convex $u(.)$.

(8) $\lim_{k\to 0_+} \dfrac{V_0(c+kx)-V_0(c)}{k} = \min_{\delta\in\Delta_0} \sum_{t=0}^{T} \delta_t u'(c(t))x(t) \leq \sum_{t=0}^{T} \delta_t^* u'(c(t))x(t) = 0,$
and strict inequality holds for some x when Δ_0 is not degenerate.

Since any discount factor in Δ_0 can be an effective selection for constant consumption, following Fisher (1930), we can consider Δ_0 to be a set of all admissible time-preferences. Then a structure of Δ_0 expresses the extent to which an agent evaluates each time as well as the extent to which an agent is time-variability averse. If we regard some $\delta^* \in \Delta_0$ as a base-line time-preference, (8) also implies that an agent *locally* dislikes mean-preserving spreads of atemporal utility indices at constant consumption, where the *mean* is defined as a weighted summation of atemporal utility indices under this base-line time-preference. In fact, non-differentiability at constant consumption implies that an agent requires a large premium in order to deviate from constant consumption.[18] This implication is analogous to that from first-order risk aversion by Segal and Spivak (1990, 1997), where the value function is not differentiable at constant consumption over states.

We now consider conditional preference relations. As opposed to states, time evolves sequentially. The disadvantage of (7) is that if we apply it for an economy with more than two periods, we may face dynamic inconsistency.[19] To resolve this difficulty, we need to restrict the behavior of conditional preference relations. Let $\mathcal{T}_t$ be a finite set of periods from time t to time T and $\mathcal{T}_{-t}$ be a finite set of periods from time 0 to time $t-1$. Define f^t as a function such that $f^t : \mathcal{T}_t \to Y$ and f^{-t} as a function such that $f^{-t} : \mathcal{T}_{-t} \to Y$. If $\mathcal{T}_{-t}$ is empty, f^t defines an act f, and vice versa. A preference relation on $\mathcal{H}$ conditional on time t is denoted by $\succeq_t$. All conditional preference relations (including the ex-ante preference relation $\succeq_0$) in the collection of $\{\succeq_t\} \equiv \{\succeq_t |t \in \mathcal{T}\}$ follow additional axioms:

A7-IH Independence of History up to $t-1$:[20] *For all $f, f', g, g' \in \mathcal{H}$ such that $f = (a^{-t}, f^t)$, $g = (b^{-t}, g^t)$, $f' = (c^{-t}, f^t)$, $g' = (d^{-t}, g^t)$, $f \succeq_t g$ if and only if $f' \succeq_t g'$.*

A8-DC Dynamic Consistency: *For all $f = (a^{-t}, y_t, f^{t+1}), g = (a^{-t}, y_t, g^{t+1}) \in \mathcal{H}$, $f \succeq_t g$ if and only if $f \succeq_{t+1} g$.*

Under the above axioms, (7) must be rewritten in the following form (for the proof, see Appendix

[18]This result is applied to explain asset pricing; see Wakai (2002).

[19]Eichberger and Kelsey (1996) utilize Machina's (1989) notion of dynamic consistency to examine a dynamically consistent updating rule for the non-additive prior model of Schmeidler (1989). They show that if an agent's preferences satisfy "strict uncertainty aversion," a dynamically consistent update rule does not produce the conditional preferences that confirm the non-additive prior model. Wakai (2003) also shows that an analogous result holds for the multiple-priors model.

[20]Independence of history might not be innocuous given the nature of consumption smoothing. We require this condition for time-variability aversion to be dynamically consistent. However, without axiomatic reasoning, time-variability aversion can be defined on utility indices that depend on historical consumption; see Section 4.

A):[21]

Proposition 1: *Suppose that an agent's preference relation on $\mathcal{H}$ satisfies A1 to A6 at time 0, and let $U_0 \equiv U$ and Δ_0 be defined as in Theorem 1. Then preference relations in $\{\succeq_t\}$ on $\mathcal{H}$ satisfy A7 and A8 if and only if there exist $\{[\alpha_t, \beta_t]\}_{1\leq t\leq T}$ such that*

(9) *For all $t \in \mathcal{T}$ and for all $f, g \in \mathcal{H}$, $f \succeq_t g$ if and only if $V_t(f) \geq V_t(g)$,*
where $\{V_t(f)\}_{0\leq t\leq T}$ are recursively defined by
$V_t(f) \equiv \min_{\delta_{t+1}\in[\alpha_{t+1},\beta_{t+1}]}[(1-\delta_{t+1})U(f(t)) + \delta_{t+1}V_{t+1}(f)]$
and $V_T(f) \equiv U(f(T))$,

(10) $0 < \alpha_t \leq \beta_t < 1$ *for all $t \in [1, T]$.*

Moreover,

(11) $[\alpha_t, \beta_t]$ *is uniquely defined.*

Given dynamic consistency, $W_t(U(f(t)), ..., U(f(T)))$ becomes $W_t(U(f(t)), V_{t+1}(f))$, which is time-dependent and recursive. The main observation in Proposition 1 is that dynamic consistency and history-independence contribute to one distinct feature of the structure of non-differentiability of V: gain/loss asymmetry. More specifically, to avoid time-variability, an agent assigns a higher discount factor for the present discounted value of future utility from $t+1$ onward when it is lower than the utility of present consumption, and vice versa. An increase from the present utility requires a lower discount factor, and a decrease from the present utility requires a higher discount factor. On the contrary, under (7), discount factors can move very smoothly, and we may not observe clear directional adjustment. In other words, to be dynamically consistent, an agent who follows our axioms *must* show gain/loss asymmetry in discount factors. For this reason, we argue that gain/loss asymmetry is *consistent* behavior, instead of an *anomaly*. Note that the gain/loss asymmetry we derive is a *recursive* one as opposed to one between two periods. It incorporates both the attitude toward an entire sequence of consumption and the attitude toward two-period gains and losses.[22]

Finally, we investigate a binary relation of "more time-variability averse." From the representation of (9), it is clear that the attitude toward risk (i.e., $u(.)$) affects *effective* intertemporal substitution. To define relative characteristics over time, we need to neutralize the effects from the attitude toward risk. Thus, we can only permit a binary relation of "more time-variability averse" for pairs of preference relations that embody the same ranking of single-period risk. Then, V_t is more time-variability averse than V'_t if and only if $U = U'$, and

[21] This result is an application of the recursive structure of the multiple-priors set defined by Epstein and Schneider (2003) and Wakai (2003).

[22] Many questions in experimental study concerning gain/loss asymmetry are framed to examine preferences between a single-period gain vs. a single-period loss. This single-period effect is then generalized to multi-period cases. However, we consider multi-period cases to be a basic framework and deduce a single-period effect from multiple-period cases.

(12) $V_t(f) \leq V'_t(f)$ for all $f \in \mathcal{H}$,

where each U and U' represents a preference relation on Y for V_t and V'_t, respectively. In fact, condition (12) at $t = 0$ is equivalent to the following:

(13) $\Delta_0 \supseteq \Delta'_0$ (equivalently, $[\alpha_t,\beta_t] \supseteq [\alpha'_t, \beta'_t]$ for all $t \in [1, T]$),

where each Δ_0 and Δ'_0 represents a set of discount factors in (7) for V_0 and V'_0, respectively.

3 Consumption Smoothing under Uncertainty

In this section, we derive the representation of intertemporal preferences of consumption smoothing under uncertainty. As in the case under certainty, we adapt the Anscombe-Aumann (1963) framework to our setting. The model has the following structure: There is a finite state space Ω with S elements; the time horizon is finite. Let $\mathcal{T} = \{0, ..., T\}$ be a finite set of periods from time 0 to time T, and let $\mathcal{T}_t = \{t, ..., T\}$ be a finite set of periods from time t to time T. The information structure is represented by $\mathcal{F} = \{\mathcal{F}_0, ..., \mathcal{F}_T\}$, where $\mathcal{F}_t$ is a partition of Ω that is finer than $\mathcal{F}_\tau$ for all $\tau \in [0, t)$ and any event in $\mathcal{F}_t$ must be a subset of some event of $\mathcal{F}_{t-1}$. Assume that $\mathcal{F}_0 = \{\Omega\}$ and $\mathcal{F}_T = \{\{1\}, ..., \{S\}\}$ with $\mathcal{F}_t(\omega)$ denoting an event in $\mathcal{F}_t$ that contains a state $\omega \in \Omega$. Y is a lottery space as defined in Section 2.2. Let f_ω be a state act such that $f_\omega : \mathcal{T} \to Y$, i.e., $f_\omega = (y_0, ..., y_T) \in Y^{T+1}$, and f^t be a time act such that $f^t : \Omega \to Y$, i.e., $f^t = (y_1, ..., y_S)' \in Y^S$. Define $\mathcal{Y}^{T+1}$ as a collection of all state acts and $\mathcal{Y}^S$ as a collection of all time acts. The primitives on which an agent forms preference relations are *adapted acts* such that $f : \mathcal{T} \times \Omega \to Y$, i.e., $f = (f^0, ..., f^T) = (f_1, ..., f_S)' \in Y^{S(T+1)}$ and f^t is measurable with respect to the σ-field generated by $\mathcal{F}_t$ for all $t \in \mathcal{T}$. Define $\mathcal{D}$ as a collection of all adapted acts. We denote $f^t(\omega) \in Y$ as a lottery in f^t at ω, $f_\omega(t) \in Y$ as a lottery in f_ω at t, and $f(t, \omega) \in Y$ as a lottery in f at (t, ω). We also define the following operation: $[\alpha f \oplus (1 - \alpha)g](t, \omega) = \alpha f(t, \omega) + (1 - \alpha)g(t, \omega)$.

We now define subsets of $\mathcal{D}$. A constant act is $l \in \mathcal{D}$ such that $l(t, \omega) = y \in Y$ for all $(t, \omega) \in \mathcal{T} \times \Omega$, that will also be denoted by l or y. $\mathcal{L}$ is a collection of all constant acts. A certainty act is $f = (f_1, ..., f_S)' \in \mathcal{D}$ such that $f_\omega = f_{\omega'}$ for all $\omega \in \Omega$, that will also be denoted by f_ω. $\mathcal{H}$ is a collection of all certainty acts. A t-recursive act is $f \in \mathcal{D}$ such that for all $\tau \in \mathcal{T}_{t+1}$, f^τ is measurable with respect to $\mathcal{F}_{t+1}$. $\mathcal{D}_t$ is a collection of all t-recursive acts. Given this definition, t-recursive acts represent adapted acts under which all future uncertainty is resolved on each of tomorrow's events. Clearly, $\mathcal{L} \subset \mathcal{H} \subset \mathcal{D}_t \subset \mathcal{D}$. The following table summarizes the allocation of lotteries within each subset. Suppose that $T = 2$, $\mathcal{F}_0 = \{\Omega\}$, $\mathcal{F}_1 = \{\mathcal{F}_{1,1}, \mathcal{F}_{1,2}\}$ with $\mathcal{F}_{1,1} = \{1, 2\}$ and $\mathcal{F}_{1,2} = \{3, 4\}$, $\mathcal{F}_2 = \{\mathcal{F}_{2,1}, \mathcal{F}_{2,2}, \mathcal{F}_{2,3}, \mathcal{F}_{2,4}\}$ with $\mathcal{F}_{2,1} = \{1\}$, $\mathcal{F}_{2,2} = \{2\}$, $\mathcal{F}_{2,2} = \{3\}$, and $\mathcal{F}_{2,2} = \{4\}$.

Information structure	$\mathcal{D}$	$\mathcal{L}$

$t=0$	$t=1$	$t=2$
$\mathcal{F}_0$	$\mathcal{F}_{1,1}$	$\mathcal{F}_{2,1}$
$\mathcal{F}_0$	$\mathcal{F}_{1,1}$	$\mathcal{F}_{2,2}$
$\mathcal{F}_0$	$\mathcal{F}_{1,2}$	$\mathcal{F}_{2,3}$
$\mathcal{F}_0$	$\mathcal{F}_{1,2}$	$\mathcal{F}_{2,4}$

$t=0$	$t=1$	$t=2$
y_0	$y_{1,1}$	$y_{2,1}$
y_0	$y_{1,1}$	$y_{2,2}$
y_0	$y_{1,2}$	$y_{2,3}$
y_0	$y_{1,2}$	$y_{2,4}$

$t=0$	$t=1$	$t=2$
y	y	y
y	y	y
y	y	y
y	y	y

$\mathcal{H}$

$t=0$	$t=1$	$t=2$
y_0	y_1	y_2
y_0	y_1	y_2
y_0	y_1	y_2
y_0	y_1	y_2

$\mathcal{D}_0$

$t=0$	$t=1$	$t=2$
y_0	$y_{1,1}$	$y_{2,1}$
y_0	$y_{1,1}$	$y_{2,1}$
y_0	$y_{1,2}$	$y_{2,2}$
y_0	$y_{1,2}$	$y_{2,2}$

Preference relations follow a structure similar to that of the previous section. An agent forms *ex-ante* preference relations at $t=0$ on each $\omega \in \Omega$, and conditional preference relations are updated from *ex-ante* preference relations (with additional axioms that define an attitude toward uncertainty at each (t,ω)). First, an agent's *ex-ante* preference relations in the collection of $\{\succeq_{(0,\omega)}\} \equiv \{\succeq_{(0,\omega)} \mid \omega \in \Omega\}$ satisfy the following axioms:

A1U-WO Weak Order: *For all* $f, g, h \in \mathcal{D}$, (i) $f \succeq_{(0,\omega)} g$ *or* $g \succeq_{(0,\omega)} f$ *and* (ii) $f \succeq_{(0,\omega)} g$ *and* $g \succeq_{(0,\omega)} h$ *imply* $f \succeq_{(0,\omega)} h$.

A2U-C Continuity: *For all* $f, g, h \in \mathcal{D}$ *with* $f \succ_{(0,\omega)} g \succ_{(0,\omega)} h$, *there exist* α, $\beta \in (0,1)$ *such that* $\alpha f \oplus (1-\alpha)h \succ_{(0,\omega)} g$ *and* $g \succ_{(0,\omega)} \beta f \oplus (1-\beta)h$.

A3U-ND Non-Degeneracy: *There exist* $f, g \in \mathcal{D}$ *such that* $f \succ_{(0,\omega)} g$.

A4U-CI Constant-Independence: *For all* $f, g \in \mathcal{D}$, $l \in \mathcal{L}$, *and for* $\alpha \in (0,1)$, $f \succ_{(0,\omega)} g$ *if and only if* $\alpha f \oplus (1-\alpha)l \succ_{(0,\omega)} \alpha g \oplus (1-\alpha)l$.

Since adding a constant act does not alter the nature of "variations" in an allocation of lotteries, we assume that it does not alter preference relations (A4U-CI). Then A1U to A4U imply that for all $l, l' \in \mathcal{L}$, there exists a function $U : Y \to R$ such that $l \succeq_{(0,\omega)} l'$ if and only if $U(y) \geq U(z)$, where $l(t,\omega) = y$ and $l'(t,\omega) = z$ for all $(t,\omega) \in \mathcal{T} \times \Omega$. Since constant acts do not involve variations in lotteries over time and states, we consider preference relations on $\mathcal{L}$ to express atemporal risk-preferences on $y \in Y$. Hence, for all $y, z \in Y$, $y \succeq_{(0,\omega)} z$ means that $l \succeq_{(0,\omega)} l'$, where $l, l' \in \mathcal{L}$ such that for all $(t,\omega) \in \mathcal{T} \times \Omega$, $l(t,\omega) = y$ and $l'(t,\omega) = z$. Given this notion of atemporal risk-preferences, we assume the following axiom under which an agent considers the "risk" of each time-state lottery separately in order to evaluate each time act:[23]

[23]Under this axiom, if $f(t,\omega) \simeq_{(0,\omega)} g(t,\omega)$ for $(t,\omega) \in \mathcal{T} \times \Omega$, then $f \simeq_{(0,\omega)} g$.

A5U-SMT Strict Monotonicity on Time: *For all $f, g \in \mathcal{D}$, if for all $\tau \in \mathcal{T}$, $f(\tau,\omega) \succeq_{(0,\omega)} g(\tau,\omega)$ for all $\omega \in \Omega$, then $f \succeq_{(0,\omega)} g$. In addition, if for some $\tau \in \mathcal{T}$, $f(\tau,\omega) \succ_{(0,\omega)} g(\tau,\omega)$ for all $\omega \in \Omega$, then $f \succ_{(0,\omega)} g$.*

However, A6-TVA cannot be applied to $\mathcal{D}$ because it is defined on a sequence of lotteries without uncertainty and does not take into account variations over states. Therefore, we can apply A6-TVA only to $\mathcal{H}$ as follows:

A6U-TVA Time-Variability Aversion on $\mathcal{H}$: *For all $f, g \in \mathcal{H}$ and for all $\alpha \in (0,1)$, $f \simeq_{(0,\omega)} g$ implies $\alpha f \oplus (1-\alpha) g \succeq_{(0,\omega)} g$.*

Under A1U to A6U, a preference relation on $\mathcal{H}$ also induces a preference relation on $\mathcal{Y}^{T+1}$: For $f_\omega, g_\omega \in \mathcal{Y}^{T+1}$, $f_\omega \succeq_{(0,\omega)} g_\omega$ means that $h \succeq_{(0,\omega)} h'$, where $h, h' \in \mathcal{H}$ such that $h_{\omega'} = f_\omega$ and $h'_{\omega'} = g_\omega$ for all $\omega' \in \Omega$.

At each (t,ω), an agent forms a conditional preference relation on $\mathcal{D}$ denoted by $\succeq_{(t,\omega)}$. All conditional preference relations (including *ex-ante* preference relations) in the collection of $\{\succeq_{(t,\omega)}\} \equiv \{\succeq_{(t,\omega)}: (t,\omega) \in \mathcal{T} \times \Omega\}$ satisfy the following axioms that define consistency over time and states:

A7U-F Indifference on $\mathcal{F}_t(\omega)$: *$\succeq_{(t,\omega)}$ is identical to $\succeq_{(t,\omega')}$ if $\mathcal{F}_t(\omega) = \mathcal{F}_t(\omega')$.*

A8U-IHA Independence of History up to $t-1$ and of Irrelevant Alternatives: *For all $f, q \in \mathcal{D}$, if $f(\tau,\omega') = g(\tau,\omega')$ for all $(\tau,\omega') \in \mathcal{T}_t \times \mathcal{F}_t(\omega)$, then $f \simeq_{(t,\omega)} g$.*

A9U-DC Dynamic Consistency: *For all $f, g \in \mathcal{D}$ with $f(\tau,\omega'') = g(\tau,\omega'')$ for all $(\tau,\omega'') \notin \mathcal{T}_{t+1} \times \mathcal{F}_{t+1}(\omega')$ for some $\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$, $f \succeq_{(t,\omega)} g$ if and only if $f \succeq_{(t+1,\omega'')} g$ for all $\omega'' \in \mathcal{F}_{t+1}(\omega')$.*

Given the structure of information, dynamic consistency should be modified as in A9U-DC. In addition, we adjust history independence as in A8U-IHA for an agent unconcerned with unrealized events. A7U-F ensures that preference relations are based on information available at (t,ω).

Next, we define attitude toward uncertainty. In this paper, we will not consider uncertainty aversion, but rather assume that an agent regards variations over states as another source of (subjective) *risk*. Naively, we can apply (9) to each state act on Ω and aggregate the utilities of the state acts with subjective probability. However, if $T > 1$, this approach violates dynamic consistency (see Appendix B) because (9) is non-time-separable. Given this result, we need to take dynamic consistency as a prime objective and return to the basic functional form implied by dynamic consistency.

Assume that an agent is at (t,ω). Define N as the number of events $\mathcal{F}_{t+1}(\omega') \in \mathcal{F}_{t+1}$ with $\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$, and choose any $\omega_n \in \mathcal{F}_t(\omega)$ to represent each of these events (i.e., $\mathcal{F}_{t+1}(\omega_n) \neq$

$\mathcal{F}_{t+1}(\omega_{n'})$ if $n \neq n'$). For $f \in \mathcal{D}$, denote y_{ω_n} as a constant act such that $y_{\omega_n} \simeq_{(t+1,\omega_n)} f$ (i.e., a certainty and non-time-variability equivalent of f at $(t+1, \omega_n)$).[24] Let f and g be adapted acts such that $f(\tau, \omega') = g(\tau, \omega')$ for all $(\tau, \omega') \notin \mathcal{T}_{t+1} \times \mathcal{F}_{t+1}(\omega_n)$ for some $\mathcal{F}_{t+1}(\omega_n) \subseteq \mathcal{F}_t(\omega)$. A7U to A9U imply that $f \succeq_{(t,\omega)} g$ if and only if $f \succeq_{(t+1,\omega')} g$ for all $\omega' \in \mathcal{F}_{t+1}(\omega_n)$, regardless of the lotteries assigned on other parts. Repeating this argument, f is equally preferred to d^f, where d^f is the t-recursive act ($d^f \in \mathcal{D}_t$) such that $d^f(t, \omega') = f(t, \omega')$ for all $(t, \omega') \in \{t\} \times \mathcal{F}_t(\omega)$ and $d^f(\tau, \omega') = y_{\omega_n}$ for all $(\tau, \omega) \in \mathcal{T}_{t+1} \times \mathcal{F}_{t+1}(\omega_n)$ for all $n \in [1, N]$. Hence, a conditional preference relation $\succeq_{(t,\omega)}$ on $\mathcal{D}_t$ can represent a conditional preference relation $\succeq_{(t,\omega)}$ on $\mathcal{D}$ by the functional form $\Psi_{(t,\omega)}(f(t,\omega), y_{\omega_1}, ..., y_{\omega_N})$. We neutralize the effects from uncertainty *within* each event $\mathcal{F}_{t+1}(\omega_n)$ in order to deduce pure preferences on uncertainty (in fact, deducing neutrality to uncertainty) *over* events $\{\mathcal{F}_{t+1}(\omega_n)\}_{n \in N}$ at $\mathcal{F}_t(\omega)$.

To be probabilistic, an agent needs to apply a version of the independence axiom on subsets of $\mathcal{D}_t$; two approaches are possible. The first is analogous to the recursive utility approach proposed by Epstein and Zin (1989). Under this approach, an agent applies the independence axiom to a collection of t-recursive acts that have an identical lottery $f(t, \omega)$ assigned for all $(t, \omega') \in \{t\} \times \mathcal{F}_t(\omega)$. In the resulting representation, an agent derives the certainty and non-time-variability equivalent y for $\{y_{\omega_1}, ..., y_{\omega_N}\}$, and aggregates y with $f(t, \omega)$ under (9). However, under A9U-DC, this operation implies that $\Psi_{(t,\omega)}(f(t,\omega), y_{\omega_1}, ..., y_{\omega_N}) > \Psi_{(t,\omega)}(f(t,\omega), y'_{\omega_1}, ..., y'_{\omega_N})$ if and only if $\Psi_{(t,\omega)}(f'(t,\omega), y_{\omega_1}, ..., y_{\omega_N}) > \Psi_{(t,\omega)}(f'(t,\omega), y'_{\omega_1}, ..., y'_{\omega_N})$. Hence, the preference ordering between $d \in \mathcal{D}_t$ that derives $\Psi_{(t,\omega)}(f(t,\omega), y_{\omega_1}, ..., y_{\omega_N})$ and $d' \in \mathcal{D}_t$ that derives $\Psi_{(t,\omega)}(f(t,\omega), y'_{\omega_1}, ..., y'_{\omega_N})$ is independent of $f(t, \omega)$. Given knowledge of the information structure, this independence of the future from today may not be preferable because it emphasizes risk reduction over states rather than consumption smoothing over time.

The second approach is to apply the independence axiom to a collection of t-recursive acts among which the operation $\oplus$ is not subject to the effects from time-variability aversion. Then we derive a subjective prior over the utility indices calculated by (9) for each state act $(f(t, \omega_n), y_{\omega_n})$. Under this approach, an agent considers time-variability aversion on an event-by-event basis, then regards uncertainty as a probabilistic phenomenon over these utility indices. Non-time-separability of (9) also implies that for some $f(t,\omega)$ and $f'(t,\omega)$, $\Psi_{(t,\omega)}(f(t,\omega), y_{\omega_1}, ..., y_{\omega_N}) > \Psi_{(t,\omega)}(f(t,\omega), y'_{\omega_1}, ..., y'_{\omega_N})$ but $\Psi_{(t,\omega)}(f'(t,\omega), y_{\omega_1}, ..., y_{\omega_N}) < \Psi_{(t,\omega)}(f'(t,\omega), y'_{\omega_1}, ..., y'_{\omega_N})$ because a level of $f(t,\omega)$ changes the nature of time-variability at each $\mathcal{F}_{t+1}(\omega_n) \subseteq \mathcal{F}_t(\omega)$. Hence, this approach closely represents the evolution of knowledge implied by $\mathcal{F}$ and emphasizes consumption smoothing over time as a prime objective.[25] Therefore, we take the second approach according to the following axioms:

[24]A *certainty and non-time-variability equivalent* of $f \in \mathcal{D}$ at (t, ω) is defined as $y \in \mathcal{L}$ such that $f \simeq_{(t,\omega)} y$. On the other hand, a *certainty equivalent* of $f \in \mathcal{D}$ at (t, ω) is defined as $h \in \mathcal{H}$ such that $f \simeq_{(t,\omega)} h$.

[25]For a more detailed comparison of these two approaches, see Section 4.2.

A10U-SMS Strict Monotonicity on States on $\mathcal{D}_t$:[26] *For all $f, g \in \mathcal{D}_t$, if $f_{\omega'} \succeq_{(t,\omega)} g_{\omega'}$ for all $\omega' \in \Omega$, then $f \succeq_{(t,\omega)} g$. In addition, if for some $\omega' \in \mathcal{F}_t(\omega)$, $f_{\omega'} \succ_{(t,\omega)} g_{\omega'}$, then $f \succ_{(t,\omega)} g$.*

A11U-IA Independence Axiom on $\mathcal{D}_t$:[27] *For all $f, g, h \in \mathcal{D}_t$, if at all $\omega' \in \Omega$,* (i) *there exist $l, l', l'' \in \mathcal{L}$ and $\alpha, \beta, \gamma \in (0,1)$ such that $\alpha f_{\omega'} \oplus (1-\alpha) l_{\omega'} \simeq_{(t,\omega)} \beta g_{\omega'} \oplus (1-\beta) l'_{\omega'} \simeq_{(t,\omega)} \gamma h_{\omega'} \oplus (1-\gamma) l''_{\omega'}$ and* (ii) *for all $\varphi, \psi \in [0,1]$, $\varphi[\alpha f_{\omega'} \oplus (1-\alpha) l_{\omega'}] \oplus (1-\varphi)[\gamma h_{\omega'} \oplus (1-\gamma) l''_{\omega'}] \simeq_{(t,\omega)} \psi[\beta g_{\omega'} \oplus (1-\beta) l'_{\omega'}] \oplus (1-\psi)[\gamma h_{\omega'} \oplus (1-\gamma) l''_{\omega'}]$, then for all $\theta \in (0,1)$, $f \succ_{(t,\omega)} g$ if and only if $\theta f \oplus (1-\theta) h \succ_{(t,\omega)} \theta g \oplus (1-\theta) h$.*

Under A10U-SMS, for all $\omega' \in \mathcal{F}_t(\omega)$, a preference relation on $\mathcal{H}$ at $(t+1, \omega')$ is based on a preference relation on $\mathcal{H}$ at (t, ω), and consumption smoothing is defined on an event-by-event basis. Then, A11U-IA ensures the existence of a subjective prior over time-aggregated utility indices. In particular, a subjective prior is constructed recursively. An agent does not form a prior over all states in Ω at any particular time because these axioms are only concerned with acts that are measurable with respect to the σ-field generated by $\mathcal{F}_{t+1}$. Also, note that the independence axiom is applicable only to the subset of $\mathcal{D}_t$. Since time-variability aversion generates a preference for consumption smoothing over time on each event, the operation $\oplus$ combines the effects from time-variability aversion and the effects from an attitude toward uncertainty. To deduce a pure attitude toward uncertainty, we must neutralize the effects from time-variability aversion on each event. Formally, time-variability aversion is defined between equally preferred acts. To determine whether time-variability aversion has any effect on a preference relation on acts generated by a convex operation, we need to adjust the level of utility for each state act by utilizing only constant acts to preserve the nature of time-variability (i.e., Condition (i) in A11U-IA).[28] Then, Condition (ii) in A11U-IA ensures that we focus on acts for which time-variability aversion does not change a preference relation under a convex operation.[29]

Given the above axioms, preference relations in the collection of $\{\succeq_{(t,\omega)}\}$ can be represented by the following formula (for the proof, see Appendix C):

Proposition 2: *Preference relations in $\{\succeq_{(0,\omega)}\}$ satisfy A1U to A6U and preference relations in $\{\succeq_{(t,\omega)}\}$ satisfy A7U to A11U if and only if there exist $\{[\alpha_t, \beta_t]\}_{1 \le t \le T}$ and* $\mathrm{E}[.|\mathcal{F}_t(\omega)]$ *such that*

[26] This axiom implies that all states are non-null.

[27] Under the following axiom, (14) has multiple priors instead of a single subjective prior: For all $f, g \in \mathcal{D}_t$ and for all $\alpha \in (0,1)$, $f \simeq_{(t,\omega)} g$ implies $\alpha f \oplus (1-\alpha) g \succeq_{(t,\omega)} g$. Alternatively, we eliminate A11U-IA and change A6U-TVA as follows to incorporate uncertainty-aversion: For all $f, g \in \mathcal{D}$ and for all $\alpha \in (0,1)$, $f \simeq_{(0,\omega)} g$ implies $\alpha f \oplus (1-\alpha) g \succeq_{(0,\omega)} g$.

[28] An act generated by a convex operation between act f and a constant act does not change the nature of time-variability in f (i.e., A4U-CI).

[29] For example, the operation $\oplus$ is not subject to time-variability aversion for $h, h' \in \mathcal{H}$ such that $h^{\tau}(\omega) \succeq_{(\tau+1,\omega)} h$ and $h'^{\tau}(\omega) \succeq_{(\tau+1,\omega)} h'$ for $\tau \in [t, T-1]$.

(14) *For all* $f, g \in \mathcal{D}$, $f \succeq_{(t,\omega)} g$ *if and only if* $V_{(t,\omega)}(f) \geq V_{(t,\omega)}(g)$,
where $\{V_{(t,\omega)}(f)\}_{(t,\omega)\in\mathcal{T}\times\Omega}$ *are recursively defined by*
$V_{(t,\omega)}(f) \equiv \mathrm{E}[\mathrm{Min}_{\delta_{t+1,\omega'}\in[\alpha_{t+1},\beta_{t+1}]}[(1-\delta_{t+1,\omega'})U(f(t,\omega')) + \delta_{t+1,\omega'}V_{(t+1,\omega')}(f)|\mathcal{F}_t(\omega)]$
and $V_{(T,\omega)}(f) \equiv U(f^T(\omega))$, *where* $U(f(t,\omega)) \equiv \mathrm{E}_{f(t,\omega)}[u(x)]$ *with* $x \in X$,

(15) $\delta^*_{t+1,\omega'} = \delta^*_{t+1,\omega''}$ *for all* $\omega'' \in \mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$,
where $\delta^*_{t+1,\omega'} \in \mathrm{argmin}_{\delta_{t+1,\omega'}\in[\alpha_{t+1},\beta_{t+1}]}[(1-\delta_{t+1,\omega'})U(f(t,\omega')) + \delta_{t+1,\omega'}V_{(t+1,\omega')}(f)]$,

(16) $0 < \alpha_t \leq \beta_t < 1$ for all $t \in [1, T]$.

Moreover,

(17) $\mathrm{E}[.|\mathcal{F}_t(\omega)]$ *and* $[\alpha_t, \beta_t]$ *are uniquely defined, and* $[\alpha_t, \beta_t]$ *is independent of states.*

(18) u: $Y \to R$ *is unique up to a positive affine transformation.*

The crucial result is that an agent first considers intertemporal substitution on each of tomorrow's events in $\mathcal{F}_{t+1}$ and then aggregates utility indices across events with probability weights. Therefore, a selection of discount factors at time t is based on tomorrow's events in $\mathcal{F}_{t+1}$ (but not on each state in $\mathcal{F}_{t+1}$). In addition, the expectation is based on a subjective prior, and α_t and β_t depend only on time. Also, $V_{(t,\omega)}(f)$ depends only on future payoffs of f from (t,ω) onward, which implies independence of history and irrelevant alternatives. This operation, $V_{(t,\omega)}(f)$, is recursively applied. Note that if there are no fluctuations in payoffs over states at every point of time, (14) becomes (9).

Now, we examine the connection between intertemporal substitution and risk aversion over states. In terms of (1), we can write (14) as follows:

$$V_{(t,\omega)}(f) = \mathrm{E}[W_t(U(f(t,\omega')), V_{t+1,\omega'}(f))|\mathcal{F}_t(\omega)]$$

On the other hand, the recursive utility approach with an identical attitude toward subjective risk (i.e., the expected utility with the same subjective prior over states) produces the following formula:

$$V_{(t,\omega)}(f) = W_t(U(f(t,\omega')),\ \mathrm{E}[V_{t+1,\omega'}(f)|\mathcal{F}_t(\omega)])$$

Since W_t is a concave function, by Jensen's inequality, our approach generates a lower value for volatile consumption sequences than that of the recursive utility approach. Clearly, (14) captures time-variability aversion more strongly because it is applied on a path-by-path basis. Moreover, the "first-order" property of time-variability aversion generates an asymmetric attitude toward gains and losses. In particular, under (14), this "first-order" property increases an effective risk attitude over tomorrow's states when some of tomorrow's states become "gains" and others become "losses" relative to today's utility level. In other words, the agent discounts utility in good states more than in bad states so that discount factors show gain/loss asymmetry over *states.* This intertemporal substitution mechanism effectively boosts relative risk aversion over tomorrow's consumption.[30] This result is also analogous to that of loss aversion by Kahneman and Tversky (1979), where in our model, today's utility implicitly becomes an endogenous reference point.

Finally, as in the case under certainty, we can only compare the attitude toward time-variability for pairs of preference relations that embody the same ranking of single-period risk and the same ranking of subjective risk (i.e., an identical subjective prior). Then $V_{(t,\omega)}$ is more time-variability averse than $V'_{(t,\omega)}$ if and only if $U = U'$, $\mathrm{E}[.|\mathcal{F}_t] = \mathrm{E}'[.|\mathcal{F}_t]$, and

(19) $\quad V_{(t,\omega)}(f) \leq V'_{(t,\omega)}(f)$ for all $f \in \mathcal{D}$,

where U and U' represent preference relations on Y for $V_{(t,\omega)}$ and $V'_{(t,\omega)}$, and $\mathrm{E}[.|\mathcal{F}_t]$ and $\mathrm{E}'[.|\mathcal{F}_t]$ are subjective priors for $V_{(t,\omega)}$ and $V'_{(t,\omega)}$, respectively. In fact, condition (19) at $t = 0$ is equivalent to

(20) $\quad [\alpha_t,\beta_t] \supseteq [\alpha'_t, \beta'_t]$ for all $t \in [1, T]$,

where $[\alpha_t,\beta_t]$ and $[\alpha'_t, \beta'_t]$ represent sets of discount factors in (14) for $V_{(0,\omega)}$ and $V'_{(0,\omega)}$, respectively.

4 Comparison with Other Intertemporal Utility Functions

4.1 Models of Time-Preferences

In this subsection, we compare the representation of (9) with other intertemporal utility functions that model time-preferences. For this purpose, let $c = (c_0, ..., c_T)$ be a sequence of consumption from time 0 to time T. First, time-preferences have been modeled through a notion of impatience following Fisher (1930). Samuelson (1937) extends Fisher's analysis into the discounted utility model for which Koopmans (1960) provides an axiomatic foundation. Epstein (1983) provides justification for impatience through an axiomatic approach that is applied to risky outcomes and derives the model for the rate of time-preferences (i.e., discount rates) that depend on historical consumption. The main departure of our model from these theories is an introduction of time-variability aversion with the "first-order" property. For example, the model by Epstein (1983) applied to c shows the following:

(21) $\quad V(c) \equiv u(c_0) + \sum_1^T u(c_t) \exp(-\sum_{\tau=0}^{t-1} v(c_\tau)),$

where $u(.)$ and $v(.)$ are continuous. If $v(.)$ is increasing, then this model can express a desire for consumption smoothing over time (as consumption increases, the discount rate will increase). However, as long as $u(.)$ and $v(.)$ are increasing, $u(.)$ and $v(.)$ are differentiable almost everywhere.

[30] See Wakai (2002) for an application to asset pricing.

Then (21) does not demonstrate "first-order" time-variability aversion, for which we require non-differentiability at any constant sequences of consumption.

Second, the representation of (9) shares a structure similar to the aggregator function by Koopmans (1960). In our approach, we impose axioms that derive a particular functional form for an aggregator function (although the function is time-dependent) that incorporates "first-order" time-variability aversion (i.e., gain/loss asymmetry). We find significance of gain/loss asymmetry once we apply (9) under uncertainty (see Section 4.2).

Third, methodologically, our model follows the directions developed in Gilboa (1989) and Shalev (1997). Gilboa (1989) applies the non-additive prior model of Schmeidler (1989) over a sequence of lotteries and derives a utility function that depends not only on consumption itself but also on the difference between adjacent consumptions:

(22) $V(c) \equiv \sum_0^T [\alpha_t u(c_t) + \beta_t |u(c_t) - u(c_{t-1})|]$ with $\beta_0 = 0$.

Shalev (1997) modifies Gilboa (1989)'s formulation and introduces different weights for positive and negative increments:

(23) $V(c) \equiv u(c_0) + \sum_0^T \{\delta_t^+ \max[u(c_t) - u(c_{t-1}), 0] + \delta_t^- \min[u(c_t) - u(c_{t-1}), 0]\}$ with $\delta_0^+ = \delta_0^- = 0$.

Under appropriate parameters, both formulas incorporate "first-order" time-variability aversion; Shalev (1997) also captures gain/loss asymmetry over time. However, under the assumption of history-independence and under the condition of $\beta_1 \neq 0$ and $\delta_1^+ \neq 0$ (or $\delta_1^- \neq 0$), neither formula satisfies dynamic consistency: The choice for $(c_1, ..., c_T)$ at time 0 might not be optimal at time 1 (also see Sarin and Wakker (1998) and Grant, Kajii, and Polak (2000)). On the other hand, the multiple-priors model of Gilboa and Schmeidler (1989) can be applied consistently over time under a fixed informational structure (i.e., filtration).[31] In addition, time-variability aversion is a dual notion of uncertainty aversion in Gilboa and Schmeidler (1989) applied over time, and an interpretation of time-variability aversion is clear under our model. On the contrary, (22) and (23) require additional conditions on parameters to show time-variability aversion, and defining a base-line time-preference to derive the "first-order" property may not be obvious. Moreover, under dynamic consistency, an agent who follows our model compares present utility with the present discounted value of all future utility. Then our formula automatically assigns a lower discount

[31] Sarin and Wakker (1998) demonstrate that under the multiple-priors model, an agent can stay in the same family of the representation that shows consistent preferences over time, as long as the agent has a recursive multiple-priors set that needs to be updated under Bayes' Rule. Epstein and Schneider (2003) formally prove the above claim. In addition, Wakai (2003) shows that given the *ex-ante* multiple-priors model, conditional preference relations must satisfy the set of axioms identical to those of the multiple-priors model if they satisfy dynamic consistency and independence of irrelevant alternatives.

factor for higher future consumption. On the other hand, for (23) to have $\delta_t^- \geq \delta_t^+$, we need to impose an additional assumption.

Fourth, to capture "first-order" time-variability aversion, we may apply other non-expected utility models over time as Segal and Spivak (1990, 1997) apply them under risk.[32] However, the direct application of non-expected theory that is concerned with *objective* risk (for example, Yaari (1987) and Gul (1991)) does not capture the notion of time-preferences. For instance, under the assumption of finitely many distinct outcomes, $x_1, ..., x_N$, the dual theory of Yaari (1987) can derive the following utility for objective risk $p = (p_1, ..., p_N)$:

$$(24) \qquad V(c) \equiv \textstyle\sum_{i=1}^{N} \{g(\sum_1^i p_j) - g(\sum_1^{i-1} p_j)\} x_i,$$

if $x_1 < x_2 < ... < x_N$. The function $g : [0, 1] \to [0, 1]$ is onto, strictly increasing, and concave, where concavity of g implies risk aversion. The application of (24) over time immediately raises the problem of how to map objective risk $p = (p_1, ..., p_N)$ over outcome $x = (x_1, ..., x_N)$ to discount factor $\delta = (\delta_1, ..., \delta_T)$ over time $T = (1, ..., T)$.

Fifth, as an application of the optimal control theory motivated by a worst-case scenario analysis, Geoffard (1996) defines the variational utility as follows:

$$(25) \qquad V(c) \equiv \min_{r_t \in \mathcal{R}} \textstyle\int_0^\infty f(c_t, B_t, r_t) dt,$$

where $\mathcal{R}$ is a set of admissible paths of rates of time-preferences, and f gives current felicity as a function of the current value of consumption c_t, discount factor B_t, and rate of time-preference r_t. Geoffard (1996) shows that the recursive utility of Epstein (1987), the discounted utility, and Uzawa utility (1968) belong to this class of the representation.[33] In our approach, we focus on the normative aspect of minimization as a consequence of consumption smoothing. For this reason, our approach may be regarded as an axiomatization of (25) under a discrete time setting with a particular emphasis on "first-order" time variability aversion and gain/loss asymmetry.

Sixth, Loewenstein and Thaler (1989) conjecture that loss aversion (gains and losses are based on the deviation from the previous level of consumption) is one reason for which an agent has a preference for increasing sequences. Their postulate is a direct application of prospect theory as developed by Kahneman and Tversky (1979): Atemporal utility depends on a reference point (here, a level of past consumption), and the attitude toward gains and the attitude toward losses from the reference point is asymmetric. On the other hand, our representation of (9) is only based on an aversion to fluctuations of payoffs over time and does not pre-specify the existence of reference points. Gain/loss asymmetry is expressed in (9) because an agent compares only two numbers at

[32]For applications of first-order risk aversion in Segal and Spivak (1990) to asset pricing, see Epstein and Zin (1990) and Bekaert, Hodrick, and Marshall (1997).

[33]Dumas, Uppal, and Wang (2000) extend Geoffard's analysis to the case with uncertainty and investigate efficient allocations under a continuous-time version of recursive utility as developed by Duffie and Epstein (1992).

each time due to a recursive structure. In other words, dynamic consistency necessitates gain/loss asymmetry. Therefore, our model can be regarded as an axiomatization of loss aversion *over time*.

Seventh, Loewenstein and Prelec (1993) develop the model that explains both a preference for spreading consumption and a preference for increasing consumption. They focus on the global nature of sequences and argue that loss aversion should be based on the comparison between the set of outcomes that are yet to occur and the set of outcomes that have already occurred. Our approach shares an idea similar to Loewenstein and Prelec (1993) because time-variability aversion explains global characteristics. However, in our approach, gain/loss asymmetry is based on the comparison between today's utility (rather than all utility accumulated until today) and the present discounted value of all future utility. Moreover, under the assumption of history-independence, the formula in Loewenstein and Prelec (1993) does not confirm dynamic consistency.

Finally, the habit formation models[34] of Sundaresan (1989) and Constantinides (1990) (also, recently applied in Campbell and Cochrane (1999, 2000)) introduce a utility function that depends on historical consumption according to the following formula:

$$(26) \qquad V(c) \equiv \textstyle\sum_{t=0}^{T} \delta^{t-1} u(c_t - x_t),$$

where $x_0 = 0$ and $x_t = f(c_0, ..., c_{t-1})$. The term $u(c_t - x_t)$ represents an idea similar to Loewenstein and Thaler (1989) under which a reference point is based on the previous level of consumption. However, unless we use loss aversion in $u(c_t - x_t)$, (26) does not satisfy "first-order" time-variability aversion because an increasing u is differentiable almost everywhere.

On logical grounds, our model is based on the assumption of history independence. Clearly, this assumption is not innocuous. However, without a formal derivation, we may extend time-variability aversion (i.e., the multiple-discount-factors model) over atemporal utility indices that depend on previous levels of consumption.

4.2 Comparison among (9), (14), and Recursive Utility

The main assumption of the recursive utility model of Kreps and Porteus (1979) and Epstein and Zin (1989) is that the ranking of future risky prospects is independent of today's consumption. To be temporally consistent (Kreps and Porteus (1979)), this separation is sensible if both early and late resolution of uncertainty is in an agent's choice set. The novelty of Epstein and Zin (1989) is an interpretation of this separability. Their model defines an attitude toward variability over time as an attitude toward a sequence of consumption without risk and an attitude toward atemporal risk as an attitude toward objective risk (over a sequence of consumption) that resolves completely in a single period. Both attitudes are combined by an aggregator function under which an agent considers intertemporal substitution between today's consumption and a certainty equivalent level

[34]See Frederick, Loewenstein, and O'Donoghue (2002) for a survey.

of future utility. However, since the attitude toward variability over time changes the distribution of time-$t+1$ utility, the ranking of risky future prospects necessarily depends on the attitude toward variability over time. Then a binary relation of “more risk averse” is permitted only for pairs of preference relations that embody the same ranking of deterministic prospects.

On the other hand, we also consider the separation of attitudes toward risk and time-variability. Our model deduces atemporal risk attitude from a preference relation among constant acts. Since receiving an identical lottery over time is not equivalent to a resolution of risk in a single period, risk attitude as deduced in our model may involve both risk and time preferences. This procedure only gains rationality when we introduce dynamic consistency under which atemporal risk-preferences at the last period become atemporal risk-preferences of earlier periods. However, when we evaluate a sequence of consumption without uncertainty, intertemporal substitution necessarily depends not only on the attitude toward time-variability but also on the attitude toward risk. Then a binary relation of “more time-variability averse” is permitted only for pairs of preference relations that embody the same ranking of single-period risk.

In addition, in our model, the ranking of uncertain future prospects depends on today’s consumption. This dependence is summarized by an apparent change of order of aggregation: An agent first considers time-variability aversion on an event-by-event basis, then aggregates utilities under a subjective prior. The structure of the model provides the justification of this construction. In our model, an agent is informed of a particular event-tree as an information structure. Then, there is no “early resolution” of uncertainty at $t+1$ over all states $\omega' \in \mathcal{F}_t(\omega)$. Thus, a certainty equivalent over all states $\omega' \in \mathcal{F}_t(\omega)$ at $t+1$ is not well-defined because an agent at $t+1$ already knows that one of the events $\mathcal{F}_{t+1}(\omega) \subseteq \mathcal{F}_t(\omega)$ has been realized. Therefore, an agent would rather consider sequences of consumption to be possibilities and forms a notion of probability over these possibilities *ex-ante* at time t.

Given the above characteristics of both models, the model developed by Epstein and Zin (1989) is suitable for the situation under which an agent faces the possibility of early resolution of uncertainty. In addition, their model can incorporate non-expected utility models, whereas our model only permits the expected utility model (although our model can incorporate the multiple-priors model if we change A11U-IA). Moreover, the recursive utility model can evaluate either objective temporal lotteries or lotteries on a state space,[35] whereas our model can evaluate lotteries only on a state space. In contrast, our model is suitable for the situation under which an agent is informed of a particular structure of information. Our model utilizes this knowledge and expresses a stronger motive for consumption smoothing over time. In particular, our model considers time-variability aversion before subjective risk over states. This operation generates dependence of the ranking of uncertain future prospects on today’s consumption and amplifies the effects of gain/loss asymmetry over time. Finally, our model provides a direct interpretation of time-preferences through the

[35]See Skiadas (1998).

model of discount factors.

4.3 Comparison between (14) and Epstein and Schneider (2003)

In our model, preference relations are defined on the same domain as in Epstein and Schneider (2003). Their model derives recursive multiple-priors over a time-state lottery domain by adapting the constant-independence axiom of Gilboa and Schmeidler (1989) under a new definition of constant acts: constant acts in their model are certainty acts in our model that involve variations over time but not over states. Therefore, their model does not consider the aversion to fluctuations in utility over time. On the other hand, our model uses constant acts that do not involve fluctuations over time and states. Thus, our approach can capture the aversion to the variations of utility over both time and states although we eliminate concerns with the variations over states by imposing the axiom that derives a subjective prior instead of multiple priors. However, our model is not a simple re-interpretation of Epstein and Schneider (2003). In order to derive a subjective prior as well as time-variability aversion in their framework, we need to exchange the role of time with that of states. Then we need to define acts that involve fluctuations over states but not over time as constant acts, which no longer satisfy measurability. Instead, we focus on t-recursive acts that summarize the uncertainty an agent faces at (t, ω), then derive a subjective prior by neutralizing effects from time-variability aversion. On technical grounds, we do not assume that each conditional preference relation satisfies the axioms of time-variability aversion. Instead, we update the *ex-ante* preference relation based on dynamic consistency, independence of history and irrelevant alternatives, and monotonicity over states.

5 Conclusion

We have axiomatized the behavioral notion of time-variability aversion and have derived the representation. Our model reaches the following conclusions: (i) Time-variability aversion is captured as a separate attitude from risk-aversion through the model of multiple discount factors, and an axiomatic derivation provides a clear picture of an agent's motives. The representation is also parsimonious; (ii) Given dynamic consistency and history independence, the "first-order" property of time-variability aversion necessarily implies gain/loss asymmetry in discount factors; and (iii) Under uncertainty, a consideration of intertemporal substitution occurs before a consideration of subjective risk. Under this operation, discount factors depend on *tomorrow's* states and exhibit gain/loss asymmetry over *states* by taking today's utility as an endogenous reference point; time-variability aversion effectively boosts risk aversion over states.

Appendix A: Proof of Proposition 1

We define time passage as an information structure, i.e., $\mathcal{F}_0 = \{\{0,..,T\}\}$, $\mathcal{F}_1 = \{\{0\},\{1,..,T\}\}$, ..., $\mathcal{F}_t = \{\{0\},\{1\},...,\{t-1\},\{t,..,T\}\}$, ..., $\mathcal{F}_T = \{\{0\},\{1\},...,\{T\}\}$. For sufficiency, between $\mathcal{F}_0$ and $\mathcal{F}_1$, by A7-IH and A8-DC, Proposition 2 of Wakai (2003) developed for the multiple-priors model implies that A1 to A6 hold for a collection of all f^1. Then by Theorem 1, for all $f, g \in \mathcal{H}$, $f \succeq_1 g$ if and only if $V_1(f) \geq V_1(g)$, where $V_1(f) = V_1(f^1)$, as defined by Theorem 1 (i.e., $\delta_0 = 0$). In addition, let $\mathcal{H}(h^t)$ be a collection of all $h' \in \mathcal{H}$ such that $h^{t\prime} = h^t$. Then by A5-SM, for all $f, g \in \mathcal{H}(h^1)$, $f \succeq_0 g$ if and only if $f(0) \succeq_0 g(0)$. Then A1 to A6 restrictively hold for a collection of $\{h(0)\}$ on $\mathcal{H}(h^1)$. Hence, Proposition 3 of Wakai (2003) implies that there exists a unique $[\alpha_1, \beta_1]$ with $0 < \alpha_1 \leq \beta_1 < 1$ such that for all $f, g \in \mathcal{H}$, $f \succeq_0 g$ if and only if $V_0(f) \geq V_0(g)$, where $V_0(f) \equiv \min_{\delta_1 \in [\alpha_1, \beta_1]}[(1-\delta_1)U(f(0))+\delta_1 V_1(f)]$. By applying the same construction for each $V_t(f)$, we derive the conclusion. For necessity, (8) immediately implies that A1 to A6 hold for $\mathcal{H}$ at time 0. A7-IH is obviously implied. A8-DC is satisfied because (9) has a recursive structure with $0 < \alpha_1 \leq \beta_1$.■

Appendix B: Problem of Dynamic Consistency under Uncertainty

We want to show that applying (9) at each state $\omega \in \Omega$ then aggregating each utility with subjective probability is inconsistent with A8U-IHA and A9U-DC if $T > 1$. We illustrate this point by the following example: Suppose that there are three periods and four states. At time 0, an agent does not have any information about states. At time 1, an agent is informed that either (state 1, state 2) or (state 3, state 4) has occurred. At time 2, an agent knows the true state. The subjective probability of state realization is (0.25, 0.25, 0.25, 0.25) at time 0 and (0.5, 0.5) on (state 1, state 2) and (state 3, state 4) at time 1. Now an agent assigns a utility index following (9) with $u(c) = c$ and $[\alpha_t, \beta_t] = [0.3, 0.7]$ for both time 0 and time 1 and, under the assumption of A8U-IHA, evaluates the following stream of consumptions:

Consumption A

State\Time	0	1	2
1	1.5	2	2
2	1.5	2	1
3	1.5	1	2
4	1.5	1	1

Consumption B

State\Time	0	1	2
1	1.5	3	1
2	1.5	3	1
3	1.5	1	2
4	1.5	1	1

Utility

Cons.\Time	0	1 at (1,2)
Cons. A	1.4325	1.65
Cons. B	1.4450	1.60

Clearly, at time 0, an agent chooses Consumption B, but after the agent realizes that the agent is at (state 1, state 2) at time 1, the agent selects Consumption A. This example violates dynamic consistency. More formally, under a neoclassical expected utility model with a subjective-prior, inconsistency emerges over time unless an agent's preferences are time additive. Since (9) is not time additive, having a dynamically consistent utility function (9) on each state does not confirm dynamic consistency.

Appendix C: Proof of Proposition 2

Sufficiency: The proof of sufficiency is based on the following Lemmas (C.1 to C.6):

Lemma C.1: *Under A7U to A10U, preference relations in $\{\succeq_{(t,\omega)}\}$ satisfy A1U to A4U and AU6 at all $(t,\omega) \in \mathcal{T} \times \Omega$. Moreover, at all $(t,\omega) \in \mathcal{T} \times \Omega$, a preference relation on $\mathcal{H}$ can be represented by $V_{(t,\omega)}$ that is identical to V_t of Proposition 1, where a set of discount factors depends only on time, not on a state (state-independence).*

Proof. By A1U-WO, A2U-C, and A4U-CI, a preference relation on $\mathcal{L}$ at $(0,\omega)$ depends only on which lottery is assigned at each (t,ω). Then there exists a function $U : Y \to R$ such that for all $l, l' \in \mathcal{L}$, $l \succeq_{(0,\omega)} l'$ if and only if $U_{(0,\omega)}(y) \geq U_{(0,\omega)}(z)$, where $l(t,\omega) = y$ and $l'(t,\omega) = z$ for all $(t,\omega) \in \mathcal{T} \times \Omega$. Given this result, $y \succeq_{(0,\omega)} y'$ means that $l \succeq_{(0,\omega)} l'$. Suppose that there are no $y, y' \in Y$ such that $y \succ_{(0,\omega)} y'$. Then by A5U-SMT, there are no $f, f' \in \mathcal{D}$ such that $f \succ_{(0,\omega)} f'$, which violates A3U-ND. Hence, there exist $y, y' \in Y$ such that $y \succ_{(0,\omega)} y'$.

Next, by A1U to A6U as well as the results in the above paragraph, a preference relation on $\mathcal{H}$ depends only on which state act is assigned on each state $\omega \in \Omega$ (in particular, $h_\omega = h_{\omega'}$ for all $\omega, \omega' \in \Omega$); thus, the representation is concerned only with a particular lottery assigned on each state act. Hence, we obtain Theorem 1 for $\mathcal{H}$ at $(0,\omega)$, where $V_{0,\omega}(h) \equiv \min_{\delta \in \Delta_{(0,\omega)}} \sum_{\tau=0}^{T} U_{0,\omega}(h(\tau,\omega))\delta_\tau$ and for all $\delta \in \Delta_{(0,\omega)}$, $\sum_{\tau=0}^{T} \delta_\tau$ =1 and $\delta_\tau > 0$ for all $\tau \in \mathcal{T}$. Note that $V_{0,\omega}(h) = \min_{\delta \in \Delta_{(0,\omega)}} \sum_{\tau=0}^{T}$ $U_{0,\omega}(h(\tau,\omega'))\delta_\tau$ for all $\omega' \in \Omega$ with $\omega' \neq \omega$ so that a selection of state acts is irrelevant.

As for A1U to A3U, first, by A7U-F, A8U-IHA, and A9U-DC, both A1U-WO and A2U-C hold for $\succeq_{(1,\omega)}$ for all $\omega \in \Omega$. Second, let $y, y' \in Y$ satisfy $y \succ_{(0,\omega)} y'$. Suppose that $y \preceq_{(1,\omega')} y'$ at some $\omega' \in \mathcal{F}_1(\omega)$. Let $f \in \mathcal{D}_0$ satisfy $f(\tau,\omega'') = y'$ for all $(\tau,\omega'') \in \mathcal{T}_1 \times \mathcal{F}_1(\omega)$ and $f(\tau,\omega'') = y$ for all $(\tau,\omega'') \notin \mathcal{T}_1 \times \mathcal{F}_1(\omega)$. Then by A10U-SMS and the result in the above paragraph, $y \succ_{(0,\omega)} f$.

However, A7U-F, A8U-IHA, and A9U-DC imply that $y \preceq_{(0,\omega)} f$, which is inconsistent with the result from A10U-SMS. Hence, for all $y, y' \in Y$, $y \succ_{(0,\omega)} y'$ implies $y \succ_{(1,\omega')} y'$ for all $\omega' \in \mathcal{F}_1(\omega)$. Conversely, the argument above implies that for each $\omega' \in \mathcal{F}_1(\omega)$, $y \succ_{(1,\omega')} y'$ implies $y \succ_{(0,\omega)} y'$ for all $y, y' \in Y$. The same result holds for all events in $\mathcal{F}_1$ that are subsets of $\mathcal{F}_0(\omega) = \Omega$. Thus, for all $y, y' \in Y$, $y \succ_{(0,\omega)} y'$ if and only if $y \succ_{(1,\omega')} y'$ for all $\omega' \in \Omega$. Since $\succeq_{(0,\omega)}$ and $\succeq_{(1,\omega')}$ contain all elements of $\mathcal{L} \times \mathcal{L}$, the argument above implies that $\succeq_{(0,\omega)}$ and $\succeq_{(1,\omega')}$ are identical on $\mathcal{L}$ for all $\omega' \in \mathcal{F}_0(\omega) = \Omega$. This result indicates that preference relations on $\mathcal{L}$ at $(1, \omega')$ can be represented by $U_{(1,\omega')}$, which is a positive affine transformation of $U_{(0,\omega)}$. Clearly, A3U-ND holds for $\succeq_{(1,\omega')}$ for all $\omega' \in \Omega$.

We now consider A4U-CI. For $f \in \mathcal{D}$, let $\mathcal{D}_{\mathcal{F}_1(\omega)}(f)$ be a collection of all acts $g \in \mathcal{D}$ such that $f(\tau, \omega') = g(\tau, \omega')$ for all $(\tau, \omega') \notin \mathcal{T}_1 \times \mathcal{F}_1(\omega)$. Let $f, g \in \mathcal{D}$ satisfy $g \in \mathcal{D}_{\mathcal{F}_1(\omega)}(f)$. By A7U-F and A9U-DC, $f \succeq_{(0,\omega)} g$ if and only if $f \succeq_{(1,\omega')} g$ for all $\omega' \in \mathcal{F}_1(\omega)$. By A4U-CI, for all $l \in \mathcal{L}$ and for all $\alpha \in (0,1)$, $f \succeq_{(0,\omega)} g$ if and only if $\alpha f \oplus (1-\alpha) l \succeq_{(0,\omega)} \alpha g \oplus (1-\alpha) l$. Since $\alpha g \oplus (1-\alpha) l \in \mathcal{D}_{\mathcal{F}_1(\omega)}(\alpha f \oplus (1-\alpha) l)$, the above result implies that at all $\omega' \in \mathcal{F}_1(\omega)$, for all $l \in \mathcal{L}$ and for all $\alpha \in (0,1)$, $f \succeq_{(1,\omega')} g$ if and only if $\alpha f \oplus (1-\alpha) l \succeq_{(1,\omega')} \alpha g \oplus (1-\alpha) l$. Given A8U-IHA, this result implies that A4U-CI holds for $\succeq_{(1,\omega')}$ for all $\omega' \in \mathcal{F}_1(\omega)$. By repeating the same argument on all events in $\mathcal{F}_1$, A4U-CI is satisfied under $\succeq_{(1,\omega')}$ for all $\omega' \in \Omega$.

To show that A6U-TVA holds for $\succeq_{(1,\omega')}$ for all $\omega' \in \Omega$, let $f, g \in \mathcal{D}_0$ satisfy $g \in \mathcal{D}_{\mathcal{F}_1(\omega)}(f)$ (i.e., $f_{\omega'} = f_{\omega''}$ and $g_{\omega'} = g_{\omega''}$ for all $\omega', \omega'' \in \mathcal{F}_1(\omega)$). Under A10U-SMS, $f \succeq_{(0,\omega)} g$ if and only if $f_{\omega'} \succeq_{(0,\omega)} g_{\omega'}$, where $f_{\omega'}, g_{\omega'} \in \mathcal{Y}^{T+1}$ at any one of $\omega' \in \mathcal{F}_1(\omega)$. By A7U-F and A9U-DC, $f \succeq_{(0,\omega)} g$ if and only if $f \succeq_{(1,\omega'')} g$ for all $\omega'' \in \mathcal{F}_1(\omega)$. By A7U-F, A8U-IHA, and A10U-SMS, $f \succeq_{(1,\omega'')} g$ for all $\omega'' \in \mathcal{F}_1(\omega)$ if and only if $f_{\omega'} \succeq_{(1,\omega'')} g_{\omega'}$ for all $\omega'' \in \mathcal{F}_1(\omega)$. Hence, for any $\omega' \in \mathcal{F}_1(\omega)$, $f_{\omega'} \succeq_{(0,\omega)} g_{\omega'}$ if and only if $f_{\omega'} \succeq_{(1,\omega'')} g_{\omega'}$ for all $\omega'' \in \mathcal{F}_1(\omega)$. Now, by A6U-TVA, for $\alpha \in (0,1)$, if $f \simeq_{(0,\omega)} g$, then $\alpha f \oplus (1-\alpha) g \succeq_{(0,\omega)} g$. Since $f, g, \alpha f \oplus (1-\alpha) g \in \mathcal{D}_{\mathcal{F}_1(\omega)}(f)$ for all $\alpha \in (0,1)$, by the above result, for $\alpha \in (0,1)$, if $f_{\omega'} \simeq_{(1,\omega'')} g_{\omega'}$, then $\alpha f_{\omega'} \oplus (1-\alpha) g_{\omega'} \succeq_{(1,\omega'')} g_{\omega'}$ for all $\omega'' \in \mathcal{F}_1(\omega)$. Since $f_{\omega'}, g_{\omega'} \in \mathcal{H}$, under A8U-IHA, this result implies that A6U-TVA holds for $\succeq_{(1,\omega')}$ for all $\omega' \in \mathcal{F}_1(\omega)$. By repeating the same argument on all events in $\mathcal{F}_1$, A6U-TVA is satisfied under $\succeq_{(1,\omega')}$ for all $\omega' \in \Omega$.

Under A7U-F, A8U-IHA, and A9U-DC, by applying the same argument at all events in $\mathcal{F}_t$ from time 2 to time T, A1U to A4U and A6U-TVA hold for $\succeq_{(t,\omega)}$ for all $(t, \omega) \in \mathcal{T}_1 \times \Omega$. In particular, for Y, at all $(t, \omega') \in \mathcal{T}_1 \times \Omega$, $U_{(t,\omega')}$ is a positive affine transformation of $U_{(0,\omega)}$.

We now show that a modified version of A5U-SMT is applied to all $(t, \omega) \in \mathcal{T}_1 \times \Omega$. At any $(t, \omega) \in \mathcal{T}_{-T} \times \Omega$, let $f, g \in \mathcal{D}_t$ satisfy $g \in \mathcal{D}_{\mathcal{F}_{t+1}(\omega)}(f)$, where $\mathcal{D}_{\mathcal{F}_{t+1}(\omega)}(f)$ is a collection of all acts $f' \in \mathcal{D}$ such that $f'(\tau, \omega') = f(\tau, \omega')$ for all $(\tau, \omega') \notin \mathcal{T}_{t+1} \times \mathcal{F}_{t+1}(\omega)$. By the results in the above paragraphs, (i) for $f_{\omega'}$ and $g_{\omega'}$ at any $\omega' \in \mathcal{F}_{t+1}(\omega)$, $f_{\omega'} \succeq_{(t,\omega)} g_{\omega'}$ if and only if $f_{\omega'} \succeq_{(t+1,\omega'')} g_{\omega'}$ for all $\omega'' \in \mathcal{F}_{t+1}(\omega)$ and (ii) for Y, $U_{(t+1,\omega'')}$ is a positive affine transformation of $U_{(0,\omega)}$. Then under A8U-IHA, the following modified version of A5U-SMT holds for all $\omega'' \in \mathcal{F}_1(\omega)$: For all $f, g \in \mathcal{H}$, if for all $\tau \in \mathcal{T}$, $f(\tau,\omega') \succeq_{(1,\omega'')} g(\tau,\omega')$ for all $\omega' \in \Omega$, then $f \succeq_{(1,\omega'')} g$. In addition, if for some $\tau \in \mathcal{T}_1$, $f(\tau,\omega') \succ_{(1,\omega'')} g(\tau,\omega')$ for all $\omega' \in \Omega$, then $f \succ_{(1,\omega'')} g$. Under A7U-F and A8U-IHA, by repeating the same argument at all events in $\mathcal{F}_t$ from time 2 to time T, the above result implies that the modified version of A5U-SMT holds at all $(t,\omega) \in \mathcal{T}_1 \times \Omega$: For all $f, g \in \mathcal{H}$, if for all $\tau \in \mathcal{T}$, $f(\tau,\omega') \succeq_{(t,\omega)} g(\tau,\omega')$ for all $\omega' \in \Omega$, then $f \succeq_{(t,\omega)} g$. In addition, if for some $\tau \in \mathcal{T}_t$, $f(\tau,\omega') \succ_{(t,\omega)} g(\tau,\omega')$ for all $\omega' \in \Omega$, then $f \succ_{(t,\omega)} g$.

Given the results above, we obtain Proposition 1 at each $(t,\omega) \in \mathcal{T}_1 \times \Omega$ on $\mathcal{H}$, where $V_{(t,\omega)}(h) = \min_{\delta \in \Delta_{(t,\omega)}} \sum_{\tau=t}^{T} U_{(t,\omega)}(h(\tau,\omega))\delta_\tau$ and for all $\delta \in \Delta_{(t,\omega)}$, $\sum_{\tau=t}^{T} \delta_\tau = 1$ and $\delta_\tau > 0$ for all $\tau \in \mathcal{T}_t$. Note that $\delta_\tau = 0$ for all $\tau \in \mathcal{T}_{-t}$ by A8U-IHA (i.e., time τ is null for all $\tau \in \mathcal{T}_{-t}$). Also, under each $\succeq_{(t,\omega)}$, we must use $U_{(t,\omega)}$ in the representation on $\mathcal{H}$ (consistency within $\succeq_{(t,\omega)}$). This result implies that at each (t,ω), the representation does not aggregate different $U_{(\tau,\omega)}$ for $\tau \in \mathcal{T}_t$ under $V_{(t,\omega)}$. In addition, preference relations on $\mathcal{L}$ are identical among all $\succeq_{(t,\omega)}$ ($U_{(t,\omega)}$ is unique up to a positive affine transformation among $\{\succeq_{(t,\omega)}\}$), and each $\succeq_{(t,\omega)}$ is a distinct binary relation. Hence, without loss of generality, the same U can represent each $\succeq_{(t,\omega)}$ on $\mathcal{L}$ at every $(t,\omega) \in \mathcal{T} \times \Omega$.

Now let $\mathcal{D}_{0,\omega,y}$ be a collection of all acts $f, g \in \mathcal{D}_0$ such that $f(\tau,\omega') = g(\tau,\omega') = y$ for all $(\tau,\omega') \notin \mathcal{T}_1 \times \mathcal{F}_1(\omega)$. By A10U-SMS, for all $f, g \in \mathcal{D}_{0,\omega,y}$, $f \succeq_{(0,\omega)} g$ if and only if $f_{\omega'} \succeq_{(0,\omega)} g_{\omega'}$ for any $\omega' \in \mathcal{F}_1(\omega)$. By the result in the previous paragraph, $f_{\omega'} \succeq_{(0,\omega)} g_{\omega'}$ if and only if $f_{\omega'} \succeq_{(1,\omega'')} g_{\omega'}$ on all $\omega'' \in \mathcal{F}_1(\omega)$. Note that both $\succeq_{(0,\omega)}$ on $\mathcal{H}$ and $\succeq_{(1,\omega'')}$ on $\mathcal{H}$ are represented by multiple-discount-factors models with $\Delta_{(0,\omega)}$ and $\Delta_{(1,\omega)}$, respectively. By A7U-F, $V_{(1,\omega'')}$ represents the same ranking on $\mathcal{H}$ as that of $V_{(1,\omega''')}$ at all $\omega''' \in \mathcal{F}_1(\omega)$ and $V_{(0,\omega)}$ represents the same ranking on $\mathcal{H}$ as that of $V_{(0,\omega')}$ for all $\omega' \in \Omega$. On the other hand, let $\mathcal{H}(h^t)$ be a collection of all $h' \in \mathcal{H}$ such that $h^{\tau\prime} = h^\tau$ for all $\tau \in \mathcal{T}_t$. Then by A5U-SMT, for all $h, h' \in \mathcal{H}(h^1)$, $h \succeq_{(0,\omega)} h'$ if and only if $h(0,\omega) \succeq_{(0,\omega)} h'(0,\omega)$. Then A1U to A6U restrictively hold for a collection of time acts $\{h^0\}$ on $\mathcal{H}(h^1)$. Then under A7U-F, A8U-IHA, and A9U-DC, Proposition 3 of Wakai (2003) is applied at all $\omega' \in \mathcal{F}_1(\omega)$ and $\Delta_{(0,\omega')} = \Lambda_{(1,\omega')} \times \Delta_{(1,\omega')}$ with $\Lambda_{(1,\omega')} = [\alpha_{1,\omega'}, \beta_{1,\omega'}]$. Repeatedly applying the same argument for all events in $\mathcal{F}_1$, $\Delta_{(0,\omega)} = \Lambda_{(1,\omega)} \times \Delta_{(1,\omega)}$ with $\Lambda_{(1,\omega)} = [\alpha_{1,\omega}, \beta_{1,\omega}]$ for all $\omega \in \Omega$. Then the uniqueness of $\Delta_{(0,\omega)}$ implies that $\Lambda_{(1,\omega)} = \Lambda_{(1,\omega')}$ and $\Delta_{(1,\omega)} = \Delta_{(1,\omega')}$ for all $\omega, \omega' \in \Omega$.

Let $\mathcal{D}_{1,\omega,y}$ be a collection of all acts $f, g \in \mathcal{D}_1$ such that $f(\tau,\omega') = g(\tau,\omega') = y$ for all $(\tau,\omega') \notin \mathcal{T}_2 \times \mathcal{F}_2(\omega)$. By applying the same argument as in the above paragraph on all $\mathcal{F}_2(\omega') \subseteq \mathcal{F}_1(\omega)$, $\Delta_{1,\omega'} = \Lambda_{2,\omega'} \times \Delta_{2,\omega'}$ with $\Lambda_{2,\omega'} = [\alpha_{2,\omega'}, \beta_{2,\omega'}]$ for all $\omega' \in \mathcal{F}_1(\omega)$, and both $\Lambda_{2,\omega} = \Lambda_{2,\omega'}$ and $\Delta_{2,\omega} = \Delta_{2,\omega'}$ for all $\omega, \omega' \in \mathcal{F}_1(\omega)$. Moreover, since $V_{(1,\omega)}$ represents the same ranking on $\mathcal{H}$ as that of $V_{(1,\omega')}$ for all $\omega' \in \Omega$, the uniqueness of $\Delta_{(1,\omega)}$ implies that $\Lambda_{2,\omega} = \Lambda_{2,\omega'}$ and $\Delta_{2,\omega} = \Delta_{2,\omega'}$ for all $\omega, \omega' \in \Omega$.

By repeatedly applying the same construction from time 2 to time T, under A7U-F, A8U-IHA, A9U-DC, and A10U-SMS, we conclude that preference relations on $\mathcal{H}$ can be represented by $V_{(t,\omega)}$ that is identical to V_t of Proposition 1 at all $(t,\omega) \in \mathcal{T} \times \Omega$, where a set of discount factors depends only on time.■

Lemma C.2: *Suppose that preference relations in $\{\succeq_{(0,\omega)}\}$ satisfy A1U to A6U and preference relations in $\{\succeq_{(t,\omega)}\}$ satisfy A7U to A10U. Then at all $(t, \omega) \in \mathcal{T}_{-T} \times \Omega$, for all $f \in \mathcal{D}$,*

(i) If $h \succ_{(t,\omega)} f \succ_{(t,\omega)} h'$ for some $h, h' \in \mathcal{H}$, then there exists $\alpha \in (0, 1)$ such that $f \simeq_{(t,\omega)} \alpha h \oplus (1 - \alpha)h'$,

(ii) There exists $g \in \mathcal{D}_t$ such that (a) $g(\tau, \omega) = f(\tau, \omega)$ for all $(\tau, \omega) \notin \mathcal{T}_{t+1} \times \Omega$, (b) $f \simeq_{(t,\omega)} g$, and (c) $f \simeq_{(t+1,\omega')} g$ for all $\omega' \in \Omega$,

(iii) There exist $h \in \mathcal{H}$ and $y \in \mathcal{L}$ such that $f \simeq_{(t,\omega)} h \simeq_{(t,\omega)} y$.

Proof. For (i), let $h, h' \in \mathcal{H}$ satisfy $h \succ_{(t,\omega)} h'$, and let $V_{(t,\omega)}$ be defined as in Lemma C.1. By Gilboa and Schmeidler (1989), $V_{(t,\omega)}$ is superadditive among $\mathcal{H}$, i.e., for all $h, h' \in \mathcal{H}$, $V_{(t,\omega)}(\alpha h \oplus (1 - \alpha)h') \geq \alpha V_{(t,\omega)}(h) + (1 - \alpha)V_{(t,\omega)}(h')$ for $\alpha \in [0, 1]$. So $V_{(t,\omega)}(.)$ is concave in $h \in \mathcal{H}$ under the operation $\oplus$. Let $\Phi(\alpha) = V_{(t,\omega)}(\alpha h \oplus (1-\alpha)h')$. Then for $\alpha, \beta, \gamma \in [0, 1]$, $\Phi(\gamma\alpha + (1-\gamma)\beta) = V_{(t,\omega)}(\gamma(\alpha h \oplus (1-\alpha)h') \oplus (1-\gamma)(\beta h \oplus (1-\beta)h')) \geq \gamma V_{(t,\omega)}(\alpha h \oplus (1-\alpha)h') + (1-\gamma)V_{(t,\omega)}(\beta h \oplus (1-\beta)h') = \gamma\Phi(\alpha) + (1 - \gamma)\Phi(\beta)$. So $\Phi(\alpha)$ is concave in $\alpha \in [0, 1]$. In addition, $\Phi(0) = V_{(t,\omega)}(h')$ and $\Phi(1) = V_{(t,\omega)}(h)$. By continuity of $V_{(t,\omega)}(.)$ in $h \in \mathcal{H}$, $\Phi(\alpha)$ is continuous in α. By compactness of $[0, 1]$, for each $\alpha^* \in (0, 1)$, there exists a collection of all points $\alpha \in [0, 1]$ denoted by $A(\alpha^*)$ such that $\Phi(\alpha) \geq \Phi(\alpha^*)$. Clearly, $A(\alpha^*)$ is compact. Let α be any element of $A(\alpha^*)$, and let $\beta, \beta' \in [0, 1]\backslash A(\alpha^*)$ satisfy $\alpha > \beta, \beta'$ and $\Phi(\beta) > \Phi(\beta')$. Suppose that $\beta' \geq \beta$. Then by concavity of Φ, $\Phi(\beta') = \Phi(\gamma\alpha + (1 - \gamma)\beta) \geq \gamma\Phi(\alpha) + (1 - \gamma)\Phi(\beta) \geq \Phi(\beta)$, where $\gamma = \dfrac{\beta' - \beta}{\alpha - \beta}$, which contradicts $\Phi(\beta) > \Phi(\beta')$. Conversely, let $\beta, \beta' \in [0, 1]\backslash A(\alpha^*)$ satisfy $\alpha > \beta > \beta'$. Then by concavity of Φ, $\Phi(\beta) = \Phi(\gamma\alpha + (1 - \gamma)\beta') \geq \gamma\Phi(\alpha) + (1 - \gamma)\Phi(\beta') > \Phi(\beta')$, where $\gamma = \dfrac{\beta - \beta'}{\alpha - \beta'} > 0$. Hence, for any $\alpha \in A(\alpha^*)$ and for any $\beta, \beta' \in [0, 1]\backslash A(\alpha^*)$ such that $\alpha > \beta$ and $\alpha > \beta'$, $\beta > \beta'$ if and only if $\Phi(\beta) > \Phi(\beta')$.

Now, let $f \in \mathcal{D}$ satisfy $h \succ_{(t,\omega)} f \succ_{(t,\omega)} h'$. Then by A2U-C from Lemma C.1, there exists $\alpha^* \in (0, 1)$ such that $\alpha^* h \oplus (1 - \alpha^*)h' \succ_{(t,\omega)} f$. Let $A(f)$ be a non-empty collection of all $\alpha \in [0, 1]$ such that $\alpha h \oplus (1 - \alpha)h' \succ_{(t,\omega)} f$, and let $\widetilde{\alpha}$ be an infimum of $A(f)$. Suppose that $\widetilde{\alpha}h \oplus (1 - \widetilde{\alpha})h' \succ_{(t,\omega)} f$. By A2U-C from Lemma C.1, there exists $\gamma \in (0, 1)$ such that $\gamma(\widetilde{\alpha}h \oplus (1 - \widetilde{\alpha})h') \oplus (1 - \gamma)h' \succ_{(t,\omega)} f$. However, $\gamma\widetilde{\alpha} < \widetilde{\alpha}$, which contradicts the definition of $\widetilde{\alpha}$. Suppose that $\widetilde{\alpha}h \oplus (1 - \widetilde{\alpha})h' \prec_{(t,\omega)} f$. By A2U-C from Lemma C.1, there exists $\gamma \in (0, 1)$ such that $\gamma(\alpha^* h \oplus (1 - \alpha^*)h') \oplus (1 - \gamma)(\widetilde{\alpha}h \oplus (1 - \widetilde{\alpha})h') = \widehat{\alpha}h \oplus (1 - \widehat{\alpha})h' \prec_{(t,\omega)} f$, where $\widehat{\alpha} = \gamma\alpha^* + (1 - \gamma)\widetilde{\alpha}$. Clearly, $\alpha^* > \widehat{\alpha} > \widetilde{\alpha}$. Let $A(\alpha^*)$ be defined as in the above paragraph. Since $A(\alpha^*)$ is a collection of all $\alpha \in [0, 1]$ such that $\Phi(\alpha) \geq \Phi(\alpha^*)$, $\alpha^* h \oplus (1 - \alpha^*)h' \succ_{(t,\omega)} f$ implies that $\widehat{\alpha}$ and $\widetilde{\alpha}$ are not in $A(\alpha^*)$. Then by the result in the above paragraph, for all $\alpha \in [0, 1]$ such that $\widehat{\alpha} > \alpha$, $\Phi(\widehat{\alpha}) > \Phi(\alpha)$. In particular, $\Phi(\widehat{\alpha}) > \Phi(\widetilde{\alpha})$. However, by the definition of $\widetilde{\alpha}$, there exists some $\alpha \in [0, 1]$ such that $\widehat{\alpha} > \alpha > \widetilde{\alpha}$ and $\Phi(\alpha) > \Phi(\widehat{\alpha})$, which is impossible. Hence, $\widetilde{\alpha}h \oplus (1 - \widetilde{\alpha})h' \simeq_{(t,\omega)} f$.

For (ii) and (iii), we apply backward induction. At $T - 1$, all acts in $\mathcal{D}$ are in $\mathcal{D}_{T-1}$. In addition, for $f \in \mathcal{D}$, by Lemma C.1, there exist h and $h' \in \mathcal{H}$ such that $h \succeq_{(T-1,\omega)} f_\omega \succeq_{(T-1,\omega)} h'$ for all $\omega \in \Omega$. Then by A7U-F, A8U-IHA, and A10U-SMS, $h \succeq_{(T-1,\omega)} f \succeq_{(T-1,\omega)} h'$ for all $\omega \in \Omega$. By (i), for each $\omega \in \Omega$, there is $h_\omega \in \mathcal{H}$ such that $h_\omega \simeq_{(T-1,\omega)} f$. By A7U-F, we can let $h_\omega = h_{\omega'}$ if $\omega' \in \mathcal{F}_{T-1}(\omega)$. For $f \in \mathcal{D}$, let $g \in \mathcal{D}_{T-2}$ satisfy $g(\tau,\omega) = f(\tau,\omega)$ at each $(\tau,\omega) \notin \mathcal{T}_{T-1} \times \Omega$, $g(\tau,\omega) = h_\omega(\tau)$ for all $(\tau,\omega) \in \mathcal{T}_{T-1} \times \Omega$, where $h_\omega \simeq_{(T-1,\omega)} f$ and $h_\omega = h_{\omega'}$ if $\omega' \in \mathcal{F}_{T-1}(\omega)$. Clearly, $g \simeq_{(T-1,\omega)} f$ for all $\omega \in \Omega$. Then by A7U-F, A8U-IHA, and A9U-DC, $g \simeq_{(T-2,\omega)} f$ on $\omega \in \Omega$. Then by A7U-F, A8U-IHA, A10U-SMS, and Lemma C.1, there exist $h, h' \in \mathcal{H}$ such that $h \succeq_{(T-2,\omega)} f \succeq_{(T-2,\omega)} h'$ for all $\omega \in \Omega$. By (i), at every $\omega \in \Omega$, there is $h_\omega \in \mathcal{H}$ such that $h_\omega \simeq_{(T-2,\omega)} f$ with $h_\omega = h_{\omega'}$ if $\omega' \in \mathcal{F}_{T-2}(\omega)$. By repeatedly applying the same argument from $T-1$ to 0, under A7U to A10U, (ii) holds for all $f \in \mathcal{D}$ at all $(t,\omega) \in \mathcal{T}_{-T} \times \Omega$: There exists $g \in \mathcal{D}_t$ such that $g(\tau,\omega) = f(\tau,\omega)$ for all $(\tau,\omega) \notin \mathcal{T}_{t+1} \times \Omega$, $f \simeq_{(t,\omega)} g$, and $f \simeq_{(t+1,\omega')} g$ for all $\omega' \in \Omega$. In addition, the above construction demonstrates that for all $f \in \mathcal{D}$ at all $(t,\omega) \in \mathcal{T}_{-T} \times \Omega$, there exists $h \in \mathcal{H}$ such that $f \simeq_{(t,\omega)} h$. Finally, since $h \in \mathcal{H} \subset \mathcal{D}$ and $y \in \mathcal{L} \subset \mathcal{H}$, under Lemma C.1, we can repeat the same argument above and obtain that for all $h \in \mathcal{H}$ and for all $(t,\omega) \in \mathcal{T}_{-T} \times \Omega$, there exists $y \in \mathcal{L}$ such that $h \simeq_{(t,\omega)} y$. Hence, (iii) holds for all $f \in \mathcal{D}$ at all $(t,\omega) \in \mathcal{T}_{-T} \times \Omega$: There exist $h \in \mathcal{H}$ and $y \in \mathcal{L}$ such that $f \simeq_{(t,\omega)} h \simeq_{(t,\omega)} y$.■

Lemma C.3: *Given $U : Y \to R$ from Lemma C.1, under A1U to A10U, at all $(t,\omega) \in \mathcal{T} \times \Omega$, there exits a unique $J_{(t,\omega)} : \mathcal{D} \to R$ such that*

(i) For all $f, g \in \mathcal{D}$, $f \succeq_{(t,\omega)} g$ if and only if $J_{(t,\omega)}(f) \geq J_{(t,\omega)}(g)$.

(ii) For all $y \in \mathcal{L}$, $J_{(t,\omega)}(y) = U(y)$.

Proof. By (iii) of Lemma C.2, the result follows from Lemma 3.2 of Gilboa and Schmeidler (1989).■

We shall use a specific $U : Y \to R$ such that there exist lotteries, $\overline{y}$, $\widehat{y}$, and $\underline{y}$, for which $U(\overline{y}) = 1$, $U(\widehat{y}) = 0$, and $U(\underline{y}) = -1$ (by A3U-ND and A5U-SMT, such a function U and lotteries $\overline{y}$, $\widehat{y}$, and $\underline{y}$ exist). We denote by B_t the space of all bounded functions from Ω to $\mathbb{R}$ such that all $\nu \in B_t$ are measurable with respect to the σ-field generated by $\mathcal{F}_{t+1}$.[36] Denote $\nu(\omega)$ as a number assigned at ω for $\nu \in B_t$. For all $f \in \mathcal{D}_t$, define $V \circ f$ to be a function on Ω such that at each $\omega \in \Omega$, $(V \circ f)(\omega) = V_{(t,\omega)}(f_\omega)$. Since each $(V \circ f)(\omega)$ is bounded and Ω is finite, $V \circ f \in B_t$ for all $f \in \mathcal{D}_t$. Let $B_t(\mathcal{D}_t)$ be a subset of B_t such that for each element of $\nu \in B_t(\mathcal{D}_t)$, there exists some $f \in \mathcal{D}_t$ such that $\nu = (V_{(t,1)}(f_1),...,V_{(t,S)}(f_S))'$. Clearly, the defined map $V \circ (.)$ from $\mathcal{D}_t$ to $B_t(\mathcal{D}_t)$ is surjective. For $\gamma \in \mathbb{R}$, let $\gamma^* \in B_t$ be the constant function on Ω with value γ.

Lemma C.4: *Under A1U to A10U, there exists a functional $I_{(t,\omega)} : B_t \to R$ such that*

(i) For all $f \in \mathcal{D}_t$, $I_{(t,\omega)}(V \circ f) = J_{(t,\omega)}(f)$ (hence, $I_{(t,\omega)}(0^*) = 0$ and $I_{(t,\omega)}(1^*) = 1$).

[36] This is a Banach space with a sup norm.

(ii) $I_{(t,\omega)}$ *is homogeneous of degree 1.*

(iii) $I_{(t,\omega)}$ *is C-independent: For any* $a \in B_t$ *and* $\gamma \in R$, $I_{(t,\omega)}(a+\gamma^*) = I_{(t,\omega)}(a) + I_{(t,\omega)}(\gamma^*) = I_{(t,\omega)}(a) + \gamma$.

(iv) $I_{(t,\omega)}$ *is monotonic: For any* $a, b \in B_t$, $a(\omega) \geq b(\omega)$ *for all* $\omega \in \Omega$ *implies* $I_{(t,\omega)}(a) \geq I_{(t,\omega)}(b)$. *In addition, if for some* $\omega' \in \mathcal{F}_{t+1}(\omega)$, $a(\omega) > b(\omega)$, *then* $I_{(t,\omega)}(a) > I_{(t,\omega)}(b)$.

Proof. First, by Lemma C.3 and A10U-SMS, we can define $I_{(t,\omega)}$ on $B_t(\mathcal{D}_t)$ by $I_{(t,\omega)}(V \circ f) = J_{(t,\omega)}(f)$ for all $f \in \mathcal{D}_t$. For (ii) and (iii), by Lemma C.1, $V_{(t,\omega)}(\alpha f_\omega + (1-\alpha)l) = \alpha V_{(t,\omega)}(f_\omega) + (1-\alpha)V_{(t,\omega)}(l)$ for all $l \in \mathcal{L}$. Then given Lemmas C.1, C.2, and C.3, we can apply Lemma 3.3 in Gilboa and Schmeidler (1989). In particular, the second equality of (iii) follows from (i) and (ii). Also, by (i) and A10U-SMS, (iv) holds for $B_t(\mathcal{D}_t)$. Then by (ii), we extend $I_{(t,\omega)}$ to all of B_t.■

Let $y \in Y$ satisfy $y \succ_{(t,\omega)} \widehat{y}$, and let $h \in \mathcal{H}$ satisfy $h^t(\omega) = \widehat{y}$ for all $\omega \in \Omega$ and $h^\tau(\omega) = y$ for all $\tau \in \mathcal{T}\backslash\{t\}$ and for all $\omega \in \Omega$. Then by Lemma C.1 and Proposition 1, $V_{(t,\omega)}(y) > V_{(t,\omega)}(h) > V_{(t,\omega)}(\widehat{y})$. Let $\varepsilon = V_{(t,\omega)}(h)$. Then again by Lemma C.1 and Proposition 1, for all $y' \in Y$ such that $y \succeq_{(t,\omega)} y' \succeq_{(t,\omega)} \widehat{y}$, $V_{(t,\omega)}(h) \geq V_{(t,\omega)}(h') \geq V_{(t,\omega)}(\widehat{y})$, where $h' \in \mathcal{H}$ assigns $h'^t(\omega) = \widehat{y}$ for all $\omega \in \Omega$ and $h'^\tau(\omega) = y'$ for all $\tau \in \mathcal{T}\backslash\{t\}$ and for all $\omega \in \Omega$. Let $\mathcal{D}_t^R([\widehat{y}, y])$ be a collection of all acts $f \in \mathcal{D}_t$ such that (i) $f(t,\omega) = \widehat{y}$ for all $\omega \in \Omega$ and (ii) for all $\omega' \in \Omega$, $y \succeq_{(t,\omega)} f(\tau,\omega') = f(\tau',\omega') \succeq_{(t,\omega)} \widehat{y}$ for all $\tau, \tau' \in \mathcal{T}_{t+1}$. Note that by Lemma C.1, for all $y', y'' \in Y$ with $y' \succeq_{(t,\omega)} y''$ and for all $\alpha \in [0,1]$, $y' \succeq_{(t,\omega)} \alpha y' + (1-\alpha)y'' \succeq_{(t,\omega)} y''$. Then for all $f, g \in \mathcal{D}_t^R([\widehat{y}, y])$ and for all $\alpha \in (0,1)$, $\alpha f \oplus (1-\alpha)g \in \mathcal{D}_t^R([\widehat{y}, y])$.

Let $B_t([0,\varepsilon])$ be a subset of B_t such that for $a \in B_t([0,\varepsilon])$, $a(\omega) \in [0,\varepsilon]$ at each $\omega \in \Omega$. Then by construction and the continuity of $V_{(t,\omega)}(.)$ in $h \in \mathcal{H}$, there exists a surjective map from $\mathcal{D}_t^R([\widehat{y}, y])$ to $B_t([0,\varepsilon])$ defined by $V \circ (.) : \mathcal{D}_t^R([\widehat{y}, y]) \to B_t([0,\varepsilon])$ (this result is implied by the following: For all $\alpha \in [0,1]$, $y \succeq_{(t,\omega)} \alpha y + (1-\alpha)\widehat{y} \succeq_{(t,\omega)} \widehat{y}$ and $V_{(t,\omega)}(\alpha h + (1-\alpha)\widehat{y}) = \alpha V_{(t,\omega)}(h) + (1-\alpha)V_{(t,\omega)}(\widehat{y})$). In addition, by Lemma C.4, for all $f, g \in \mathcal{D}_t^R([\widehat{y}, y])$, $f \succeq_{(t,\omega)} g$ if and only if $I_{(t,\omega)}(V \circ f) \geq I_{(t,\omega)}(V \circ g)$. Let $\succeq_{(t,\omega)}^V$ be a preference relation on $B_t([0,\varepsilon])$ induced by a preference relation on $\mathcal{D}_t^R([\widehat{y}, y])$ under $I_{(t,\omega)}$: For $\upsilon, \upsilon' \in B_t([0,\varepsilon])$, $\upsilon \succeq_{(t,\omega)}^V \upsilon'$ if and only if $I_{t,\omega}(V \circ f) \geq I_{t,\omega}(V \circ g)$, where $V \circ f = \upsilon$ and $V \circ g = \upsilon'$ for some $f, g \in \mathcal{D}_t^R([\widehat{y}, y])$.

Lemma C.5: *Under A1U to A11U, for all* $(t,\omega) \in \mathcal{T} \times \Omega$, *a preference relations* $\succeq_{(t,\omega)}^V$ *on* $B_t([0,\varepsilon])$ *satisfy the following:*

(i) *For all* $\upsilon, \nu \in B_t([0,\varepsilon])$ *and for all* $\alpha \in [0,1]$, $\alpha\upsilon + (1-\alpha)\nu \in B_t([0,\varepsilon])$.

(ii) *For all* $\upsilon, \nu \in B_t([0,\varepsilon])$, $\upsilon \succeq_{(t,\omega)}^V \nu$ *or* $\nu \succeq_{(t,\omega)}^V \upsilon$.

(iii) *For all* $\upsilon, \nu, \mu \in B_t([0,\varepsilon])$, $\upsilon \succ_{(t,\omega)}^V \nu$ *and* $\nu \succ_{(t,\omega)}^V \mu$ *imply that* $\upsilon \succ_{(t,\omega)}^V \mu$.

(iv) *There exist* $\upsilon, \nu \in B_t([0,\varepsilon])$ *such that* $\upsilon \succ_{(t,\omega)}^V \nu$ *(in particular,* $\varepsilon^* \succ_{(t,\omega)}^V 0^*$*).*

(v) For all $\upsilon, \nu, \mu \in B_t([0,\varepsilon])$ and for all $\alpha \in (0,1)$, $\upsilon \succ^V_{(t,\omega)} \nu$ if and only if $\alpha\upsilon + (1-\alpha)\mu \succ^V_{(t,\omega)} \alpha\nu + (1-\alpha)\mu$.

(vi) For all $\upsilon, \nu, \mu \in B_t([0,\varepsilon])$ with $\upsilon \succ^V_{(t,\omega)} \nu \succ^V_{(t,\omega)} \mu$, there exist $\alpha, \beta \in (0,1)$ such that $\alpha\upsilon + (1-\alpha)\nu \succ^V_{(t,\omega)} \mu$ and $\mu \succ^V_{(t,\omega)} \beta\upsilon + (1-\beta)\mu$.

Moreover, for all $(t,\omega) \in \mathcal{T}_{-T} \times \Omega$,

(vii) For all $\upsilon \in B_t([0,\varepsilon])$, let $\nu, \mu \in B_t([0,\varepsilon])$ satisfy $\upsilon(\omega'') = \nu(\omega'') = \mu(\omega'')$ for all $\omega'' \notin \mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$, and let $\nu', \mu' \in B_t([0,\varepsilon])$ satisfy $\upsilon(\omega''') = \nu(\omega''') = \mu(\omega''')$ for all $\omega''' \notin \mathcal{F}_{t+1}(\widetilde{\omega}) \subseteq \mathcal{F}_t(\omega)$. Assume that $\nu(\omega'') = \nu(\omega''')$ and $\mu(\omega'') = \mu(\omega''')$ for all $\omega'' \in \mathcal{F}_{t+1}(\omega')$ and for all $\omega''' \in \mathcal{F}_{t+1}(\widetilde{\omega})$. If $\nu \succ^V_{(t,\omega)} \mu$, then $\nu' \succ^V_{(t,\omega)} \mu'$.

Proof. (i) is obvious. (ii) and (iii) are implied by the definition of $\succeq^V_{(t,\omega)}$: For $\upsilon, \upsilon' \in B_t([0,\varepsilon])$, $\upsilon \succeq^V_{(t,\omega)} \upsilon'$ if and only if $I_{(t,\omega)}(\upsilon) \geq I_{(t,\omega)}(\upsilon')$. Given this definition of $\succeq^V_{(t,\omega)}$, (iv) here follows from (iv) of Lemma C.4. For (vii), note that $a \in B_t([0,\varepsilon])$ is measurable with respect to the σ-field generated by $\mathcal{F}_{t+1}$. Then (vii) follows from the definition of $\succeq^V_{(t,\omega)}$ and (iv) of Lemma C.4.

As for (v), consider the following: For all $f, g \in \mathcal{D}^R_t([\widehat{y}, y])$ and for all $\theta \in (0,1)$, at each $\omega' \in \Omega$,

$$
\begin{aligned}
&V_{(t,\omega)}(\theta f_{\omega'} \oplus (1-\theta) g_{\omega'}) \\
&= (1-\delta^*_{t+1})U(\theta f(t,\omega') + (1-\theta)g(t,\omega')) + \delta^*_{t+1} V_{(t+1,\omega)}(\theta f_{\omega'} \oplus (1-\theta)g_{\omega'}) \\
&= (1-\delta^*_{t+1})[\theta U(f(t,\omega')) + (1-\theta)U(g(t,\omega'))] \\
&\quad + \delta^*_{t+1}[U_{(t+1,\omega)}(\theta f(t+1,\omega') + (1-\theta)g(t+1,\omega'))] \\
&= (1-\delta^*_{t+1})[\theta U(f(t,\omega')) + (1-\theta)U(g(t,\omega'))] \\
&\quad + \delta^*_{t+1}[\theta U_{(t+1,\omega)}(f(t+1,\omega')) + (1-\theta)U_{(t+1,\omega)}(g(t+1,\omega'))] \\
&= \theta[(1-\delta^*_{t+1})U(f(t,\omega')) + \delta^*_{t+1}U_{(t+1,\omega)}(f(t+1,\omega'))] \\
&\quad + (1-\theta)[(1-\delta^*_{t+1})U(g(t,\omega')) + \delta^*_{t+1}U_{(t+1,\omega)}(g(t+1,\omega'))] \\
&= \theta[(1-\delta^*_{t+1})U(f(t,\omega')) + \delta^*_{t+1}V_{(t+1,\omega)}(f_{\omega'})] + (1-\theta)[(1-\delta^*_{t+1})U(g(t,\omega')) + \delta^*_{t+1}V_{(t+1,\omega)}(g_{\omega'})],
\end{aligned}
$$

where δ^*_{t+1} is the effective selection of the discount factor for $\theta f_{\omega'} \oplus (1-\theta)g_{\omega'}$. Since $U(\theta f(t,\omega') + (1-\theta)g(t,\omega')) \leq U_{(t+1,\omega)}(\theta f(t+1,\omega') \oplus (1-\theta)g(t+1,\omega'))$, $\alpha_{t+1} \in \{\delta^*_{t+1}\}$. In addition, both a set of effective selection of discount factors for $f_{\omega'}$ and a set for $g_{\omega'}$ contain α_{t+1}. Then $V_{(t,\omega)}(\theta f_{\omega'} \oplus (1-\theta)g_{\omega'}) = \theta V_{(t,\omega)}(f_{\omega'}) + (1-\theta)V_{(t,\omega)}(g_{\omega'})$. This result implies that for all $f, g \in \mathcal{D}^R_t([\widehat{y}, y])$ and for all $\theta \in (0,1)$, $V \circ (\theta f \oplus (1-\alpha)\theta) = \theta V \circ (f) + (1-\theta) V \circ (g)$. Also, by Lemma C.1 and Proposition 1, for all $y' \in \mathcal{L}$ and for all $\eta \in (0,1)$, $V_{(t,\omega)}(\eta(\theta f_{\omega'} \oplus (1-\theta)g_{\omega'}) \oplus (1-\eta)y') = \eta V_{(t,\omega)}(\theta f_{\omega'} \oplus (1-\theta)g_{\omega'}) + (1-\eta)U_{(t,\omega)}(y')$.

We now consider A11U-IA. First, for each $f \in \mathcal{D}^R_t([\widehat{y}, y])$, at each $\omega' \in \Omega$, by Lemma C.1 and Proposition 1, $V_{(t,\omega)}(\alpha f_{\omega'} \oplus (1-\alpha)y') = \alpha V_{(t,\omega)}(f_{\omega'}) + (1-\alpha)U_{(t,\omega)}(y')$ for all $y' \in \mathcal{L}$ and for all $\alpha \in (0,1)$. Then, for all $f, g, h \in \mathcal{D}^R_t([\widehat{y}, y])$, at each $\omega' \in \Omega$, there exist $y', y'', y''' \in \mathcal{L}$ such that (i) $y \succeq_{(t,\omega)} y', y'', y''' \succeq_{(t,\omega)} \widehat{y}$ and (ii) $\alpha f_{\omega'} \oplus (1-\alpha)y' \simeq_{(t,\omega)} \beta g_{\omega'} \oplus (1-\beta)y'' \simeq_{(t,\omega)} \gamma h_{\omega'} \oplus (1-\gamma)y'''$ for some $\alpha, \beta, \gamma \in (0,1)$. Second, for (ii) of A11U-IA, $\varphi[\alpha f_{\omega'} \oplus (1-\alpha)y'] \oplus (1-\varphi)[\gamma h_{\omega'} \oplus (1-\gamma)y'''] =$

$\widetilde{\eta}(\widetilde{\theta} f_{\omega'} \oplus (1-\widetilde{\theta})h_{\omega'}) \oplus (1-\widetilde{\eta})\widetilde{y}$, where $\widetilde{\theta} = \frac{\varphi\alpha}{\widetilde{\eta}}$, $\widetilde{\eta} = \varphi\alpha + (1-\varphi)\gamma$, and $\widetilde{y} = \frac{\varphi(1-\alpha)}{(1-\widetilde{\eta})}y' \oplus \frac{(1-\varphi)(1-\gamma)}{(1-\widetilde{\eta})}y'''$. Given the result in the above paragraph, $V_{(t,\omega)}(\widetilde{\eta}(\widetilde{\theta} f_{\omega'} \oplus (1-\widetilde{\theta})h_{\omega'}) \oplus (1-\widetilde{\eta})\widetilde{y})$ $= \widetilde{\eta}V_{(t,\omega)}(\widetilde{\theta} f_{\omega'} \oplus (1-\widetilde{\theta})h_{\omega'}) + (1-\widetilde{\eta})U_{(t,\omega)}(\widetilde{y}) = \widetilde{\eta}\{\widetilde{\theta}V_{(t,\omega)}(f_{\omega'}) + (1-\widetilde{\theta})V_{(t,\omega)}(h_{\omega'})\} + (1-\widetilde{\eta})U_{(t,\omega)}(\widetilde{y})$. The same result holds for $\psi[\beta g_{\omega'} \oplus (1-\beta)y''] \oplus (1-\psi)[\gamma h_{\omega'} \oplus (1-\gamma)y''']$. Therefore, at each $\omega' \in \Omega$, for all $\varphi, \psi \in [0,1]$, $\varphi[\alpha f_{\omega'} \oplus (1-\alpha)y'] \oplus (1-\varphi)[\gamma h_{\omega'} \oplus (1-\gamma)y'''] \simeq_{(t,\omega)} \psi[\beta g_{\omega'} \oplus (1-\beta)y''] \oplus (1-\psi)[\gamma h_{\omega'} \oplus (1-\gamma)y''']$, where $y', y'', y''' \in \mathcal{L}$ and $\alpha, \beta, \gamma \in (0,1)$ are defined above. Since any $h \in \mathcal{D}_t^R([\widehat{y}, y])$ can be used for this operation, by A11U-IA, for all $f, g, h \in \mathcal{D}_t^R([\widehat{y}, y])$, for all $\theta \in (0,1)$, $f \succ_{(t,\omega)} g$ if and only if $\theta f \oplus (1-\theta)h \succ_{(t,\omega)} \theta g \oplus (1-\theta)h$. Then for all $\theta \in (0,1)$, $I_{(t,\omega)}(V \circ (f)) > I_{(t,\omega)}(V \circ (g))$ if and only if $I_{(t,\omega)}(V \circ (\theta f \oplus (1-\theta)h)) = I_{(t,\omega)}(\theta V \circ (f) + (1-\theta)V \circ (h))$ $> I_{(t,\omega)}(V \circ (\theta g \oplus (1-\theta)h)) = I_{(t,\omega)}(\theta V \circ (g) + (1-\theta)V \circ (h))$. Since $\theta f \oplus (1-\theta)h, \theta g \oplus (1-\theta)h$ $\in \mathcal{D}_t^R([\widehat{y}, y])$, this result implies that for all $V \circ (f)$, $V \circ (g)$, $V \circ (h) \in B_t([0,\varepsilon])$ and for all $\theta \in (0,1)$, $V \circ (f) \succ^V_{(t,\omega)} V \circ (g)$ if and only if $\theta V \circ (f) + (1-\theta)V \circ (h) \succ^V_{(t,\omega)} \theta V \circ (g) + (1-\theta)V \circ (h)$. Note that by (i), $\theta V \circ (f) + (1-\theta)V \circ (h)$ and $\theta V \circ (g) + (1-\theta)V \circ (h)$ are in $B_t([0,\varepsilon])$; hence, (v) holds.

To show (vi), by A2U-C from Lemma C.1, for all $f, g, h \in \mathcal{D}_t^R([\widehat{y}, y])$ with $f \succ_{(t,\omega)} g \succ_{(t,\omega)} h$, there exist $\alpha, \beta \in (0,1)$ such that $\alpha f \oplus (1-\alpha)h \succ_{(t,\omega)} g \succ_{(t,\omega)} \beta f \oplus (1-\beta)h$. Then the calculation above implies that $I_{t,\omega}(V \circ (\alpha f \oplus (1-\alpha)h)) = I_{t,\omega}(\alpha V \circ (f) + (1-\alpha)V \circ (h)) > I_{t,\omega}(V \circ (g)) > I_{t,\omega}(V \circ (\beta f \oplus (1-\beta)h)) = I_{t,\omega}(\beta V \circ (f) + (1-\beta)V \circ (h))$. Since $\alpha f \oplus (1-\alpha)h, \beta f \oplus (1-\beta)h \in \mathcal{D}_t^R([\widehat{y}, y])$, this result implies that for all $V \circ (f), V \circ (g), V \circ (h) \in B_t([0,\varepsilon])$ with $V \circ (f) \succ^V_{(t,\omega)} V \circ (g) \succ^V_{(t,\omega)} V \circ (h)$, there exist $\alpha, \beta \in (0,1)$ such that $\alpha V \circ (f) + (1-\alpha)V \circ (h) \succ^V_{(t,\omega)} V \circ (g) \succ^V_{(t,\omega)} \beta V \circ (f) + (1-\beta)V \circ (h)$. Note that by (i), $\alpha V \circ (f) + (1-\alpha)V \circ (h)$ and $\beta V \circ (f) + (1-\beta)V \circ (h)$ are in $B_t([0,\varepsilon])$; hence, (vi) holds.■

Lemma C.6: *Under A1U to 11U, at each $(t,\omega) \in \mathcal{T}_{-T} \times \Omega$, there exists a strictly positive weight $\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))$ for all $\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$ with $\sum_{\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)} \mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega)) = 1$ such that for all $f, g \in \mathcal{D}_t$, $f \succeq_{(t,\omega)} g$ if and only if $V_{(t,\omega)}(f) \geq V_{(t,\omega)}(g)$, where $V_{(t,\omega)}(f) \equiv \sum_{\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)} V_{(t,\omega')}(f_{\omega'})\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))$. Moreover, $\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))$ is unique and $V_{(t,\omega)}$ on $\mathcal{H}$ is defined by Lemma C.1.*

Proof. We follow the proof in Kreps (1988). First, by (i) of Lemma C.5, $B_t([0,\varepsilon])$ with $\psi_\alpha^t(\upsilon, \upsilon') = \alpha\upsilon + (1-\alpha)\upsilon'$ forms a mixture space. Then given (ii), (iii), (v), and (vi) of Lemma C.5, by the mixture space theorem, there exists a functional $F_{(t,\omega)} : B_t([0,\varepsilon]) \to R$ such that the following holds:

(27) For all $\upsilon, \upsilon' \in B_t([0,\varepsilon])$, $\upsilon \succeq^V_{(t,\omega)} \upsilon'$ if and only if $F_{(t,\omega)}(\upsilon) \geq F_{(t,\omega)}(\upsilon')$,

(28) $F_{(t,\omega)}(\psi_\alpha^t(\upsilon, \upsilon')) = F_{(t,\omega)}(\alpha\upsilon + (1-\alpha)\upsilon') = \alpha F_{(t,\omega)}(\upsilon) + (1-\alpha)F_{(t,\omega)}(\upsilon')$.

(29) Moreover, $F_{(t,\omega)}$ is unique up to a positive affine transformation.

Let $\overline{\upsilon}$ be some fixed element in $B_t([0,\varepsilon])$, and let $\upsilon(\mathcal{F}_{t+1}(\omega))$ be an element of $B_t([0,\varepsilon])$ such that $\upsilon(\omega') = \overline{\upsilon}(\omega')$ for all $\omega' \notin \mathcal{F}_{t+1}(\omega)$ and $\upsilon(\omega') = \upsilon \in [0, \varepsilon]$ for all $\omega' \in \mathcal{F}_{t+1}(\omega)$. Again, note that $a \in B_t([0, \varepsilon])$ is measurable only with respect to the σ-field generated by $\mathcal{F}_{t+1}$. In addition, let N_t be the number of events in $\mathcal{F}_{t+1}$. Then by the proof in Kreps (1988) at P.104-105, $F_{(t,\omega)}$ can be rewritten as follows:

$$(30) \qquad F_{(t,\omega)}(\upsilon) = \textstyle\sum_{\mathcal{F}_{t+1}(\omega')\subseteq\Omega} F_{\mathcal{F}_{t+1}(\omega')}(\{\upsilon(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}),$$

$$(31) \qquad F_{\mathcal{F}_{t+1}(\omega')}(\{\upsilon(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}) = F_{(t,\omega)}(\upsilon(\mathcal{F}_{t+1}(\omega'))) - \frac{N_t - 1}{N_t} F_{(t,\omega)}(\overline{\upsilon}).$$

Now, let $\upsilon^* \in B_t([0, \varepsilon])$ satisfy $\upsilon^*(\omega) = \upsilon^*(\omega')$ for all $\omega' \in \Omega$. Note that $\{\upsilon(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}$ is a point on the line between $\{0^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}$ and $\{\varepsilon^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}$. Let $\upsilon(\omega'') = \dfrac{\upsilon(\omega'')}{\varepsilon}\varepsilon^*(\omega'') + (1 - \dfrac{\upsilon(\omega'')}{\varepsilon})0^*(\omega'')$ for all $\omega'' \in \mathcal{F}_{t+1}(\omega')$. Pick one of ω'' in $\mathcal{F}_{t+1}(\omega')$ and define $\dfrac{\upsilon(\omega''|\mathcal{F}_{t+1}(\omega'))}{\varepsilon} \equiv \dfrac{\upsilon(\omega'')}{\varepsilon}$. Then, from (28), (30), and (31),

$$\begin{aligned}
&F_{\mathcal{F}_{t+1}(\omega')}(\{\upsilon(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}) \\
&= \frac{\upsilon(\omega''|\mathcal{F}_{t+1}(\omega'))}{\varepsilon} F_{\mathcal{F}_{t+1}(\omega')}(\{\varepsilon^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}) \\
&\quad + (1 - \frac{\upsilon(\omega''|\mathcal{F}_{t+1}(\omega'))}{\varepsilon}) F_{\mathcal{F}_{t+1}(\omega')}(\{0^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}) \\
&= F_{\mathcal{F}_{t+1}(\omega')}(\{0^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}) \\
&\quad + \frac{\upsilon(\omega''|\mathcal{F}_{t+1}(\omega'))}{\varepsilon} [F_{\mathcal{F}_{t+1}(\omega')}(\{\varepsilon^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}) - F_{\mathcal{F}_{t+1}(\omega')}(\{0^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')})].
\end{aligned}$$

Let $a_{\mathcal{F}_{t+1}(\omega')} \equiv F_{\mathcal{F}_{t+1}(\omega')}(\{\varepsilon^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}) - F_{\mathcal{F}_{t+1}(\omega')}(\{0^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')})$. Then (30) is rewritten as follows:

$$\begin{aligned}
(32) \qquad & F_{(t,\omega)}(\upsilon) \\
&= \textstyle\sum_{\mathcal{F}_{t+1}(\omega')\subseteq\Omega} \dfrac{\upsilon(\omega''|\mathcal{F}_{t+1}(\omega'))}{\varepsilon} a_{\mathcal{F}_{t+1}(\omega')} + \textstyle\sum_{\mathcal{F}_{t+1}(\omega')\subseteq\Omega} F_{\mathcal{F}_{t+1}(\omega')}(\{0^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')}).
\end{aligned}$$

Let $\widetilde{F}_{(t,\omega)}(\upsilon) = \varepsilon F_{(t,\omega)}(\upsilon) - \varepsilon \sum_{\mathcal{F}_{t+1}(\omega')\subseteq\Omega} F_{\mathcal{F}_{t+1}(\omega')}(\{0^*(\omega'')\}_{\omega''\in\mathcal{F}_{t+1}(\omega')})$ (by construction, $\varepsilon > 0$). Since $F_{(t,\omega)}$ is unique up to a positive affine transformation, $\widetilde{F}_{(t,\omega)}(\upsilon)$ can replace $F_{(t,\omega)}(\upsilon)$. Then (32) becomes the following:

$$(33) \qquad \widetilde{F}_{(t,\omega)}(\upsilon) = \textstyle\sum_{\mathcal{F}_{t+1}(\omega')\subseteq\Omega} \upsilon(\omega''|\mathcal{F}_{t+1}(\omega')) a_{\mathcal{F}_{t+1}(\omega')}.$$

By (vii) of Lemma C.5, all $a_{\mathcal{F}_{t+1}(\omega')}$ with $\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$ must have the same sign. Then by (iv) of Lemma C.5, $a_{\mathcal{F}_{t+1}(\omega')} > 0$ for all $\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$. In addition, under A8U-IHA, all events in $\mathcal{F}_{t+1}$ with $\mathcal{F}_{t+1}(\omega') \nsubseteq \mathcal{F}_t(\omega)$ are null (i.e., $a_{\mathcal{F}_{t+1}(\omega')} = 0$) at (t, ω). Then (33) is rewritten as

$$(34) \qquad \widetilde{F}_{(t,\omega)}(\upsilon) = \textstyle\sum_{\mathcal{F}_{t+1}(\omega')\subseteq\mathcal{F}_t(\omega)} \upsilon(\omega''|\mathcal{F}_{t+1}(\omega'))\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))A, \text{ where}$$

$$(35) \qquad A = \textstyle\sum_{\mathcal{F}_{t+1}(\omega')\subseteq\mathcal{F}_t(\omega)} a_{\mathcal{F}_{t+1}(\omega')}, \text{ and}$$

(36) $\quad \mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega)) = \dfrac{a_{\mathcal{F}_{t+1}(\omega')}}{A}$ and $\sum_{\mathcal{F}_{t+1}(\omega')\subseteq\mathcal{F}_t(\omega)} \mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega)) = 1.$

Note that $\upsilon(\omega''|\mathcal{F}_{t+1}(\omega')) = \upsilon(\omega''')$ for all $\omega''' \in \mathcal{F}_{t+1}(\omega')$. Thus, we let $\upsilon(\omega''|\mathcal{F}_{t+1}(\omega')) \equiv \upsilon(\omega')$ for $\mathcal{F}_{t+1}(\omega')$. Again, under the uniqueness of $\widetilde{F}_{(t,\omega)}$ up to a positive affine transformation, we define the following representation:

(37) $\quad V^V_{(t,\omega)}(\upsilon) \equiv \dfrac{\widetilde{F}_{(t,\omega)}(\upsilon)}{A} = \sum_{\mathcal{F}_{t+1}(\omega')\subseteq\mathcal{F}_t(\omega)} \upsilon(\omega')\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega)).$

This proves the existence of a subjective prior μ over $\mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$. Since $V^V_{(t,\omega)}(\upsilon)$ is unique up to a positive affine transformation, given (iv) and (vii) of Lemma C.5, μ is uniquely determined. Note that $\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))$ is a weight for $\mathcal{F}_{t+1}(\omega')$, not for each $\omega'' \in \mathcal{F}_{t+1}(\omega')$.

Now, by the definition of $\succeq^V_{(t,\omega)}$, $\upsilon \succeq^V_{(t,\omega)} \upsilon'$ if and only if $I_{(t,\omega)}(\upsilon) \geq I_{(t,\omega)}(\upsilon')$, which implies that $I_{(t,\omega)}(\upsilon) \geq I_{(t,\omega)}(\upsilon')$ if and only if $V^V_{(t,\omega)}(\upsilon) \geq V^V_{(t,\omega)}(\upsilon')$. Hence, there exists a strictly increasing ϕ, where $I_{(t,\omega)}(\upsilon) = \phi(V^V_{(t,\omega)}(\upsilon))$. In addition, $I_{(t,\omega)}(0^*) = J_{(t,\omega)}(\widehat{y}) = 0 = V^V_{(t,\omega)}(0^*) = \phi(0)$ (because $\widehat{y} \in \mathcal{D}^R_t([\widehat{y}, y]))$ and $I_{(t,\omega)}(\varepsilon^*) = \phi(V^V_{(t,\omega)}(\varepsilon^*)) = \phi(\varepsilon)$. Also, by the linearity of $V^V_{(t,\omega)}(\upsilon)$ in υ (i.e., (37)) with $V^V_{(t,\omega)}(0^*) = 0$, $V^V_{(t,\omega)}(\upsilon)$ is homogeneous of degree one. Then by (ii) of Lemma C.4 (homogeneity of $I_{(t,\omega)}$), for $\alpha > 0$ such that $\upsilon, \alpha\upsilon \in B_t([0,\varepsilon])$, $I_{(t,\omega)}(\upsilon) = \phi(V^V_{(t,\omega)}(\upsilon)) = \phi(a)$ and $I_{(t,\omega)}(\alpha\upsilon) = \phi(V^V_{(t,\omega)}(\alpha\upsilon)) = \phi(\alpha V^V_{(t,\omega)}(\upsilon)) = \phi(\alpha a) = \alpha I_{(t,\omega)}(\upsilon) = \alpha\phi(a)$. So ϕ is homogeneous of degree one in $[0,\varepsilon]$. Also, since ϕ has one argument in $[0,\varepsilon]$, $\phi(a) = ka$ for some $k > 0$ in $[0,\varepsilon]$ (clearly if $\dfrac{\phi(a)}{a} = k$ and $\dfrac{\phi(a')}{a'} = k'$, then ϕ violates homogeneity). Then $I_{(t,\omega)}(\upsilon) = kV^V_{(t,\omega)}(\upsilon)$. By (iii) of Lemma C.2, there exists $y \in Y$ with $U(y) = \varepsilon$ and a constant act $l \in \mathcal{L}$ that assigns y for all (t,ω) such that $I_{(t,\omega)}(\varepsilon^*) = J_{(t,\omega)}(y) = U(y) = \varepsilon$. This result implies that $I_{(t,\omega)}(\varepsilon^*) = kV^V_{(t,\omega)}(\varepsilon^*) = k\varepsilon = \varepsilon$. Hence, $k = 1$. Therefore, $I_{(t,\omega)}(\upsilon) = V^V_{(t,\omega)}(\upsilon)$ for all $\upsilon \in B_t([0,\varepsilon])$. By (ii) of Lemma C.4 (homogeneity of $I_{(t,\omega)}$) and homogeneity of $V^V_{(t,\omega)}$, $I_{(t,\omega)}(\upsilon) = V^V_{(t,\omega)}(\upsilon)$ for all $\upsilon \in B_t([0,\infty))$. Let γ^* be a constant function such that $I_{(t,\omega)}(\varepsilon^* + \gamma^*) < 0$ (such γ^* exists by construction). Then by (iii) of Lemma C.4 (C-independence of $I_{(t,\omega)}$), for any $\upsilon \in B_t$ and $\gamma \in R$, $I_{(t,\omega)}(\upsilon + \gamma^*) = I_{(t,\omega)}(\upsilon) + I_{(t,\omega)}(\gamma^*) = I_{(t,\omega)}(\upsilon) + \gamma$. In addition, by (37), for any $\upsilon \in B_t$ and $\gamma \in R$, $V^V_{(t,\omega)}(\upsilon + \gamma^*) = V^V_{(t,\omega)}(\upsilon) + V^V_{(t,\omega)}(\gamma^*) = V^V_{(t,\omega)}(\upsilon) + \gamma$. Hence, $I_{(t,\omega)}(\upsilon) = V^V_{(t,\omega)}(\upsilon)$ for all $\upsilon \in B_t([0,\varepsilon])$ implies that $I_{(t,\omega)}(\upsilon) = V^V_{(t,\omega)}(\upsilon)$ for all $\upsilon \in B_t([\gamma, \varepsilon + \gamma])$ with $\varepsilon + \gamma < 0$. By (ii) of Lemma C.4 (homogeneity of $I_{(t,\omega)}$) and homogeneity of $V^V_{(t,\omega)}$, $I_{(t,\omega)}(\upsilon) = V^V_{(t,\omega)}(\upsilon)$ for all $\upsilon \in B_t((-\infty, 0))$. Therefore, $I_{(t,\omega)}(\upsilon) = V^V_{(t,\omega)}(\upsilon)$ for all $\upsilon \in B_t$.

Finally, since $I_{(t,\omega)}$ on B_t represents $\succeq_{(t,\omega)}$ on $\mathcal{D}_t$, under $I_{(t,\omega)}(\upsilon) = V^V_{(t,\omega)}(\upsilon)$ for all $\upsilon \in B_t$, the following represents a preference relation on $\mathcal{D}_t$:

(38) $\quad$ For all $f, g \in \mathcal{D}_t$, $f \succeq_{(t,\omega)} g$ if and only if $V^D_{(t,\omega)}(f) \geq V^D_{(t,\omega)}(g)$,

(39) $\quad V^D_{(t,\omega)}(f) = \sum_{\mathcal{F}_{t+1}(\omega')\subseteq\mathcal{F}_t(\omega)} V_{(t,\omega)}(f_{\omega'})\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega)).$

By Lemma C.1, $V_{(t,\omega)} = V_{(t,\omega')}$ for all $\omega, \omega' \in \Omega$. Then $V^D_{(t,\omega)}(f)$ is rewritten as

(40) $V^D_{(t,\omega)}(f) = \sum_{\mathcal{F}_{t+1}(\omega')\subseteq\mathcal{F}_t(\omega)} V_{(t,\omega')}(f_{\omega'})\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))$

Note that under $V^D_{(t,\omega)}(.)$, $V^D_{(t,\omega)}(h) = V_{(t,\omega)}(h)$ for all $h \in \mathcal{H}$. Hence, we define $V_{(t,\omega)}(f)$ as $V^D_{(t,\omega)}(f)$ for all $f \in \mathcal{D}_t$.■

The conclusion of the proof of sufficiency in Proposition 2:

We prove Proposition 2 by backward induction. At $(T-1,\omega)$ for all $\omega \in \Omega$, Lemma C.6 is equivalent to Proposition 2. At $T-2$, for all $f \in \mathcal{D}$, there exists $d^f \in \mathcal{D}_{T-2}$ that satisfies the conditions stated in (ii) of Lemma C.2. In particular, $d^f \simeq_{(T-2,\omega)} f$ for all $\omega \in \Omega$. By (ii) of A1U-WO (transitivity) from Lemma C.1, for all $f \in \mathcal{D}$, we define $V_{(T-2,\omega)}(f) \equiv V_{(T-2,\omega)}(d^f)$, where $V_{(T-2,\omega)}(d^f)$ is defined in Lemma C.6. Then by Lemmas C.1 and C.6,

(41) $V_{(T-2,\omega)}(f) \equiv V_{(T-2,\omega)}(d^f) = \sum_{\mathcal{F}_{T-1}(\omega')\subseteq\mathcal{F}_{T-2}(\omega)} V_{(T-2,\omega')}(d^f_{\omega'})\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))$,

(42) $V_{(T-2,\omega')}(d^f_{\omega'}) = \text{Min}_{\delta_{T-1,\omega'}\in[\alpha_{T-1},\beta_{T-1}]}[(1-\delta_{T-1,\omega'})U(f(T-2,\omega'))+\delta_{T-1,\omega'}V_{(T-1,\omega')}(d^f_{\omega'})]$.

Since $d^f_{\omega'} \simeq_{(T-1,\omega')} f$, $V_{(T-1,\omega')}(f) = V_{(T-1,\omega')}(d^f_{\omega'})$. Note that $V_{(T-1,\omega')}(f)$ has already been defined above. Then,

(43) $V_{(T-2,\omega)}(f)$
$= \sum_{\mathcal{F}_{T-1}(\omega')\subseteq\mathcal{F}_{T-2}(\omega)}\{[\text{Min}_{\delta_{T-1,\omega'}\in[\alpha_{T-1},\beta_{T-1}]}[(1-\delta_{T-1,\omega'})U(f(T-2,\omega'))$
$+ \delta_{T-1,\omega'}V_{(T-1,\omega')}(f)\ ]]\mu(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega))\}$.

By repeatedly applying Lemmas C.1 and C.6 from $T-1$ to 0 under (ii) of A1U-WO (transitivity) from Lemma C.1,

(44) For all $f, g \in \mathcal{D}$, $f \succeq_{(t,\omega)} g$ if and only if $V_{(t,\omega)}(f) \geq V_{(t,\omega)}(g)$,

(45) $V_{(t,\omega)}(f)$
$\equiv \sum_{\mathcal{F}_{t+1}(\omega')\subseteq\mathcal{F}_t(\omega)}\{[\text{Min}_{\delta_{t+1,\omega'}\in[\alpha_{t+1},\beta_{t+1}]}[(1-\delta_{t+1,\omega'})U(f(t,\omega'))$
$+\delta_{t+1,\omega'}V_{(t+1,\omega')}(f)]\mu(\mathcal{F}_{t+1}(\omega')\ |\mathcal{F}_t(\omega))\}$,

(46) $V_{(T,\omega)}(f) \equiv U(f(T,\omega))$ and $U(f(t,\omega)) \equiv \text{E}_{f(t,\omega)}[u(x)]$ over $x \in X$,

(47) μ is defined by Lemma C.6,

(48) u is unique up to a positive affine transformation.

Note that U is defined as a function that represents a preference relation on $y \in \mathcal{L}$ at time 0, and it is shared by all $V_{(t,\omega)}$. The uniqueness of u up to a positive affine transformation follows from Lemma C.1.

Next, we define $p(\omega)$ as follows:

(49) $p(\omega) = \mu(\mathcal{F}_T(\omega)|\mathcal{F}_{T-1}(\omega))...\mu(\mathcal{F}_2(\omega)|\mathcal{F}_1(\omega))\mu(\mathcal{F}_1(\omega)|\mathcal{F}_0(\omega))$.

Then, $p(\omega)$ satisfies the following:

$$\text{(50)} \qquad \textstyle\sum_{\omega\in\Omega} p(\omega) = 1 \text{ and } p(\omega) > 0 \text{ for all } \omega \in \Omega,$$

$$\text{(51)} \qquad p(\mathcal{F}_{t+1}(\omega')|\mathcal{F}_t(\omega)) = \frac{\sum_{\omega''\in\mathcal{F}_{t+1}(\omega')} p(\omega'')}{\sum_{\omega''\in\mathcal{F}_t(\omega)} p(\omega'')}.$$

This definition satisfies the condition for $p(\omega)$ to be a probability measure with conditional probability as in (51). By Lemma C.6, this measure is unique. Let $\mathrm{E}[.|\mathcal{F}_t(\omega)]$ be a conditional expectation operator on $\mathcal{F}_t(\omega)$. Then we can use $\mathrm{E}[.|\mathcal{F}_t(\omega)]$ to write $V_{(t,\omega)}(f)$ as

$$\text{(52)} \qquad V_{(t,\omega)}(f) = \mathrm{E}[\min_{\delta_{t+1,\omega'}\in[\alpha_{t+1},\beta_{t+1}]}[(1-\delta_{t+1,\omega'})U(f(t,\omega')) + \delta_{t+1,\omega'}V_{(t+1,\omega')}(f)]|\mathcal{F}_t(\omega)],$$

$$\text{(53)} \qquad \delta^*_{t+1,\omega''} = \delta^*_{t+1,\omega'''} \text{ for all } \omega'',\omega''' \in \mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega),$$

$$\text{(54)} \qquad \delta^*_{t+1,\omega'} \in \mathrm{argmin}_{\delta_{t+1,\omega'}\in[\alpha_{t+1},\beta_{t+1}]}[(1-\delta_{t+1,\omega'})U(f(t,\omega')) + \delta_{t+1,\omega'}V_{(t+1,\omega')}(f)].$$

In order for (52) to be consistent with (45), the representation assigns identical discount factors for all $\omega'',\omega''' \in \mathcal{F}_{t+1}(\omega') \subseteq \mathcal{F}_t(\omega)$. This condition is necessary when $U(f(t,\omega')) = V_{(t+1,\omega')}(f)$, where a set of effective selection of discount factors for f_ω consist of multiple $\delta^*_{t+1,\omega'}$. Otherwise, $\delta^*_{t+1,\omega'}$ is uniquely determined on $\mathcal{F}_{t+1}(\omega')$. However, even when $U(f(t,\omega')) = V_{(t+1,\omega')}(f)$, assigning different discount factors on each $\omega'',\omega''' \in \mathcal{F}_{t+1}(\omega')$ will lead to the same value of $V_{(t,\omega)}(f)$.■

Necessity: A1U-WO, A4U-CI, A7U-F, A8U-IHA, and A9U-DC are immediate. Since $f \in \mathcal{D}$ is a collection of lotteries and each lottery is represented by the expected utility form in $V_{(t,\omega)}$, $V_{(t,\omega)}$ is continuous in $U \circ f$. Then A2U-C (α close to 1 and β close to 0) is satisfied. A3U-ND follows by the uniqueness of $\Delta_{(0,\omega)}$ (see Gilboa and Schmeidler (1989)). For certainty acts, the representation satisfies A6U-TVA.

Assume that for $f,g \in \mathcal{D}$, for all $\tau \in \mathcal{T}$, $f(\tau,\omega) \succeq_{(0,\omega)} g(\tau,\omega)$ for all $\omega \in \Omega$. Note that under fixed discount factors, $V_{(0,\omega)}$ on $f \in \mathcal{D}$ is a weighted summation of $U(f(t,\omega))$, where weights are based on a subjective prior and discount factors. Let $V^\delta_{(0,\omega)}$ be (52) without minimization under a fixed set of discount factors δ (δ a collection of all fixed $\delta_{t,\omega}$ for all $(t,\omega) \in \mathcal{T}_1 \times \Omega$, and a collection of all δ is compact). Then under any choice of discount factors δ, $V^\delta_{(0,\omega)}(f) \geq V^\delta_{(0,\omega)}(g)$. Since $V_{(0,\omega)}(f)$ is attained as a minimum of $V^\delta_{(0,\omega)}(f)$, this result implies that $V_{(0,\omega)}(f) \geq V_{(0,\omega)}(g)$. In particular, if for some $\tau \in \mathcal{T}$, $f(\tau,\omega) \succ_{(0,\omega)} g(\tau,\omega)$ for all $\omega \in \Omega$, then $V^\delta_{(0,\omega)}(f) > V^\delta_{(0,\omega)}(g)$ for any δ. Then $V_{(0,\omega)}(f) > V_{(0,\omega)}(g)$, and A5U-SMT is satisfied.

Finally, for $\mathcal{D}_t$, the representation is equivalent to a state-by-state application of multiple discount factors of $V_{(t,\omega)}$ weighted by a subjective prior. Then A10U-SMS holds. In addition, under $V_{(t,\omega)}$, for any $f,g,h \in \mathcal{D}_t$ and at each $\omega' \in \Omega$, we can find $l,l',l'' \in \mathcal{L}$ and $\alpha,\beta,\gamma \in (0,1)$ that satisfy (i) of A11U-IA. Moreover, (ii) of A11U-IA considers acts in $\mathcal{D}_t$ such that at each $\omega' \in \Omega$ and for all $\varphi \in [0,1]$, $\varphi[\alpha f_{\omega'} \oplus (1-\alpha)l_{\omega'}] \oplus (1-\varphi)[\gamma h_{\omega'} \oplus (1-\gamma)l''_{\omega'}] \simeq_{(t,\omega)} \alpha f_{\omega'} \oplus (1-\alpha)l_{\omega'} \simeq_{(t,\omega)}$

$\gamma h_{\omega'} \oplus (1-\gamma)l''_{\omega'}$. By Gilboa and Schmeidler (1989), $V_{(t,\omega)}$ is superadditive among $\mathcal{H}$, i.e., for all $h, h' \in \mathcal{H}$, $V_{(t,\omega)}(\eta h \oplus (1-\eta)h') \geq \eta V_{(t,\omega)}(h) + (1-\eta)V_{(t,\omega)}(h')$ for $\eta \in [0,1]$. Then,

$$(55) \qquad V_{(t,\omega)}(\varphi[\alpha f_{\omega'} \oplus (1-\alpha)l_{\omega'}] \oplus (1-\varphi)[\gamma h_{\omega'} \oplus (1-\gamma)l''_{\omega'}]) \\ \geq \varphi\alpha V_{(t,\omega)}(f_{\omega'}) + (1-\varphi)\gamma V_{(t,\omega)}(h_{\omega'}) + \varphi(1-\alpha)V_{(t,\omega)}(l_{\omega'}) + (1-\varphi)(1-\gamma)V_{(t,\omega)}(l''_{\omega'}),$$

and

$$(56) \qquad V_{(t,\omega)}(\varphi[\alpha f_{\omega'} \oplus (1-\alpha)l_{\omega'}] \oplus (1-\varphi)[\gamma h_{\omega'} \oplus (1-\gamma)l''_{\omega'}]) \\ = V_{(t,\omega)}(\alpha f_{\omega'} \oplus (1-\alpha)l_{\omega'}) = \alpha V_{(t,\omega)}(f_{\omega'}) + (1-\alpha)V_{(t,\omega)}(l_{\omega'}) \\ = V_{(t,\omega)}(\gamma h_{\omega'} \oplus (1-\gamma)l''_{\omega'}) = \gamma V_{(t,\omega)}(h_{\omega'}) + (1-\gamma)V_{(t,\omega)}(l''_{\omega'}).$$

Note that $V_{(t,\omega)}$ on $f \in \mathcal{H}$ is a weighted summation of $U_{(t,\omega)}(h(\tau,\omega))$ under some fixed discount factors. Therefore, in order for $f, h \in \mathcal{D}_t$ to satisfy (56) (i.e., (ii) of A11U-IA), the sets of effective selection of discount factors for $f_{\omega'}$ and $h_{\omega'}$ must have a non-empty intersection, and the set for $\varphi\alpha f_{\omega'} \oplus (1-\varphi)\gamma h_{\omega'}$ only includes elements in this intersection. This result and superadditivity of $V_{(t,\omega)}$ on $\mathcal{H}$ implies that for all $\theta \in (0,1)$, $V_{(t,\omega)}(\theta f_{\omega'} \oplus (1-\theta)h_{\omega'})$ attains the minimum value when we use any such element (i.e., $V_{(t,\omega)}(\theta f_{\omega'} \oplus (1-\theta)h_{\omega'}) = \theta V_{(t,\omega)}(f_{\omega'}) + (1-\theta)V_{(t,\omega)}(h_{\omega'})$). Hence, among $f, g, h \in \mathcal{D}_t$ that satisfy (i) and (ii) of A11U-IA, the intersection of the sets of effective selection of discount factors for $f_{\omega'}$ and $h_{\omega'}$ (also, that for $g_{\omega'}$ and $h_{\omega'}$) are non-empty for all $\omega' \in \Omega$, and for all $\theta \in (0,1)$, $V_{(t,\omega)}(f) > V_{(t,\omega)}(g)$ if and only if $V_{(t,\omega)}(\theta f \oplus (1-\theta)h) = \theta V_{(t,\omega)}(f) + (1-\theta)V_{(t,\omega)}(h) > V_{(t,\omega)}(\theta g \oplus (1-\theta)h) = \theta V_{(t,\omega)}(g) + (1-\theta)V_{(t,\omega)}(h)$. Then, A11U-IA holds.■

References

1. Anscombe, F.J., and R.J. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34(1), 199-205.

2. Bekaert, G., R. Hodrick, and D. Marshall (1997): "The Implications of First-order Risk Aversion for Asset Market Risk Premiums," *Journal of Monetary Economics*, 40(1), 3-39.

3. Campbell, J., and J. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205-251.

4. Campbell, J., and J. Cochrane (2000): "Explaining the Poor Performance of Consumption-based Asset Pricing Models," *Journal of Finance*, 55(5), 2863-2878.

5. Constantinides, G. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98(3), 519-543.

6. Duffie, D., and L. Epstein (1992): "Stochastic Differential Utility," *Econometrica*, 60(2), 353-394.

7. Dumas, B., R. Uppal, and T. Wang (2000): "Efficient Intertemporal Allocations With Recursive Utility," *Journal of Economic Theory*, 93(2), 240-59.

8. Eichberger, J., and D. Kelsey (1996): "Uncertainty Aversion and Dynamic Consistency," *International Economic Review*, 37(3), 625-640.

9. Epstein, L. (1983): "Stationary Cardinal Utility and Optimal Growth under Uncertainty," *Journal of Economic Theory*, 31(1), 133-152.

10. Epstein, L. (1987): "The Global Stability of Efficient Intertemporal Allocations," *Econometrica*, 55(2), 329-355.

11. Epstein, L., and M. Schneider (2003): "Recursive Multiple-Priors," *Journal of Economic Theory*, 113(1), 1-31.

12. Epstein, L., and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 63(3), 283-322.

13. Epstein, L., and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: a Theoretical Framework," *Econometrica*, 57(4), 937-969.

14. Epstein, L., and S. Zin (1990): "'First-order' Risk Aversion and the Equity Premium Puzzle," *Journal of Monetary Economics*, 26(3), 387-407.

15. Fisher, I. (1930): *The Theory of Interest*, Macmillan.

16. Frederic, S., G. Loewenstein, and T. O'Donogue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40(2), 351-401.

17. Geoffard, P. (1996): "Discounting and Optimizing: Capital Accumulation Problems as Variational Minmax Problems," *Journal of Economic Theory*, 69(1) 53-70.

18. Gilboa, I. (1989): "Expectation and Variation in Multi-Period Decisions," *Econometrica,* 57(5), 1153-1169.

19. Gilboa, I., and D. Schmeidler (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18(2), 141-153.

20. Grant S., A. Kajii, and B. Polak (2000): "Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models," *Econometrica*, 68(2), 425-434.

21. Gul, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59(3), 667-687.

22. Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263-291.

23. Koopmans, T. (1960): "Stationary Ordinal Utility and Impatience," *Econometrica*, 28(2), 287-309.

24. Kreps, D. (1988): *Note on the Theory of Choice*, Western Press.

25. Kreps, D, and E. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185-200.

26. Loewenstein G., and D. Prelec (1993): "Preferences for Sequences of Outcomes," *Psychological Review*, 100(1), 91-108.

27. Loewenstein G., and R. Thaler (1989): "Anomalies: Intertemporal Choice," *Journal of Economic Perspectives*, 3(4), 181-193.

28. Machina, M. (1989): "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature*, 27(4), 1622-1668.

29. Samuelson, P. (1937): "A Note on Measurement of Utility," *Review of Economic Studies*, 4(2), 155-161.

30. Sarin B., and P. Wakker (1998): "Dynamic Choice and NonExpected Utility," *Journal of Risk and Uncertainty*, 17(2), 87-119.

31. Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57(3), 571-587.

32. Segal, U., and A. Spivak (1990): "First Order versus Second Order Risk Aversion," *Journal of Economic Theory*, 51(1), 111-125.

33. Segal, U., and A. Spivak (1997): "First-Order Risk Aversion and Non-Differentiability," *Economic Theory*, 9(1), 179-183.

34. Shalev, J. (1997): "Loss Aversion in a Multiple-Period Model," *Mathematical Social Sciences,* 33(3), 203-226.

35. Shi, S., and L. Epstein (1993): "Habit and Time Preference," *International Economic Review*, 34(1), 61-84.

36. Skiadas, C. (1998): "Recursive Utility and Preferences for Information," *Economic Theory*, 12(2), 293-312.

37. Sundaresan, S. (1989): "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies*, 2(1), 73-89.

38. Uzawa, H. (1968): "Time Preference, the Consumption Function, and Optimum Asset Holdings," in J.N. Wolfe, ed., *Capital and Growth: Papers in Honour of Sir John Hicks* (Chicago: Aldine, 1968).

39. Wakai, K. (2002): "Linking Behavioral Economics, Axiomatic Decision Theory and General Equilibrium Theory," Ph.D. Dissertation, Yale University.

40. Wakai, K. (2003): "Dynamic Consistency and Multiple Priors," SUNY at Buffalo.

41. Yaari, M. (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 95-115.