Supplementary Appendix to Dynamic Random Utility

Mira Frick, Ryota Iijima, and Tomasz Strzalecki

F Proof of Theorem 0

F.1 Preliminaries

In this section we prove Theorem 0, which extends the characterizations of REU representations in Gul and Pesendorfer (2006) and Ahn and Sarver (2013) to allow for an arbitrary separable metric space X of outcomes. Refer to section 2.1 of the main text for all relevant notation and terminology. Throughout, we fix some $y^* \in X$ and let $\tilde{\mathbb{R}}^X = \{0\} \times \mathbb{R}^{X \smallsetminus \{y^*\}}$ denote the set of utility functions u in $\mathbb{R}^X$ that are normalized by $u(y^*) = 0$.

We first define the static analog of S-based representations introduced in Appendix A:

Definition 13. An *S-based REU representation* of ρ is a tuple $(S, \mu, \{U_s, \tau_s\}_{s\in S})$ such that

(i). S is a finite state space and μ is a probability measure on S such that $\text{supp}(\mu) = S$

(ii). for each $s \in S$, the utility $U_s \in \tilde{\mathbb{R}}^X$ is nonconstant and $U_s \not\approx U_{s'}$ for $s \neq s'$

(iii). for each $s \in S$, the tie-breaking rule τ_s is a proper finitely-additive probability measure on $\tilde{\mathbb{R}}^X$ endowed with the Borel σ-algebra

(iv). for all $p \in \Delta(X)$ and $A \in \mathcal{A}$,

$$\rho(p; A) = \sum_{s\in S} \mu(s)\tau_s(p, A),$$

where $\tau_s(p, A) := \tau_s(\{u \in \tilde{\mathbb{R}}^X : p \in M(M(A, U_s), u)\})$.

Analogous arguments as for the DREU part of Proposition A.1 yield the equivalence of S-based REU representations and static REU representations.

Proposition F.1. Let ρ be a stochastic choice rule on $\mathcal{A}$. Then ρ admits an REU representation if and only if it admits an S-based REU representation.

Proof. Analogous to Proposition A.1 (i). ■

Thus, Theorem 0 is equivalent to the following result, which we prove throughout the rest of this section.

Theorem F.1. The stochastic choice rule ρ on $\mathcal{A}$ satisfies Axiom 0 if and only if ρ admits an S-based REU representation $(S, \mu, \{U_s, \tau_s\}_{s\in S})$.

Note that because X may be infinite, continuity of each U_s in the representation is not directly implied by linearity. However, the following additional axiom ensures this. As in Section 3.3, let $\mathcal{A}^*$ denote the collection of *menus without ties*, i.e., the set of all $A \in \mathcal{A}$ such that for any $p \in A$ and any sequences $p^n \to^m p$ and $B^n \to^m A \smallsetminus \{p\}$, we have $\lim_{n\to\infty} \rho(p^n; B^n \cup \{p^n\}) = \rho(p; A)$.

Axiom F.1 (Continuity)**.** $\rho : \mathcal{A}^* \to \Delta(\Delta(X))$ is continuous.

Here $\mathcal{A}$ is endowed with the Hausdorff topology induced by the Prokhorov metric π on $\Delta(X)$, and $\mathcal{A}^*$ with the relative topology. We have the following proposition.

Proposition F.2. Suppose ρ admits an S-based REU representation $(S, \mu, \{U_s, \tau_s\}_{s\in S})$. Then ρ satisfies Axiom F.1 if and only if each utility U_s is continuous.

Proof. See Section F.5. ■

Additional notation: For any $Y \subseteq X$, let $\mathcal{A}(Y) := \{A \in \mathcal{A} : \forall p \in A, \mathrm{supp}(p) \subseteq Y\} \subseteq \mathcal{A}$ denote the space of all menus consisting only of lotteries with support in Y. Note that for each $A \in \mathcal{A}$, there is a finite Y such that $A \in \mathcal{A}(Y)$. We denote by ρ^Y the restriction of ρ to $\mathcal{A}(Y)$, which can be seen as a map from $\mathcal{A}(Y)$ to $\Delta(\Delta(Y))$. If $y^* \in Y$, we write $\tilde{\mathbb{R}}^Y := \{0\} \times \mathbb{R}^{Y \smallsetminus \{y^*\}}$.

For any $A \in \mathcal{A}(Y)$ and $p \in \Delta(X)$, let $N_Y(A,p) := \{u \in \tilde{\mathbb{R}}^Y : p \in M(A,u)\}$ and let $N^+_Y(A,p) := \{u \in \tilde{\mathbb{R}}^Y : \{p\} = M(A,u)\}$. Note that $N_Y(\{p\},p) = N^+_Y(\{p\},p) = \tilde{\mathbb{R}}^Y$ and that $N_Y(A,p) = N^+_Y(A,p) = \varnothing$ if $p \notin A$. Let $\mathcal{N}(Y) := \{N_Y(A,p) : A \in \mathcal{A}(Y) \text{ and } p \in \Delta(X)\}$, $\mathcal{N}^+(Y) := \{N^+_Y(A,p) : A \in \mathcal{A}(Y) \text{ and } p \in \Delta(X)\}$.

We will consider both the Borel σ-algebra on $\tilde{\mathbb{R}}^Y$ and its subalgebra $\mathcal{F}(Y)$ that is generated by $\mathcal{N}(Y) \cup \mathcal{N}^+(Y)$. A finitely-additive probability measure ν^Y on either of these algebras is called *proper* if $\nu^Y(N_Y(A,p)) = \nu^Y(N^+_Y(A,p))$ for any $A \in \mathcal{A}(Y)$ and $p \in \Delta(X)$. Whenever $Y = X$, we omit Y from the description of $N_Y(A,p)$, $N^+_Y(A,p)$, $\mathcal{N}(Y)$, $\mathcal{N}^+(Y)$, and $\mathcal{F}(Y)$.

F.2 Proof of Theorem F.1: Sufficiency

F.2.1 Outline

The proof proceeds as follows:

(i). In section F.2.2, we use conditions (i)–(iv) of Axiom 0 and Theorem 2 in Gul and Pesendorfer (2006) to construct, for each *finite* $Y \subseteq X$, a proper finitely-additive probability measure ν^Y on $\mathcal{F}(Y)$ representing ρ^Y, in the sense that $\rho^Y(p; A) = \nu^Y(N_Y(A,p))$ for all A, p. Given the fact that each ρ^Y is derived from the same ρ, it is easy to check that the family $\{\mathcal{F}(Y), \nu^Y\}$ is Kolmogorov consistent. We can then find a proper finitely-additive probability measure ν on $\mathcal{F}$ extending all the ν^Y (and hence representing ρ).

(ii). The support of ν is defined by

$$\mathrm{supp}(\nu) := \left(\bigcup\{V \in \mathcal{F} : V \text{ is open and } \nu(V) = 0\}\right)^c.$$

In section F.2.3, we use part (v) of Axiom 0 to show that $\mathrm{supp}\,\nu$ is finite (up to positive affine transformation of utilities) and contains at least one non-constant utility function. While Axiom 0 (v) is similar to the finiteness axiom in Ahn and Sarver (2013), this step requires more work in our setting. A key technical challenge is that unlike in Ahn and Sarver, it is not clear in our infinite outcome space setting how to normalize utilities to ensure that $N(A,p)$-sets are compact. Compact sets C have the useful property (used repeatedly by Ahn and Sarver) that if $C \cap \mathrm{supp}\,\nu = \varnothing$, then $\nu(C) = 0$. Lemma F.5 exploits the geometry of $N(A,p)$-sets to show that this property continues to hold for $N(A,p)$-sets in our setting, even though they are not compact.

(iii). In section F.2.4, we proceed in a similar way to the proof of Theorem S3 in Ahn and Sarver (2013)(again using Lemma F.5 to circumvent technical difficulties). Letting $S := \{s_1, \ldots, s_L\}$ denote the equivalence classes of nonconstant utilities in $\operatorname{supp} \nu$, we find separating neighborhoods $B_s \in \mathcal{F}$ of each s such that $\nu(B_s) > 0$. We then define $\mu(s) = \nu(B_s)$ and $\tau_s(V) = \frac{\nu(V \cap B_s)}{\nu(B_s)}$ and show that this yields an S-based REU representation of ρ.

F.2.2 Construction of ν

In this section, we construct a proper finitely-additive probability measure ν on $\mathcal{F}$ that represents ρ, i.e., such that for all $A \in \mathcal{A}$ and $p \in A$, we have

$$\rho(p; A) = \nu(N(A, p)) = \nu(N^+(A, p))).$$

First consider any finite $Y \subseteq X$ with $y^* \in Y$. By Axiom 0 (i)–(iv) (Regularity, Linearity, Extremeness, and Mixture Continuity), Theorem 2 in Gul and Pesendorfer (2006) ensures that there is a proper finitely-additive probability measure ν^Y on $\mathcal{F}^Y$ such that

$$\rho^Y(p; A) = \nu^Y(N_Y(A, p)) = \nu^Y(N_Y^+(A, p))$$

for all $A \in \mathcal{A}(Y)$ and $p \in A$.

Claim 4. For any finite $Y' \supseteq Y \ni y^*$, $(\nu^{Y'}, \mathcal{F}(Y'))$ and $(\nu^Y, \mathcal{F}(Y))$ are Kolmogorov consistent, i.e., for any $E \in \mathcal{F}(Y)$, we have

$$\nu^{Y'}(E \times \mathbb{R}^{Y' \smallsetminus Y}) = \nu^Y(E). \tag{24}$$

Proof. To see this, note first that the LHS of (24) is well-defined, since $E \times \mathbb{R}^{Y' \smallsetminus Y} \in \mathcal{F}^{Y'}$ by Lemma F.4 (iv). Note next that by Lemma F.4 (iii), E is of the form $\bigcup_{i=1}^n N_Y(A_i, p_i) \cap N_Y^+(B_i, q_i)$ for some finite n and $A_i, B_i \in \mathcal{A}(Y)$. Let E' be obtained from E by replacing each $N_Y(A_i, p_i)$ with $N_Y^+(A_i, p_i)$. By Lemma F.4 (ii), $E' = \bigcup_{i=1}^n N_Y^+(C_i, r_i)$ for some family $\{C_i\} \subseteq \mathcal{A}(Y)$. Moreover, since both ν^Y and $\nu^{Y'}$ are proper, we have that $\nu^Y(E) = \nu^Y(E')$ and $\nu^{Y'}(E \times \mathbb{R}^{Y' \smallsetminus Y}) = \nu^{Y'}(E' \times \mathbb{R}^{Y' \smallsetminus Y})$. Hence, it suffices to prove that $\nu^{Y'}(E' \times \mathbb{R}^{Y' \smallsetminus Y}) = \nu^Y(E')$. For this, it is enough to show that for any collection of sets $N_1^+, \ldots, N_n^+ \in \mathcal{N}^+(Y) := \{N^+(A, p) : A \in \mathcal{A}(Y)\}$, we have $\nu^Y(\bigcup_{i=1}^n N_i^+) = \nu^{Y'}(\bigcup_{i=1}^n N_i^+ \times \mathbb{R}^{Y' \smallsetminus Y})$. We prove this by induction. For the base case, note that for any $N^+(A, p) \in \mathcal{N}^+(Y)$, we have

$$\nu^{Y'}(N^+(A, p)) \times \mathbb{R}^{Y' \smallsetminus Y}) = \rho^{Y'}(p, A) := \rho(p; A) =: \rho^Y(p; A) = \nu^Y(N^+(A, p)).$$

Suppose next that the claim is true whenever $m < n$. Then

$$\begin{aligned}
&\nu^Y(\bigcup_{i=1}^{m+1} N_i^+) = \nu^Y(\bigcup_{i=1}^{m} N_i^+) + \nu^Y(N_{m+1}^+) - \nu^Y(\bigcup_{i=1}^{m}(N_i^+ \cap N_{m+1}^+)) = \\
&\nu^{Y'}(\bigcup_{i=1}^{m} N_i^+ \times \mathbb{R}^{Y' \smallsetminus Y}) + \nu^{Y'}(N_{m+1}^+ \times \mathbb{R}^{Y' \smallsetminus Y}) - \nu^{Y'}(\bigcup_{i=1}^{m}(N_i^+ \cap N_{m+1}^+) \times \mathbb{R}^{Y' \smallsetminus Y}) = \\
&\nu^{Y'}(\bigcup_{i=1}^{m+1} N_i^+ \times \mathbb{R}^{Y' \smallsetminus Y}),
\end{aligned}$$

where the second equality follows from the inductive hypothesis and the fact that $N_i^+ \cap N_{m+1}^+ \in \mathcal{N}^+(Y)$ by Lemma F.4 (ii). ■

Now define ν on $\mathcal{F}$ by setting $\nu(E) := \nu^Y(\text{proj}_{\tilde{\mathbb{R}}^Y} E)$ for any finite $Y \ni y^*$ such that $E = \text{proj}_{\tilde{\mathbb{R}}^Y} E \times \mathbb{R}^{X \setminus Y}$ and $\text{proj}_{\tilde{\mathbb{R}}^Y} E \in \mathcal{F}^Y$. By Lemma F.4 (iv) such a Y exists. Moreover, given Kolmogorov consistency of the family $\{\nu^Y\}_{Y \subseteq X}$, this is well-defined. Finally, it is immediate that ν is a proper finitely-additive probability measure and that ν represents ρ.

F.2.3 Finiteness of $\operatorname{supp} \nu$

The support of a finitely-additive probability measure ν is defined by

$$\operatorname{supp}(\nu) := \left(\bigcup \{V \in \mathcal{F} : V \text{ is open and } \nu(V) = 0\} \right)^c.$$

The next lemma invokes Axiom 0 (v) (Finiteness) to show that the support of ν constructed in the previous section contains finitely many equivalence classes of utility functions and contains at least one nonconstant function. We use 0 to denote the unique constant utility function in $\tilde{\mathbb{R}}^X$.

Lemma F.1. Let K be as in the statement of the Finiteness Axiom and let $Pref(\Delta(X))$ denote the set of all preferences over $\Delta(X)$. Then

$$\#\{\succsim \in Pref(\Delta(X)) : \succsim \text{ is represented by some } u \in \operatorname{supp}(\nu) \setminus \{0\}\} = L,$$

where $1 \leq L \leq K$.

Proof. We first show that $L \leq K$. If not, then we can find utilities $\{u_1, ..., u_{K+1}\} \subseteq \operatorname{supp}(\nu)$ such that each u_i is non-constant over X and $u_i \not\approx u_j$ for all $i \neq j$. By Lemma E.2, we can find a menu $A = \{p^i : i = 1, \ldots, K+1\} \in \mathcal{A}$ such that $u_i \in N^+(A, p^i)$ for each i. Take any $B \subseteq A$ with $|B| \leq K$. Then $p^i \notin B$ for some i.

Fix any sequences $p_n^i \to^m p^i$ and $B_n \to^m B$. By definition, this means that there exists $r \in \Delta(X)$ and $\alpha_n \to 0$ such that $p_n^i = \alpha_n r + (1-\alpha)p^i$ for all n, and that for each $q \in B$ there exists $B_q \in \mathcal{A}$ and $\beta_n(q) \to 0$ such that $B_n = \bigcup_{q \in B} (\beta_n(q) B_q + (1 - \beta_n(q))\{q\})$ for all n. Now, B and each B_q are finite, and u_i is linear with $u_i \cdot p^i > u_i \cdot q$ for all $q \in B$. Hence, there is N such that for all $n \geq N$, $u_i \cdot p_n^i > u \cdot q_n$ for all $q_n \in B_n$. Thus, $u_i \in N^+(\{p_n^i\} \cup B_n, p_n^i)$ for all $n \geq N$. But since $u_i \in \operatorname{supp}(\nu)$ and $N^+(\{p_n^i\} \cup B_n, p_n^i)$ is an open set in $\mathcal{F}$, the definition of $\operatorname{supp}(\nu)$ then implies that $\nu(N^+(\{p_n^i\} \cup B_n, p_n^i)) > 0$ for all $n \geq N$. But then $\rho(p_n^i; \{p_n^i\} \cup B_n) = \nu(N^+(\{p_n^i\} \cup B_n, p_n^i)) > 0$ for all $n \geq N$, contradicting Finiteness.

Next we show that $L \geq 1$. Indeed, if $L = 0$, then for any $A \in \mathcal{A}$ with $|A| \geq 2$ and for any $p \in A$, we have $(N(p, A) \setminus \{0\}) \cap \operatorname{supp} \nu = \varnothing$. By Lemma F.5 below, this implies that $\nu(N^+(p, A)) = 0$ for any $p \in A$. But since ν represents ρ, $\rho(p; A) = \nu(N^+(p, A))$ for any $p \in A$, so we have $\sum_{p \in A} \rho(p; A) = 0$, which is a contradiction. ∎

F.2.4 Constructing the REU Representation

Let $\succsim_1, \ldots, \succsim_L$ denote all the preferences represented by some non-constant utility in $\operatorname{supp}(\nu)$, where by Lemma F.1 we know that L is finite and $L \geq 1$. For each $i = 1, \ldots, L$, pick some $u_i \in \operatorname{supp} \nu$ representing $\succsim_i$. For any $u \in \tilde{\mathbb{R}}^X$, let $[u] := \{u' \in \tilde{\mathbb{R}}^X : u' \approx u\}$. By Lemma E.2, we can find $A := \{p_1, \ldots, p_L\} \in \mathcal{A}$ such that $u_i \in N^+(A, p_i)$ for all $i = 1, \ldots, L$. Let $B_{u_i} := N^+(A, p_i)$ for all i. By construction, $[u_i] \subseteq B_{u_i}$ and $B_{u_i} \cap B_{u_j} = \varnothing$ for $j \neq i$. Moreover, by the definition of $\operatorname{supp}(\nu)$, we have $\nu(B_{u_i}) > 0$ for each i, since $B_{u_i} \in \mathcal{F}$ is open and $u_i \in B_{u_i} \cap \operatorname{supp}(\nu) \neq \varnothing$.

Let $S := \{u_1, \ldots, u_L\}$ and define the function $\mu : S \to [0, 1]$ by

$$\mu(s) = \nu(B_s) \text{ for each } s \in S.$$

We claim that μ defines a full-support probability measure on S. For this it remains to show that $\sum_s \mu(s) = 1$. Since $\sum_s \mu(s) = \sum_s \nu(B_s) = \nu(\bigcup_{s\in S} B_s)$, it suffices to prove the following claim:

Lemma F.2. $\nu(\bigcup_{s\in S} B_s) = 1$.

Proof. It suffices to prove that $\nu(\tilde{\mathbb{R}}^X \setminus \bigcup_{s\in S} B_s) = 0$. Note that $\tilde{\mathbb{R}}^X = \bigcup_{i=1}^L N(A, p_i)$, since $A = \{p_i, \dots, p_L\}$. Thus,

$$\tilde{\mathbb{R}}^X \setminus \bigcup_{s\in S} B_s \subseteq \bigcup_{i=1}^L (N(A, p_i) \setminus N^+(A, p_i)).$$

By finite additivity of ν, this implies that

$$\nu(\tilde{\mathbb{R}}^X \setminus \bigcup_{s\in S} B_s) \leq \sum_{i=1}^L \nu(N(A, p_i) \setminus N^+(A, p_i)) = 0,$$

where the last inequality follows from properness of ν. ∎

Next, we define a set function $\tau_s : \mathcal{F} \to \mathbb{R}_+$ for each $s \in S$ by setting

$$\tau_s(V) := \frac{\nu(V \cap B_s)}{\nu(B_s)}$$

for each $V \in \mathcal{F}$. Since $\nu(B_s) > 0$ for all $s \in S$, this is well-defined. Moreover, since ν is a proper finitely-additive probability measure on $\mathcal{F}$, so is τ_s.

Note that for all $A \in \mathcal{A}$ and $p \in \Delta(X)$, $\{u \in \tilde{\mathbb{R}}^X : p \in M(M(A, s), u)\} = N(M(A, s), p) \in \mathcal{F}$, so $\tau_s(\{u \in \tilde{\mathbb{R}}^X : p \in M(M(A, s), u)\})$ is well-defined. The next lemma will allow us to complete the representation:

Lemma F.3. For each $s \in S$, $A \in \mathcal{A}$, and $p \in A$,

$$\nu(N(A, p)) = \sum_{s\in S} \mu(s)\tau_s(\{u \in \tilde{\mathbb{R}}^X : p \in M(M(A, s), u)\}).$$

Proof. We first show that for each $s \in S$, $\operatorname{supp} \tau_s \setminus \{0\} = [s]$. To see that $[s] \subseteq \operatorname{supp} \tau_s \setminus \{0\}$, consider any $u \in [s]$ and any open $V \in \mathcal{F}$ such that $u \in V$. By Lemma F.4 (iii), V is a finite union of finite intersections of sets in $\mathcal{N} \cup \mathcal{N}^+$. Hence, since each element of $\mathcal{N} \cup \mathcal{N}^+$ is closed under positive affine transformations so is V. Thus, $u \in V$ implies $s \in V$. But then $V \cap B_s \in \mathcal{F}$ is open and contains s, and hence $\nu(V \cap B_s) > 0$ since $s \in \operatorname{supp} \nu$. This proves $u \in \operatorname{supp} \tau_s \setminus \{0\}$.

To see that $\operatorname{supp} \tau_s \setminus \{0\} \subseteq [s]$, consider any $u \neq 0$ such that $u \notin [s]$. It suffices to show that there exists an open $V \in \mathcal{F}$ such that $u \in V$ and $\tau_s(V) = 0$. If $u \approx s'$ for some $s' \in S \setminus \{s\}$, then $V = B_{s'}$ is as required since $B_{s'} \cap B_s = \varnothing$ and $u \in B_{s'}$. If there is no $s' \in S \setminus \{s\}$ such that $u \approx s'$, then $u \notin \operatorname{supp} \nu$. But then there exists an open $V \in \mathcal{F}$ such that $u \in V$ and $\nu(V) = 0$, so also $\tau_s(V) = 0$.

By Lemma F.6 below, this implies that $\tau_s(N(A, p)) = \tau_s(N(M(A, s), p))$ for any $A \in \mathcal{A}$ and $p \in A$.

This implies that for any $A \in \mathcal{A}$ and $p \in A$

$$\begin{aligned}
\sum_{s \in S} \mu(s)\tau_s(\{u \in \tilde{\mathbb{R}}^X : p \in M(M(A,s),u)\}) &= \sum_{s \in S} \mu(s)\tau_s(N(M(A,s),p)) \\
&= \sum_{s \in S} \mu(s)\tau_s(N(A,p)) \\
&= \sum_{s \in S} \nu(N(A,p) \cap B_s) \\
&= \nu(N(A,p) \cap \bigcup_{s \in S} B_s) \\
&= \nu(N(A,p),
\end{aligned}$$

where the last equality follows from Lemma F.2. ∎

For any $s \in S = \{u_1, \ldots, u_L\}$, we write $U_s := s$. We claim that $(S, \mu, \{U_s, \tau_s\}_{s \in S})$ is an S-based REU representation of ρ. Indeed, by construction, U_s is non-constant for all s, $U_s \not\approx U_{s'}$ for any distinct $s, s' \in S$, and μ is a full-support probability measure on S. Moreover, each τ_s is a proper finitely-additive probability measure on $\tilde{\mathbb{R}}^X$ endowed with the algebra $\mathcal{F}$. By standard arguments (cf. Rao and Rao (2012)), we can extend τ_s to a proper finitely-additive probability measure on the Borel σ-algebra on $\tilde{\mathbb{R}}^X$. Finally, Lemma F.3 and the fact that ν represents ρ implies that for all $A \in \mathcal{A}$ and $p \in A$, we have $\rho(p; A) = \sum_{s \in S} \mu(s)\tau_s(p, A)$, as required.

F.3 Proof of Theorem F.1: Necessity

Suppose that ρ admits an S-based REU representation $(S, \mu, \{U_s, \tau_s\}_{s \in S})$. We show that ρ satisfies Axiom 0. Observe first that for any finite $Y \subseteq X$ with $y^* \in Y$, $(S, \mu, \{U_s \restriction_Y, \tau_s \restriction_Y\}_{s \in S})$ constitutes an S-based REU representation of ρ^Y, where $U_s \restriction_Y$ denotes the restriction of U_s to Y and $\tau_s \restriction_Y$ is given by $\tau_s \restriction_Y (B) = \tau_s(B \times \mathbb{R}^{X \smallsetminus Y})$ for any Borel set B on $\mathbb{R}^Y$. Thus, by Theorem S3 in Ahn and Sarver (2013), ρ^Y satisfies Regularity, Linearity, Extremeness, and Mixture Continuity.

To show that ρ satisfies Regularity, consider any $p \in A \subseteq A'$. Pick a finite $Y \subseteq X$ with $y^* \in Y$ such that $A, A' \in \mathcal{A}(Y)$. By definition, $\rho(p; A) = \rho^Y(p; A)$ and $\rho(p; A') = \rho^Y(p; A')$. Hence, by Regularity for ρ^Y, we have $\rho(p; A) \geq \rho(p; A')$, as required. Similarly, we can show that ρ satisfies Linearity, Extremeness, and Mixture Continuity by using the fact that for each finite Y, each ρ^Y satisfies these axioms.

Finally, to show that ρ satisfies Finiteness, let $K := |S|$ and consider any $A \in \mathcal{A}$. For each $s \in S$, pick any $q_s \in M(A, U_s)$, and define $B := \{q_s : s \in S\}$. Note that $|B| \leq K$. If $B = A$, then Finiteness is trivially satisfied. If $B \subsetneq A$, then pick any $p \in A \smallsetminus B$. We can pick a large enough finite $Y \subseteq X$ such that each U_s is non-constant on Y and $U_s \restriction_Y \not\approx U_{s'} \restriction_Y$ for any distinct $s, s' \in S$. Let $r \in \Delta(Y)$ be given by $r(y) := \frac{1}{|Y|}$ for each $y \in Y$. For each $s \in Y$, pick any $y_s \in \text{argmax}_{y \in Y} U_s(y)$. Note that $U_s(y_s) > U_s(r)$. Define $B^n := \frac{n-1}{n} B + \frac{1}{n}\{y_s : s \in S\}$ and $p^n := \frac{n-1}{n} p + \frac{1}{n} r$. Then $B^n \to^m B$ and $p_n \to^m p$. Moreover, for all large enough n, we have $U_s(\frac{n-1}{n} q_s + \frac{1}{n} y_s) > U_s(p^n)$ for each $s \in S$. Thus, $\rho(p^n; \{p^n\} \cup B^n) = 0$, proving Finiteness.

F.4 Additional Lemmas for Section F

F.4.1 Properties of $N(A,p)$ Sets

Lemma F.4. Fix any $X' \subseteq X$ with $y^* \in X$. For any collection $\mathcal{S}$, we let $\mathcal{U}(\mathcal{S})$ denote the set of all finite unions of elements of $\mathcal{S}$.

(i). If $E \in \mathcal{N}(X')$ (resp. $E \in \mathcal{N}^+(X')$), then $E^c \in \mathcal{U}(\mathcal{N}^+(X'))$ (resp. $E^c \in \mathcal{U}(\mathcal{N}(X'))$.

(ii). If $E_1, E_2 \in \mathcal{N}(X')$ (resp. $E_1, E_2 \in \mathcal{N}^+(X')$), then $E_1 \cap E_2 \in \mathcal{N}(X')$ (resp. $E_1 \cap E_2 \in \mathcal{N}^+(X')$).

(iii). $\mathcal{F}(X')$ is the set of all E such that $E = \bigcup_{\ell \in L} M_\ell \cap N_\ell$ for some finite index set L and $M_\ell \in \mathcal{N}(X')$, $N_\ell \in \mathcal{N}^+(X')$ for each $\ell \in L$.

(iv). $\mathcal{F}(X')$ is the set of all E for which there exists a finite $Y \subseteq X'$ with $y^* \in Y$ and $E^Y \in \mathcal{F}(Y)$ such that $E = E^Y \times \mathbb{R}^{X \setminus Y}$.

Proof.

(i): If $E = N(A,p) \in \mathcal{N}(X')$, then $E^c = \bigcup_{q \in A \setminus \{p\}} N^+(\{p,q\},q) \in \mathcal{U}(\mathcal{N}^+(X'))$ if $p \in A$ and $E^c = \tilde{\mathbb{R}}^{X'} \in \mathcal{U}(\mathcal{N}^+(X'))$ if $p \notin A$. Similarly, if $E = N^+(A,p) \in \mathcal{N}^+(X')$, then $E^c = \bigcup_{q \in A \setminus \{p\}} N(\{p,q\},q) \in \mathcal{U}(\mathcal{N}(X'))$ if $p \in A$ and $E^c = \tilde{\mathbb{R}}^{X'} \in \mathcal{U}(\mathcal{N}(X'))$ if $p \notin A$.

(ii): If $N(A_1,p_1), N(A_2,p_2) \in \mathcal{N}(X')$, then $N(A_1,p_1) \cap N(A_2,p_2) = N(\frac{1}{2}A_1 + \frac{1}{2}A_2, \frac{1}{2}p_1 + \frac{1}{2}p_2) \in \mathcal{N}(X')$. The same argument goes through replacing all instances of N with N^+.

(iii): By standard results, $\mathcal{F}(X')$ can be described as follows: Let $\mathcal{F}_0(X')$ denote the set of all elements of $\mathcal{N}(X') \cup \mathcal{N}^+(X')$ and their complements. Let $\mathcal{F}_1(X')$ denote the set of all finite intersections of elements of $\mathcal{F}_0(X')$. Then $\mathcal{F}(X')$ is the set of all finite unions of elements of $\mathcal{F}_1(X')$. By part (i), $\mathcal{F}_0(X) = \mathcal{U}(N(X)) \cup \mathcal{U}(N(X'))$ is the collection of all finite unions of elements of $\mathcal{N}(X')$ and of all finite unions of elements of $\mathcal{N}^+(X')$. By part (ii), $\mathcal{F}_1(X') = \mathcal{F}_0(X) \cup \mathcal{I}(X')$, where $\mathcal{I}(X')$ consists of all finite unions of the form $\bigcup_{\ell \in L} M_\ell \cap N_\ell$, where $M_\ell \in \mathcal{N}(X')$ and $N_\ell \in \mathcal{N}^+(X')$ for each $\ell \in L$. Note that $\tilde{\mathbb{R}}^{X'} \in \mathcal{N}(X') \cap \mathcal{N}^+(X')$, since $\tilde{\mathbb{R}}^{X'} = N_{X'}(\{p\},p) = N^+_{X'}(\{p\},p)$ for any $p \in \Delta(X')$. Thus, $\mathcal{F}_0(X) = \mathcal{U}(N(X)) \cup \mathcal{U}(N(X')) \subseteq \mathcal{I}(X)$. Hence, $\mathcal{F}_1(X) = \mathcal{I}(X) = \mathcal{F}(X)$.

(iv): Note first that for any $N_{X'}(A,p) \in \mathcal{N}(X')$ (resp. $N^+_{X'}(A,p) \in \mathcal{N}^+(X')$) and any finite $Y \subseteq X'$ with $y^* \in Y$ and $A \in \mathcal{A}(Y)$, we have $N_{X'}(A,p) = N_Y(A,p) \times \mathbb{R}^{X' \setminus Y}$ (resp. $N^+_{X'}(A,p) = N^+_Y(A,p) \times \mathbb{R}^{X' \setminus Y}$). Now fix any $E \in \mathcal{F}(X')$. By part (iv), we have a finite index set L and $M_\ell \in \mathcal{N}(X')$, $N_\ell \in \mathcal{N}^+(X')$ for each $\ell \in L$ such that $E = \bigcup_{\ell \in L} M_\ell \cap N_\ell$. By the first sentence, we can then pick a finite $Y \subseteq X'$ with $y^* \in Y$ such that for each ℓ, we have $M_\ell = M^Y_\ell \times \mathbb{R}^{X' \setminus Y}$ and $N_\ell = N^Y_\ell \times \mathbb{R}^{X' \setminus Y}$, where $M^Y_\ell \in \mathcal{N}(Y)$ and $N^Y_\ell \in \mathcal{N}^+(Y)$. Then $E = E^Y \times \mathbb{R}^{X \setminus Y}$, where $E^Y := \bigcup_{\ell \in L} M^Y_\ell \cap N^Y_\ell \in \mathcal{F}(Y)$. Conversely, if $E^Y \in \mathcal{F}(Y)$, then by part (iv), E^Y is of the form $\bigcup_{\ell \in L} M^Y_\ell \cap N^Y_\ell \in \mathcal{F}(Y)$ for some finite collection of $M^Y_\ell \in \mathcal{N}(Y)$ and $N^Y_\ell \in \mathcal{N}^+(Y)$. Then by the first sentence, $M_\ell := M^Y_\ell \times \mathbb{R}^{X' \setminus Y} \in \mathcal{N}(X')$ and $N_\ell = N^Y_\ell \times \mathbb{R}^{X' \setminus Y} \in \mathcal{N}^+(X')$, so $E^Y \times \mathbb{R}^{X' \setminus Y} = \bigcup_{\ell=1}^{L} M_\ell \cap N_\ell) \in \mathcal{F}(X')$ by part (iv). ■

F.4.2 Properties of Proper Finitely-Additive Probability Measures on $\mathcal{F}$

Lemma F.5. Let ν be a proper finitely-additive probability measure on $\mathcal{F}$ and suppose that $(N(p,A) \setminus \{0\}) \cap \operatorname{supp} \nu = \varnothing$ for some $A \in \mathcal{A}$ and $p \in A$, where 0 denotes the unique constant utility in $\tilde{\mathbb{R}}^X$. Then $\nu(N^+(A,p)) = \nu(N(A,p)) = 0$.

Proof. Since $(N(A,p) \setminus \{0\}) \cap \operatorname{supp} \nu = \varnothing$, we have

$$N(A,p) \setminus \{0\} \subseteq (\operatorname{supp} \nu)^c := \bigcup \{V \in \mathcal{F} : V \text{ open and } \nu(V) = 0\}.$$

Thus, for some possibly infinite index set I, there exists a family $\{V_i\}_{i\in I}$, with $V_i \in \mathcal{F}$ open and $\nu(V_i) = 0$ for each i such that

$$N(A,p) \setminus \{0\} \subseteq \bigcup_{i\in I} V_i.$$

We now show that there is a finite subset $\{i_1, \dots, i_n\} \subseteq I$ such that

$$N(A,p) \setminus \{0\} \subseteq \bigcup_{j=1}^{n} V_{i_j}.$$

To see this, define $L(A,p) := (N(A,p) \cap [-1,1]^X) \setminus \{0\}$. Note that since $[-1,1]^X$ is compact in $\mathbb{R}^X$ (by Tychonoff's theorem) and $N(A,p)$ is closed in $\tilde{\mathbb{R}}^X$, $C(A,p)$ is compact in the relative topology on $\tilde{\mathbb{R}}^X \setminus \{0\}$. Hence, since $L(A,p) \subseteq N(A,p) \setminus \{0\}$ is covered by $\bigcup_{i\in I} V_i$ and each V_i is open, it has a finite subcover $\bigcup_{j=1}^n V_{i_j}$.

We claim that $N(A,p) \setminus \{0\}$ is also covered by $\bigcup_{j=1}^n V_{i_j}$. To see this, consider any $u^* \in N(A,p) \setminus \{0\}$. We can find a finite $Y \subseteq X$ such that $y^* \in Y$, $u^* \restriction_Y$ is not constant, $N(A,p) = N_Y(A,p) \times \mathbb{R}^{X\setminus Y}$, and for each $j = 1, \dots, n$, $V_{i_j} = V^Y_{i_j} \times \mathbb{R}^{X\setminus Y}$ for some $V^Y_{i_j} \in \mathcal{F}^Y$ (see Lemma F.4 (iv)).

Since Y is finite, there exists $\alpha > 0$ small enough such that $\alpha u^*(y) \in [-1,1]$ for all $y \in Y$. Define $u \in \tilde{\mathbb{R}}^X$ by $u \restriction_Y = \alpha u^* \restriction_Y$ and $u(x) = 0$ for all $x \in X \setminus Y$. Note that $u \in N(A,p)$: Indeed, $u^* \in N(A,p) = N_Y(A,p) \times \mathbb{R}^{X\setminus Y}$, $u \restriction_Y = \alpha u^* \restriction_Y$, and $N_Y(A,p)$ is closed under positive scaling. Moreover, u is not constant, since $u^* \restriction_Y$ is not constant. Finally, $u \in [-1,1]^X$. This shows $u \in L(A,p)$. Since $L(A,p)$ is covered by $\bigcup_{j=1}^n V_{i_j}$, there exists j such that $u \in V_{i_j} = V^Y_{i_j} \times \mathbb{R}^{X\setminus Y}$. But note that $V^Y_{i_j}$ is closed under positive scaling, since by Lemma F.4 (iii) it is a finite union of sets which are closed under positive scaling. Since $u \restriction_Y = \alpha u^* \restriction_Y$, this implies $u^* \in V_{i_j}$.

The above shows that $N(A,p) \setminus \{0\}$ is covered by $\bigcup_{j=1}^n V_{i_j}$, and hence so is $N^+(A,p)$. But since $\nu(V_{i_j}) = 0$ for all $j = 1, \dots, n$ and ν is finitely additive, it follows that $\nu(N^+(A,p)) = 0$. Moreover, by properness of ν, this implies $\nu(N(A,p)) = 0$. ∎

Lemma F.6. Suppose ν is a proper finitely-additive probability measure on $\mathcal{F}$ and $\operatorname{supp}\nu \setminus \{0\} = [u]$ for some $u \in \tilde{\mathbb{R}}^X$. Then for any $A \in \mathcal{A}$ and $p \in A$, we have $\nu(N(A,p)) = \nu(N(M(A,u),p))$.

Proof. Fix any $A \in \mathcal{A}$ and $p \in A$. Note first that for any $q \in A$,

$$q \notin M(A,u) \Rightarrow \nu(N(A,q)) = 0. \tag{25}$$

Indeed, if $q \notin M(A,u)$, then $\varnothing = [u] \cap N(A,q) = (N(A,q) \setminus \{0\}) \cap \operatorname{supp}\nu$. But then Lemma F.5 implies that $\nu(N(A,q)) = 0$, as claimed.

Suppose now that $p \notin M(A,u)$. Then (25) implies that $\nu(N(A,p)) = 0$. Moreover, $N(B,p) := \varnothing$ if $p \notin B$, so also $\nu(N(M(A,u),p)) = 0$, as required.

Suppose next that $p \in M(A,u)$. Then

$$N(A,p) \subseteq N(M(A,u),p) \subseteq N(A,p) \cup \bigcup_{q\in A\setminus M(A,u)} N(A,q),$$

so that

$$\nu(N(A,p)) \leq \nu(N(M(A,u),p)) \leq \nu(N(A,p)) + \sum_{q\in A\setminus M(A,u)} \nu(N(A,q)) = \nu(N(A,p)),$$

where the last equality follows from (25). This again shows that $\nu(N(A,p)) = \nu(N(M(A,u),p))$, as required. ■

F.5 Proof of Proposition F.2

"Only if" direction: We prove the contrapositive. Suppose that there exists some $s' \in S$ and $x \in X$ such that $\lim_n U_{s'}(x_n) \neq U_{s'}(x)$ for some sequence $x_n \to x$. Since S is finite, by taking an appropriate subsequence of $\{x_n\}$, we can assume that $\lim_n U_s(x_n)$ exists (allowing for $\pm\infty$) for every $s \in S$.

Let $S_+ := \{s \in S : \lim_n U_s(x_n) < U_s(x)\}$, $S_- := \{s \in S : \lim_n U_s(x_n) > U_s(x)\}$, and $S_0 := S \setminus (S_+ \cup S_-)$. Then there exist $\gamma > 0$ and N such that for all $n \geq N$, $U_s(x_n) + 2\gamma < U_s(x)$ for all $s \in S_+$ and $U_s(x_n) > U_s(x) + 2\gamma$ for all $s \in S_-$. Let $p = \alpha\delta_x + (1-\alpha)\delta_{x_N}$. By setting α sufficiently large, we can guarantee that for all $n \geq N$, $U_s(x_n) + \gamma < U_s(p)$ for all $s \in S_+$ and $U_s(x_n) > U_s(p) + \gamma$ for all $s \in S_-$. Note also that $U_s(x) > U_s(p) + 2\gamma(1-\alpha)$ for all $s \in S_+$ and $U_s(x) + 2\gamma(1-\alpha) < U_s(p)$ for all $s \in S_-$.

Since S is finite and each U_s is non-constant, we can assume that $U_s(p) \neq U_s(x)$ for all $s \in S$. (Otherwise, we can replace p with a lottery that is obtained by mixing an appropriate lottery to p, without violating the above construction). This implies that there exist $\gamma' > 0$ and N' such that for all $s \in S_0$, either $\min\{U_s(x_n), U_s(x)\} > U_s(p) + \gamma'$ for all $n \geq N'$ or $\max\{U_s(x_n), U_s(x)\} + \gamma' < U_s(p)$ for all $n \geq N'$. Let S_{0-} be the set of states in S_0 that satisfy the former inequality, and S_{0+} be the set of states in S_0 that satisfy the latter inequality.

Let $m := |S|$. By Lemma E.2 we can find distinct lotteries $\{q_1, ..., q_m\}$ such that $U_{s_i} \in N^+(\{q_1, .., q_m\}, q_i)$ for each $s_i \in S$. Define $p_i = (1-\varepsilon)p + \varepsilon q_i$ for each $s_i \in S$, and $A := \{p_1, ..., p_m, \delta_x\}$ and $A_n := \{p_1, ..., p_m, \delta_{x_n}\}$. By construction, if we take ε sufficiently small, then for all $n \geq \max\{N, N'\}$,

$$[U_{s_i} \in N^+(A_n, p_i) \cap N^+(A, \delta_x), \forall s_i \in S_+], \quad [U_{s_i} \in N^+(A_n, \delta_{x_n}) \cap N^+(A, p_i), \forall s_i \in S_-],$$

$$[U_{s_i} \in N^+(A_n, \delta_{x_n}) \cap N^+(A, \delta_x), \forall s_i \in S_{0-}], \quad [U_{s_i} \in N^+(A_n, p_i) \cap N^+(A, p_i), \forall s_i \in S_{0+}].$$

By Lemma E.3, $A, A_n \in \mathcal{A}^*$ for all $n \geq \max\{N, N'\}$. Note that $S_+ \cup S_- \neq \varnothing$ by assumption. Take any $s_i \in S_+ \cup S_-$. If $s_i \in S_+$, then $\rho(p_i; A_n) = \mu(s_i)$ for every $n \geq \max\{N, N'\}$ and $\rho(p_i; A) = 0$. If $s_i \in S_-$, then $\rho(p_i; A_n) = 0$ for every $n \geq \max\{N, N'\}$ and $\rho(p_i; A) = \mu(s_i)$. In either case, Axiom F.1 is violated.

"If" direction: Suppose each U_s is continuous. Take any sequence $A_n \to A$ of menus that converge under the Hausdorff metric such that $A, A_n \in \mathcal{A}^*$ for each n. Enumerate the elements in A by $A = \{p_1, ..., p_m\}$, where we can assume up to relabeling that for some $k \leq m$ we have $\rho(p_i; A) > 0$ for each $i = 1, ..., k$ and $\rho(p_i; A) = 0$ for each $i = k+1, ..., m$. For each $i = 1, ..., k$, define $S_i := \{s \in S : M(A, U_s) = \{p_i\}\}$. Note that by Lemma E.3, $S = \cup_i S_i$ since $A \in \mathcal{A}^*$.

Take any B that is a continuity set under $\rho(\cdot; A)$. For each $i = 1, ..., k$, we have either $p_i \in \text{int}B$ or $p_i \in \text{int}(\Delta(X) \setminus B)$. We can pick $\varepsilon > 0$ sufficiently small such that:

(i). $B_\varepsilon(p_i) \subseteq \text{int}B$ if $p_i \in \text{int}B$, and $B_\varepsilon(p_i) \subseteq \text{int}(\Delta(X) \setminus B)$ if $p_i \in \text{int}(\Delta(X) \setminus B)$

(ii). for any $i, j = 1, ..., m$ with $i \neq j$, we have $B_\varepsilon(p_i) \cap B_\varepsilon(p_j)$

(iii). for any $i = 1, ..., k$, $j = 1, ..., m$ with $i \neq j$, $q_i \in B_\varepsilon(p_i)$, and $q_j \in B_\varepsilon(p_j)$, we have $U_{s_i}(q_i) > U_{s_i}(q_j)$ for all $s_i \in S_i$.

Here $B_\varepsilon(\cdot)$ denotes ε-neighborhoods with respect to the Prokhorov metric π, and (iii) holds by the assumption that each U_s is continuous. Since $A_n \to A$, there exists N such that for all $n \geq N$, we have the following: (a) for each $q \in A_n$, there exists $i = 1, ..., m$ such that $q \in B_\varepsilon(p_i)$; and (b) for each $i = 1, ..., m$, there exists $q \in A_n$ such that $q \in B_\varepsilon(p_i)$. For such $n \geq N$, we then have $M(A_n, U_{s_i}) \in B_\varepsilon(p_i)$ for each $i = 1, ..., k$ and $s_i \in S_i$. Thus $\rho(B; A_n) = \sum_{i=1}^k \mu(S_i) = \rho(B; A)$. By the Portmanteau theorem, this guarantees that $\rho(\cdot; A_n) \to \rho(\cdot; A)$ under weak convergence, as claimed. ■

G Proofs for Section 5

G.1 Proof of Proposition 1

The first part is immediate from the i.i.d. full-support assumption on ε. To show the second part, suppose that $v_1(z_1) < v_1(z_1')$. We consider the equivalent problem of scaling v terms by $\alpha := \frac{1}{\lambda} > 0$ while fixing ε terms. That is, we write

$$U_0(z_0, A_1^{\text{big}}) = \alpha v_0(z_0) + \varepsilon_0^{(z_0, A_1^{\text{big}})} + \delta \mathbb{E}[\max\{\alpha v_1(z_1) + \varepsilon_1^{z_1}, \alpha v_1(z_1') + \varepsilon_1^{z_1'}\}]$$

$$U_0(z_0, A_1^{\text{small}}) = \alpha v_0(z_0) + \varepsilon_0^{(z_0, A_1^{\text{small}})} + \delta \alpha v_1(z_1),$$

where the second line used the fact that $\varepsilon_1^{z_1}$ has mean zero.

By the i.i.d. full-support assumption on ε_0, the desired claim follows if we show that the difference $U_0(z_0, A_1^{\text{big}}) - U_0(z_0, A_1^{\text{small}})$ is decreasing in α. To show this, suppose without loss of generality that $v_0(z_0) = 0$. Then for all α, the derivatives of the utilities satisfy

$$\frac{dU(z_0, A_1^{\text{big}})}{d\alpha} = \delta \left(\rho_1(z_1, A_1^{\text{big}}) v_1(z_1) + \rho_1(z_1', A_1^{\text{big}}) v_1(z_1') \right), \quad \frac{dU(z_0, A_1^{\text{small}})}{d\alpha} = \delta v_1(z_1'),$$

where we can suppress the dependence on histories in ρ_1 since ε shocks are i.i.d. Moreover, letting f denote the density of the ε shocks and setting $\kappa(\varepsilon_1^{z_1'}) := \alpha(v_1(z_1') - v_1(z_1)) + \varepsilon_1^{z_1'}$, we have that $\rho_1(z_1, A_1^{\text{big}}) = \int_{-\infty}^{\infty} \int_{\kappa(\varepsilon_1^{z_1'})}^{\infty} f(\varepsilon_1^{z_1}) d\varepsilon_1^{z_1} f(\varepsilon_1^{z_1'}) d\varepsilon_1^{z_1'}$ and $\rho_1(z_1', A_1^{\text{big}}) = 1 - \rho_1(z_1, A_1^{\text{big}})$. Note that both choice probabilities are strictly positive since the ε_1 shocks are i.i.d. with full support. Thus, $v_1(z_1) < v_1(z_1')$ implies $\frac{dU(z_0, A_1^{\text{big}})}{d\alpha} < \frac{dU(z_0, A_1^{\text{small}})}{d\alpha}$ for all α, as required. ■

G.2 Proof of Proposition 2

BEU: For BEU, we have

$$U_0(x, A_1^{\text{early}}) = \mathbb{E}[\max\{\mathbb{E}[u_2(y)|\mathcal{F}_1], \mathbb{E}[u_2(z)|\mathcal{F}_1]\}|\mathcal{F}_0]$$
$$U_0(x, A_1^{\text{late}}) = \mathbb{E}[\mathbb{E}[\max\{u_2(y), u_2(z)\}|\mathcal{F}_1]|\mathcal{F}_0].$$

By the conditional Jensen inequality and convexity of the max operator, $U_0(x, A_1^{\text{early}}) \leq U_0(x, A_1^{\text{late}})$. Moreover, this inequality is strict at ω as long as there exist $\omega', \omega'' \in \mathcal{F}_0(\omega)$ with $\mathcal{F}_1(\omega') = \mathcal{F}_1(\omega'')$ such that $u_2(y) - u_2(z)$ changes sign on $\{\omega', \omega''\}$.

i.i.d. DDC: For i.i.d. DDC, to simplify the notation we assume $v_0(x) = v_1(x) = 0$ without loss of generality. Take a measurable function $\sigma : \mathbb{R}^2 \to [0, 1]$ such that

$$\sigma(\varepsilon^y, \varepsilon^z) \in \text{argmax}_{\alpha \in [0,1]} \alpha(v_2(y) + \varepsilon^y) + (1 - \alpha)(v_2(z) + \varepsilon^z)$$

for all $(\varepsilon^y, \varepsilon^z) \in \mathbb{R}^2$.[83] Then $U_0(x, A_1^{\text{late}}) - \varepsilon_0^{(x, A_1^{\text{late}})}$ is equal to

$$\begin{aligned}
& \delta^2 \mathbb{E}[\max\{v_2(y) + \varepsilon_2^y, v_2(z) + \varepsilon_2^z\}] \\
= \;& \delta^2 \mathbb{E}[\sigma(\varepsilon_2^y, \varepsilon_2^z)(v_2(y) + \varepsilon_2^y) + (1 - \sigma(\varepsilon_2^y, \varepsilon_2^z))(v_2(z) + \varepsilon_2^z)] \\
= \;& \delta^2(\alpha^* v_2(y) + (1 - \alpha^*) v_2(z)) + \delta^2 \mathbb{E}[\sigma(\varepsilon_2^y, \varepsilon_2^z)\varepsilon_2^y + (1 - \sigma(\varepsilon_2^y, \varepsilon_2^z))\varepsilon_2^z]
\end{aligned}$$

where $\alpha^* := \mathbb{E}[\sigma(\varepsilon_2^y, \varepsilon_2^z)]$. Since ε_2^y and ε_2^z have mean zero, $\delta^2(\alpha^* v_2(y) + (1 - \alpha^*) v_2(z))$ in the last line is equal to the expected value the agent would obtain from A_1^{late} if in period 2 she chooses y with probability α^* regardless of the realization of ε_2. Since such a decision rule is strictly suboptimal at A_1^{late} under the full support assumption on ε_2, the term $\delta^2 \mathbb{E}[\sigma(\varepsilon_2^y, \varepsilon_2^z)\varepsilon_2^y + (1 - \sigma(\varepsilon_2^y, \varepsilon_2^z))\varepsilon_2^z]$ in the last line is strictly positive. At the same time, $U_0(x, A_1^{\text{early}}) - \varepsilon_0^{(x, A_1^{\text{early}})}$ is equal to

$$\begin{aligned}
& \delta \mathbb{E}[\max\{\delta v_2(y) + \varepsilon_1^{(x,\{y\})}, \delta v_2(z) + \varepsilon_1^{(x,\{z\})}\}] \\
\geq \;& \delta \mathbb{E}[\sigma(\varepsilon_1^{(x,\{y\})}, \varepsilon_1^{(x,\{z\})})(\delta v_2(y) + \varepsilon_1^{(x,\{y\})}) + (1 - \sigma(\varepsilon_1^{(x,\{y\})}, \varepsilon_1^{(x,\{z\})}))(\delta v_2(z) + \varepsilon_1^{(x,\{z\})})] \\
= \;& \delta^2(\alpha^* v_2(y) + (1 - \alpha^*) v_2(z)) + \delta \mathbb{E}[\sigma(\varepsilon_2^y, \varepsilon_2^z)\varepsilon_2^y + (1 - \sigma(\varepsilon_2^y, \varepsilon_2^z))\varepsilon_2^z]
\end{aligned}$$

where the inequality follows since the value in the second line is the expected payoff if the agent follows the decision rule σ at A_1^{early}. The equality holds by the i.i.d. assumption on ε_1 and ε_2. Since $\delta \in (0, 1)$, it follows that $U_0(x, A_1^{\text{early}}) - \varepsilon_0^{(x, A_1^{\text{early}})} > U_0(x, A_1^{\text{late}}) - \varepsilon_0^{(x, A_1^{\text{late}})}$. Thus, the desired claim follows from the i.i.d. assumption on ε_0.

"Moreover" part: We consider the equivalent problem in which we scale v terms by a scaling factor $\alpha := \frac{1}{\lambda} > 0$ while fixing ε terms. Assume $v_2(y) > v_2(z)$ without loss of generality. Then:

$$U_0(x, A_1^{\text{early}}) = \varepsilon_0^{(x, A_1^{\text{early}})} + \delta \mathbb{E}[\max\{\delta \alpha v_2(y) + \varepsilon_1^{x,\{y\}}, \delta \alpha v_2(z) + \varepsilon_1^{x,\{z\}}\}]$$

$$U_0(x, A_1^{\text{late}}) = \varepsilon_0^{(x, A_1^{\text{late}})} + \delta^2 \mathbb{E}[\max\{\alpha v_2(y) + \varepsilon_2^y, \alpha v_2(z) + \varepsilon_2^z\}]$$

By the i.i.d full-support assumption on ε_0, the desired claim follows if we show that $U_0(x, A_1^{\text{early}}) - U_0(x, A_1^{\text{late}})$ is strictly decreasing in α. As in the proof of Proposition 1, the derivatives of utilities with respect to α satisfy

$$\frac{dU_0(x, A_1^{\text{early}})}{d\alpha} = \delta^2 \left(\rho_1((x, \{y\}); A_1^{\text{early}}) v_2(y) + \rho_1((x, \{z\}); A_1^{\text{early}}) v_2(z) \right),$$

$$\frac{dU_0(x, A_1^{\text{late}})}{d\alpha} = \delta^2 \left(\rho_2(y; \{y, z\}) v_2(y) + \rho_2(z; \{y, z\}) v_2(z) \right),$$

where we can again suppress the dependence of choice probabilities on histories due to the i.i.d. ε assumption. But note that

$$\rho_1((x, \{y\}); A_1^{\text{early}}) = \Pr[\delta(v_2(y) - v_2(z)) \geq \varepsilon_1^{x,\{z\}} - \varepsilon_1^{x,\{y\}}] < \Pr[v_2(y) - v_2(z) \geq \varepsilon_2^z - \varepsilon_2^y] = \rho_2(y; \{y, z\}),$$

where the inequality holds since $\delta < 1$, $v_2(y) > v_2(z)$ and by the i.i.d. full support assumption on ε. Thus, $\frac{dU_0(x, A_1^{\text{early}})}{d\alpha} < \frac{dU_0(x, A_1^{\text{late}})}{d\alpha}$, as required. ■

[83] The existence of such a function follows by the measurable selection theorem.

G.3 Proof of Proposition 3

Let G denote the cdf of the difference $\varepsilon - \varepsilon'$ of two shocks $\varepsilon, \varepsilon'$ that are independently drawn from F.

Proof of Proposition 3. Because the density of ε is symmetric and unimodal around 0, G dominates F in terms of the peakedness order by Theorem 3.D.4 in Shaked and Shanthikumar (2007). Thus, $F(\gamma) \geq G(\gamma)$ for any $\gamma > 0$ and $F(\gamma) \leq G(\gamma)$ for any $\gamma < 0$ by Theorem 3.D.1 in Shaked and Shanthikumar (2007); moreover, the inequalities are strict because the distribution F has full support.

We express choice probabilities of a in each period as functions of parameters (w, δ), where we can suppress the dependence on histories by the i.i.d. assumption on shocks. That is, for each model $\mathrm{M} = \mathrm{DDC}, \mathrm{BEU}$, let $\rho_0^{\mathrm{M}}(w, \delta) := \rho_0^{\mathrm{M}}(a; A_0)$ and $\rho_1^{\mathrm{M}}(w, \delta) := \rho_1^{\mathrm{M}}(a; A_1)$ for each (w, δ). Let $V(w) := \mathbb{E}[\max\{w + \varepsilon_1^a, \varepsilon_1^b\}]$. Note that $V(w) \geq 0$ since shocks have mean zero, and the inequality is strict because of the full support assumption. We have $\rho_1^{\mathrm{DDC}}(w, \delta) = \rho_1^{\mathrm{BEU}}(w, \delta) = \Pr(w + \varepsilon_1^a \geq \varepsilon_1^b) = G(w)$. Moreover, $\rho_0^{\mathrm{DDC}}(w, \delta) = \Pr(w + \varepsilon_0^a \geq \delta V(w) + \varepsilon_0^{A_1}) = 1 - G(\delta V(w) - w)$. Finally, $\rho_0^{\mathrm{BEU}}(w, \delta) = \Pr(w + \varepsilon_0^a \geq \delta V(w)) = 1 - F(\delta V(w) - w)$.

For each model M, we consider the maximization problem

$$\max_{(\hat{\omega}, \hat{\delta}) \in \Theta} \quad \rho_0(a; A_0) \log[\rho_0^{\mathrm{M}}(\hat{w}, \hat{\delta})] + (1 - \rho_0(a; A_0)) \log[1 - \rho_0^{\mathrm{M}}(\hat{w}, \hat{\delta})]$$
$$+ (1 - \rho_0(a; A_0)) \left(\rho_1(a; A_1) \log[\rho_1^{\mathrm{M}}(\hat{w}, \hat{\delta})] + (1 - \rho_1(a; A_1)) \log[1 - \rho_1^{\mathrm{M}}(\hat{w}, \hat{\delta})] \right).$$

By the assumption that ρ is compatible, for each model M = DDC, BEU, there exists $(\hat{w}^{\mathrm{M}}, \hat{\delta}^{\mathrm{M}}) \in \Theta$ such that

$$\rho_0(a; A_0) = \rho_0^{\mathrm{M}}(\hat{w}^{\mathrm{M}}, \hat{\delta}^{\mathrm{M}}) \text{ and } \rho_1(a; A_1) = \rho_1^{\mathrm{M}}(\hat{w}^{\mathrm{M}}, \hat{\delta}^{\mathrm{M}}) \tag{26}$$

hold. By Gibbs' inequality, $(\hat{w}^{\mathrm{M}}, \hat{\delta}^{\mathrm{M}})$ achieves the maximum of the above maximization problem. The latter condition in (26) implies $\hat{w}^{\mathrm{DDC}} = \hat{w}^{\mathrm{BEU}} = G^{-1}(\rho_1(a, A_1)) =: \hat{w}^*$ (the value is unique as G is strictly increasing). Then the first condition in (26) implies $1 - \rho_0(a; A_0) = G(\hat{\delta}^{\mathrm{DDC}} V(\hat{w}^*) - \hat{w}^*) = F(\hat{\delta}^{\mathrm{BEU}} V(\hat{w}^*) - \hat{w}^*)$ and the corresponding values of $\hat{\delta}^{\mathrm{DDC}}, \hat{\delta}^{\mathrm{BEU}}$ are uniquely determined (as F, G are strictly increasing and $V(\cdot) > 0$). If $\rho_0(a; A_0) > 0.5$, then $\hat{\delta}^{\mathrm{DDC}} V(\hat{w}^*) - \hat{w}^*, \hat{\delta}^{\mathrm{BEU}} V(\hat{w}^*) - \hat{w}^* < 0$. By the observation in the first paragraph, this implies $\hat{\delta}^{\mathrm{DDC}} V(\hat{w}^*) < \hat{\delta}^{\mathrm{BEU}} V(\hat{w}^*)$. Thus $\hat{\delta}^{\mathrm{DDC}} < \hat{\delta}^{\mathrm{BEU}}$ since $V(\hat{w}^*) > 0$. If $\rho_0(a; A_0) < 0.5$, a symmetric argument yields $\hat{\delta}^{\mathrm{DDC}} > \hat{\delta}^{\mathrm{BEU}}$.

By standard results (e.g., Theorem 2 in White (1982)) the maximum likelihood estimates $(\hat{w}_n^{\mathrm{M}}, \hat{\delta}_n^{\mathrm{M}})$ for each model M converge almost surely to $(\hat{w}^{\mathrm{M}}, \hat{\delta}^{\mathrm{M}})$. This completes the proof. ■

In Proposition 3, we assumed that distribution F has a symmetric and unimodal density around 0. While this assumption is satisfied by several commonly used distributions including the probit model, it rules out other instances such as the logit model. The following proposition accommodates such distributions under the assumption that F and G have finite crossings, i.e., $|\{\gamma : F(\gamma) = G(\gamma)\}| < \infty$.

Proposition G.1. Suppose that ρ is compatible with both models. If F and G have finite crossings, then there exist $\overline{\alpha}, \underline{\alpha} \in (0, 1)$ such that almost surely

(i). $\lim_n \hat{w}_n^{\mathrm{DDC}} = \lim_n \hat{w}_n^{\mathrm{BEU}}$

(ii). $\lim_n \hat{\delta}_n^{\mathrm{DDC}} < \lim_n \hat{\delta}_n^{\mathrm{BEU}}$ if $\rho_0(a; A_0) > \overline{\alpha}$ and $\lim_n \hat{\delta}_n^{\mathrm{DDC}} > \lim_n \hat{\delta}_n^{\mathrm{BEU}}$ if $\rho_0(a; A_0) < \underline{\alpha}$.

The proposition shows that the same conclusion as in Proposition 3 holds as long as period 0 choice probabilities are relatively extreme. The proof is identical to Proposition 3 except for modifying the first paragraph in the following manner. Note that F and G cross at least once since they have the same mean. By the finite crossing assumption, we can take $\overline{\gamma}$ and $\underline{\gamma}$ to be the largest and smallest crossing points of F and G. Since ε has mean zero, G is a mean-preserving spread of F by construction. Thus, since their means are finite, $\int_0^p F^{-1}(q)dq \geq \int_0^p G^{-1}(q)dq$ and $\int_p^1 F^{-1}(q)dq \leq \int_p^1 G^{-1}(q)dq$ hold for any $p \in (0,1)$ (Theorem 3.A.5 in Shaked and Shanthikumar (2007)). This implies $F(\gamma) < G(\gamma)$ for all $\gamma < \underline{\gamma}$ and $F(\gamma) > G(\gamma)$ for all $\gamma > \overline{\gamma}$. Based on this modification, the remaining proof goes through by defining $\overline{\alpha} := F(\overline{\gamma})$ and $\underline{\alpha} := F(\underline{\gamma})$.

Finally, while we have assumed that shocks to each option are identically distributed according to F, this assumption is also not crucial. In particular, suppose that the shock distribution can depend on both the option and the time period; i.e., for each $x \in \{a, b, A_1\}$ and $t \in \{0, 1\}$, ε_t^x follows some mean-zero distribution F_t^x with full-support density and all shocks are independent. In this more general case, the same argument as above yields the same predictions as Proposition G.1 as long as F_0^a and $F_0^{A_1}$ have finite crossings.

H Proofs for Section 6

We use the following preliminary lemma in the proofs.

Lemma H.1. Take any finite set of non-constant utilities $\{u^1, .., u^m\} \subseteq \mathbb{R}^Z$ and a convex set $D \subseteq \mathbb{R}^Z$ such that $\{u^1, .., u^m\} \cap [D] \neq \varnothing$. Suppose there exist $\overline{\ell}, \underline{\ell} \in \Delta(Z)$ such that $u^i(\overline{\ell}) > u^i(\underline{\ell})$ for each $i = 1, ..., m$. Then there exists a finite set $L \subseteq \Delta(Z)$ and $\ell^* \in \text{int}\Delta(Z)$ such that (i) $|M(L, u^i)| = 1$ for all u^i, (ii) $M(L, u^i) = \{\ell^*\}$ if and only if $u^i \in [D]$.

Proof. We suppose $\{u^1, .., u^m\} \not\subseteq [D]$, because otherwise we can take any lottery $\ell^* \in \text{int}\Delta(Z)$ and set $L = \{\ell^*\}$. For convenience, we relabel the utilities such that $u^i \in [D]$ for $i = 1, ..., k$ and $u^i \notin [D]$ for $i = k+1, .., m$. By the affine aggregation theorem (e.g., Theorem 2 in Fishburn (1984)), for any $u \in \mathbb{R}^Z$, the following statements are equivalent:

(i). for any $w \in \mathbb{R}^Z$ such that $\sum_{z \in Z} w(z) = 0$,

$$[\forall i = 1, ..., k, u^i \cdot w \leq 0] \Rightarrow u \cdot w \leq 0$$

(ii). $u \in [\text{co}\{u^1, ..., u^k\}]$.

Note that by definition for any $i = k+1, ..., m$, u^i does not belong to $[\text{co}\{u^1, ..., u^k\}] \subseteq [D]$. Thus, by the above equivalence result, for each $i = k+1, ..., m$, we can find a vector $w^i \in \mathbb{R}^Z$ with $\sum_{z \in Z} w^i(z) = 0$ such that $u^i \cdot w^i > 0 \geq u^j \cdot w^i$ for any $j = 1, ..., k$. Fix any $\ell \in \text{int}\Delta(Z)$. For each $i = k+1, ..., m$, we construct $\ell(i) \in \Delta(Z)$ such that the vector $\ell(i) - \ell$ (in $\mathbb{R}^Z$) is proportional to w^i. Note that such a construction is possible because ℓ is in the interior of $\Delta(Z)$. Thus $u^j(\ell) \geq \max_{i=k+1,...,m} u^j(\ell(i))$ for each $j = 1, ..., k$ and $u^i(\ell) < u^i(\ell(i))$ for each $i = k+1, ..., m$.

Let $\ell^* := \ell + \varepsilon(\overline{\ell} - \underline{\ell})$, where $\varepsilon > 0$ is small enough so that the lottery is well-defined (this is possible because ℓ is in the interior). By choosing ε small, we can guarantee that $u^j(\ell^*) > \max_{i=k+1,...,m} u^j(\ell(i))$ for each $j = 1, ..., k$ and $u^i(\ell^*) < u^i(\ell(i))$ for each $i = k+1, ..., m$. Let $L := \ell^* \cup \{\ell(i) : i = k+1, ..., m\}$. Since each utility is non-constant, up to perturbing lotteries in L, we can assume without loss that $|M(L, u^i)| = 1$ for each $i = 1, ..., m$ while preserving the above strict inequalities. This completes the proof as $M(L, u^j) = \{\ell^*\}$ for each $j = 1, ..., k$ and $M(L, u^i) \neq \{\ell^*\}$ for each $i = k+1, ..., m$. ■

H.1 Proof of Proposition 4

"If" direction: Consider any $L_0 \in \mathcal{L}_0^*$, $L_1 \in \mathcal{A}_1^*$ with $L_1 \subseteq L_0$ such that $\rho_0^Z(\ell; L_0), \rho_0^Z(\ell'; L_0) > 0$.

Let $\mathcal{U}_0(\ell) := \{u_0(\omega) : \omega \in C(L_0, \ell)\}$ and $\mathcal{U}_0(\ell') := \{u_0(\omega) : \omega \in C(L_0, \ell')\}$. Note that since L_1 features no ties, Lemma E.3 implies $C(L_1, \ell) = \{\omega : \ell \in M(L_1, u_1(\omega))\}$ by the representation in the atemporal domain. Hence

$$\rho_1^Z(\ell; L_1|L_0, \ell) = \mu(\{\ell \in M(L_1, u_1)\}|C(L_0, \ell)) \geq \min_{u \in \mathcal{U}_0(\ell)} \mu(\{p \in M(L_1, u_1)\}|\{u_0 \approx u\}). \tag{27}$$

Likewise,

$$\rho_1^Z(\ell; L_1|L_0, \ell) = \mu(\{\ell \in M(L_1, u_1)\}|C(L_0, \ell')) \leq \max_{u' \in \mathcal{U}_0(\ell')} \mu(\{\ell \in M(L_1, u_1)\}|\{u_0 \approx u'\}). \tag{28}$$

Pick $u \in \mathcal{U}_0(\ell)$ (respectively, $u' \in \mathcal{U}_0(\ell')$) which achieve the min (respectively max) in (27) (respectively, in (28)). Let $\{u_1^1, ..., u_1^m\} := \{u_1(\omega) : \omega \in C(L_0, \ell) \cup C(L_0, \ell') \text{ and } \ell \in M(L_1, u_1(\omega))\}$ and let $D := \text{co}\{u, u_1^1, ..., u_1^m\}$. Note that since $L_0 \supseteq L_1$, we have $\ell \in M(L_1, u)$. Hence, $\{\omega : u_0(\omega) \approx u, \ell \in M(L_1, u_1(\omega))\} = \{\omega : u_0(\omega) \approx u, u_1(\omega) \in [D]\}$, and likewise $\{\omega : u_0(\omega) \approx u', \ell \in M(L_1, u_1(\omega))\} = \{\omega : u_0(\omega) \approx u', u_1(\omega) \in [D]\}$. Thus,

$$\begin{aligned}&\mu(\{\ell \in M(L_1, u_1)\}| \cap \{u_0 \approx u\}) = \mu([D]|\{u_0 \approx u\}) \geq \\ &\mu([D]|\{u_0 \approx u'\}) = \mu(\{\ell \in M(L_1, u_1)\}|\{u_0 \approx u'\}),\end{aligned} \tag{29}$$

where the inequality holds by assumption. Combining (27), (28), and (29) yields $\rho_1^Z(\ell; L_1|L_0, \ell) \geq \rho_1^Z(\ell; L_1|L_0, \ell')$, as required.

"Only if" direction: We prove the contrapositive. Suppose that for some $u, u' \in \mathbb{R}^Z$ and convex $D \subseteq \mathbb{R}^Z$ with $u \in D$ such that $\mu(\{u_0 \approx u\}), \mu(\{u_0 \approx u'\}) > 0$, we have

$$\mu(\{u_1 \in [D]\}|\{u_0 \approx u\}) < \mu(\{u_1 \in [D]\}|\{u_0 \approx u'\}). \tag{30}$$

Let $\mathcal{U}_1$ be the set of possible realizations of u_1 conditional on the event $\{\omega : u_0 \approx u \text{ or } u_0 \approx u'\}$. Let $\mathcal{U}_0$ be the set of possible realizations of u_0. Enumerate $\{u_1^1, ..., u_1^m\} := \mathcal{U}_1 \cap [D]$, which is nonempty by (30).

By Condition 1, $u_t(\overline{\ell}) > u_t(\underline{\ell})$ for each $t = 1, 2$ and any possible realization u_t. Thus we can apply Lemma H.1 so that there exist some menu L_1 and ℓ^* such that (i) $M(L_1, u) = M(L_1, u_1^i) = \{\ell^*\}$ for all $i = 1, ..., m$, (ii) $|M(L_1, u_1)| = 1$ and $M(L_1, u_1) \not\ni \ell^*$ for each $u_1 \in \mathcal{U}_1 \setminus [D]$. Subject to perturbations of the lotteries in L_1, we can assume without loss that $|M(L_1, u_t)| = 1$ for each $t = 1, 2$ and any possible realization of u_t (since every such realization is non-constant). Thus $L_1 \in \mathcal{A}_1^*$ by Lemma E.3.

By construction of L_1, we have

$$\{\ell^* \in M(u_1, L_1)\} \cap \{u_0 \approx u\} = \{u_1 \in [D]\} \cap \{u_0 \approx u\}$$

Let $\{[u_0^1], ..., [u_0^k]\}$ denote the collection of equivalence classes of utilities in $\mathcal{U}_0$, and assume without loss that $u \in [u_0^1]$. By Lemma E.2, we construct a collection of consumption lotteries $\{\ell(h) : h = 1, ..., k\}$ such that $u_0(\ell(h)) > u_0(\ell(h'))$ for any distinct $h, h' = 1, ..., k$ with $u_0 \in [u_0^h]$.

Pick $\varepsilon' > 0$ sufficiently small such that $\ell^* + \varepsilon'(\ell(h) - \ell(1)) \in \Delta(Z)$ for all $h = 2, \ldots, k$; the construction is possible since ℓ^* is in the interior of $\Delta(Z)$. Define a menu L_0 by

$$L_0 := L_1 \cup \{\ell^* + \varepsilon'(\ell(h) - \ell(1)) : h = 2, ..., k\}.$$

For each $h = 2, ..., k$ and $u_0 \in [u_0^h]$, $u_0(\ell^* + \varepsilon'(\ell(h) - \ell(1)))$ is non-constant in ε'; therefore, for small enough $\varepsilon' > 0$, $M(L_0, u_0)$ is either $\{\ell^* + \varepsilon'(\ell(h) - \ell(1))\}$ or a singleton included in L_1. Furthermore,

$M(L_0, u_0) = \{\ell^*\}$ for each $u_0 \in [u_0^1]$. This ensures that $L_0 \in \mathcal{L}_0^*$ by Lemma E.3. Furthermore,

$$\{\ell^* \in M(u_0, L_0)\} = \{u_0 \approx u\}.$$

Since $u \not\approx u'$, there is a lottery $\ell_0 \in L_0$ different from ℓ^* such that $M(L_0, u') = \{\ell_0\}$.

By the previous observations, we have $L_0 \supseteq L_1$ and $\rho_1^Z(\ell^*; L_1|L_0, \ell^*) = \mu(u_1 \in [D]|u_0 \approx u)$ and $\rho_1^Z(\ell^*; L_1|L_0, \ell_0) = \mu(u_1 \in [D]|u_0 \approx u')$. But then (30) implies that $\rho_1^Z(\ell^*; L_1|L_0, \ell_0) < \rho_1^Z(\ell^*; L_1|L_0, \ell_0)$, which is a violation of consumption persistence. ■

H.2 Proof of Proposition 5

For each menu L of consumption lotteries and $\ell \in L$, recall the notation $N(L, \ell) = \{u \in \mathbb{R}^Z : u \cdot \ell \geq u \cdot \ell', \forall \ell' \in L\}$. Note that $N(L, \ell)$ is convex with $N(L, \ell) = [N(L, \ell)]$.

"If" direction: $\rho_0^Z = \hat{\rho}_0^Z$ follows directly from the condition that $\mu(\{u_0 \approx u\}) = \hat{\mu}(\{\hat{u}_0 \approx u\})$ for each $u \in \mathbb{R}^Z$.

Take any $L_0 \in \mathcal{L}_0^*$, $L_1 \in \mathcal{A}_1^*$ and $\ell \in L_0$ such that $L_0 \supseteq L_1$. Let $\{[u_0^1], ..., [u_0^k]\}$ denote the set of possible consumption preferences in period 0 that can realize with positive probabilities under μ or $\hat{\mu}$ and belong to $[N(L_0, \ell)]$. Since there is no tie in L_0 and L_1, we can write

$$\begin{aligned}
\rho_1^Z(\ell, L_1|L_0, \ell) &= \frac{\sum_{i=1}^k \mu(\{u_0 \approx u_0^i\})\mu(\{u_1 \in N(L_1, \ell)\}|\{u_0 \approx u_0^i\})}{\sum_{i=1}^k \mu(\{u_0 \approx u_0^i\})} \\
&\geq \frac{\sum_{i=1}^k \hat{\mu}(\{\hat{u}_0 \approx u_0^i\})\mu(\{\hat{u}_1 \in N(L_1, \ell)\}|\{\hat{u}_0 \approx u_0^i\})}{\sum_{i=1}^k \hat{\mu}(\{\hat{u}_0 \approx u_0^i\})} = \hat{\rho}_1^Z(\ell, L_1|L_0, \ell)
\end{aligned}$$

where the inequality follows from the condition that $\mu(\{u_0 \approx u_0^i\}) = \hat{\mu}(\{\hat{u}_0 \approx u_0^i\})$ and $\mu(\{u_1 \in N(L_1, \ell)\}|\{u_0 \approx u_0^i\}) \geq \mu(\{\hat{u}_1 \in N(L_1, \ell)\}|\{\hat{u}_0 \approx u_0^i\})$ for each $i = 1, ..., k$.

"Only if" direction: For each $u \in \mathbb{R}^Z$, $\mu(\{u_0 \approx u\}) = \hat{\mu}(\{u_0 \approx \hat{u}\})$ follows directly from $\rho_0^Z = \hat{\rho}_0^Z$. To complete the remaining part, we suppose to the contrary that there exist $u \in \mathbb{R}^Z$ and a convex set $D \ni u$ such that $\mu(\{u_1 \in [D]\}|\{u_0 \approx u\}) < \hat{\mu}(\{\hat{u}_1 \in [D]\}|\{\hat{u}_0 \approx u\})$.

Let $\{[u_0^1], ..., [u_0^k]\}$ and $\{[u_1^1], ..., [u_1^m]\}$ denote the set of possible consumption preferences that can realize with positive probabilities under μ or $\hat{\mu}$ in periods 0 and 1, respectively. Note that by the joint uniformly ranked pair condition, $u_1^i(\overline{\ell}) > u_1^i(\underline{\ell})$ for each $i = 1, ..., m$. Thus, by Lemma H.1, there exist a lottery ℓ^* and a menu L_1 such that (i) $M(L_1, u) = M(L_1, u_1^i) = \{\ell^*\}$ for each $i = 1, ..., m$ with $u_1^i \in [D]$, and (ii) $M(L_1, u_1^i) \not\ni \ell^*$ and $|M(L_1, u_1^i)| = 1$ for each $i = 1, ..., m$ with $u_1^i \notin [D]$. Thus $L_1 \in \mathcal{A}_1^*$.

Moreover, following the same construction as in the proof of Proposition 4, we construct a menu $L_0 \supseteq L_1$ such that (i) $M(L_0, u) = \{\ell^*\}$ and (ii) $M(L_0, u_0^i) \not\ni \ell^*$ and $|M(L_0, u_0^i)| = 1$ for each $i = 1, ..., k$ with $u_0^i \notin [u]$. Thus, $L_0 \in \mathcal{L}_0^*$.

Based on this, we can write the choice probabilities as

$$\begin{aligned}
\rho_1^Z(\ell^*, L_1|L_0, \ell^*) &= \mu(\{u_1 \in \{[u_1^1], ..., [u_1^k]\} \cap [D]\}|\{u_0 \approx u\}) \\
&< \hat{\mu}(\{\hat{u}_1 \in \{[u_1^1], ..., [u_1^k]\} \cap [D]\}|\{\hat{u}_0 \approx u\}) = \hat{\rho}_1^Z(\ell^*, L_1|L_0, \ell^*),
\end{aligned}$$

which contradicts the fact that ρ^Z features more consumption persistence than $\hat{\rho}^Z$. ■

H.3 Proof of Corollary 1

(i) $\implies$ (ii):

We consider the case $m \geq 2$ as otherwise the desired statement trivially holds with any α. Observe first that for any distinct indices $i, j \in \{1, \dots, m\}$, consumption persistence and its characterization (Proposition 4) imply

$$M_{ii} = \mu(\{u_1 \in [u^i]\}|u_0 = u^i) \geq \mu(\{u_1 \in [u^i]\}|u_0 = u^j) = M_{ji} \tag{31}$$

by taking $D = \{u^i\}$. (Note that by definition both u^i and u^j arise with positive probability in period 0). Moreover, if $D = \text{co}\{u^i, u^j\}$, then by the non-collinearity assumption there is no $k \notin \{i, j\}$ such that $u^k \in [D]$. Thus, by consumption persistence and its characterization (Proposition 4),

$$M_{ii} + M_{ij} = \mu(\{u_1 \in [D]\}|u_0 = u^i) = \mu(\{u_1 \in [D]\}|u_0 = u^j) = M_{jj} + M_{ji}. \tag{32}$$

Suppose first that $m = 2$. Since $1 = M_{11}+M_{12} = M_{22}+M_{21}$, we have $M_{11}-M_{21} = M_{22}-M_{12} := \alpha$, which is nonnegative by (31). Since the Markov chain is irreducible, $M_{21}, M_{12} > 0$, which also ensures $\alpha < 1$. One can verify the desired form by setting $\nu(u^1) = \frac{M_{21}}{1-\alpha}$ and $\nu(u^2) = \frac{M_{12}}{1-\alpha}$.

Suppose next that $m \geq 3$. Take any distinct $i, j, k \in \{1, ..., m\}$ and let $D' = \text{co}\{u^i, u^j, u^k\}$. By non-collinearity, there is no $l \notin \{i, j, k\}$ such that $u^l \in [D']$. Thus, by consumption persistence and its characterization (Proposition 4),

$$M_{ii} + M_{ij} + M_{ik} = \mu(\{u_1 \in [D']\}|u_0 = u^i) = \mu(\{u_1 \in [D']\}|u_0 = u^j) = M_{jj} + M_{ji} + M_{jk}.$$

Combined with (32), this implies that $M_{ik} = M_{jk}$ for any distinct i, j, k. Thus, for any k, we can define $\beta_k := M_{ik}$ for some arbitrary $i \neq k$. Here $\beta_k > 0$, because otherwise $\sum_{i \text{ s.t. } i \neq k} M_{ik} = 0$, contradicting irreducibility of the Markov chain. By (32), $M_{ii} - M_{ji} = M_{jj} - M_{ij}$ for any i, j, and thus $M_{ii} - \beta_i = M_{jj} - \beta_j =: \alpha$ for any i, j. By (31) $\alpha \geq 0$, and $\alpha < 1$ as $\beta_k > 0$ for all k. Thus, setting $\nu(u^j) = \frac{\beta_j}{1-\alpha}$ for each j yields to the desired form.

(ii) $\implies$ (i):

Take any pair $u, u' \in \mathbb{R}^Z$ of possible realizations of period 0 felicities. Then for any convex set $D \subseteq \mathbb{R}^Z$ with $u \in D$, by (ii) we have

$$\mu(\{u_1 \in [D]\}|u_0 \approx u) = \alpha + (1-\alpha) \sum_{u^j \in [D]} \nu(u^j) \geq \alpha I_{u' \in [D]} + (1-\alpha) \sum_{u^j \in [D]} \nu(u^j) = \mu(\{u_1 \in [D]\}|u_0 \approx u').$$

Thus, ρ features consumption persistence by Proposition 4.

(ii) $\implies$ (iii):

Note that for any $L = \{\ell^1, ..., \ell^m\} \in \mathcal{L}_0^*$ and distinct indices i, j, we have $\rho_1^Z(\ell^i; L|, L, \ell^i) = \alpha + (1-\alpha) \sum_{u^k \in N(L, \ell^i)} \nu(u^k) = \alpha + (1-\alpha)\rho_0^Z(\ell^i; L)$ and $\rho_1^Z(\ell^j; L|L, \ell^i) = (1-\alpha) \sum_{u^k \in N(L, \ell^j)} \nu(u^k) = (1-\alpha)\rho_0^Z(\ell^j; L)$.

(iii) $\implies$ (ii):

Since $\{u^1, ..., u^m\}$ are ordinally distinct, Lemma E.2 yields $L = \{\ell^1, ..., \ell^m\}$ such that $M(L, u^i) = \{\ell^i\}$ for each i. Then by the Markov representation we have $\rho_1^Z(\ell^j; L|L, \ell^i) = M_{ij}$ and $\rho_0^Z(\ell^i; L) = \nu(u^i)$ for all indices i, j. Thus, by (iii), there exists $\beta \in [0, 1)$ such that $M_{ii} = \beta + (1-\beta)\nu(u^i)$ and $M_{ij} = (1-\beta)\nu(u^j)$ for all $i \neq j$, which verifies (ii). ■

H.4 Proof of Corollary 2

Since ρ and $\hat{\rho}$ admit sticky i.i.d. representations, for each $\ell \in L_1 \subseteq L_0$ with $L_0 \in \mathcal{L}_0^*$ and $L_1 \in \mathcal{A}_1^*$, choice probabilities satisfy:

$$\rho_0^Z(\ell; L_0) = \sum_{u^i \in N(L_0,\ell)} \nu(u^i), \quad \hat{\rho}_0^Z(\ell; L_0) = \sum_{u^i \in N(L_0,\ell)} \hat{\nu}(u^i),$$

$$\rho_1^Z(\ell; L_1|L_0,\ell) = \alpha + (1-\alpha) \sum_{u^i \in N(L_1,\ell)} \nu(u^i), \quad \hat{\rho}_1^Z(\ell; L_1|L_0,\ell) = \hat{\alpha} + (1-\hat{\alpha}) \sum_{u^i \in N(L_1,\ell)} \hat{\nu}(u^i).$$

The "if" direction is immediate from these expressions. For the "only if" direction, the existence of the bijection ϕ follows from the fact that ρ^Z and $\hat{\rho}^Z$ coincide on period 0 consumption choices and the assumption that in each representation all felicities are ordinally distinct. To show that $\alpha \geq \hat{\alpha}$, consider any $\ell \in L_1 \subseteq L_0$ (with $L_0 \in \mathcal{L}_0^*$ and $L_1 \in \mathcal{A}_1^*$) such that $\sum_{u^i \in N(L_1,\ell)} \nu(u^i) = \sum_{u^i \in N(L_1,\ell)} \hat{\nu}(u^i) < 1$. Then $\rho_1^Z(\ell; L_1|L_0,\ell) \geq \hat{\rho}_1^Z(\ell; L_1|L_0,\ell)$ implies $\alpha \geq \hat{\alpha}$. ■

H.5 Proof of Proposition 6

Necessity:

Take any $L \in \mathcal{L}_0^*$ and $\ell, \ell' \in L$ with $\{\ell, \ell'\} \in \mathcal{A}_1^*$. If $\rho_0^Z(\ell; L) > 0$, then there exists $u \in N(L,\ell)$ such that $\mu(\{u_0 = u\}) > 0$. This implies $u(\ell) > u(\ell')$. Then by (2), there exists some $u' \in \mathbb{R}^Z$ with $\mu(\{u_1 = u'\}|\{u_0 = u\}) > 0$ such that $u'(\ell) > u'(\ell')$. This ensures $\rho_1^Z(\ell; \{\ell,\ell'\}|L,\ell) > 0$ because $\mu(\{u_1 \approx u'\}|\{u_0 \in N(L,\ell)\}) > 0$.

Sufficiency:

Take a BEU representation $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, u_t, \delta_t, W_t))$ of ρ^Z. Let $\hat{\mathcal{F}}_0$ be the sigma algebra generated by the random equivalence class $[u_0]$, i.e., $\hat{\mathcal{F}}_0$ is induced by the finest partition over Ω such that $u_0(\cdot)$ corresponds to the same preference within each cell. Likewise, let $\hat{\mathcal{F}}_1$ be the sigma algebra generated by the random sequence of equivalence classes $([u_0],[u_1])$. Note that $\hat{\mathcal{F}}_0 \subseteq \mathcal{F}_0$ and $\hat{\mathcal{F}}_1 \subseteq \mathcal{F}_1$. For each $t = 0,1$, construct an $\hat{\mathcal{F}}_t$-measurable function $\hat{u}_t$ such that $\hat{u}_t(\omega) \approx u_t(\omega)$ and $\sum_z \hat{u}_t(\omega)(z) = 0$ for each ω.

We consider a tuple $(\Omega, \mathcal{F}^*, \mu, (\hat{\mathcal{F}}_t, \hat{U}_t, \hat{u}_t, \delta_t, \hat{W}_t))$, where $(\hat{U}_t)$ is induced from $(\mu, (\hat{\mathcal{F}}_t, \hat{u}_t, \delta_t))$ by equation (1), and $(\hat{W}_t)$ is any $\mathcal{F}^*$-measurable tiebreaker that satisfies the properness condition with respect to $(\mu, (\hat{\mathcal{F}}_t))$. This tuple is clearly a BEU representation of ρ^Z, since (u_t) and $(\hat{u}_t)$ are ordinally equivalent at every state.[84]

Next we fix any $\hat{u} \in \mathbb{R}^Z$ such that $\mu(\{\hat{u}_0 = \hat{u}\}) > 0$, and let $\mathcal{U}_{\hat{u}} := \{\hat{u}' \in \mathbb{R}^Z : \mu(\{\hat{u}_1 = \hat{u}'\}|\{\hat{u}_0 = \hat{u}\}) > 0\}$. We now use Axiom 11 (consumption inertia) to show that $\hat{u} \in \mathrm{co}\mathcal{U}_{\hat{u}}$. By the affine aggregation theorem (e.g., Theorem 2 in Fishburn (1984)), it suffices to establish that for all $\ell, \ell' \in \Delta(Z)$, we have

$$[\hat{u}'(\ell') \geq \hat{u}'(\ell), \forall \hat{u}' \in \mathcal{U}_{\hat{u}}] \Rightarrow \hat{u}(\ell') \geq \hat{u}(\ell).$$

Suppose to the contrary that $[\hat{u}'(\ell') \geq \hat{u}'(\ell), \forall \hat{u}' \in \mathcal{U}_{\hat{u}}]$ and $\hat{u}(\ell') < \hat{u}(\ell)$ for some ℓ, ℓ'. By the Uniformly Ranked Pair condition, we have $\hat{u}'(\overline{\ell}) > \hat{u}'(\underline{\ell})$ for all $\hat{u}' \in \mathcal{U}_{\hat{u}}$. Thus, by mixing $\overline{\ell}$ with ℓ' (resp. $\underline{\ell}$ with ℓ) with a small weight on $\overline{\ell}$ (resp. $\underline{\ell}$), we can assume without loss that $[\hat{u}'(\ell') > \hat{u}'(\ell), \forall \hat{u}' \in \mathcal{U}_{\hat{u}}]$ and $\hat{u}(\ell) > \hat{u}(\ell')$. In addition, since the relevant inequalities are all strict, we can assume that $\ell, \ell' \in \mathrm{int}\Delta(Z)$ and $\{\ell, \ell'\} \in \mathcal{A}_1^*$. Take a menu of consumption lotteries $L \in \mathcal{L}_0^*$ such that $\ell, \ell' \in L$, $M(L, \hat{u}) = \{\ell\}$, and $M(L, \hat{u}'') \not\ni \ell$ for all other period 0 felicities $\hat{u}'' \neq \hat{u}$ that can realize with

[84] Note that the exact specification of $(\hat{W}_t)$ is irrelevant in this argument because we restrict attention to menus without ties.

positive probability under μ.[85] For this menu L, it follows that $\rho_0^Z(\ell; L) = \mu(\{\hat{u}_0 = \hat{u}\}) > 0$ and $\rho_1^Z(\ell; \{\ell, \ell'\}|L, \ell) = \mu(\{\hat{u}_1(\ell) > \hat{u}_1(\ell')\}|\{\hat{u}_0 = u\}) = 0$, contradicting consumption inertia.

The observation in the previous paragraph implies that for each $\hat{u} \in \mathbb{R}^Z$ such that $\mu(\{\hat{u}_0 = \hat{u}\}) > 0$, there exist constants $(\alpha_{\hat{u},\hat{u}'})_{\hat{u}' \in \mathcal{U}_{\hat{u}}} \geq 0$ and $\beta_{\hat{u}} \in \mathbb{R}$ such that

$$\hat{u} = \sum_{\hat{u}' \in \mathcal{U}_{\hat{u}}} \alpha_{\hat{u},\hat{u}'} \hat{u}' + \beta_{\hat{u}}. \tag{33}$$

Since by construction $\sum_z \hat{u}_t(\omega)(z) = 0$ at every state ω and period t, we must have $\beta_{\hat{u}} = 0$.

Define $\hat{u}_0'(\omega) := \hat{u}_0(\omega)$ and $\hat{u}_1'(\omega) := \frac{\alpha_{\hat{u}_0(\omega),\hat{u}_1(\omega)}}{\mu(E_1(\omega)|E_0(\omega))}\hat{u}_1(\omega)$ for each ω, where $E_t(\cdot)$ denotes each cell of the partition that generates $\hat{\mathcal{F}}_t$ for $t = 1, 2$. Note that each $\hat{u}_t'$ is $\hat{\mathcal{F}}_t$-measurable. We consider the tuple $(\Omega, \mathcal{F}^*, \mu, (\hat{\mathcal{F}}_t, \hat{U}_t', \hat{u}_t', \delta_t, \hat{W}_t))$, where $(\hat{U}_t')$ is induced from $(\mu, (\hat{\mathcal{F}}_t, \hat{u}_t', \delta_t))$ by equation (1). This tuple is still a BEU representation of ρ^Z, since $(\hat{u}_t')$ and $(\hat{u}_t)$ are ordinally equivalent at every state.

To conclude that the representation is BEB, we verify that (2) holds with $\tilde{u} := \hat{u}_1'$. That is, for each ω

$$\mathbb{E}[\hat{u}_1'|\hat{\mathcal{F}}_0(\omega)] = \sum_{E_1 \subseteq E_0(\omega)} \mu(E_1|E_0(\omega))\hat{u}_1'(E_1) = \sum_{E_1 \subseteq E_0(\omega)} \alpha_{\hat{u}_0(\omega),\hat{u}_1(E_1)}\hat{u}_1(E_1) = \hat{u}_0(\omega) = \hat{u}_0'(\omega)$$

where the second and fourth equalities hold by definition of $\hat{u}_t'$ and the third equality uses (33) with $\beta_{\hat{u}_0(\omega)} = 0$ for each ω. ■

I Additional Results

I.1 Identification

The following proposition provides identification results for our representations (see Remark 1 for the discussion).

Proposition I.1. Suppose ρ and $\hat{\rho}$ admit DREU representations $\mathcal{D} = (\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t))$ and $\hat{\mathcal{D}} = (\hat{\Omega}, \hat{\mathcal{F}}^*, \hat{\mu}, (\hat{\mathcal{F}}_t, \hat{U}_t, \hat{W}_t))$, with partitions Π_t and $\hat{\Pi}_t$ generating $\mathcal{F}_t$ and $\hat{\mathcal{F}}_t$, respectively. Then $\rho = \hat{\rho}$ if and only if for each t there exists a bijection $\phi_t : \Pi_t \to \hat{\Pi}_t$ and $\mathcal{F}_t$-measurable functions $\alpha_t : \Omega \to \mathbb{R}_{++}$ and $\beta_t : \Omega \to \mathbb{R}$ such that for all $\omega \in \Omega$:

(i). $\mu(\mathcal{F}_0(\omega)) = \hat{\mu}(\phi_0(\mathcal{F}_0(\omega)))$ and $\mu(\mathcal{F}_t(\omega)|\mathcal{F}_{t-1}(\omega)) = \hat{\mu}(\phi_t(\mathcal{F}_t(\omega))|\phi_{t-1}(\mathcal{F}_{t-1}(\omega)))$ if $t \geq 1$;

(ii). $U_t(\omega) = \alpha_t(\omega)\hat{U}_t(\hat{\omega}) + \beta_t(\omega)$ whenever $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$;

(iii). $\mu[\{W_t \in B_t(\omega)\}|\mathcal{F}_t(\omega)] = \hat{\mu}[\{\hat{W}_t \in B_t(\omega)\}|\phi_t(\mathcal{F}_t(\omega))]$ for any $B_t(\omega)$ such that $B_t(\omega) = \{w \in \mathbb{R}^X : p_t \in M(M(A_t, U_t(\omega)), w)\}$ for some $p_t \in A_t \in \mathcal{A}_t$.

[85] To see why such a construction is possible, first note that all the possible realizations of period 0 felicities $\hat{u}_0(\cdot)$ are ordinally distinct by construction. Take a set of consumption lotteries $\bar{L}$ that separates all the period 0 felicities $\hat{u}_0(\cdot)$ by Lemma E.2. Here we can assume that the sup-norm distance among these lotteries is bounded by ε by mixing them to a common lottery if necessary, where $\varepsilon := \min_{z \in Z}\{\ell(z), 1 - \ell(z)\} > 0$. Let $\bar{\ell} \in \bar{L}$ be the lottery that strictly maximizes $\hat{u}$ in $\bar{L}$. Then we define $\bar{L}^* := \{\ell + \bar{\ell}'' - \bar{\ell} : \bar{\ell}'' \in \bar{L}\}$. This is a well-defined set of lotteries by the construction of ε. Note that this set also separates all period 0 felicities. Then the desired set L can be constructed by adding ℓ' to $\bar{L}^*$ such that there is no tie (that is guaranteed by slightly perturbing lotteries if necessary).

If $(\mathcal{D}, (u_t, \delta_t))$ is a BEU representation of ρ, then $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ is a BEU representation of ρ if and only if (i)-(iii) hold and additionally, for all $t = 0, \dots, T$:

(iv). $\alpha_t(\omega) = \alpha_0(\omega) \prod_{\tau=0}^{t-1} \frac{\hat{\delta}_\tau(\hat{\omega})}{\delta_\tau(\omega)}$ whenever $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$

(v). $u_t(\omega) = \alpha_t(\omega)\hat{u}_t(\hat{\omega}) + \gamma_t(\omega)$ whenever $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$, where $\gamma_T(\omega) := \beta_T(\omega)$ and $\gamma_t(\omega) := \beta_t(\omega) - \delta_t(\omega)\mathbb{E}_\mu[\beta_{t+1}|\mathcal{F}_t(\omega)]$ if $t \leq T-1$.

If $(\mathcal{D}, (u_t, \delta_t))$ is a BEB representation of ρ that satisfies Condition D.1, then $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ is a BEB representation of ρ if and only if (i)-(v) hold and additionally, for all $t = 0, \dots, T-1$:

(vi). $\delta_t(\omega) = \hat{\delta}_t(\hat{\omega})$ for all $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$

(vii). $\gamma_t(\omega) = \mathbb{E}_\mu[\beta_T|\mathcal{F}_t(\omega)]$ for all ω.

Proof. See Appendix J.3. ■

I.2 Markov Evolving Utility

Definition 14. A *(stationary) Markov evolving utility* representation is a BEU representation $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t, u_t, \delta_t))$ for which there exists a finite set of felicities $\mathcal{U} = \{u^1, u^2, \dots, u^m\} \subseteq \mathbb{R}^Z$, with $u^i \not\approx u^j$ for all $i \neq j$, along with a stationary distribution $\xi \in \Delta^\circ(\mathcal{U})$ and a right stochastic transition matrix $\Pi = (\Pi_{i,j})_{i,j=1,\dots,m}$ such that

(i). $\mu(u_t(\omega) \approx u^i) = \xi(u^i)$ for all $t = 0, \dots, T$ and $i = 1, \dots, m$;

(ii). $\mu(u_{t+1}(\omega) \approx u_{t+1}|u_0(\omega) \approx u_0, \dots, u_{t-1}(\omega) \approx u_{t-1}, u_t(\omega) \approx u_t) = \mu(u_{t+1}(\omega) \approx u_{t+1}|u_t(\omega) \approx u_t)$ for all $t = 0, \dots, T-1$ and $u_0, \dots, u_{t+1} \in \mathcal{U}$;

(iii). $\mu(u_{t+1}(\omega) \approx u^j|u_t(\omega) \approx u^i) = \Pi_{i,j}$ for all $t = 0, 1, \dots, T-1$ and $i, j = 1, \dots, m$.

We assume that ρ admits a BEU representation. As in Section 6, we consider the restriction ρ^Z of ρ to atemporal consumption problems without ties; this is well-defined given the assumption that ρ admits a BEU representation. For each $\ell_{T-1} \in \Delta(Z)$ and $L_{T-1}, L_T \in \mathcal{K}(\Delta(Z))$, we define the lottery $(\ell_{T-1}, L_T) := (\delta_{\ell_{T-1}}, \delta_{L_T})$ and menu $(L_{T-1}, L_T) := \{(\ell'_{T-1}, L_T) : \ell'_{T-1} \in L_{T-1}\}$. Recursively, for each $t \leq T-2$, $\ell_t \in \Delta(Z)$, and $L_t, ..., L_{T-1} \in \mathcal{K}(\Delta(Z))$, we define the lottery $(\ell_t, L_{T-1}, ..., L_T) := (\delta_{\ell_t}, \delta_{(L_{t+1}, ..., L_T)})$ and menu $(L_t, ..., L_T) := \{(\ell'_t, L_{t+1}, ..., L_T) : \ell'_t \in L_t\}$.

Let $\mathcal{L}_0^* \subseteq \mathcal{K}(\Delta(Z))$ denote the set of period 0 *consumption menus without ties*, which consists of all L_0 such that $(L_0, L_1) \in \mathcal{A}_0^*$ for all $L_1 \in \mathcal{K}(\Delta(Z))$. For any $L_0 \in \mathcal{L}_0^*$ and $\ell_0 \in L_0$, define $\rho_0^Z(\ell_0; L_0) := \rho_0((\ell_0, L_1); (L_0, L_1))$ for an arbitrary choice of L_1. This induces the set of all period 0 *consumption histories without ties*, i.e., sequences $h_Z^0 = (L_0, \ell_0)$ such that $\rho_0^Z(\ell_0, L_0) > 0$ and $L_0 \in \mathcal{L}_0^*$. Recursively, for each period $t-1$ consumption history without ties $h_Z^{t-1} = (L_0, \ell_0, ..., L_{t-1}, \ell_{t-1})$, we denote by $\mathcal{L}_t^*(h_Z^{t-1})$ the set of period t consumption menus without ties conditional on h_Z^{t-1}, which consists of all L_t such that $(L_t, L_{t+1}, .., L_T) \in \mathcal{A}_t^*(h^{t-1})$ for all $L_{t+1}, .., L_T$, where $h^{t-1} = (A_0, p_0, ..., A_{t-1}, p_{t-1})$ is given by $A_\tau = (L_\tau, L_{\tau+1}, ..., L_T)$ and $p_\tau = (\ell_\tau, L_{\tau+1}, ..., L_T)$ for each $\tau = 0, ..., t-1$. Given any such h_Z^{t-1} and h^{t-1}, we define $\rho_t^Z(\ell_t, L_t|h_Z^{t-1}) := \rho_t((\ell_t, L_{t+1}, ..., L_T), (L_t, ..., L_T)|h^{t-1})$ for each $L_t^* \in \mathcal{L}_t(h_Z^{t-1})$ and $\ell_t \in L_t$; if $\rho_t^Z(\ell_t, L_t|h_Z^{t-1}) > 0$ then we say that the sequence $(L_0, \ell_0, ..., L_t, \ell_t)$ is a consumption history without ties in period t. Finally, we say that a consumption history without ties is *degenerate* if the corresponding L_τ's are all singleton.

Axiom I.1 (Unconditional Stationarity)**.** For all degenerate consumption histories d_Z^{t-1}, $L \in \mathcal{L}_t^*(d_Z^{t-1}) \cap \mathcal{L}_0^*$, and $\ell \in L$, we have $\rho_0^Z(\ell, L) = \rho_t^Z(\ell, L|d_Z^{t-1})$.

A *consumption atom* is a pair (L, ℓ) with $L \in \mathcal{L}_0^*$ and $\ell \in \Delta^\circ(Z)$ such that

(i). $\rho_0^Z(\ell, L) > 0$;

(ii). $\rho_0^Z(\ell, L') \in \{\rho_0^Z(\ell, L), 0\}$ for all $L' \in \mathcal{L}_0^*$ with $L' \supseteq L$.

Axiom I.2 (Markov)**.** For any consumption atom (L, ℓ) and consumption history h_Z^{t-1} without ties, we have $\rho_1^Z(\cdot|L, \ell) = \rho_{t+1}^Z(\cdot|h_Z^{t-1}, L, \ell)$.

Proposition I.2. Suppose that ρ admits a BEU representation that satisfies Condition D.1 (Uniformly Ranked Pair). Then ρ^Z satisfies Axioms I.1 and I.2 if and only if it admits a Markov evolving utility representation.

Proof. See Appendix J.4. ■

I.3 Approximate Contraction History Independence

Consider the following strengthening of Axiom 1:

Axiom I.3 (Approximate Contraction History Independence)**.** If $(A_0, p_0), (B_0, p_0) \in \mathcal{H}_0(A_1)$ and $B_0 \supseteq A_0$, then for all $p_1 \in A_1$,

$$\Big|\rho_1(p_1; A_1|A_0, p_0) - \rho_1(p_1; A_1|B_0, p_0)\Big| \leq 2\Big|1 - \frac{\rho_0(p_0; B_0)}{\rho_0(p_0; A_0)}\Big|.$$

Clearly, Axiom I.3 implies Axiom 1, but it also captures that whenever $\rho_0(p_0; B_0)$ and $\rho_0(p_0; A_0)$ are close, then period-1 choice probabilities following (A_0, p_0) and (B_0, p_0) must be close. The following proposition shows that Axiom I.3 remains necessary under DREU.[86]

Proposition I.3. Suppose $T = 1$ and ρ admits a DREU representation. Then ρ satisfies Axiom I.3.

Proof. Define the following subsets of Ω:

$$E := C(p_1, A_1), \qquad F := C(p_0, A_0), \qquad G := C(p_0, B_0).$$

Note that $G \subseteq F$ and let $H := F \smallsetminus G$. We have

$$\mu(E|F) = \mu(E|G)\mu(G|F) + \mu(E|H)\mu(H|F).$$

Thus,

$$\begin{aligned} |\rho_1(p_1; A_1|A_0, p_0) - \rho_1(p_1; A_1|B_0, p_0)| &= \Big|\mu(E|F) - \mu(E|G)\Big| \\ &\leq \Big|1 - \mu(G|F)\Big| + \mu(H|F) \\ &= 2\Big|1 - \frac{\mu(G)}{\mu(F)}\Big| = 2\Big|1 - \frac{\rho_0(p_0; B_0)}{\rho_0(p_0; A_0)}\Big|. \end{aligned}$$

■

J Proofs for Sections A, E, and I

J.1 Proof of Proposition A.1

The following three subsections prove Proposition A.1, that is, the equivalence between DREU, BEU, BEB and their respective S-based analogs.

[86] Note that the proof does not exploit expected utility; thus, the axiom holds under any dynamic random utility model.

J.1.1 DREU

"If" direction: Suppose ρ admits an S-based DREU representation $(S_t, \{\mu_t^{s_{t-1}}\}_{s_{t-1}\in S_{t-1}}, \{U_{s_t}, \tau_{s_t}\}_{s_t \in S_t})_{t=0,\dots,T}$. We will construct a DREU representation $(\hat{\Omega}, \hat{\mathcal{F}}^*, \hat{\mu}, (\hat{\mathcal{F}}_t, \hat{U}_t, \hat{W}_t))$.

Consider the space $G := \prod_{t=0}^T \left(S_t \times \mathbb{R}^{X_t}\right)$ of all sequences of states and tie-breaking utilities. Let $\hat{\Omega} := \{(s_0, W_0, \dots, s_T, W_T) \in G : \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k) > 0\}$. Let $\hat{\mathcal{F}}^*$ be the restriction to $\hat{\Omega}$ of the product sigma-algebra of the discrete sigma-algebra on $\prod_{t=0}^T S_t$ and the product Borel sigma-algebra on $\prod_{t=0}^T \mathbb{R}^{X_t}$. For each $K = (\{s_0\}, K_0, ..., \{s_T\}, K_T) \in \hat{\mathcal{F}}^*$, let $\hat{\mu}(K) = \prod_{t=0}^T \mu_t^{s_{t-1}}(s_t)\tau_{s_t}(K_t)$; by finiteness of $\prod_{t=0}^T S_t$, $\hat{\mu}$ extends to a finitely-additive probability measure on $\hat{\Omega}$ in the natural way.

Let Π_t be the finite partition of $\hat{\Omega}$ whose cells are all the cylinders $C(s_0, \dots, s_t) := \{\hat{\omega} \in \hat{\Omega} : \text{proj}_{S_0\times\ldots\times S_t}(\hat{\omega}) = (s_0, \dots, s_t)\}$. Let $\hat{\mathcal{F}}_t$ be the sigma-algebra generated by Π_t; by definition of $\hat{\Omega}$, $\mu(\hat{\mathcal{F}}_t(\hat{\omega})) > 0$ for all $\hat{\omega} \in \hat{\Omega}$. Also, $\hat{\mathcal{F}}_t(\hat{\omega}) = \bigcup_{\hat{\omega}' \in \hat{\mathcal{F}}_t(\hat{\omega})} \hat{\mathcal{F}}_{t+1}(\hat{\omega}')$, so $(\hat{\mathcal{F}}_t)_{0\le t\le T} \subseteq \hat{\mathcal{F}}^*$ is a filtration. Define $\hat{U}_t : \hat{\Omega} \to \mathbb{R}^{X_t}$ by $\hat{U}_t(\hat{\omega}) = U_{s_t}$ where $\text{proj}_{S_t}(\hat{\omega}) = s_t$. Note that $(\hat{U}_t)$ is adapted to $(\hat{\mathcal{F}}_t)$ and that $\hat{U}_t(\hat{\omega})$ is nonconstant for each $\hat{\omega}$ since each U_{s_t} is nonconstant. Finally, if $\mathcal{F}_{t-1}(\hat{\omega}) = \mathcal{F}_{t-1}(\hat{\omega}')$ and $\mathcal{F}_t(\hat{\omega}) \neq \mathcal{F}_t(\hat{\omega}')$, then $\text{proj}_{S_{t-1}}(\hat{\omega}) = \text{proj}_{S_{t-1}}(\hat{\omega}) = s_{t-1}$ and $\text{proj}_{S_t}(\hat{\omega}) = s_t \neq s'_t = \text{proj}_{S_t}(\hat{\omega}')$ for some $s_{t-1} \in S_{t-1}$ and $s_t, s'_t \in \text{supp}\, \mu_t^{s_{t-1}}$. By DREU1 (a), this implies $\hat{U}_t(\hat{\omega}) := U_{s_t} \not\approx U_{s'_t} =: \hat{U}_t(\hat{\omega}')$. Thus, $(\mathcal{F}_t, U_t)$ are simple.

Define $\hat{W}_t : \hat{\Omega} \to \mathbb{R}^{X_t}$ by $\hat{W}_t(\hat{\omega}) = W_t$ where $\text{proj}_{\mathbb{R}^{X_t}}(\hat{\omega}) = W_t$. Note that for all A_t, $\hat{\mu}(\{\hat{\omega} \in \hat{\Omega} : |M(A_t, \hat{W}_t)| = 1\}) = \sum_{(s_0,\dots,s_T)} \left(\prod_{k=0}^T \mu_k^{s_{k-1}}(s_k)\right) \tau_{s_t}(\{W_t \in \mathbb{R}^{X_t} : |M(A_t, W_t)| = 1\}) = 1$, since each τ_{s_t} is proper. Thus, $(\hat{W}_t)$ satisfies part (i) of the properness requirement for DREU. Moreover, for any $\hat{\mathcal{F}}_T(\hat{\omega}) = C(s_0, \dots, s_T)$ and any sequence (B_t) of Borel sets $B_t \subseteq \mathbb{R}^{X_t}$, the definition of $\hat{\mu}$ implies

$$\hat{\mu}\left(\bigcap_{t=0}^T \{\hat{W}_t \in B_t\} | C(s_0, \dots, s_T)\right) = \prod_{t=0}^T \tau_{s_t}(B_t) = \prod_{t=0}^T \hat{\mu}\left(\{\hat{W}_t \in B_t\} \mid C(s_0, \dots, s_t)\right). \tag{34}$$

Since $\hat{\mathcal{F}}_T(\hat{\omega}) = C(s_0, \dots, s_T)$ implies $\hat{\mathcal{F}}_t(\hat{\omega}) = C(s_0, \dots, s_t)$ for all $t \le T$, this shows that $(\hat{W}_t)$ also satisfies parts (ii) and (iii) of the properness requirement.

Finally, to see that $(\hat{\Omega}, \hat{\mathcal{F}}^*, \hat{\mu}, (\hat{\mathcal{F}}_t, \hat{U}_t, \hat{W}_t))$ represents ρ, fix any $h^t = (A_0, p_0, ..., A_t, p_t) \in \mathcal{H}_t$. Then

$$\hat{\mu}(C(h^t)) = \hat{\mu}\left(\bigcap_{k=0}^t \{\hat{\omega} \in \hat{\Omega} : p_k \in M(M(A_k, \hat{U}_k(\hat{\omega})), \hat{W}_k(\hat{\omega}))\}\right) =$$
$$\sum_{C(s_0,...,s_t)\in\Pi_t} \hat{\mu}(C(s_0, ..., s_t))\hat{\mu}\left(\bigcap_{k=0}^t \{\hat{\omega} \in \hat{\Omega} : p_k \in M(M(A_k, \hat{U}_k), \hat{W}_k)\} | C(s_0, ..., s_t)\right) =$$
$$\sum_{(s_0,...,s_t)\in S_0\times\ldots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\hat{\mu}\left(\bigcap_{k=0}^t \{\hat{\omega} \in \hat{\Omega} : p_k \in M(M(A_k, U_{s_k}), \hat{W}_k)\} | C(s_0, ..., s_t)\right) =$$
$$\sum_{(s_0,...,s_t)\in S_0\times\ldots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)$$

where the third equality follows from the definition of $\hat{\mu}$ and $\hat{U}$, and the final equality follows from (34). Thus, as required, we have

$$\hat{\mu}(C(p_t, A_t)|C(h^{t-1})) = \frac{\hat{\mu}(C(h^t))}{\hat{\mu}(C(h^{t-1})} = \frac{\sum_{(s_0,\dots,s_t)} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)}{\sum_{(s_0,\dots,s_{t-1})} \prod_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)} = \rho_t(p_t; A_t|h^{t-1}),$$

where the final equality holds by DREU2.

"Only if" direction: Take any DREU representation $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t))$ of ρ. We will construct an S-based DREU representation $(S_t, \{\hat{\mu}_t^{s_{t-1}}\}_{s_{t-1}\in S_{t-1}}, \{\hat{U}_{s_t}, \tau_{s_t}\}_{s_t\in S_t})_{t=0,\dots,T}$.

For each t, let $S_t := \{\mathcal{F}_t(\omega) : \omega \in \Omega\}$ denote the partition generating $\mathcal{F}_t$, which is finite since $(\mathcal{F}_t)$ is simple. Each $\hat{\mu}^{s_t}_{t+1}$ is defined to be the one-step-ahead conditional of μ, i.e., $\hat{\mu}_0(s_0) := \mu(s_0)$ for all $s_0 \in S_0$ and $\hat{\mu}^{s_t}_{t+1}(s_{t+1}) := \mu(s_{t+1}|s_t)$ for all $s_t \in S_t$, $s_{t+1} \in S_{t+1}$. This is well-defined since $\mu(\mathcal{F}_t(\omega)) > 0$ for all ω. For each $s_t \in S_t$, define $\hat{U}_{s_t} := U_t(\omega)$ if $\omega \in s_t$; this is well-defined as (U_t) is $\mathcal{F}_t$-adapted and each U_{s_t} is nonconstant since each $U_t(\omega)$ is nonconstant. Finally, for any Borel set $B_t \subseteq \mathbb{R}^{X_t}$, define $\tau_{s_t}(B_t) := \mu(\{W_t \in B_t\}|s_t)$. This is well-defined since W_t is $\mathcal{F}^*$-measurable. Moreover, because $\mu(\{\omega \in \Omega : |M(A_t, W_t(\omega)| = 1\} = 1$ for all A_t and $|S_t|$ is finite, it follows that $\tau_{s_t}(N(A_t, p_t)) = \tau_{s_t}(N^+(A_t, p_t))$ for all p_t, i.e., τ_{s_t} is proper. Thus, each $(S_t, \mu^{s_{t-1}}_t, \{U_{s_t}, \tau_{s_t}\}_{s_t \in S_t})$ is an REU form on X_t.

Moreover, (a) for any distinct $s_t, s'_t \in \operatorname{supp}(\mu^{s_{t-1}}_t)$, we have ω, ω' such that $\mathcal{F}_{t-1}(\omega) = s_{t-1} = \mathcal{F}_{t-1}(\omega')$ and $\mathcal{F}_t(\omega) = s_t \neq \mathcal{F}_t(\omega') = s'_t$. Thus, $\hat{U}_{s_t} = U_t(\omega) \not\approx U_t(\omega') = \hat{U}_{s'_t}$, since $(U_t, \mathcal{F}_t)$ is simple. Also, since $(\mathcal{F}_t)$ is adapted, the partition S_t refines the partition S_{t-1}, so that (b) for any distinct s_{t-1}, s'_{t-1}, we have $\operatorname{supp}(\hat{\mu}^{s_{t-1}}_t) \cap \operatorname{supp}(\hat{\mu}^{s'_{t-1}}_t) = \varnothing$. Since additionally $\mu(s_t) > 0$ for all $s_t \in S_t$, we have (c) $\bigcup_{s_{t-1} \in S_{t-1}} \operatorname{supp} \hat{\mu}^{s_{t-1}}_t = S_t$. Thus, DREU1 is satisfied.

To see that DREU2 holds, observe that for each $h^t = (A_0, p_0, ..., A_t, p_t) \in \mathcal{H}_t$, we have

$$\begin{aligned}
\mu(C(h^t)) &= \textstyle\sum_{s_T \in S_T} \mu(s_T)\mu\left(C(h^t)|s_T\right) \\
&= \textstyle\sum_{s_T \in S_T} \mu(s_T)\mu\left(\bigcap_{k=0}^t \{\omega \in \Omega : p_k \in M(M(A_k, U_k), W_k)\}|s_T\right) \\
&= \sum_{\substack{(s_0,\dots,s_T) \\ \exists \omega \in \Omega \forall t:\, s_t = \mathcal{F}_t(\omega)}} \mu(s_T)\mu\left(\bigcap_{k=0}^t \{p_k \in M(M(A_k, U_{s_k}), W_k)\}|s_t\right) \\
&= \sum_{\substack{(s_0,\dots,s_T) \\ \exists \omega \in \Omega \forall t:\, s_t = \mathcal{F}_t(\omega)}} \mu(s_T) \prod_{k=0}^t \mu\left(\{p_k \in M(M(A_k, U_{s_k}), W_k)\}|s_k\right) \\
&= \sum_{\substack{(s_0,\dots,s_t) \\ \exists \omega \in \Omega \forall k \le t:\, s_k = \mathcal{F}_k(\omega)}} \prod_{k=0}^t \mu^{s_{k-1}}_k(s_k) \prod_{k=0}^t \tau_{s_k}(p_k, A_k) \\
&= \textstyle\sum_{(s_0,\dots,s_t) \in S_0 \times \dots \times S_t} \prod_{k=0}^t \mu^{s_{k-1}}_k(s_k) \prod_{k=0}^t \tau_{s_k}(p_k, A_k),
\end{aligned}$$

where the third equality follows from the fact that (U_t) is $\mathcal{F}_t$-adapted, the fourth equality follows from parts (ii) and (iii) of the properness assumption on (W_t), the final equality follows from the fact that $\prod_{k=0}^t \mu^{s_{k-1}}_k(s_k) = 0$ whenever $(s_0, \dots, s_t) \neq (\mathcal{F}_0(\omega), \dots, \mathcal{F}_t(\omega))$ for all ω, and the remaining equalities hold by definition. Since $\rho_t(p_t; A_t|h^{t-1}) = \frac{\mu(C(h^t))}{\mu(C(h^{t-1})}$ by (3), this shows that DREU2 holds.

J.1.2 BEU

"If" direction: Suppose ρ admits an S-based BEU representation $(S_t, \{\mu^{s_{t-1}}_t\}_{s_{t-1} \in S_{t-1}}, \{U_{s_t}, u_{s_t}, \tau_{s_t}\}_{s_t \in S_t})_{t=0,\dots,T}$. Let $(\hat{\Omega}, \hat{\mathcal{F}}^*, \hat{\mu}, (\hat{\mathcal{F}}_t, \hat{U}_t, \hat{W}_t))$ denote the corresponding DREU representation of ρ obtained in the "if" direction for DREU. In addition, define $\hat{u}_t : \hat{\Omega} \to \mathbb{R}^Z$ for each t by $\hat{u}_t(\hat{\omega}) := u_{s_t}$ whenever $\operatorname{proj}_{S_t}(\hat{\omega}) = s_t$. Note that the process $(\hat{u}_t)$ is $\hat{\mathcal{F}}_t$-adapted. Moreover, for each $\hat{\omega} = (s_0, W_0, ..., s_T, W_T)$, we have $\hat{U}_T(\hat{\omega}) = U_{s_T} = u_{s_T} = \hat{u}_T(\hat{\omega})$ and for each $t \le T - 1$ and (z_t, A_{t+1})

$$\begin{aligned}
\hat{U}_t(\hat{\omega})(z_t, A_{t+1}) &= U_{s_t}(z_t, A_{t+1}) \\
&= u_{s_t}(z_t) + \sum_{s_{t+1} \in S_{t+1}} \mu^{s_t}_{t+1}(s_{t+1}) \max_{p_{t+1} \in A_{t+1}} U_{s_{t+1}}(p_{t+1}) \\
&= \hat{u}_t(\hat{\omega})(z_t) + \sum_{s_{t+1} \in S_{t+1}} \hat{\mu}(s_{t+1}|s_t) \max_{p_{t+1} \in A_{t+1}} U_{s_{t+1}}(p_{t+1}) \\
&= \hat{u}_t(\hat{\omega})(z_t) + \mathbb{E}[\max_{p_{t+1} \in A_{t+1}} \hat{U}_{t+1}(p_{t+1})|\hat{\mathcal{F}}_t(\hat{\omega})],
\end{aligned}$$

where we let $\hat{\mu}(s_{t+1}|s_t) := \hat{\mu}\left(C(s_0,\dots,s_{t+1}) \mid C(s_0,\dots,s_t)\right)$. Thus we constructed a BEU representation with $\delta_t(\cdot)=1$ for every t.

"Only if" direction: Suppose ρ admits a BEU representation $(\Omega,\mathcal{F}^*,\mu,(\mathcal{F}_t,U_t,u_t,\delta_t,W_t))$. We construct another tuple $(\Omega,\mathcal{F}^*,\mu,(\mathcal{F}_t,U'_t,u'_t,\delta'_t,W_t))$ by setting $U'_t(\omega) := \prod_{\tau=0}^{t-1}\delta_\tau(\omega)U_t(\omega)$, $u'_t(\omega) := \prod_{\tau=0}^{t-1}\delta_\tau(\omega)u_t(\omega)$, and $\delta'_t(\omega)=1$ for each t and ω, which are all $\mathcal{F}_t$-measurable. By Proposition I.1, $(\Omega,\mu,(\mathcal{F}_t,U'_t,W_t))$ is still a DREU representation of ρ. Furthermore, for each ω (omitting its notational dependence),

$$U'_t(z_t,A_{t+1}) = \prod_{\tau=0}^{t-1}\delta_\tau U_t(z_t,A_{t+1}) = \prod_{\tau=0}^{t-1}\delta_\tau\left(u_t(z)+\delta_t\mathbb{E}[\max_{A_{t+1}}U_{t+1}|\mathcal{F}_t]\right) = u'_t(z)+\delta_t\mathbb{E}[\max_{A_{t+1}}U'_{t+1}|\mathcal{F}_t]$$

for every (z_t,A_{t+1}). Thus $(\Omega,\mathcal{F}^*,\mu,(\mathcal{F}_t,U'_t,u'_t,\delta'_t,W_t))$ is still a BEU representation of ρ. Based on this tuple, let $(S_t,\{\hat{\mu}_t^{s_{t-1}}\}_{s_{t-1}\in S_{t-1}},\{\hat{U}_{s_t},\tau_{s_t}\}_{s_t\in S_t})_{t=0,\dots,T}$ denote the corresponding S-based DREU representation obtained in the "only if" direction for DREU. In addition, for each s_t, define $\hat{u}_{s_t}\in\mathbb{R}^Z$ by $\hat{u}_{s_t}=u'_t(\omega)$ for any $\omega\in s_t$; this is well-defined as (u'_t) is $\mathcal{F}_t$-adapted. Reversing the argument in the previous part, we can verify that $\hat{u}_{s_T}=\hat{U}_{s_T}$ for each s_T and $\hat{U}_{s_t}(z_t,A_{t+1}) = \hat{u}_{s_t}(z_t)+\sum_{s_{t+1}}\hat{\mu}_{t+1}^{s_t}(s_{t+1})\max_{p_{t+1}\in A_{t+1}}\hat{U}_{s_{t+1}}(p_{t+1})$ for each s_t with $t\le T-1$.

J.1.3 BEB

"If" direction: Suppose ρ admits an S-based BEB representation $(S_t,\{\mu_t^{s_{t-1}}\}_{s_{t-1}\in S_{t-1}},\{U_{s_t},u_{s_t},\tau_{s_t},\delta_{s_t}\}_{s_t\in S_t})_{t=0,\dots,T}$. Let $(\hat{\Omega},\hat{\mathcal{F}}^*,\hat{\mu},(\hat{\mathcal{F}}_t,\hat{U}_t,\hat{u}_t,1,\hat{W}_t))$ denote the corresponding BEU representation obtained in the "if" direction for BEU. In addition, define $\hat{\delta}_t:\hat{\Omega}\to\mathbb{R}$ for each t by $\hat{\delta}_t(\hat{\omega}) := \delta_{s_t}$ whenever $\text{proj}_{S_t}(\hat{\omega})=s_t$. Note that for each $\hat{\omega}=(s_0,W_0,..,s_T,W_T)$ and $t\le T-1$, we have $\hat{u}_t(\hat{\omega}) = u_{s_t} = \frac{1}{\delta_{s_t}}\sum_{s_{t+1}}\mu_{t+1}^{s_t}(s_{t+1})u_{s_{t+1}} = \frac{1}{\hat{\delta}_t(\hat{\omega})}\mathbb{E}[\hat{u}_{t+1}|\hat{\mathcal{F}}_t(\hat{\omega})]$. Iterating expectations, this yields $\hat{u}_t(\hat{\omega}) = \mathbb{E}[\prod_{\tau=t}^{T-1}\hat{\delta}_\tau^{-1}\hat{u}_T|\hat{\mathcal{F}}_t(\hat{\omega})] = \mathbb{E}[\prod_{\tau=t}^{T-1}\hat{\delta}_\tau^{-1}\hat{U}_T|\hat{\mathcal{F}}_t(\hat{\omega})]$. Replace $\hat{U}_t(\hat{\omega})$ with $\hat{U}'_t(\hat{\omega}) := \mathbb{E}[\prod_{\tau=t}^{T-1}\hat{\delta}_\tau|\hat{\mathcal{F}}_t(\hat{\omega})]\hat{U}_t(\hat{\omega})$ for each t and $\hat{\omega}$. By Proposition I.1, $(\hat{\Omega},\hat{\mathcal{F}}^*,\hat{\mu},(\hat{\mathcal{F}}_t,\hat{U}'_t,\hat{W}_t))$ is still a DREU representation of ρ. Moreover, for each $t\le T-1$, we have

$$\begin{aligned}\hat{U}'_t(\hat{\omega})(z_t,A_{t+1}) &= \mathbb{E}[\prod_{\tau=t}^{T-1}\hat{\delta}_\tau|\hat{\mathcal{F}}_t(\hat{\omega})]\hat{u}_t(\hat{\omega})(z_t)+\mathbb{E}[\prod_{\tau=t}^{T-1}\hat{\delta}_\tau\max_{p_{t+1}\in A_{t+1}}\hat{U}_{t+1}(p_{t+1})|\hat{\mathcal{F}}_t(\hat{\omega})]\\ &= \mathbb{E}[\hat{U}'_T(z_t)\mid\hat{\mathcal{F}}_t(\hat{\omega})]+\hat{\delta}_t(\hat{\omega})\mathbb{E}[\max_{p_{t+1}\in A_{t+1}}\hat{U}'_{t+1}(p_{t+1})|\hat{\mathcal{F}}_t(\hat{\omega})].\end{aligned}$$

Thus, $(\hat{\Omega},\hat{\mathcal{F}}^*,\hat{\mu},(\hat{\mathcal{F}}_t,\hat{U}'_t,\hat{\delta}_t,\hat{W}_t))$ is a BEB representation of ρ.

"Only if" direction: Suppose that ρ admits a BEB representation $(\Omega,\mu,(\mathcal{F}_t,U_t,\delta_t,W_t))$. Let $(\Omega,\mathcal{F}^*,\mu,(\mathcal{F}_t,U'_t,u'_t,\delta'_t,W_t))$ and $(S_t,\{\hat{\mu}_t^{s_{t-1}}\}_{s_{t-1}\in S_{t-1}},\{\hat{U}'_{s_t},\hat{u}'_{s_t},\tau_{s_t}\}_{s_t\in S_t})_{t=0,\dots,T}$ respectively denote the corresponding BEU representation and S-based BEU representation of ρ obtained in the "only if" direction for BEU. In addition, define $\hat{\delta}_{s_t} := \delta_t(\omega)$ for $\mathcal{F}_t(\omega)=s_t$. Then for each $t\le T-1$ and ω with $\mathcal{F}_t(\omega)=s_t$, we have

$$\hat{u}'_{s_t} = u'_t(\omega) = \prod_{\tau=0}^{t-1}\delta_\tau(\omega)u_t(\omega) = \prod_{\tau=0}^{t-1}\delta_\tau(\omega)\mathbb{E}[u_{t+1}(\omega)|\mathcal{F}_t(\omega)] = \frac{1}{\delta_t(\omega)}\mathbb{E}[u'_{t+1}|\mathcal{F}_t(\omega)] = \frac{1}{\hat{\delta}_{s_t}}\sum_{s_{t+1}}\hat{\mu}_{t+1}^{s_t}(s_{t+1})\hat{u}'_{s_t}$$

where the first and last equality used the construction of S-based BEU, and the second and fourth equality used the construction of BEU. Thus $(S_t,\{\hat{\mu}_t^{s_{t-1}}\}_{s_{t-1}\in S_{t-1}},\{\hat{U}'_{s_t},\hat{u}'_{s_t},\hat{\delta}_{s_t},\tau_{s_t}\}_{s_t\in S_t})_{t=0,\dots,T}$ is an S-based BEB representation of ρ. ■

J.2 Proofs for Appendix E

This appendix presents proofs of all lemmas from Appendix E.

J.2.1 Proof of Lemma E.1

By standard arguments, for any separable metric space (Y, d): (a) the set $\mathcal{P}(Y)$ of Borel probability measures on Y endowed with the topology of weak convergence is a separable metric space metrized by the Prokhorov metric π_d induced by d (e.g., Theorem 15.12 in Aliprantis and Border (2006)); (b) the set $\mathcal{K}_C(Y)$ of nonempty compact subsets of Y endowed with the Hausdorff distance induced by d is a separable metric space (e.g., Khamsi and Kirk (2011) p. 40); (c) every dense subspace of Y is separable.

We now prove the claim inductively, working backwards from period T. Since $X_T := Z$ is finite, the claim is immediate. Consider $t < T$ and suppose that X_τ is a separable metric space for all $\tau \geq t+1$. By (a) above, $\mathcal{P}(X_{t+1})$ endowed with the induced Prokhorov metric is separable, so since $\Delta(X_{t+1})$ is dense in $\mathcal{P}(X_{t+1})$ (e.g., Theorem 15.10 in Aliprantis and Border (2006)) $\Delta(X_{t+1})$ is also separable (by (c)). Then by (b) above, $\mathcal{K}_C(\Delta(X_{t+1}))$ endowed with the induced Hausdorff metric is separable, so since $\mathcal{A}_{t+1} := \mathcal{K}(\Delta(X_{t+1}))$ is dense in $\mathcal{K}_C(\Delta(X_{t+1}))$ (e.g., Lemma 0 in Gul and Pesendorfer (2001)), $\mathcal{A}_{t+1}$ is also separable. Finally, $X_t := Z \times \mathcal{A}_{t+1}$ endowed with the product of the discrete metric and the Hausdorff metric is separable, as required. ■

J.2.2 Proof of Lemma E.2

By the finiteness of S, there is a finite set $Y' \subseteq Y$ such that for each s the restriction $U_s \restriction_{Y'}$ to Y' is nonconstant and for any distinct s, s', $U_s \restriction_{Y'} \not\approx U_{s'} \restriction_{Y'}$ (that is, there exists $p, q \in \Delta(Y')$ such that $U_s(p) \geq U_s(q)$ and $U_{s'}(p) < U_{s'}(q)$). By Lemma 1 in Ahn and Sarver (2013), there is a collection of lotteries $\{p^s : s \in S\} \subseteq \Delta(Y')$ such that $U_s(p^s) = U_s \restriction_{Y'} (p^s) > U_s \restriction_{Y'} (p^{s'}) = U_s(p^{s'})$ for any distinct s, s'. ■

J.2.3 Proof of Lemma E.3

(i) $\Longrightarrow$ (ii): We prove the contrapositive. Suppose that there is $s_{t-1} \in S(h^{t-1})$ and $s_t \in \operatorname{supp} \mu_t^{s_{t-1}}$ such that $|M(A_t, U_{s_t})| > 1$. Pick any $p_t \in M(A_t, U_{s_t})$ such that $\tau_{s_t}(p_t, A_t) > 0$. Since U_{s_t} is nonconstant, we can find lotteries $\underline{r}, \overline{r} \in \Delta(X_t)$ such that $U_{s_t}(\underline{r}) < U_{s_t}(\overline{r})$. Fix any sequence $\alpha_n \in (0,1)$ with $\alpha_n \to 0$. Let $p_t^n := \alpha_n \underline{r} + (1-\alpha_n)p_t$. For every $q_t \in A_t \smallsetminus \{p_t\}$, let $\underline{q}_t^n := \alpha_n \underline{r} + (1-\alpha_n)q_t$ and $\overline{q}_t^n := \alpha_n \overline{r} + (1-\alpha_n)q_t$. Let $\underline{B}_t^n := \{\underline{q}_t^n : q_t \in A_t \smallsetminus \{p_t\}\}$, let $\overline{B}_t^n := \{\overline{q}_t^n : q_t \in A_t \smallsetminus \{p_t\}\}$, and let $B_t^n := \underline{B}_t^n \cup \overline{B}_t^n$. Then $B_t^n \to^m A_t \smallsetminus \{p_t\}$ and $p_t^n \to^m p_t$.

Moreover, since $|M(A_t, U_{s_t})| > 1$, there exists $q_t \in A_t \smallsetminus \{p_t\}$ such that $U_{s_t}(\alpha_n \overline{r} + (1-\alpha_n)q_t) > U_{s_t}(p_t^n)$ for all n, so that $\tau_{s_t}(p_t^n, B_t^n \cup \{p_t^n\}) = 0$. Furthermore, note that for all $s_t' \in S_t \smallsetminus \{s_t\}$, we have $N(M(A_t, U_{s_t'}), p_t) = N(M(\underline{B}_t^n \cup \{p_t^n\}, U_{s_t'}), p_t^n) \supseteq N(M(B_t^n \cup \{p_t^n\}, U_{s_t'}), p_t^n)$, so that $\tau_{s_t'}(p_t, A_t) \geq \tau_{s_t'}(p_t^n, B_t^n \cup \{p_t^n\})$ for all n. Letting $\operatorname{pred}(s_{t-1}) = (s_0, \ldots, s_{t-2})$, Lemma E.5 then implies that for all

n,

$$\rho_t(p_t; A_t|h^{t-1}) - \rho_t(p_t^n; B_t^n \cup \{p_t^n\}|h^{t-1}) =$$

$$\frac{\sum_{s'_0,\ldots,s'_t} \prod_{k=0}^{t-1} \mu_k^{s'_{k-1}}(s'_k)\tau_{s'_k}(p_k, A_k)\mu_t^{s'_{t-1}}(s'_t)\left(\tau_{s'_t}(p_t, A_t) - \tau_{s'_t}(p_t^n, B_t^n \cup \{p_t^n\})\right)}{\sum_{s'_0,\ldots,s'_{t-1}} \prod_{k=0}^{t-1} \mu_k^{s'_{k-1}}(s'_k)\tau_{s'_k}(p_k, A_k)} \geq$$

$$\frac{\prod_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)\mu_t^{s_{t-1}}(s_t)\tau_{s_t}(p_t, A_t)}{\sum_{s'_0,\ldots,s'_{t-1}} \sum_{s'_0,\ldots,s'_{t-1}} \prod_{k=0}^{t-1} \mu_k^{s'_{k-1}}(s'_k)\tau_{s'_k}(p_k, A_k)} > 0.$$

Since the last line does not depend on n, this implies $\lim_{n\to\infty} \rho_t(p_t^n; B_t^n \cup \{p_t^n\}|h^{t-1}) < \rho_t(p_t; A_t|h^{t-1})$. By definition of $\mathcal{A}_t^*$, this means $A_t \notin \mathcal{A}_t^*(h^{t-1})$.

(ii) $\implies$ (i): Suppose A_t satisfies (ii). Consider any $p_t \in A_t$, $p_t^n \to^m p_t$, $B_t^n \to^m A_t \setminus \{p_t\}$. Consider any $s_{t-1} \in S(h^{t-1})$ and $s_t \in \operatorname{supp} \mu_t^{s_{t-1}}$. By (ii), we either have $M(A_t, U_{s_t}) = \{p_t\}$ or $p_t \notin M(A_t, U_{s_t})$. In the former case, $U_{s_t}(p_t) > U_{s_t}(q_t)$ for all $q_t \in A_t \setminus \{p_t\}$. But then, for all n large enough, linearity of U_{s_t} implies $U_{s_t}(p_t^n) > U_{s_t}(q_t^n)$ for all $q_t^n \in B_t^n$, i.e., $\tau_{s_t}(p_t, A_t) = \lim_n \tau_{s_t}(p_t^n, B_t^n \cup \{p_t^n\}) = 1$. In the latter case, $U_{s_t}(p_t) < U_{s_t}(q_t)$ for some $q_t \in A_t \setminus \{p_t\}$. But then, for all n large enough, linearity of U_{s_t} implies $U_{s_t}(p_t^n) < U_{s_t}(q_t^n)$ for all $q_t^n \in B_t^n$ such that $q_t^n \to^m q_t$, i.e., $\tau_{s_t}(p_t, A_t) = \lim_n \tau_{s_t}(p_t^n, B_t^n \cup \{p_t^n\}) = 0$.

Thus, for all $s_{t-1} \in S(h^{t-1})$ and $s_t \in \operatorname{supp} \mu_t^{s_{t-1}}$, we have $\tau_{s_t}(p_t, A_t) = \lim_n \tau_{s_t}(p_t^n, B_t^n \cup \{p_t^n\})$. Hence, the representation in Lemma E.5 implies that for all n sufficiently large,

$$\rho_t(p_t^n; B_t^n \cup \{p_t^n\}|h^{t-1}) = \rho_t(p_t; A_t|h^{t-1}),$$

as required. ■

J.2.4 Proof of Lemma E.4

Let $k := \max\{n = 0\ldots, t-1 : q_n \neq \hat{q}_n\}$ be the last entry at which d^{t-1} and $\hat{d}^{t-1}$ differ, where we set $k = -1$ if $q_n = \hat{q}_n$ for all $n = 0, \ldots, t-1$. We prove the claim by induction on k.

Suppose first that $k = -1$, i.e., that $d^{t-1} = \hat{d}^{t-1}$. If $\lambda_0 > \hat{\lambda}_0$, then the 0-th entry of $\lambda h^{t-1} + (1-\lambda)d^{t-1}$ can be written as an appropriate mixture of the 0-th entry of $\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}$ with (A_0, p_0); if $\lambda_0 \leq \hat{\lambda}_0$, then the 0-th entry of $\lambda h^{t-1} + (1-\lambda)d^{t-1}$ can be written as an appropriate mixture of the 0-th entry of $\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}$ with $(\{q_0\}, q_0)$. In either case, Axiom B.2 implies that $\rho_t(\cdot; A_t|\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1})$ is unaffected after replacing the 0-th entry of $\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}$ with the 0-th entry of $\lambda h^{t-1} + (1-\lambda)d^{t-1}$. Continuing this way, we can successively apply Axiom B.2 to replace each entry of $\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}$ with the corresponding entry of $\lambda h^{t-1} + (1-\lambda)d^{t-1}$ without affecting ρ_t. This yields the desired conclusion.

Suppose the claim holds whenever $k \leq m-1$ for some $0 \leq m \leq t-1$. We show that the claim continues to hold for $k = m$. Note first that we can assume that

$$\begin{aligned}
&\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}, \frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1} \in \mathcal{H}_{t-1}(A_t);\\
&\frac{2}{3}B_m + \frac{1}{3}\{\hat{q}_m\}, \{\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m\} \in \operatorname{supp} q_{m-1}^A;\\
&\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \{\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m\} \in \operatorname{supp} \hat{q}_{m-1}^A,
\end{aligned} \tag{35}$$

where $B_m := \frac{1}{2}A_m + \frac{1}{2}\{q_m\}$,$\hat{B}_m := \frac{1}{2}A_m + \frac{1}{2}\{\hat{q}_m\}$, $r_m := \frac{1}{2}p_m + \frac{1}{2}q_m$, and $\hat{r}_m := \frac{1}{2}p_m + \frac{1}{2}\hat{q}_m$.

Indeed, we can find a sequence of lotteries $(\ell_n)_{n=0}^{t-1}$ such that for all $n = 1, \ldots, t-1$

$$\lambda_n A_n + (1-\lambda_n)\{o_n\}, \frac{1}{2}A_n + \frac{1}{2}\{o_n\}, \hat{\lambda}_n A_n + (1-\hat{\lambda}_n)\{\hat{o}_n\}, \frac{1}{2}A_n + \frac{1}{2}\{\hat{o}_n\}, \{o_n\} \in \operatorname{supp} \ell_{n-1}^A;$$
$$\frac{2}{3}B_m + \frac{1}{3}\{\hat{o}_m\}, \frac{2}{3}\hat{B}_m + \frac{1}{3}\{o_m\}, \{\frac{1}{2}o_m + \frac{1}{2}\hat{o}_m\} \in \operatorname{supp} \ell_{m-1}^A,$$

where $o_n := \frac{1}{2}q_n + \frac{1}{2}\ell_n$ and $\hat{o}_n := \frac{1}{2}\hat{q}_n + \frac{1}{2}\ell_n$. Letting $c^{t-1} := (\{o_n\}, o_n)_{n=0}^{t-1}$ and $\hat{c}^{t-1} := (\{\hat{o}_n\}, \hat{o}_n)_{n=0}^{t-1}$, we have that $c^{t-1}, \hat{c}^{t-1} \in \mathcal{D}_{t-1}$, $\lambda h^{t-1} + (1-\lambda)c^{t-1}, \hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{c}^{t-1} \in \mathcal{H}_{t-1}(A_t)$, and the last entry at which c^{t-1} and $\hat{c}^{t-1}$ differ is m. Moreover, repeated application of Axiom B.2 implies

$$\rho_t(\cdot; A_t|\lambda h^{t-1} + (1-\lambda)d^{t-1}) = \rho_t(\cdot; A_t|\lambda h^{t-1} + (1-\lambda)c^{t-1});$$
$$\rho_t(\cdot; A_t|\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}) = \rho_t(\cdot; A_t|\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{c}^{t-1}).$$

Thus, we can replace d^{t-1} and $\hat{d}^{t-1}$ with c^{t-1} and $\hat{c}^{t-1}$ if need be and guarantee that (35) is satisfied.

Given (35), $\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}, \frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1} \in \mathcal{H}_{t-1}(A_t)$, so the base case of the proof implies

$$\begin{aligned} \rho_t(\cdot; A_t|\lambda h^{t-1} + (1-\lambda)d^{t-1}) &= \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}); \\ \rho_t(\cdot; A_t|\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}) &= \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1}). \end{aligned} \tag{36}$$

Also, (35) guarantees that $((\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1})_{-m}, (\frac{2}{3}B_m + \frac{1}{3}\{\hat{q}_m\}, \frac{2}{3}r_m + \frac{1}{3}\hat{q}_m))$ and $((\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1})_{-m}, (\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \frac{2}{3}\hat{r}_m + \frac{1}{3}q_m))$ are well-defined histories in $\mathcal{H}_{t-1}(A_t)$. Thus, by Axiom B.2

$$\begin{aligned} \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}) &= \rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1})_{-m}, (\frac{2}{3}B_m + \frac{1}{3}\{\hat{q}_m\}, \frac{2}{3}r_m + \frac{1}{3}\hat{q}_m)); \\ \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1}) &= \rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1})_{-m}, (\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \frac{2}{3}\hat{r}_m + \frac{1}{3}q_m)). \end{aligned} \tag{37}$$

But note that

$$\begin{aligned} \left(\frac{2}{3}B_m + \frac{1}{3}\{\hat{q}_m\}, \frac{2}{3}r_m + \frac{1}{3}\hat{q}_m\right) &= \left(\frac{1}{3}A_m + \frac{2}{3}\{\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m\}, \frac{1}{3}p_m + \frac{2}{3}(\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m)\right) \\ &= \left(\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \frac{2}{3}\hat{r}_m + \frac{1}{3}q_m\right). \end{aligned}$$

Thus, $((\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1})_{-m}, (\frac{2}{3}B_m + \frac{1}{3}\{\hat{q}_m\}, \frac{2}{3}r_m + \frac{1}{3}\hat{q}_m))$ is an entry-wise mixture of h^{t-1} with the degenerate history $e^{t-1} := ((d^{t-1})_{-m}, (\{\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m\}, \frac{1}{2}q_m + \frac{1}{2}\hat{q}_m)$ and similarly $((\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1})_{-m}, (\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \frac{2}{3}\hat{r}_m + \frac{1}{3}q_m))$ is an entry-wise mixture of h^{t-1} with the degenerate history $\hat{e}^{t-1} := ((\hat{d}^{t-1})_{-m}, (\{\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m\}, \frac{1}{2}q_m + \frac{1}{2}\hat{q}_m)$. But the last entry at which e^{t-1} and $\hat{e}^{t-1}$ differ is strictly smaller than m. Hence, applying the inductive hypothesis, we obtain

$$\begin{aligned} &\rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1})_{-m}, (\frac{2}{3}B_m + \frac{1}{3}\{q_m\}, \frac{2}{3}r_m + \frac{1}{3}q_m)) = \\ &\quad \rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1})_{-m}, (\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \frac{2}{3}\hat{r}_m + \frac{1}{3}q_m)). \end{aligned} \tag{38}$$

Combining (36), (37), and (38) yields the required equality

$$\rho_t(\cdot; A_t|\lambda h^{t-1} + (1-\lambda)d^{t-1}) = \rho_t(\cdot; A_t|\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}).$$

Finally, let $\hat{d}^{t-1}$ and $\hat{\lambda} \in (0,1]$ be the choices from Definition 10 such that $\rho_t^{h^{t-1}}(\cdot; A_t) := \rho_t(\cdot; A_t|\hat{\lambda}h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1})$. Then the above implies that $\rho_t^{h^{t-1}}(\cdot; A_t) = \rho_t(\cdot; A_t|\lambda h^{t-1} + (1-\lambda)d^{t-1})$, as claimed. ■

J.2.5 Proof of Lemma E.5

If $h^{t'-1} \in \mathcal{H}_{t'-1}(A_{t'})$, the claim is immediate from DREU2. So suppose $h^{t'-1} \notin \mathcal{H}_{t'-1}(A_{t'})$. Let $\lambda \in (0,1)$ and $d^{t'-1} = (\{q_\ell\}, q_\ell)_{\ell=0}^{t'-1} \in \mathcal{D}_{t'-1}$ be the choices from Definition 11 such that $\lambda h^{t'-1} + (1-\lambda)d^{t'-1} \in \mathcal{H}_{t'-1}(A_{t'})$ and $\rho_{t'}(p_{t'}, A_{t'}|h^{t'-1}) := \rho_{t'}(p_{t'}, A_{t'}|\lambda h^{t'-1} + (1-\lambda)d^{t'-1})$.

Note that for all $k \leq t'$, $s_k \in S_k$, and $w \in \mathbb{R}^{X_k}$, we have $p_k \in M(M(A_k, U_{s_k}), w)$ if and only if $\lambda p_k + (1-\lambda)q_k \in M(M(\lambda A_k + (1-\lambda)\{q_k\}, U_{s_k}), w)$. Hence, $\tau_{s_k}(p_k, A_k) = \tau_{s_k}(\lambda p_k + (1-\lambda)q_k, \lambda A_k + (1-\lambda)\{q_k\})$. Thus, the claim follows from DREU2 applied to the history $\lambda h^{t'-1} + (1-\lambda)d^{t'-1} \in \mathcal{H}_{t'-1}(A_{t'})$. ■

J.2.6 Proof of Lemma E.6

Let $S_t(s_{t-1}) := \mathrm{supp}\mu_t^{s_{t-1}}$. By DREU1, we can find a finite $Y_t \subseteq X_t$ such that (i) for any $s_t \in S_t(s_{t-1})$, U_{s_t} is non-constant over Y_t; (ii) for any distinct $s_t, s_t' \in S_t(s_{t-1})$, $U_{s_t} \not\approx U_{s_t'}$ over Y_t; and (iii) $\bigcup_{p_t \in A_t} \mathrm{supp} p_t \subseteq Y_t$. By (i) and (ii) and Lemma E.2, we can find a menu $D_t := \{q_t^{s_t} : s_t \in S_t(s_{t-1})\} \subseteq \Delta(Y_t)$ such that $M(D_t, U_{s_t}) = \{q_t^{s_t}\}$ for all $s_t \in S_t(s_{t-1})$. Define $b_t := \sum_{y \in Y_t} \frac{1}{|Y_t|}\delta_y \in \Delta(Y)$. For each $s_t \in S_t(s_{t-1})$, pick $z^{s_t} \in \mathrm{argmax}_{y \in Y} U_{s_t}$ and let $g_t^{s_t} := \delta_{z^{s_t}}$. By (i), we have $U_{s_t}(g_t^{s_t}) > U_{s_t}(b_t)$ for all $s_t \in S_t(s_{t-1})$. Hence, there exists $\alpha \in (0,1)$ small enough such that for all $s_t \in S_t(s_{t-1})$, we have $U_{s_t}(\hat{q}^{s_t}) > U_{s_t}(b_t)$, where $\hat{q}^{s_t} := \alpha q_t^{s_t} + (1-\alpha)g_t^{s_t}$. Note that setting $\hat{D} := \{\hat{q}_t^{s_t} : s_t \in S_t(s_{t-1})\}$, we still have $M(\hat{D}_t, U_{s_t}) = \{\hat{q}_t^{s_t}\}$.

For each $s_t \in S_t(s_{t-1})$, pick some $p_t(s_t) \in M(A_t, U_{s_t})$. For the "moreover" part, we can ensure that $p_t(s_t^*) = p_t^*$. Fix any sequence (ε_n) from $(0,1)$ such that $\varepsilon_n \to 0$. For each n and $s_t \in S_t(s_{t-1})$, let $p_t^n(s_t) := (1-\varepsilon)p_t(s_t) + \varepsilon\hat{q}_t^{s_t}$. And for each $r_t \in A_t$, let $r_t^n := (1-\varepsilon)r_t + \varepsilon b_t$. Finally, let $A_t^n := \{p_t^n(s_t) : s_t \in S_t(s_{t-1})\} \cup \{r_t^n : r_t \in A_t\}$. Note that $A_t^n \to^m A_t$. Moreover, by construction, for all $s_t \in S_t(s_{t-1})$ and n, we have $M(A_t^n, U_{s_t}) = \{p_t^n(s_t)\}$: Indeed, $U_{s_t}(p_t^n(s_t)) > U_{s_t}(r_t^n)$ for all $r_t \in A_t$ since $U_{s_t}(p_t(s_t)) \geq U_{s_t}(r_t)$ and $U_{s_t}(\hat{q}_t^{s_t}) > U_{s_t}(b_t)$; and $U_{s_t}(p_t^n(s_t)) > U_{s_t}(p_t^n(s_t'))$ for all $s_t' \neq s_t$, since $U_{s_t}(p_t(s_t)) \geq U_{s_t}(p_t(s_t'))$ and $U_{s_t}(\hat{q}_t^{s_t}) > U_{s_t}(\hat{q}_t^{s_t'})$.

Since s_{t-1} is the only state consistent with h^{t-1}, Lemma E.3 implies that $A_t^n \in \mathcal{A}_t^*(h^{t-1})$, as required. Finally, for the "moreover" part, note that we ensured that $p_t(s_t^*) = p_t^*$. Hence $p_t^n(s_t^*)$ constructed above has the desired property that $p_t^n(s_t^*) \to^m p_t^*$ and $\mathcal{U}_{s_t}(A_t^n, p_t^n(s_t^*)) = \{U_{s_t^*}\}$ for all n. ■

J.3 Proof of Proposition I.1

J.3.1 "If" directions:

DREU: Consider any $h^t = (p_0, A_0, ..., p_t, A_t) \in \mathcal{H}_t$. Then

$$\begin{aligned}
\mu(C(h^t)) &= \sum_{\mathcal{F}_T(\omega)\in\Pi_T} \mu(\mathcal{F}_T(\omega))\mu\left(\bigcap_{k=0}^{t}\{p_k \in M(M(A_k, U_k), W_k)\}|\mathcal{F}_T(\omega)\right) \\
&= \sum_{\mathcal{F}_t(\omega)\in\Pi_t} \prod_{k=0}^{t} \mu(\mathcal{F}_k(\omega) \mid \mathcal{F}_{k-1}(\omega))\mu\left(\{W_k \in N(M(A_k, U_k(\omega)), p_k)\}|\mathcal{F}_k(\omega)\right) \\
&= \sum_{\mathcal{F}_t(\omega)\in\Pi_t} \prod_{k=0}^{t} \hat{\mu}(\phi_k(\mathcal{F}_k(\omega)) \mid \phi_{k-1}(\mathcal{F}_{k-1}(\omega)))\hat{\mu}\left(\{\hat{W}_k \in N(M(A_k, U_k(\omega)), p_k)\}|\phi_k(\mathcal{F}_k(\omega))\right) \\
&= \sum_{\hat{\mathcal{F}}_t(\hat{\omega})\in\hat{\Pi}_t} \prod_{k=0}^{t} \hat{\mu}(\hat{\mathcal{F}}_k(\hat{\omega}) \mid \hat{\mathcal{F}}_{k-1}(\omega))\hat{\mu}\left(\{\hat{W}_k \in N(M(A_k, \hat{U}_k(\hat{\omega})), p_k)\}|\hat{\mathcal{F}}_k(\hat{\omega})\right) \\
&= \sum_{\hat{\mathcal{F}}_T(\hat{\omega})\in\hat{\Pi}_T} \hat{\mu}(\hat{\mathcal{F}}_T(\hat{\omega}))\left(\bigcap_{k=0}^{t}\{p_k \in M(M(A_k, \hat{U}_k), \hat{W}_k)\} \mid \hat{\mathcal{F}}_T(\hat{\omega})\right) = \hat{\mu}(\hat{C}(h^t)),
\end{aligned}$$

where the second equality follows from properness of (W_t) and $\mathcal{F}_t$-adaptedness of (U_t), the third equality follows from assumptions (i) and (iii), the fourth equality from the fact that ϕ_t is a bijection and assumption (ii), the fifth equality from the properness of $(\hat{W}_t)$ and $\hat{\mathcal{F}}_t$-adaptedness of $(\hat{U}_t)$, and the first and last equalities hold by definiton. Since $\mathcal{D}$ represents ρ and $\hat{\mathcal{D}}$ represents $\hat{\rho}$, this implies $\rho_t(p_t, A_t|h^{t-1}) = \frac{\mu(C(h^t))}{\mu(C(h^{t-1}))} = \frac{\hat{\mu}(\hat{C}(h^t))}{\hat{\mu}(\hat{C}(h^{t-1}))} = \hat{\rho}_t(p_t, A_t|h^{t-1})$. Thus, $\hat{\rho} = \rho$, as required.

BEU: By the "if" direction for DREU, $\hat{\mathcal{D}}$ is a DREU representation of ρ. It remains to show that $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ satisfies (1). From assumptions (ii), (iv), and (v) it is immediate that $\hat{U}_T = \hat{u}_T$. Moreover, for all $t \leq T-1$, and $\omega \in \Omega$, $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$, we have

$$\begin{aligned}
\alpha_t(\omega)\hat{U}_t(\hat{\omega})(z, A_{t+1}) &= U_t(\omega)(z, A_{t+1}) - \beta_t(\omega) = u_t(\omega)(z) - \beta_t(\omega) + \delta_t(\omega)\mathbb{E}_\mu[\max_{p_{t+1}\in A_{t+1}} U_{t+1}(p_{t+1}) \mid \mathcal{F}_t(\omega)] \\
&= \alpha_t(\omega)\hat{u}_t(\hat{\omega})(z) - \delta_t(\omega)\mathbb{E}_\mu[\beta_{t+1}|\mathcal{F}_t(\omega)] + \delta_t(\omega)\mathbb{E}_{\hat{\mu}}[\alpha_{t+1}\max_{p_{t+1}\in A_{t+1}} \hat{U}_{t+1}(p_{t+1}) \mid \hat{\mathcal{F}}_t(\hat{\omega})] + \delta_t(\omega)\mathbb{E}_\mu[\beta_{t+1}|\mathcal{F}_t(\omega)] \\
&= \alpha_t(\omega)\left(\hat{u}_t(\hat{\omega})(z) + \hat{\delta}_t(\hat{\omega})\mathbb{E}_{\hat{\mu}}[\max_{p_{t+1}\in A_{t+1}} \hat{U}_{t+1}(p_{t+1}) \mid \hat{\mathcal{F}}_t(\hat{\omega})]\right)
\end{aligned}$$

where the first equality follows from (ii), the second from (1) for $(\mathcal{D}, (u_t, \delta_t))$, the third from (i), (ii), and (v) (and the fact ϕ_t is a bijection), and the fourth by (iv). Thus, $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ satisfies (1).

BEB: By the "if" direction for BEU, $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ is an BEU representation of ρ. It remains to show that $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ satisfies (2). For all $t \leq T-1$ and $\omega \in \Omega$, $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$, we have

$$\alpha_0(\omega)\hat{u}_t(\hat{\omega}) + \gamma_t(\omega) = u_t(\omega) = \mathbb{E}_\mu[U_T|\mathcal{F}_t(\omega)] = \alpha_0(\omega)\mathbb{E}_{\hat{\mu}}[\hat{U}_T|\hat{\mathcal{F}}_t(\hat{\omega})] + \mathbb{E}_\mu[\beta_T|\mathcal{F}_t(\omega)],$$

where the first equality follows from (iv), (v), and (vi), the second from (2) for $(\mathcal{D}, (u_t, \delta_t))$, and the third from (i), (ii), (iv), (vi) (and the fact that ϕ_t is a bijection). But since $\gamma_t(\omega) = \mathbb{E}_\mu[\beta_T|\mathcal{F}_t(\omega)]$ by (vii), the above implies that $\hat{u}_t(\hat{\omega}) = \mathbb{E}_{\hat{\mu}}[\hat{U}_T|\hat{\mathcal{F}}_t(\hat{\omega})]$, whence $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ satisfies (2) with $\hat{\hat{u}} := \hat{U}_T$.

J.3.2 "Only if" directions:

DREU: Throughout the proof, for any t and $E_t = \mathcal{F}_t(\omega) \in \Pi_t$, we let $U_t(E_t)$ denote $U_t(\omega)$ and likewise for $\hat{U}$; this is well-defined by adaptedness. We construct the sequence $(\phi_t, \alpha_t, \beta_t)$ inductively, dealing with the base case $t = 0$ and the inductive step simultaneously.

Suppose $t \geq 0$ and that we have constructed $(\phi_{t'}, \alpha_{t'}, \beta_{t'})$ satisfying (i)–(iii) for all $t' < t$ (disregard the latter assumption if $t = 0$). If $t > 0$, fix any $E_{t-1} = \mathcal{F}_{t-1}(\omega^*) \in \Pi_{t-1}$, let $\hat{E}_{t-1} := \phi_{t-1}(E_{t-1})$, and let $\Pi_t(E_{t-1}) := \{E_t = \mathcal{F}_t(\omega) \in \Pi_t : \mathcal{F}_{t-1}(\omega) = E_{t-1}\}$ and $\hat{\Pi}_t(\hat{E}_{t-1}) := \{\hat{E}_t = \hat{\mathcal{F}}_t(\hat{\omega}) \in \hat{\Pi}_t : \hat{\mathcal{F}}_{t-1}(\hat{\omega}) = \hat{E}_{t-1}\}$. As in the proof of Lemma B.2, we can repeatedly apply Lemma E.2 to find a separating history for $E_{t-1} = \mathcal{F}_{t-1}(\omega^*)$, i.e., a history $h^{t-1} = (B_0, q_0, \ldots, B_{t-1}, q_{t-1}) \in \mathcal{H}^*_{t-1}$ such that $\{\omega \in \Omega : q_k \in M(B_k, U_k(\omega))\} = \mathcal{F}_k(\omega^*)$ for all $k = 0, \ldots, t-1$. By inductive hypothesis h^{t-1} is then also a separating history for $\hat{E}_{t-1}$. Thus, by Lemma E.3 (and the translation to S-based DREU in Proposition A.1), $C(h^{t-1}) = E_{t-1}$ and $\hat{C}(h^{t-1}) = \hat{E}_{t-1}$. If $t = 0$, then in the following we let $E_{t-1} := \Omega$, $\hat{E}_{t-1} := \hat{\Omega}$, $\Pi_t(E_{t-1}) := \Pi_0$, $\hat{\Pi}_t(E_{t-1}) := \hat{\Pi}_0$, and we disregard all references to the separating history.

Enumerate $\Pi_t(E_{t-1}) = \{E^i_t : i = 1, \ldots, m\}$ with corresponding utilities $U^i_t := U_t(E^i_t)$ and $\hat{\Pi}_t(\hat{E}_{t-1}) = \{\hat{E}^j_t : j = 1, \ldots, \hat{m}\}$ with corresponding utilities $\hat{U}^j_0 := \hat{U}_t(\hat{E}^j_t)$. Since $(\mathcal{F}_t, U_t)$ and $(\hat{\mathcal{F}}_t, \hat{U}_t)$ are both simple, we have $\mu(E^i_t) > 0$ for all i and $U^i_t \not\approx U^{i'}_t$ for $i \neq i'$, and likewise $\hat{\mu}(\hat{E}^j_t) > 0$ for all j and $\hat{U}^j_t \not\approx \hat{U}^{j'}_t$ for $j \neq j'$. Note that for every j there exists a unique $i(j)$ such $U^{i(j)}_t \approx \hat{U}^j_t$. Indeed, if such an $i(j)$ exists it is unique because all the U^i_t represent different preferences. And the desired $i(j)$ exists, since otherwise by Lemma E.2, we can find a menu $B_t = \{q^i_t : i = 1, \ldots, m\} \cup \{\hat{q}^j_t\}$ such that $M(B_t, U^i_t) = \{q^i_t\}$ for each i and $M(B_t, \hat{U}^j_t) = \{\hat{q}^j_t\}$. We can additionally assume (by replacing h^{t-1} with an appropriate mixture if need be) that $h^{t-1} \in \mathcal{H}^*_{t-1}(B_t)$. Since $\mathcal{D}$ and $\hat{\mathcal{D}}$ both represent ρ, we obtain

$$0 = \mu[C(\hat{q}^j_t, B_t)|E_{t-1}] = \rho_t(\hat{q}^j_t; B_t|h^{t-1}) = \hat{\mu}[\hat{C}(\hat{q}^j_t, B_t)|\hat{E}_{t-1}) \geq \hat{\mu}(\hat{E}^j_t|\hat{E}_{t-1}) > 0,$$

a contradiction. Similarly, for every i, there exists a unique $j(i)$ such that $\hat{U}^{j(i)}_t \approx U^i_t$. Thus, defining $\phi_t : \Pi_t(E_{t-1}) \to \hat{\Pi}_t(\hat{E}_{t-1})$ by $\phi_t(E^i_t) = \hat{E}^{j(i)}_t$ yields a bijection. By construction, $U_t(E^i_t) \approx \hat{U}_t(\phi_t(E^i_t))$ for all i, so we can find $\alpha_t(E^i_t) \in \mathbb{R}_{++}$ and $\beta(E^i_t) \in \mathbb{R}$ such that $U_t(E^i_t) = \alpha_t(E^i_t)\hat{U}_t(\phi_t(E^i_t)) + \beta(E^i_t)$. Defining $\alpha(\omega) = \alpha(\mathcal{F}_t(\omega))$ and $\beta(\omega) = \beta(\mathcal{F}_t(\omega))$ this yields $\mathcal{F}_t$-measurable maps $\alpha_t, \beta_t : E_{t-1} \to \mathbb{R}$ such that (ii) holds for all $\omega \in E_{t-1}$. Moreover, applying Lemma E.2 again, we can find a menu $D_t = \{r^i_t : i = 1, \ldots, n\}$ such that $M(D_t, U^i_t) = \{r^i_t\}$ for each i. Again, slightly perturbing the separating history h^{t-1} for E_{t-1} if need be, we can assume that $h^{t-1} \in \mathcal{H}^*_{t-1}(D_t)$. Then by the representation, $\mu(E^i_t|E_{t-1}) = \rho_t(r^i_t; D^i_t|h^{t-1}) = \hat{\mu}(\phi_t(E^i_t)|\hat{E}_{t-1})$ for all i, yielding (i).

To show (iii), consider any $p_t \in A_t$, where we can again assume $h^{t-1} \in \mathcal{H}^*_{t-1}(\frac{1}{2}A_t + \frac{1}{2}D_t)$. Let $B^i_t := \{w \in \mathbb{R}^{X_t} : p_t \in M(M(A_t, U_t(E^i_t)), w)\}$. Note that by (ii), $B^i_t = \{w \in \mathbb{R}^{X_t} : p_t \in M(M(A_t, \hat{U}_t(\phi_t(E^i_t))), w)\}$. Thus, $\mu(\{W_t \in B_t\}|E^i_t) = \mu(C(p_t, A_t)|E^i_t)$ and $\hat{\mu}(\{\hat{W}_t \in B_t\}|\phi_t(E^i_t)) = \hat{\mu}(\hat{C}(p_t, A_t)|\phi_t(E^i_t))$. But since $\mathcal{D}$ and $\hat{\mathcal{D}}$ both represent ρ and by choice of D_t,

$$\begin{aligned}\mu(E^i_t|E_{t-1})\mu[C(p_t, A_t)|E^i_t] &= \mu[C(\frac{1}{2}p_t + \frac{1}{2}r^i_t, \frac{1}{2}A_t + \frac{1}{2}D_t)|E_{t-1}] = \\ &\rho_t(\frac{1}{2}p_t + \frac{1}{2}r^i_t; \frac{1}{2}A_t + \frac{1}{2}D_t|h^{t-1}) = \\ \hat{\mu}[\hat{C}(\frac{1}{2}p_t + \frac{1}{2}r^i_t, \frac{1}{2}A_t + \frac{1}{2}D_t)|\hat{E}_{t-1}] &= \hat{\mu}(\phi_t(E^i_t)|\hat{E}_{t-1})\hat{\mu}[\hat{C}(p_t, A_t)|\phi_t(E^i_t)],\end{aligned}$$

which implies $\mu[C(p_t, A_t)|E^i_t] = \hat{\mu}[\hat{C}(p_t, A_t)|\phi_t(E^i_t)]$, since by (i) we have $\mu(E^i_t|E_{t-1}) = \hat{\mu}(\phi_t(E^i_t)|\hat{E}_{t-1})$.

Thus, $\mu(\{W_t \in B_t\}|E_t^i) = \hat{\mu}(\{\hat{W}_t \in B_t\}|\phi_t(E_t^i))$, as required.

Finally, note that the collection $\{\Pi_t(E_{t-1}) : E_{t-1} \in \Pi_{t-1}\}$ partitions Π_t, and similarly $\{\hat{\Pi}_t(\hat{E}_{t-1}) : \hat{E}_{t-1} \in \hat{\Pi}_{t-1}\}$ partitions $\hat{\Pi}_t$. Thus, applying the above construction for every $E_{t-1} \in \Pi_{t-1}$ yields a bijection $\phi_t : \Pi_t \to \hat{\Pi}_t$ and $\mathcal{F}_t$-measurable maps $\alpha_t : \Omega \to \mathbb{R}_{++}$ and $\beta_t : \Omega \to \mathbb{R}$ such that (i)–(iii) are satisfied.

BEU: The "only if" part for DREU yields sequences $(\phi_t, \alpha_t, \beta_t)$ such that (i)–(iii) are satisfied. It remains to show that (iv) and (v) hold. Throughout the proof, for any $E_t = \mathcal{F}_t(\omega) \in \Pi_t$, we sometimes use $U_t(E_t)$, $\delta_t(E_t)$, $\alpha_t(E_t)$, $\beta_t(E_t)$ to denote $U_t(\omega)$, $\delta_t(\omega)$ $\alpha_t(\omega)$, $\beta_t(\omega)$; this is well-defined since they are $\mathcal{F}_t$-measurable. We also let $\mathcal{F}_{t-1}(E_t) := \mathcal{F}_{t-1}(\omega)$; this is well-defined since $\mathcal{F}_t(\omega) = \mathcal{F}_t(\omega')$ implies $\mathcal{F}_{t-1}(\omega) = \mathcal{F}_{t-1}(\omega')$, as $(\mathcal{F}_t)$ is a filtration.

For (iv), fix any ω and $t \leq T-1$. Let $E_t := \mathcal{F}_t(\omega)$ and pick any A_{t+1}, B_{t+1} and z_t. Then

$$\begin{aligned}
U_t(E_t)(z_t, A_{t+1}) - U_t(E_t)(z_t, B_{t+1}) &= \alpha_t(E_t)(\hat{U}_t(\phi_t(E_t))(z_t, A_{t+1}) - \hat{U}_t(\phi_t(E_t))(z_t, B_{t+1})) \\
&= \alpha_t(E_t)\hat{\delta}_t(\phi_t(E_t)) \sum_{\hat{E}_{t+1} \in \hat{\Pi}_{t+1}} \hat{\mu}(\hat{E}_{t+1}|\phi_t(E_t))[\max_{A_{t+1}} \hat{U}_{t+1}(\hat{E}_{t+1}) - \max_{B_{t+1}} \hat{U}_{t+1}(\hat{E}_{t+1})] \\
&= \alpha_t(E_t)\hat{\delta}_t(\phi_t(E_t)) \sum_{E_{t+1} \in \Pi_{t+1}} \hat{\mu}(\phi_{t+1}(E_{t+1})|\phi_t(E_t))[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))] \\
&= \alpha_t(E_t)\hat{\delta}_t(\phi_t(E_t)) \sum_{E_{t+1} \in \Pi_{t+1}} \mu(E_{t+1}|E_t)[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))] \\
&= \alpha_t(E_t)\hat{\delta}_t(\phi_t(E_t)) \sum_{E_{t+1} \text{ s.t. } \mathcal{F}_t(E_{t+1}) = E_t} \mu(E_{t+1}|E_t)[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))],
\end{aligned} \tag{39}$$

where the first equality holds by (ii), the second equality follows from $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}))$ being a BEU representation, the third equality from the fact that ϕ_t is a bijection, the fourth equality from (i), and the fifth equality from the fact that $\mu(\mathcal{F}_{t+1}(\omega')|E_t) > 0$ iff $\mathcal{F}_t(\omega') = E_t$.

At the same time, we have

$$\begin{aligned}
&U_t(E_t)(z_t, A_{t+1}) - U_t(E_t)(z_t, B_{t+1}) \\
&= \delta_t(E_t) \sum_{E_{t+1} \in \Pi_{t+1}} \mu(E_{t+1}|E_t)[\max_{A_{t+1}} U_{t+1}(E_{t+1}) - \max_{B_{t+1}} U_{t+1}(E_{t+1}) \\
&= \delta_t(E_t) \sum_{E_{t+1} \in \Pi_{t+1}} \mu(E_{t+1}|E_t)\alpha_{t+1}(E_{t+1})[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))] \\
&= \delta_t(E_t) \sum_{E_{t+1} \text{ s.t. } \mathcal{F}_t(E_{t+1}) = E_t} \mu(E_{t+1}|E_t)\alpha_{t+1}(E_{t+1})[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))],
\end{aligned} \tag{40}$$

where the first equality follows from $(\mathcal{D}, (u_t, \delta_t))$ being a BEU representation, the second equality from (ii), and the third equality from the fact that $\mu(\mathcal{F}_{t+1}(\omega')|E_t) > 0$ iff $\mathcal{F}_t(\omega') = E_t$.

Combining (39) and (40), we have that for all A_{t+1} and B_{t+1},

$$\begin{aligned}&\hat{\delta}_t(\phi_t(E_t)) \sum_{E_{t+1} \text{ s.t. } \mathcal{F}_t(E_{t+1})=E_t} \mu(E_{t+1}|E_t)\alpha_t(E_t)[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))] \\ &= \delta_t(E_t) \sum_{E_{t+1} \text{ s.t. } \mathcal{F}_t(E_{t+1})=E_t} \mu(E_{t+1}|E_t)\alpha_{t+1}(E_{t+1})[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))].\end{aligned} \tag{41}$$

Since $(\hat{\mathcal{F}}_t, \hat{U}_t)$ is simple and ϕ_t is a bijection, $\hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) \not\approx \hat{U}_{t+1}(\phi_{t+1}(E'_{t+1}))$ for all distinct E_{t+1}, E'_{t+1} with $\mathcal{F}_t(E_{t+1}) = E_t = \mathcal{F}_t(E'_{t+1})$. So by Lemma E.2, we can find a menu $A_{t+1} := \{q_{t+1}^{E_{t+1}} : \mathcal{F}_t(E_{t+1}) = E_t\}$ such that for all E_{t+1} with $\mathcal{F}_t(E_{t+1}) = E_t$ we have $M(A_{t+1}, \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) = \{q_{t+1}^{E_{t+1}}\}$. Let $E^*_{t+1} := \mathcal{F}_{t+1}(\omega)$ and let $B_{t+1} = A_{t+1} \setminus \{q_{t+1}^{E^*_{t+1}}\}$. Then in (41), $[\max_{A_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1})) - \max_{B_{t+1}} \hat{U}_{t+1}(\phi_{t+1}(E_{t+1}))] \neq 0$ iff $E_{t+1} = E^*_{t+1}$. Hence, (41) implies $\frac{\hat{\delta}_t(\phi_t(E_t))}{\delta_t(E_t)}\alpha_t(\omega) = \frac{\hat{\delta}_t(\phi_t(E_t))}{\delta_t(E_t)}\alpha_t(E_t) = \alpha_{t+1}(E^*_{t+1}) = \alpha_{t+1}(\omega)$. Since this is true for all $t \leq T-1$, (iv) follows.

For (v), note that the claim for T is immediate from (ii) and the fact that $U_T = u_T$, $\hat{U}_T = \hat{u}_T$. Next, fix any $\omega \in \Omega$, $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$, $t \leq T-1$, and $(z, \{p_{t+1}\})$. Then

$$\begin{aligned}U_t(\omega)(z, \{p_{t+1}\}) &= u_t(\omega)(z) + \delta_t(\omega)\mathbb{E}_\mu[U_{t+1}(p_{t+1})|\mathcal{F}_t(\omega)] \\ &= u_t(\omega)(z) + \alpha_t(\omega)\hat{\delta}_t(\hat{\omega})\mathbb{E}_{\hat{\mu}}[\hat{U}_{t+1}(p_{t+1})|\hat{\mathcal{F}}_t(\hat{\omega})] + \delta_t(\omega)\mathbb{E}_\mu[\beta_{t+1}|\mathcal{F}_t(\omega)],\end{aligned} \tag{42}$$

where the first equality follows from $(\mathcal{D}, (u_t, \delta_t))$ being an evolving utility representation and the second equality from (i), (ii), (iv) (and the fact that ϕ_t is a bijection). At the same time, we have

$$\begin{aligned}U_t(\omega)(z, \{p_{t+1}\}) &= \alpha_t(\omega)\hat{U}_t(\hat{\omega})(z, \{p_{t+1}\}) + \beta_t(\omega) \\ &= \alpha_t(\omega)\hat{u}_t(\omega)(z) + \alpha_t(\omega)\hat{\delta}_t(\hat{\omega})\mathbb{E}_{\hat{\mu}}[\hat{U}_{t+1}(p_{t+1})|\hat{\mathcal{F}}_t(\hat{\omega})] + \beta_t(\omega),\end{aligned} \tag{43}$$

where the first equality follows from (ii) and the second equality from $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ being an evolving utility representation. Combining (42) and (43) yields the desired claim.

BEB: The "only if" part for BEU yields sequences $(\phi_t, \alpha_t, \beta_t)$ such that (i)–(v) are satisfied. It remains to show that (vi) and (vii) hold.

For (vi), Fix any $\omega \in \Omega$ and t. Take $\overline{\ell}, \underline{\ell}$ from Condition D.1 (Uniform Ranked Pair). Then based on the representation one can verify that $u_t(\omega)(\overline{\ell}) > u_t(\omega)(\underline{\ell})$ holds by following the similar line as in Lemma D.1.

Note that by (2) and iterated expectations, we have

$$U_t(\omega)(\ell_t, \ell_{t+1}, A_{t+2}) = u_t(\omega)(\ell_t) + \delta_t(\omega)\left(u_t(\omega)(\ell_{t+1}) + \mathbb{E}[\delta_{t+1} \max_{A_{t+2}} U_{t+2}|\mathcal{F}_t(\omega)]\right)$$

for any $(\ell_t, \ell_{t+1}, A_{t+2})$. Hence $U_t(\omega)(\overline{\ell}, \underline{\ell}, A_{t+2}) - U_t(\omega)(\eta\overline{\ell} + (1-\eta)\underline{\ell}, \eta\overline{\ell} + (1-\eta)\underline{\ell}, A_{t+2}) = 0$ if and only if $\eta = \frac{1}{1+\delta_t(\omega)}$.

Now pick any $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$. Then since $(\hat{\mathcal{D}}, (\hat{u}_t, \hat{\delta}_t))$ is also a BEU representation, by the same reasoning as above we have that $\hat{U}_t(\hat{\omega})(\overline{\ell}, \underline{\ell}, A_{t+2}) - \hat{U}_t(\hat{\omega})(\eta\overline{\ell} + (1-\eta)\underline{\ell}, \eta\overline{\ell} + (1-\eta)\underline{\ell}, A_{t+2}) = 0$ if and only if $\eta = \frac{1}{1+\hat{\delta}_t(\hat{\omega})}$. By (ii), this implies that $\delta_t(\omega) = \hat{\delta}_t(\hat{\omega})$, proving (vi).

Finally (vii) is verified by observing that for any t, ω, and $\hat{\omega} \in \phi_t(\mathcal{F}_t(\omega))$,

$$\gamma_t(\omega) = u_t(\omega) - \alpha_t(\omega)\hat{u}_t(\hat{\omega}) = \mathbb{E}_\mu[u_T|\mathcal{F}_t(\omega)] - \alpha_t(\omega)\mathbb{E}_{\hat{\mu}}[\hat{u}_T|\hat{\mathcal{F}}_t(\hat{\omega})] = \mathbb{E}_\mu[\beta_T|\mathcal{F}_t(\omega)],$$

where the first equality uses (v), the second uses (2), and the third uses (i), (v) and $\alpha_t(\omega) = \alpha_T(\omega)$ (which follows from (iv) and (vi)). ■

J.4 Proof of Proposition I.2

(i) $\implies$ (ii): Suppose that ρ^Z admits a BEU representation $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t, u_t, \delta_t))$ and satisfies Axioms I.1 and I.2. For each t, we can pick a finite collection $\mathcal{U}_t = \{u_t^1, \dots, u_t^{m_t}\}$ of ordinally distinct felicities such that $[\mathcal{U}_t] = [\{u_t(\omega) : \omega \in \Omega\}]$. Condition D.1 (Uniformly ranked pairs) ensures that these felicities are non-constant. Let $\mathcal{U} := \{u^1, \dots, u^m\}$, where $m = m_0$ and $u^i = u_0^i$ for all $i = 1, \dots, m$. Define $\xi \in \Delta^\circ(\mathcal{U})$ by $\xi(u^i) := \mu(u_0(\omega) \approx u^i)$ for all i.

By Axiom I.1, for each degenerate consumption history d_Z^{t-1}, ρ_0^Z and $\rho_t^Z(\cdot|d_Z^{t-1})$ represent the same static stochastic choice rule over finite menus of consumption lotteries without ties. Hence, the same argument in the proof of Proposition I.1 implies that after suitable relabeling we can assume that $m_t = m$ and $u_t^i \approx u^i$ for all i and $\mu(u_t(\omega) \approx u^i) = \xi(u^i)$. Thus, property (i) of the Markov evolving utility representation is satisfied.

Next, we construct a menu $L = \{\ell^1, \dots, \ell^m\} \in \mathcal{L}_0^*$ such that $u^i(\ell^i) > u^i(\ell^j)$ for all $i \neq j$ and such that each (L, ℓ^i) is a consumption atom. Indeed, since the u^i are nonconstant and ordinally distinct, Lemma E.2 yields a menu $L = \{\ell^1, \dots, \ell^m\} \in \mathcal{L}_0^*$ such that $u^i(\ell^i) > u^i(\ell^j)$ for all $i \neq j$; moreover, up to mixing all ℓ^i with some full-support lottery $\ell \in \Delta^\circ(Z)$, we can assume that $L \subseteq \Delta^\circ(Z)$. By the representation, $\rho_0(\ell^i, L) = \mu(u_0(\omega) \approx u^i) > 0$ for all i. Finally, suppose that $L' \in \mathcal{L}^*$ and $L' \supseteq L$. Then either $\ell^i \in M(L', u^i)$, in which case $\rho_0(\ell^i, L') = \mu(u_0(\omega) \approx u^i) = \rho_0(\ell^i, L)$; or $\ell^i \notin M(L', u^i)$, in which case $\rho_0(\ell^i, L') = 0$ since $u^j(\ell^i) < u^j(\ell^j)$ for all $j \neq i$. Thus, each (L, ℓ^i) is a consumption atom.

Now, consider any $t \leq T - 1$ and any $u_0, \dots, u_{t+1} \in \mathcal{U}$. For each $s = 0, \dots, t+1$, let ℓ_s denote the maximizer of u_s in menu L that we constructed in the previous paragraph. Then for any degenerate consumption history d_Z^{t-1}, we have

$$\begin{aligned} \mu(u_{t+1}(\omega) \approx u_{t+1}|u_0(\omega) \approx u_0, \dots, u_{t-1}(\omega) \approx u_{t-1}, u_t(\omega) \approx u_t) = \\ \rho_{t+1}^Z(\ell_{t+1}, L \mid L, \ell_0, \dots, L, \ell_{t-1}, L, \ell_t) = \\ \rho_1^Z(\ell_{t+1}, L \mid L, \ell_t) = \rho_{t+1}^Z(\ell_{t+1}, L \mid d_Z^{t-1}, L_t, \ell_t) = \\ \mu(u_{t+1}(\omega) \approx u_{t+1}|u_t(\omega) \approx u_t), \end{aligned}$$

where the first and fourth equality hold by the BEU representation of ρ together with the fact that $M(L, u^i) = \{\ell^i\}$ for all i, and the second and third equality follow from Axiom I.2 and the fact that (L, ℓ_t) is a consumption atom. This establishes property (ii) of the Markov evolving utility representation.

Finally, set $\Pi_{i,j} := \mu(u_1(\omega) \approx u^j|u_0(\omega) \approx u^i)$ for all $i, j = 1, \dots, m$. Note that this yields a right stochastic matrix Π, because $\sum_j \Pi_{i,j} = 1$ by part (i) of Definition 14 that we established above. Consider any $i, j = 1, \dots, m$. Then letting L, ℓ^i and ℓ^j be as constructed in the third paragraph, we have for any degenerate consumption history d_Z^{t-1} that

$$\begin{aligned} \mu(u_{t+1}(\omega) \approx u^j|u_t(\omega) \approx u^i) = \rho_{t+1}^Z(\ell^j, L \mid d_Z^{t-1}, L, \ell^i) = \\ \rho_1^Z(\ell^j, L \mid L, \ell^i) = \mu(u_1(\omega) \approx u^j|u_0(\omega) \approx u^i) = \Pi_{i,j} \end{aligned}$$

where the first and third equalities again follow from the representation and the construction of L and the second equality holds by Axiom I.2 and the fact that (L, ℓ^i) is a consumption atom. This proves property (iii) of the Markov evolving utility representation.

(ii) $\Longrightarrow$ (i): Suppose that ρ admits a Markov evolving utility representation. To show that Axiom I.1 holds, consider any degenerate consumption history d_Z^{t-1}, $L \in \mathcal{L}_0^* \cap \mathcal{L}_t^*(d_Z^{t-1})$, $\ell \in L$. Then

$$\rho_0^Z(\ell, L) = \mu\{\omega : \ell \in M(L, u_0(\omega))\} = \\ \xi\{u^i \in \mathcal{U} : \ell \in M(L, u^i)\} = \mu\{\omega : \ell \in M(L, u_t(\omega))\} = \rho_t^Z(\ell, L|d_Z^{t-1}),$$

where the first and final equalities hold by the BEU representation and the fact that L is without ties and d_Z^{t-1} is degenerate, and the second and third equalities hold by property (i) of the Markov evolving utility representation.

To establish Axiom I.2, consider any consumption atom (L, ℓ). We first show that there exists $u^i \in \mathcal{U}$ such that $\mu(\ell \in M(L, u_t(\omega)) = \mu(u_t(\omega) \approx u^i)$ for all t. Since L is without ties, it suffices to show that there is a unique $i \in \{1, \dots, m\}$ such that $\ell \in M(L, u^i)$. To see this, note that since $\mu(\ell \in M(L, u_0(\omega)) = \rho_0(\ell, L) > 0$, there exists u^i such that $\ell \in M(L, u^i)$. Suppose for a contradiction that $\ell \in M(L, u^j)$ for some $j \neq i$. Since $u^i \not\approx u^j$, we can find $m \in \Delta(Z)$ such that $u^i(\ell) > u^i(m)$ and $u^j(\ell) < u^j(m)$.[87] Then, letting $M = L \cup \{m\}$, we have that $\xi(u^i) \leq \rho_0(\ell, M) \leq \rho_0(\ell, L) - \xi(u^j)$. Thus, $\rho_0(\ell, M) \notin \{\rho_0(\ell, L), 0\}$, contradicting the fact that (L, ℓ) is a consumption atom.

Now, consider any consumption history h_Z^{t-1} without ties. For any $L' \in \mathcal{L}_t^*(h_Z^{t-1})$ and $\ell' \in L'$, we have

$$\rho_1^Z(\ell', L'|L, \ell) = \mu(\ell' \in M(L', u_1(\omega)) \mid u_0(\omega) \approx u^i) = \\ \sum_{\{j : \ell' \in M(L', u^j)\}} \mu(u_1(\omega) \approx u^j \mid u_0(\omega) \approx u^i) = \sum_{\{j : \ell' \in M(L', u^j)\}} \Pi_{i,j} \\ = \sum_{\{j : \ell' \in M(L', u^j)\}} \mu(u_{t+1}(\omega)) \approx u^j \mid u_t(\omega) \approx u^i) = \mu(\ell' \in M(L', u_{t+1}(\omega)) \mid u_t(\omega) \approx u^i),$$

where the second and final equality follow from property (i) of the Markov evolving utility representation and the fact that L' is without ties, and the third and fourth equality follow from property (iii) of the Markov evolving utility representation.

Moreover, letting $\mathcal{U}(h_Z^{t-1})$ denote the set of all sequences of felicity realizations from $\mathcal{U}$ that are consistent with history h_Z^{t-1},[88] we have

$$\rho_{t+1}^Z(\ell', L'|h_Z^{t-1}, L, \ell) = \mu(\ell' \in M(L', u_{t+1}(\omega)) \mid u_t(\omega) \approx u^i, \omega \in C(h_Z^{t-1})) = \\ \frac{\sum_{(u_0,\dots,u_{t-1}) \in \mathcal{U}(h_Z^{t-1})} \mu(\ell' \in M(L', u_{t+1}(\omega)) \mid u_t(\omega) \approx u^i, \bigcap_{s=0}^{t-1}\{u_s(\omega) \approx u_s\}) \mu(u_t(\omega) \approx u^i, \bigcap_{s=0}^{t-1}\{u_s(\omega) \approx u_s\})}{\sum_{(u_0,\dots,u_{t-1}) \in \mathcal{U}(h_Z^{t-1})} \mu(u_t(\omega) \approx u^i, \bigcap_{s=0}^{t-1}\{u_s(\omega) \approx u_s\}} \\ = \mu(\ell' \in M(L', u_{t+1}(\omega)) \mid u_t(\omega) \approx u^i),$$

where the third equality follows from property (ii) of the Markov evolving utility representation. Com-

[87] Indeed, since $u^i \not\approx u^j$, we can find ℓ^i, ℓ^j such that $u^i(\ell^i) > u^i(\ell^j)$ and $u^j(\ell^j) > u^j(\ell^i)$. Then for small enough $\varepsilon > 0$, $m := \ell + \varepsilon(\ell^j - \ell^i)$ is a well-defined consumption lottery in $\Delta(Z)$, as $\ell \in \Delta^\circ(Z)$. Moreover, $u^i(\ell) > u^i(m)$ and $u^j(\ell) < u^j(m)$, as required.

[88] More formally, since h_Z^{t-1} is a consumption history without ties and by property (i) of the Markov representation, we can find $\mathcal{U}(h_Z^{t-1}) \subseteq \mathcal{U}^t$ such that $C(h_Z^{t-1}) = \{\omega : \exists (u_0, \dots, u_{t-1}) \in \mathcal{U}(h^{t-1})$ with $u_s(\omega) \approx u_s$ for all $s = 0, \dots, t-1\}$.

bining the previous two paragraphs, we have $\rho_1^Z(\ell', L'|L, \ell) = \rho_{t+1}^Z(\ell', L'|h_Z^{t-1}, L, \ell)$. This establishes Axiom I.2. ■