Dynamic Random Utility*

Mira Frick   Ryota Iijima   Tomasz Strzalecki

Abstract

Under dynamic random utility, an agent (or population of agents) solves a dynamic decision problem subject to evolving private information. We analyze the fully general and non-parametric model, axiomatically characterizing the implied dynamic stochastic choice behavior. A key new feature relative to static or i.i.d. versions of the model is that when private information displays serial correlation, choices appear *history dependent*: different sequences of past choices reflect different private information of the agent, and hence typically lead to different distributions of current choices. Our axiomatization imposes discipline on the form of history dependence that can arise under arbitrary serial correlation. Dynamic stochastic choice data lets us distinguish central models that coincide in static domains, in particular private information in the form of utility shocks vs. learning, and to study inherently dynamic phenomena such as choice persistence. We relate our model to specifications of utility shocks widely used in empirical work, highlighting new modeling tradeoffs in the dynamic discrete choice literature. Finally, we extend our characterization to allow past consumption to directly affect the agent's utility process, accommodating models of habit formation and experimentation.

*This version: 21 June 2017. Frick: Yale University (mira.frick@yale.edu); Iijima: Yale University (ryota.iijima@yale.edu); Strzalecki: Harvard University (tomasz_strzalecki@harvard.edu). This research was supported by the National Science Foundation grant SES-1255062. We thank David Ahn, Jose Apesteguia, Miguel Ballester, Dirk Bergemann, Jetlir Duraj, Drew Fudenberg, Daria Khromenkova, Yves Le Yaouanq, Jay Lu, Ariel Pakes, Larry Samuelson, Michael Whinston, as well as audiences at ASU Theory Conference, Barcelona GSE Workshop on Stochastic Choice, Berkeley, Bocconi, BU, Caltech Choice Conference, Harvard–MIT, LMU, LSE, Northwestern, Oxford, QMUL, Rochester, RUD (London Business School), Stanford, UCL, and Yale. A joint file, including both the main text and supplementary appendix, is available at https://drive.google.com/file/d/0B-372Fn5SRUAM3BYVmNrR2diR1E/view

1 Introduction

Random utility models are widely used throughout economics. In the static model, the agent chooses from her choice set after observing the realization of a random utility function U. In the dynamic model, the agent solves a dynamic decision problem, subject to a stochastic process (U_t) of utilities. The key feature of the model is an informational asymmetry between the agent (who knows her realized utility) and the analyst (who does not). In both the static and dynamic setting, this asymmetry gives rise to choice behavior that appears stochastic to the analyst but is deterministic from the point of view of the agent.[1]

In the dynamic setting, the informational asymmetry has an additional key implication: If (U_t) displays serial correlation, then choices will appear *history dependent* to the analyst. For example, we expect the agent's probability of voting Republican in 2020 to be different conditional on voting Republican in 2016 than conditional on voting Democrat in 2016. This is because her past voting behavior reveals relevant information about her past political preferences, which we expect to be at least somewhat persistent. History dependence due to serially correlated private information is pervasive in applications, from education and career choices in labor economics to consumer brand choices in marketing. Recognizing that "ignoring serial correlation in unobservables [...] can lead to serious misspecification errors" (Norets, 2009), the dynamic discrete choice literature studying these settings has developed and estimated a number of models that can accommodate history dependent choices. However, as highlighted by Pakes (1986), a limitation of these models is that they rely on specific parametric forms of serial correlation, making it "difficult to determine the robustness of the conclusions to the stochastic assumptions chosen."

This paper provides the first analysis of the fully general and non-parametric model of dynamic random utility. Our contribution is threefold: First, we axiomatically characterize the implied dynamic stochastic choice behavior, imposing discipline on the form of history dependence that can arise under arbitrary serially correlated private information. Our axiomatization answers for the dynamic model a question that has given rise to an extensive literature in the static setting (see Section 8.1) while overcoming a number of challenges that are new to the dynamic domain. Second, dynamic stochastic choice data allows us to distinguish central models that coincide in static domains, in particular private information in the form of utility shocks vs. learning; and to study important new distinctions inherent to dynamic settings, in particular the difference between history dependence due to serially correlated private information and *consumption dependence*, where past consumption affects current choices by directly shaping the agent's utility process. Finally, our analysis sheds new light on modeling tradeoffs in the dynamic discrete choice literature.

[1] An equivalent interpretation of the model is that the analyst observes a fixed population of heterogeneous individuals. Throughout the paper, we use "the agent" to refer to both interpretations.

Our model generalizes the static random expected utility framework of Gul and Pesendorfer (2006) to decision trees as defined by Kreps and Porteus (1978). Each period t, the agent chooses from a menu of lotteries over current consumptions and continuation menus by maximizing a random vNM utility U_t. A *history* $h^{t-1} = (A_0, p_0, \ldots, A_{t-1}, p_{t-1})$ summarizes that the agent chose lottery p_0 from menu A_0, then was faced with A_1 and chose p_1, and so on. Observed behavior at t is given by a history dependent choice distribution $\rho_t(\cdot|h^{t-1})$, specifying the choice frequency $\rho_t(p_t, A_t|h^{t-1})$ of p_t from any menu A_t that can arise after h^{t-1}.

Turning to the axiomatic characterization, our first main insight is the following: The fact that history dependence arises purely as a result of serial correlation in (U_t) entails two history *independence* conditions. Each condition identifies simple equivalence classes of histories that reveal the same private information to the analyst; if h^{t-1} and g^{t-1} are equivalent, then $\rho_t(\cdot|h^{t-1})$ and $\rho_t(\cdot|g^{t-1})$ are required to coincide. The first condition, *contraction history independence*, imposes equivalence if h^{t-1} can be obtained from g^{t-1} by eliminating some options that are irrelevant to choices along the history g^{t-1}. The second condition, *linear history independence*, imposes equivalence if h^{t-1} and g^{t-1} are "linear combinations" of each other. Theorem 1 shows that our most general model, *dynamic random expected utility* (DREU), is fully characterized by these two independence conditions along with a continuity condition and Gul and Pesendorfer's (2006) axioms that ensure *static* random utility maximization at each history.

In DREU, the stochastic process (U_t) is unrestricted. We next study the important special case of a dynamically sophisticated agent who has separable preferences over current consumption and continuation problems and is forward-looking with a correct assessment of option value. This allows us to distinguish *evolving utility*, where the agent faces taste shocks that evolve randomly over time, from its special case, *gradual learning*, in which the agent learns over time about her fixed but unknown tastes—two forms of private information that are indistinguishable in the static setting. To this end, we introduce a novel incomplete and history dependent revealed preference relation that infers from the agent's choices her preference conditional on any particular realization of her private information. Evolving utility is then characterized by adapting axioms from the menu-preference literature: separability, preference for flexibility, and dynamic sophistication (Theorem 2). The additional behavioral content of gradual learning is encapsulated by a consumption stationarity axiom, reflecting the martingale property of beliefs, along with a constant intertemporal tradeoff axiom (Theorem 3).

Proposition 1 establishes identification results for the three representations. In DREU, the agent's ordinal private information is uniquely pinned down; evolving utility, and more so gradual learning, impose discipline on the cardinal private information (U_t); additionally, gradual learning allows for unique identification of the discount factor.

A key challenge throughout our analysis is the following "limited observability" problem: In contrast with the static setting, where the analyst observes choices from all possible menus, in the dynamic setting each history of past choices restricts the set of current and future choice problems. Over time, this severely limits the history-dependent choice data on which axioms can be imposed and from which (U_t) can be inferred. We overcome this problem by means of the following extrapolation procedure (Definition 3): For any menu A_t and history h^{t-1} that does not lead to A_t, we define the agent's counterfactual choice distribution from A_t following h^{t-1} by extrapolating from the situation where the agent makes the sequence of choices captured by h^{t-1}, but knows that, with some exogenous probability, another sequence of choices that *does* lead to menu A_t will be implemented instead. Invoking linear history independence, the latter situation can be specified such that it reveals the same private information as the original choice sequence h^{t-1}, thus justifying the extrapolation. This extrapolation procedure relies crucially on the inclusion of lotteries as choice objects. We discuss the connection with similar uses of plausibly exogenous randomization to perform counterfactual analyses in empirical and experimental work.

Section 6 discusses the relationship with the dynamic discrete choice (DDC) literature. The uniqueness results that we develop are complementary to identification results in the DDC literature. Moreover, we contrast our evolving utility representation with the i.i.d. DDC model, which is a workhorse model for structural estimation. While also a special case of DREU, the latter is incompatible with evolving utility, as the two make opposite predictions about option value. In the evolving utility model the agent has a positive option value: she likes bigger menus, as they provide her with more flexibility, and wants to make her decisions as late as she can to condition on as much information as possible. On the other hand, the i.i.d. DDC agent sometimes prefers to commit to smaller menus and, more often than not, prefers to make her decisions as early as possible, thus displaying a negative option value. This points to a modeling tradeoff between the desirable statistical properties of the i.i.d. DDC model and a key feature of Bayesian rationality, positive option value.

Finally, in Section 7, we extend our model to additionally allow past consumption to *directly* influence the agent's current behavior by shaping her current preferences. We refer to this as *consumption dependence*, while reserving the term history dependence for the phenomenon discussed so far, where observed current behavior depends on past choices (rather than actual consumption) because different choices reflect different private information. Prominent examples of consumption dependence include habit formation, where consuming a certain good in the past may make the agent like it more in the present; and active learning/experimentation, where the agent's consumption provides information to her about some payoff-relevant state of the world. Making use of the fact that each chosen lottery can result in multiple consumption outcomes, we adapt our characterization to this setting, providing behavioral foundations for these models and distinguishing history from consumption dependence.

2 Static vs. Dynamic Random Utility

For any set Y, denote by $\mathcal{K}(Y)$ the set of all nonempty finite subsets of Y and by $\Delta(Y)$ the set of all simple (i.e., finite support) lotteries on Y; henceforth, all references to lotteries are to simple lotteries. Whenever Y is a separable metric space, we endow $\Delta(Y)$ with the induced Prokhorov metric and $\mathcal{K}(Y)$ with the Hausdorff metric. Let $\mathbb{R}^Y$ denote the set of vNM utility indices over Y, which is endowed with the product topology and its induced Borel sigma-algebra. For any $U, U' \in \mathbb{R}^Y$, write $U \approx U'$ if U and U' represent the same preference on $\Delta(Y)$. For any finite set of lotteries $A \in \mathcal{K}(\Delta(Y))$, let $M(A, U) := \operatorname{argmax}_{p \in A} U(p)$ denote the set of lotteries in A that maximize U, where $U(p) := \sum_{y \in \operatorname{supp}(p)} U(y)p(y)$ denotes the expected utility of any $p \in \Delta(Y)$. For any $A, B \in \mathcal{K}(\Delta(Y))$ and $\alpha \in [0, 1]$, define the α-mixture of A and B by $\alpha A + (1 - \alpha)B := \{\alpha p + (1 - \alpha)q : p \in A, q \in B\} \in \mathcal{K}(\Delta(Y))$.

2.1 Static Random Utility

We first briefly review the static model of random expected utility that will serve as the building block of our dynamic representation at each history. As mentioned in the Introduction, there are two equivalent interpretations of the model: a single agent with a random utility function or a population of agents with heterogeneous utilities. The model is based on Gul and Pesendorfer (2006), but allows for an infinite outcome space; this extension is necessary for our purposes, because in the dynamic setting the period-t outcome space X_t, consisting of all pairs of current consumptions and continuation menus, will be infinite in all but the final period.

2.1.1 Agent's problem

Let X be an arbitrary separable metric space of outcomes. The agent makes choices from menus, which are finite sets of lotteries over X; the set of all menus is $\mathcal{A} := \mathcal{K}(\Delta(X))$. Denote a typical menu by A and a typical lottery by p. Let $(\Omega, \mathcal{F}^*, \mu)$ be a finitely-additive probability space. In each state of the world, the agent's choices maximize her expected utility subject to her private information. Her payoff-relevant private information is captured by a sigma-algebra $\mathcal{F} \subseteq \mathcal{F}^*$ and an $\mathcal{F}$-measurable random vNM utility index $U : \Omega \to \mathbb{R}^X$. In case of indifference, ties are broken by a random vNM index $W : \Omega \to \mathbb{R}^X$, which is measurable with respect to $\mathcal{F}^*$. Thus, when faced with menu A, the agent chooses lottery p in state ω if and only if p maximizes $U(\omega)$ in A and, in case of ties, additionally maximizes $W(\omega)$ among the $U(\omega)$-maximizers; that is, $p \in M(M(A, U(\omega)), W(\omega))$.

For tractability, we follow Ahn and Sarver (2013) in assuming that the agent's payoff-relevant private informaton $(\mathcal{F}, U)$ is *simple*, i.e., (i) $\mathcal{F}$ is generated by a finite partition such that $\mu(\mathcal{F}(\omega)) > 0$ for every $\omega \in \Omega$, where $\mathcal{F}(\omega)$ denotes the cell of the partition that contains ω; and (ii) each $U(\omega)$ is nonconstant and $U(\omega) \not\approx U(\omega')$ whenever $\mathcal{F}(\omega) \neq \mathcal{F}(\omega')$. Moreover, the tie-breaker W is *proper*,[2] ensuring that under W ties occur with probability 0 in each menu; that is, $\mu(\{\omega \in \Omega : |M(A, W(\omega))| = 1\}) = 1$ for all $A \in \mathcal{A}$.

2.1.2 Analyst's problem

The analyst does not observe the agent's private information and thus cannot condition on events in $\mathcal{F}$ (equivalently, in the population interpretation, the analyst does not observe the identities of the agents, just aggregate choice frequencies). Because of this informational asymmetry, the agent's choices appear stochastic to the analyst.[3] His observations are summarized by a *stochastic choice rule* on $\mathcal{A}$, i.e., a map $\rho : \mathcal{A} \to \Delta(\Delta(X))$ such that $\sum_{p\in A} \rho(p, A) = 1$ for all $A \in \mathcal{A}$. Here $\rho(p, A)$ denotes the frequency with which the agent chooses lottery p when faced with menu A. If the agent behaves as in the previous section, then the event that the agent chooses p from A is $C(p, A) := \{\omega \in \Omega : p \in M(M(A, U(\omega)), W(\omega))\}$. Thus, the analyst's observations are consistent with the previous section if $\rho(p, A) = \mu(C(p, A))$ for all p and A.

Definition 1. A *static random expected utility (REU)* representation of the stochastic choice rule ρ is a tuple $(\Omega, \mathcal{F}^*, \mu, \mathcal{F}, U, W)$ such that $(\Omega, \mathcal{F}^*, \mu)$ is a finitely-additive probability space, the sigma-algebra $\mathcal{F} \subseteq \mathcal{F}^*$ and the $\mathcal{F}$-measurable utility $U : \Omega \to \mathbb{R}^X$ are simple, the $\mathcal{F}^*$-measurable tiebreaker $W : \Omega \to \mathbb{R}^X$ is proper, and $\rho(p, A) = \mu(C(p, A))$ for all p and A.

2.1.3 Characterization

For finite outcome spaces X, static REU representations have been characterized by Gul and Pesendorfer (2006) and Ahn and Sarver (2013). As a preliminary technical contribution, we extend their characterization to arbitrary separable metric spaces X. The first four conditions of the following axiom are the same as in Gul and Pesendorfer (2006). The fifth condition is a slight modification of the finiteness condition in Ahn and Sarver (2013).

Axiom 0. (Random Expected Utility)

(i). *Regularity*: If $A \subseteq A'$, then for all $p \in A$, $\rho(p; A) \geq \rho(p; A')$.

(ii). *Linearity*: For any A, $p \in A$, $\lambda \in (0, 1)$, and q, $\rho(p; A) = \rho(\lambda p+(1-\lambda)q; \lambda A+(1-\lambda)\{q\})$.

(iii). *Extremeness*: For any A, $\rho(\text{ext}A; A) = 1$.[4]

(iv). *Mixture Continuity*: $\rho(\cdot; \alpha A + (1 - \alpha)A')$ is continuous in α for all A, A'.

[2]This property is sometimes called "regular" in the literature; we use the term "proper" to avoid confusion with the Regularity axiom (Axiom 0 (i)) below.

[3]If the analyst observed the true state, choices would appear deterministic and could be summarized by a vNM preference $\succsim_\omega$.

[4]Here ext$A$ denotes the set of extreme points of A.

(v). *Finiteness*: There is $K > 0$ such that for all A, there is $B \subseteq A$ with $|B| \leq K$ such that for every $p \in A \smallsetminus B$, there are sequences $p^n \to^m p$ and $B^n \to^m B$ with $\rho(p^n; \{p^n\} \cup B^n) = 0$ for all n.

For condition (iv), $\alpha \mapsto \rho(\cdot; \alpha A + (1-\alpha)A')$ is viewed as a map from $[0,1]$ to $\Delta(\Delta(X))$, where $\Delta(\Delta(X))$ is endowed with the topology of weak convergence induced by the Prokhorov metric on $\Delta(X)$. For condition (v), *convergence in mixture*, denoted $\to^m$, on $\Delta(X)$ and $\mathcal{A}$ is defined as follows: For any $p \in \Delta(X)$ and sequence $\{p^n\}_{n\in\mathbb{N}} \subseteq \Delta(X)$, we write $p^n \to^m p$ if there exists $q \in \Delta(X)$ and a sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ with $\alpha_n \to 0$ such that $p^n = \alpha_n q + (1-\alpha_n)p$ for all n. Similarly, for any sequence $\{B^n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$, we write $B^n \to^m p$ if there exists $B \in \mathcal{A}$ and a sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ with $\alpha_n \to 0$ such that $B^n = \alpha_n B + (1-\alpha_n)\{p\}$ for all n. Finally, for any $A \in \mathcal{A}$ and sequence $(A^n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$, we write $A^n \to^m A$ if for each $p \in A$, there is a sequence $\{B^n_p\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that $B^n_p \to^m p$ and $A^n = \cup_{p\in A} B^n_p$ for all n.

Theorem 0. The stochastic choice rule ρ on $\mathcal{A}$ admits an REU representation if and only if ρ satisfies Axiom 0.

Proof. See Supplementary Appendix F. ■

2.2 Dynamic Random Utility

In many economic applications, the agent solves a dynamic decision problem subject to evolving, and in general serially correlated, private information. As in the static model, an equivalent interpretation is in terms of a population of agents with heterogeneous and evolving utilities, and we will move freely between the two interpretations.[5] The following two examples illustrate some new features of the dynamic setting on which our formal analysis will expand.

Example 1 (Brand choice dynamics)**.** A large marketing literature studies consumer brand choice dynamics. A widely documented phenomenon is *consumption persistence*, whereby consumers who chose brand z yesterday are more likely to choose z again today than consumers who chose z' yesterday. Our analysis is helpful in distinguishing various explanations of this phenomenon.

First, analogous to the voting example in the introduction, one possible explanation is that consumers' tastes (u_t) display persistence, so that consumers who preferred z yesterday are also more likely to prefer it today. Here past consumption has no *causal* connection with current consumption; it simply provides information about a consumer's preferences to the analyst. We refer to this phenomenon as *history dependence*. Our axioms in Section 3.1 impose limits on

[5]A special case of the model where all information is resolved at the beginning of time corresponds to a population of heterogeneous agents with fully persistent preferences, or a single agent with random but fully persistent preferences.

how much history dependence the analyst can observe as a consequence of serially correlated utilities. In addition, Section 5.2 characterizes precisely which form of taste persistence in (u_t) gives rise to consumption persistence.

An alternative explanation is *consumption dependence*, where consuming z yesterday *directly* affects the consumer's utility today, for example due to a process of habit formation. Our baseline model excludes this direct effect and assumes that utility does not depend on past consumption. However, in Section 7 we extend our characterization to additionally allow for consumption dependence, identifying precisely when the analyst must attribute a causal role to past consumption.

A further question concerns different interpretations of serially correlated utilities. The most general possibility is that the agent is subject to arbitrary correlated taste shocks. An important special case of this (e.g., Erdem and Keane 1996) is a consumer with a fixed but unknown utility $\tilde{u}$, about which she learns over time; in this case, u_t represents her expectation of $\tilde{u}$ given period t information. On static domains, a consumer with random taste shocks is indistinguishable from one that receives signals about a fixed state of the world. However, we show that in the dynamic setting the learning model has additional behavioral implications, which we identify in Section 4. Moreover, in the case of learning, Section 7 again enables us to distinguish between learning that is independent of past consumption (e.g., learning from advertising in Erdem and Keane 1996) and active learning/experimentation, where past consumption itself is the source of information (e.g., learning from experience in Erdem and Keane 1996). ▲

In many dynamic settings, the agent's choices today also affect her opportunity sets tomorrow, giving rise to additional questions.

Example 2 (School choice)**.** A growing literature in labor economics studies individuals' school and curriculum choices, recognizing the importance of such choices for eventual labor market outcomes. As a stylized example, consider Figure 1 (left). Here a parent first decides whether to enroll her child in one of two elementary schools, which differ along many decision-relevant dimensions. Upon enrolling, the parent must decide between a number of after-school care options. In either case, she could enroll the child in a high quality but high cost private after-school center (P) or stay at home/leave the child with relatives (H); additionally, school 1, unlike school 2, offers its own (more basic and lower cost) after-school program (S).

In this setting, history dependence can take the form of a dynamic selection effect, whereby parents' after-school choices differ across schools because parents with different preferences select into different schools. Failure to account for such dynamic selection may lead to misspecified models, including *spurious violations of random utility*. For example, suppose choice frequencies for each after-school option are as in Figure 1 (left); in particular, the share of parents choosing the private program is larger at school 1 (30%) than school 2 (20%). Ignoring history dependence, this behavior appears inconsistent with static random utility maximization

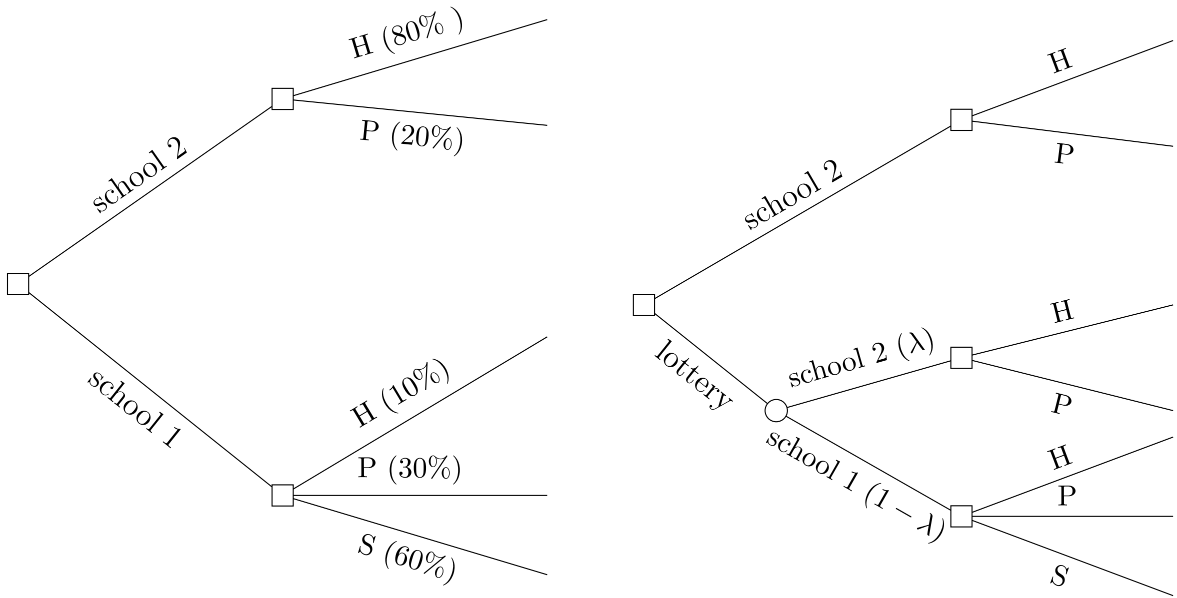

Figure 1: School choice.

as it violates Regularity, Axiom 0 (i). However, it is entirely consistent with dynamic random utility maximization, because under serially correlated private information the preferences of parents at each school will differ. For example, it could be that parents for whom option H is relatively more costly select disproportionately into school 1 because it expands their set of outside-the-home options and some of these parents subsequently choose the private program after obtaining additional information; or a preference for the specific characteristics of school 2 may happen to be strongly correlated with a preference for at-home child care. In Section 3 we will characterize precisely what kind of choice behavior is consistent with dynamic random (expected) utility, showing in particular that the static REU axioms are valid as long as the analyst conditions on past histories.

Another implication of history dependence is *limited observability*: Unlike in the static setting, where the analyst has (at least in principle) access to choice frequencies from all menus, in the dynamic setting certain past choices rule out certain future menus, so that over time the analyst observes choices from a more and more limited domain. For example, we cannot observe the counterfactual frequencies with which parents at school 1 would choose from the set $\{\mathrm{H}, \mathrm{P}\}$ if S were not available to them; and given dynamic selection, we cannot simply infer these from the corresponding choice probabilities at school 2. In practice, however, many schools ration their seats via lotteries, a fact that is widely exploited in the empirical literature on school choice to generate quasi-experimental variation.[6] This is illustrated in Figure 1 (right), where

[6]E.g., Abdulkadiroglu, Angrist, Narita, and Pathak (forthcoming); Angrist, Hull, Pathak, and Walters (forthcoming); Deming (2011); Deming, Hastings, Kane, and Staiger (2014).

each application to school 1 is successful with probability λ and the parent must select school 2 otherwise. We will see in Section 3.1 that the analyst can (under expected utility) extrapolate the choices of school 1 parents from the set $\{\mathrm{H}, \mathrm{P}\}$ by looking at choices of parents who applied to school 1 but were rejected by the lottery. ▲

In what follows, we develop and analyze a general model of dynamic random utility that encompasses these and similar examples.

2.2.1 Agent's problem

The agent faces a decision tree, as defined by Kreps and Porteus (1978). There are finitely many periods $t = 0, 1, \dots, T$. There is a finite set Z of instantaneous consumptions. Each period t, the agent chooses from a period-t menu, which is a finite set of lotteries over the period-t outcome space X_t. The spaces X_t are defined recursively. The final period outcome space $X_T := Z$ is just the space of instantaneous consumptions; the set of all period-T menus is $\mathcal{A}_T := \mathcal{K}(\Delta(X_T))$. In all earlier periods $t \leq T-1$, the outcome space $X_t := Z \times \mathcal{A}_{t+1}$ consists of all pairs of current period consumptions and next period continuation menus; the set of period-t menus is $\mathcal{A}_t := \mathcal{K}(\Delta(X_t))$.[7] Denote a typical period-t lottery by $p_t \in \Delta(X_t)$ and a typical menu by $A_t \in \mathcal{A}_t$. The agent's choice of $p_t \in A_t$ determines both her instantaneous consumption z_t and the menu A_{t+1} from which she will choose next period; let $p_t^Z \in \Delta(Z)$ and $p_t^A \in \Delta(\mathcal{A}_{t+1})$ denote the respective marginal distributions.

As in the static model, let $(\Omega, \mathcal{F}^*, \mu)$ be a finitely-additive probability space. Under *dynamic random expected utility (DREU)*, in each state of the world and in each period, the agent's choices maximize her expected utility subject to her dynamically evolving private information. The agent's payoff-relevant private information is captured by a filtration $(\mathcal{F}_t)_{0 \leq t \leq T} \subseteq \mathcal{F}^*$ and an $\mathcal{F}_t$-adapted process of random vNM utility indices $U_t : \Omega \to \mathbb{R}^{X_t}$ over X_t. This allows for arbitrary serial correlation of utilities, but does not allow the utility process to depend on past consumption; Section 7 relaxes the latter restriction. In case of indifference, ties at each t are broken by a random $\mathcal{F}^*$-measurable vNM utility index $W_t : \Omega \to X_t$, where we impose dynamic analogs of simplicity and properness that we define at the end of this section. Thus, as before, when faced with menu A_t in period t, the agent chooses lottery p_t in state ω if and only if $p_t \in M(M(A_t, U_t(\omega)), W_t(\omega))$.

DREU is a very general model because it imposes no particular structure on the family (U_t). This is the most parsimonious setting in which to isolate the behavioral implications of serially correlated private information. DREU could also accommodate various behavioral effects, such as temptation or certain forms of "mistakes," which in the static setting are indistinguishable

[7] A small technical difference from Kreps and Porteus (1978) is that they use Borel instead of simple lotteries and compact instead of finite menus, but as in their setting we can verify recursively that each X_t is a separable metric space under the appropriate topologies (see Lemma 12).

from random utility maximization. However, the following important special case rules out these possibilities.

Evolving utility captures a dynamically sophisticated agent who correctly takes into account the evolution of her future preferences. There is an $\mathcal{F}_t$-adapted process of random felicity functions $u_t : \Omega \to \mathbb{R}^Z$ over instantaneous consumptions and a discount factor $\delta > 0$ such that $U_T = u_T$ and U_t for $t \leq T$ is given by the Bellman equation

$$U_t(z_t, A_{t+1}) = u_t(z_t) + \delta \mathbb{E}\left[\max_{p_{t+1} \in A_{t+1}} U_{t+1}(p_{t+1}) | \mathcal{F}_t\right]. \tag{1}$$

Finally, as discussed in Example 1, an important special case of evolving utility arises when the agent has a fixed but unknown felicity about which she learns over time. In this *gradual learning* model there is a $\mathcal{F}^*$-measurable random felicity $\tilde{u} : \Omega \to \mathbb{R}^Z$ such that for all t[8]

$$u_t = \mathbb{E}[\tilde{u}|\mathcal{F}_t]. \tag{2}$$

For all three models, we impose the following dynamic analogs of simplicity and properness. The pair $(\mathcal{F}_t, U_t)_{0 \leq t \leq T}$ is *simple*, i.e., (i) each $\mathcal{F}_t$ is generated by a finite partition such that $\mu(\mathcal{F}_t(\omega)) > 0$ for every $\omega \in \Omega$, where $\mathcal{F}_t(\omega)$ again denotes the cell of the partition that contains ω; and (ii) each $U_t(\omega)$ is nonconstant, and $U_t(\omega) \not\approx U_t(\omega')$ whenever $\mathcal{F}_t(\omega) \neq \mathcal{F}_t(\omega')$ and $\mathcal{F}_{t-1}(\omega) = \mathcal{F}_{t-1}(\omega')$.[9] The tiebreakers $(W_t)_{0 \leq t \leq T}$ are *proper*, i.e., (i) $\mu(\{\omega \in \Omega : |M(A_t, W_t(\omega))| = 1\}) = 1$ for all $A_t \in \mathcal{A}_t$; (ii) conditional on $\mathcal{F}_T(\omega)$, $W_0, \ldots, W_T$ are independent; and (iii) $\mu(W_t \in B_t | \mathcal{F}_T(\omega)) = \mu(W_t \in B_t | \mathcal{F}_t(\omega))$ for all t and measurable B_t.[10]

2.2.2 Analyst's problem

As in the static setting, the agent's choices in each period t appear stochastic to the analyst, because he does not have access to the agent's private information. The novel feature of the dynamic setting is that the analyst can observe the agent's past choices. With serially correlated utilities, these choices convey some information about the payoff-relevant private information $\mathcal{F}_t$, so that the agent's behavior additionally appears *history dependent* to the analyst.

This is captured by a *dynamic stochastic choice rule* ρ, which for any period t and history of past choices summarizes the observed choice frequencies from any menu A_t that can arise after this history. We define choice frequencies and histories recursively. Choice frequencies

[8]Gradual learning is a model of passive learning, because the agent's choices do not affect her filtration $\mathcal{F}_t$. The more general model in Section 7 accommodates as a special case active learning/experimentation, where each period the agent obtains additional information from her consumption z_t.

[9]For $t = 0$, we let $\mathcal{F}_{t-1}(\omega) := \Omega$ for all ω.

[10](ii) rules out additional serial correlation of tiebreakers, over and above the serial correlation inherent in the agent's payoff-relevant private information $\mathcal{F}_T(\omega)$. (iii) ensures that to the extent that period-t tie breaking relies on payoff-relevant private information, it can rely only on the information $\mathcal{F}_t(\omega)$ available at t.

in period 0 are given by a (static) stochastic choice rule $\rho_0 : \mathcal{A}_0 \to \Delta(\Delta(X_0))$ on $\mathcal{A}_0$; thus, $\sum_{p_0 \in A_0} \rho_0(p_0; A_0) = 1$ for all A_0 and $\rho_0(p_0; A_0)$ denotes the frequency with which the agent chooses lottery p_0 when faced with menu A_0. The choices that occur with strictly positive probability under ρ_0 define the set of all period 0 *histories* $\mathcal{H}_0 := \{(A_0, p_0) : \rho_0(p_0, A_0) > 0\}$. For any history $h^0 = (A_0, p_0) \in \mathcal{H}_0$, let $\mathcal{A}_1(h^0) := \operatorname{supp} p_0^A$ denote the set of period 1 menus that follow h^0 with positive probability.

For $t \geq 1$ the objects $\mathcal{H}_t$ and $\mathcal{A}_{t+1}(h^t)$ are defined recursively. For any history $h^{t-1} \in \mathcal{H}_{t-1}$, choice frequencies following h^{t-1} are given by a stochastic choice rule $\rho_t(\cdot|h^{t-1}) : \mathcal{A}_t(h^{t-1}) \to \Delta(\Delta(X_t))$ on the set $\mathcal{A}_t(h^{t-1})$ of period t menus that follow h^{t-1} with positive probability; thus, $\sum_{p_t \in A_t} \rho_t(p_t; A_t \mid h^{t-1}) = 1$ for all $A_t \in \mathcal{A}_t(h^{t-1})$ and $\rho_t(p_t; A_t \mid h^{t-1})$ denotes the frequency with which the agent chooses p_t when faced with menu A_t after history h^{t-1}. The set of period-t histories is $\mathcal{H}_t := \{(h^{t-1}, A_t, p_t) : h^{t-1} \in \mathcal{H}_{t-1}$ and $A_t \in \mathcal{A}_t(h^{t-1})$ and $\rho_t(p_t; A_t|h^{t-1}) > 0\}$; this contains all sequences $(A_0, p_0, \ldots, A_t, p_t)$ of choices up to time t that arise with positive probability. Finally, for each $t \leq T-1$, the set of period $t+1$ menus that follow history $h^t = (h^{t-1}, A_t, p_t)$ with positive probability is $\mathcal{A}_{t+1}(h^t) := \operatorname{supp} p_t^A$ and the set of period-t histories that lead to A_{t+1} with positive probability is $\mathcal{H}_t(A_{t+1}) := \{h^t \in \mathcal{H}_t : A_{t+1} \in \mathcal{A}_{t+1}(h^t)\}$.

Two features of the primitive are worth noting: First, reflecting *limited observability*, for each $t \geq 1$ and history $h^{t-1} \in \mathcal{H}_{t-1}$, the stochastic choice rule $\rho_t(\cdot|h^{t-1})$ is defined only on the subset $\mathcal{A}_t(h^{t-1}) \subseteq \mathcal{A}_t$ of period t menus that arise with positive probability after h^{t-1}—typically very few menus. Nevertheless, Section 3.2 will show that under DREU the analyst can extrapolate from $\rho_t(\cdot|h^{t-1})$ to a well-defined stochastic choice rule on the whole of $\mathcal{A}_t$. Second, histories only summarize the agent's past choices of p_k from A_k and do not keep track of realized consumptions $z_k \in \operatorname{supp} p_k^Z$. This is without loss in the current model where utilities are not affected by past consumption, but will be relaxed in the model with consumption-dependence in Section 7.

Under DREU, the private information revealed to the analyst by history $h^{t-1} = (A_0, p_0, \ldots, A_{t-1}, p_{t-1})$ is given by the event $C(h^{t-1}) := \bigcap_{k=0}^{t-1} C(p_k, A_k)$, where for each k the event $C(p_k, A_k) := \{\omega \in \Omega : p_k \in M(M(A_k, U_k(\omega)), W_k(\omega))\}$ that the agent chooses p_k when faced with A_k is defined as in the static model.[11] Thus, the analyst's observations are consistent with DREU if the frequency with which the agent chooses p_t from A_t following history h^{t-1} is equal to the conditional probability $\mu[C(p_t, A_t)|C(h^{t-1})]$ of the event $C(p_t, A_t)$ given $C(h^{t-1})$.

The following definition summarizes the dynamic model:

Definition 2. A *dynamic random expected utility (DREU)* representation of the dynamic stochastic choice rule ρ is a tuple $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t)_{0 \leq t \leq T})$ such that $(\Omega, \mathcal{F}^*, \mu)$ is a finitely-additive probability space, the filtration $(\mathcal{F}_t) \subseteq \mathcal{F}^*$ and the $\mathcal{F}_t$-adapted utility process

[11] Note that $C(h^{t-1})$ does not keep track of the random realizations of menus $A_k \in \operatorname{supp} p_k^A$ along the sequence h^{t-1}, as this exogenous randomness does not reveal any information about the agent's private information.

$U_t : \Omega \to \mathbb{R}^{X_t}$ are simple, the $\mathcal{F}^*$-measurable tiebreaking process $W_t : \Omega \to \mathbb{R}^{X_t}$ is proper, and for all $p_t \in A_t$ and $h^{t-1} \in \mathcal{H}_{t-1}(A_t)$,

$$\rho_t(p_t; A_t|h^{t-1}) = \mu\left[C(p_t, A_t)|C(h^{t-1})\right], \tag{3}$$

where for $t = 0$, we abuse notation by letting $C(h^{t-1}) := \Omega$ and $\rho_0(p_0; A_0|h^{-1}) := \rho_0(p_0; A_0)$.

An *evolving utility* representation is a DREU representation along with an $\mathcal{F}_t$-adapted process of felicities $u_t : \Omega \to \mathbb{R}^Z$ and a discount factor $\delta > 0$ such that (1) holds. A *gradual learning* representation is an evolving utility representation along with an $\mathcal{F}^*$-measurable felicity $\tilde{u} : \Omega \to \mathbb{R}^Z$ such that (2) holds.

2.2.3 Discussion

Lotteries as choice objects: In addition to allowing us to model choice behavior under risk, including lotteries in the domain of choice simplifies our analysis, as it allows us to rely on the static framework of Gul and Pesendorfer (2006) instead of the more complicated one of Falmagne (1978). Lotteries play a similar technical role in the original work of Kreps and Porteus (1978), by letting them rely on the vNM framework.[12] From a conceptual point of view, we will see in Section 3.2 that lotteries are crucial in overcoming the aforementioned limited observability problem and we illustrate the availability of lotteries for this purpose with examples from experimental and empirical work. Relatedly, lotteries are key in inferring the agent's history dependent revealed preference in Section 4.1 and in disentangling history dependence from consumption dependence in Section 7.

Interpretation of data: As with static stochastic choice, the dynamic stochastic choice rule ρ admits two equivalent interpretations: The analyst either (i) repeatedly observes a single agent solve each decision tree;[13] or (ii) observes a large population of agents (with heterogeneous and evolving utilities) solve each decision tree once. In either case, ρ captures the limiting choice frequencies as the number of observations/population size tends to infinity. Abstracting from the sampling error in this manner is also typical in the econometric analysis of identification. In any application, the data set will of course be finite. However, studying behavior on the full domain is an important step in uncovering all the assumptions that are behind the model; moreover, statistical tests are often directly inspired by axioms.[14]

Dynamic stochastic choice vs. ex ante preference: In our framework, the analyst

[12]Likewise, the ambiguity aversion literature extensively relies on the Anscombe and Aumann (1963) framework rather than the more complicated one of Savage (1972); the notable exceptions include Gilboa (1987) and Epstein (1999). Similarly, the menu-preference literature uses lotteries (e.g. Dekel, Lipman, and Rustichini, 2001) to improve upon the uniqueness and comparative statics results of Kreps (1979).

[13]Here, the agent's utilities are assumed to evolve according to the same process U_t at each observation.

[14]For example Hausman and McFadden (1984) develop a test of the IIA axiom that characterizes the logit model. Likewise, Kitamura and Stoye (2016) develop axiom-based tests of the static random utility model.

observes the distribution of choices at each node of each decision tree; as we pointed out, the randomness in choice comes from an informational asymmetry between the agent and the analyst that occurs in each period. By contrast, a widespread approach in the existing dynamic decision theory literature (e.g., Gul and Pesendorfer, 2004; Krishna and Sadowski, 2014) is to only study a deterministic preference over decision trees at a hypothetical ex ante stage that features no informational asymmetry[15] or abstracts away from other forces (e.g., temptation) that the agent anticipates to affect her choices in actual decision trees.[16] Compared with this literature, our approach does not require such a hypothetical stage, and thus the primitive is closer to actual data economists can observe. Moreover, considering choice behavior in each period, not just at the beginning of time, allows us to study new phenomena such as history dependence and consumption persistence. In Section 4.1 we show how to extract deterministic preference relations from the stochastic choice data.

Role of axioms: In addition to their usual positive and normative role, we view our axioms as serving an equally important purpose as conceptual tools that elucidate key properties of any dynamic random utility model and facilitate comparisons between different versions of the model. For example, our axioms in Section 3.1 clarify the nature of history dependence that can arise under any dynamic random expected utility model; our axioms in Section 4.3 identify the additional behavioral content of gradual learning relative to evolving utility; and our comparison of the evolving utility and i.i.d. DDC model in Section 6 draws on the axioms to uncover that the two make opposite predictions about option value.

3 Characterization of DREU

DREU is characterized by four axioms, which we present in the following subsections. First, we present two history independence axioms that capture the key new implications of the dynamic model relative to the static one. Building on this, the next subsection shows how the analyst can extrapolate from each $\rho_t(\cdot|h^{t-1})$ to an extended choice rule on the whole of $\mathcal{A}_t$, thus overcoming the limited observability problem. The final subsection then imposes the static REU conditions as well as a technical history continuity axiom on this extended choice rule.

[15] Ahn and Sarver (2013) study a two-period model with a deterministic menu preference in the first period and random choice from menus in the second period. Here too there is no informational asymmetry in the first period.

[16] In the context of temptation, one exception is Noor (2011), but his is a stationary environment with no informational asymmetry and the analyst observes deterministic choices at each node of the decision tree.

3.1 History Independence Axioms

Our first two axioms identify two cases in which histories h^{t-1} and g^{t-1} reveal the same information to the analyst. Capturing the fact that history dependence arises in DREU only through the private information revealed by past choices, the axioms require that period-t choice behavior be the same after two such histories.

Given $h^{t-1} = (A_0, p_0, ..., A_{t-1}, p_{t-1}) \in \mathcal{H}_{t-1}$, let $(h^{t-1}_{-k}, (A'_k, p'_k))$ denote the sequence of the form $(A_0, p_0, ..., A'_k, p'_k, ..., A_{t-1}, p_{t-1})$.[17] We say that $g^{t-1} \in \mathcal{H}_{t-1}$ is *contraction equivalent* to h^{t-1} if for some k, we have $g^{t-1} = (h^{t-1}_{-k}, (B_k, p_k))$, where $A_k \subseteq B_k$ and $\rho_k(p_k, A_k|h^{k-1}) = \rho_k(p_k, B_k|h^{k-1})$.[18] That is, g^{t-1} and h^{t-1} differ only in period k, where under g^{t-1}, the agent chooses lottery p_k from menu B_k, while under h^{t-1}, she chooses the same lottery p_k from the contraction $A_k \subseteq B_k$; moreover, conditional on h^{k-1}, the choice of p_k from A_k and the choice of p_k from B_k occur with the same probability. Axiom 1 requires that choice behavior be the same after h^{t-1} and g^{t-1}.

Axiom 1 (Contraction History Independence)**.** For all $t \leq T$, if $g^{t-1} \in \mathcal{H}_{t-1}(A_t)$ is contraction equivalent to $h^{t-1} \in \mathcal{H}_{t-1}(A_t)$, then $\rho_t(\cdot, A_t|h^{t-1}) = \rho_t(\cdot, A_t|g^{t-1})$.

To see the idea, suppose for simplicity that $T = 1$, in which case the axiom requires that for any $p_0 \in A_0 \subseteq B_0$ if $\rho_0(p_0, A_0) = \rho_0(p_0, B_0)$, then $\rho_1(\cdot, A_1|A_0, p_0) = \rho_1(\cdot, A_1|B_0, p_0)$ for any $A_1 \in \operatorname{supp} p_0^A$. In general, the event that p_0 is the best element of menu B_0 is a subset of the event that p_0 is the best element of the smaller menu $A_0 \subseteq B_0$; thus, observing $g^0 = (B_0, p_0)$ may reveal more information about the agent's possible period-0 preferences than $h^0 = (A_0, p_0)$. However, since we additionally know that $\rho_0(p_0, A_0) = \rho_0(p_0, B_0)$, the event that p_0 is best in A_0 but not in B_0 must have probability 0; in other words, we must put zero probability on any preference that selects p_0 from A_0 but not from B_0. Given this, h^0 and g^0 reveal the same information, and hence call for the same predictions for period-1 choices. The following example illustrates this with a concrete story in the population setting.

Example 3. There are two convenience stores, A and B, each carrying three types of milk (whole, 2%, and 1%). Each store has a stable set of weekly customers whose stochastic process of preferences is identical at both stores.[19] Suppose that in week 0, store A's delivery of 1% milk breaks down unexpectedly. The purchasing shares at each store are given in Tables 1 and 2. Consider a customer of store A, Alice, and a customer of store B, Barbara, who both buy whole milk in week 0. If in week 1 all types of milk are available again at both stores, then Contraction History Independence implies that Alice and Barbara's choice probabilities will be the same. This is true because we have the same information about Alice and Barbara. Since

[17] In general this is not a history, but it is if $A'_k \in \operatorname{supp} p^A_{k-1}$ and $A_{k+1} \in \operatorname{supp} p'^A_k$ and $\rho_k(p'_k, A'_k|h^{k-1}) > 0$.

[18] This induces an equivalence relation on $\mathcal{H}_{t-1}$ by taking the symmetric and transitive closure.

[19] For simplicity, we assume in the following that all preferences are strict.

product	market share
whole	40%
2%	60%

Table 1: Market shares at store A

product	market share
whole	40%
2%	35%
1%	25%

Table 2: Market shares at store B

at store A only whole and 2% milk were available in week 0, the possible week-0 preferences of Alice are $w \succ 2 \succ 1$ or $w \succ 1 \succ 2$ or $1 \succ w \succ 2$. By contrast, since store B stocked all three types of milk, Barbara's possible preferences are $w \succ 2 \succ 1$ or $w \succ 1 \succ 2$. However, since we additionally know that the share of customers purchasing whole milk in week 0 was the same at both stores, $\rho_0(w, \{w, 1, 2\}) = \rho_0(w, \{w, 2\}) = 0.4$, we can also conclude that no customers had the ranking $1 \succ w \succ 2$ in week 0. Therefore, the analyst's prediction is the same, since the stochastic process that governs the transition from week-0 to week-1 preferences is the same for Barbara and Alice and in both cases the analyst conditions on exactly the same week-0 event $\{w \succ 2 \succ 1, w \succ 1 \succ 2\}$. ▲

The second history independence axiom takes into account that the agent is an expected utility maximizer. Under expected utility maximization, choosing p_k from A_k reveals the same information about the agent's utility as choosing $\lambda p_k + (1-\lambda)q_k$ from $\lambda A_k + (1-\lambda)\{q_k\}$. More generally, for a menu B_k, if we know that the agent chose *some* option of the form $\lambda p_k + (1-\lambda)q_k$ from $\lambda A_k + (1-\lambda)B_k$ but we do not know what q_k was, this again reveals the same information as choosing p_k from A_k. This suggests the following notion of equivalence: We say that a finite set of histories $G^{t-1} \subseteq \mathcal{H}_{t-1}$ is *linearly equivalent* to $h^{t-1} = (A_0, p_0, ..., A_{t-1}, p_{t-1}) \in \mathcal{H}_{t-1}$ if

$$G^{t-1} = \{(h^{t-1}_{-k}, (\lambda A_k + (1-\lambda)B_k, \lambda p_k + (1-\lambda)q_k)) : q_k \in B_k\}$$

for some k, B_k, and $\lambda \in (0, 1]$. That is, G^{t-1} is the collection of histories that differ from h^{t-1} only at period k: Under h^{t-1}, the agent chooses p_k from menu A_k, while G^{t-1} summarizes all possible choices of the form $\lambda p_k + (1-\lambda)q_k$ from the menu $\lambda A_k + (1-\lambda)B_k$. By the above reasoning, G^{t-1} reveals the same information about the agent as h^{t-1}. Thus, Axiom 2 requires period-t choice behavior following the set of histories G^{t-1} to be the same as conditional on h^{t-1}. To state this formally, define the choice distribution from A_t following $G^{t-1} \subseteq \mathcal{H}_{t-1}(A_t)$,

$$\rho_t(\cdot, A_t | G^{t-1}) := \sum_{g^{t-1} \in G^{t-1}} \rho_t(\cdot, A_t | g^{t-1}) \frac{\rho(g^{t-1})}{\sum_{f^{t-1} \in G^{t-1}} \rho(f^{t-1})},$$

to be the weighted average of all choice distributions $\rho_t(\cdot, A_t | g^{t-1})$ following histories in G^{t-1}, where for each $g^{t-1} = (\hat{A}_0, \hat{p}_0, \ldots, \hat{A}_{t-1}, \hat{p}_{t-1})$ its weight $\rho(g^{t-1}) := \prod_{k=0}^{t-1} \rho_k(\hat{p}_k, \hat{A}_k | g^{k-1})$ corresponds to the probability of the sequence of choices summarized by g^{t-1}.[20]

Axiom 2 (Linear History Independence)**.** For all $t \leq T$, if $G^{t-1} \subseteq \mathcal{H}_{t-1}(A_t)$ is linearly equivalent to $h^{t-1} \in \mathcal{H}_{t-1}(A_t)$, then $\rho_t(\cdot, A_t|h^{t-1}) = \rho_t(\cdot, A_t|G^{t-1})$.

3.2 Limited Observability

Recall that unlike the static setting, where the analyst observes choices from all possible menus, the dynamic setting presents a limited observability problem: At each history h^{t-1} of past choices, $\rho_t(\cdot|h^{t-1})$ is only defined on the set $\mathcal{A}_t(h^{t-1})$ of menus that occur with positive probability after h^{t-1}—typically very few menus. For the rest of the paper, it is key to overcome this problem: Otherwise we do not have enough data to verify whether observed choices at history h^{t-1} are consistent with random utility maximization or to identify whether the agent's utility process belongs to the evolving utility class or the more specific gradual learning class.

The inclusion of lotteries among the agent's choice objects allows us to do so. The idea is to use Linear History Independence to formalize the "linear extrapolation" procedure illustrated in the school choice example (Example 2). Consider any menu A_t (e.g., the two-option menu $\{H, P\}$ in the example) and some history h^{t-1} that does not lead to A_t (e.g., choosing school 1). We define the agent's counterfactual choice distribution from A_t following h^{t-1} by extrapolating from the situation where she makes the sequence of choices captured by h^{t-1}, but knows that with some probability another sequence of choices (e.g., enroll in school 2) that *does* lead to menu A_t will be implemented instead. More precisely, we consider a degenerate sequence $d^{t-1} = (\{q_0\}, q_0, \dots, \{q_{t-1}\}, q_{t-1})$ such that $A_t \in \operatorname{supp} q^A_{t-1}$ and replace $h^{t-1} = (A_0, p_0, \dots, A_{t-1}, p_{t-1})$ with $g^{t-1} := \lambda h^{t-1} + (1-\lambda)d^{t-1}$ where[21] at every period $k \leq t-1$, the agent faces menu $\lambda A_k + (1-\lambda)\{q_k\}$ and chooses lottery $\lambda p_k + (1-\lambda)q_k$.

As discussed preceding Linear History Independence, under expected utility maximization the latter sequence of choices reveals the same information about the agent as h^{t-1}. This motivates extrapolating from g^{t-1} to define choice behavior following h^{t-1}. Define the set of degenerate period-$(t-1)$ histories by $\mathcal{D}_{t-1} := \{d^{t-1} \in \mathcal{H}_{t-1} : d^{t-1} = (\{q_k\}, q_k)_{k=0}^{t-1} \text{ where } q_k \in \Delta(X_k)\, \forall k \leq t-1\}$.

Definition 3. For any $t \geq 1$, $A_t \in \mathcal{A}_t$, and $h^{t-1} \in \mathcal{H}_{t-1}$, define

$$\rho_t^{h^{t-1}}(\cdot; A_t) := \rho_t(\cdot; A_t|\lambda h^{t-1} + (1-\lambda)d^{t-1}). \tag{4}$$

[20]Note that $\rho(g^{t-1})$ does not keep track of the probabilities $\hat{p}^A_k(\hat{A}_{k+1})$, since these pertain to exogenous randomization and do not reveal any private information.

[21]In order for $\lambda h^{t-1} + (1-\lambda)d^{t-1} := (\lambda A_k + (1-\lambda)\{q_k\}, \lambda p_k + (1-\lambda)q_k)_{k=0}^{t-1}$ to be a well-defined history, it suffices that $\lambda A_k + (1-\lambda)\{q_k\} \in \operatorname{supp} q^A_{k-1}$ for all $k = 1, \dots, t-1$. This can be ensured by appropriately choosing each q_k, working backwards from period $t-1$.

for some $\lambda \in (0,1]$ and $d^{t-1} \in \mathcal{D}_{t-1}$ such that $\lambda h^{t-1} + (1-\lambda) d^{t-1} \in \mathcal{H}_{t-1}(A_t)$.

It follows from Axiom 2 (Linear History Independence) that $\rho_t^{h^{t-1}}(\cdot; A_t)$ is well-defined: Lemma 15 shows that the RHS of (4) does not depend on the specific choice of λ and d^{t-1}. Moreover, $\rho_t^{h^{t-1}}(\cdot; A_t)$ coincides with $\rho_t(\cdot; A_t|h^{t-1})$ whenever $h^{t-1} \in \mathcal{H}_{t-1}(A_t)$. In the following, we do not distinguish between the extended and nonextended version of ρ_t and use $\rho_t(\cdot; A_t|h^{t-1})$ to denote both.

As illustrated by Example 2 in the context of school choice, random assignment is prevalent in many real-world economic environments and is an important tool to obtain quasi-experimental variation in the empirical literature. While this literature typically leverages such random variation to identify the causal effect of current choices on next-period *outcomes* (e.g., test scores in the case of school choice), Definition 3 suggests exploiting it to make counterfactual inferences about next-period *choices*. Even more readily, lotteries over next-period choice problems can be generated in the laboratory, and a growing literature in experimental economics makes use of this to perform extrapolation procedures akin to Definition 3.[22]

3.3 History-Dependent REU and History Continuity Axioms

For each h^{t-1}, the extended choice distribution $\rho_t(\cdot|h^{t-1})$ from Definition 3 is a stochastic choice rule on the whole of $\mathcal{A}_t$. The next axiom imposes the standard static REU conditions from Axiom 0 on each $\rho_t(\cdot|h^{t-1})$. Note that conditioning ρ_t on past histories is key here; without controlling for past choices, choice behavior at time t will in general violate the REU axioms, as illustrated in Example 2.

Axiom 3 (History-dependent REU)**.** For all $t \leq T$ and h^{t-1}, $\rho_t(\cdot|h^{t-1})$ satisfies Axiom 0.[23]

Our final axiom reflects the way in which tie-breaking can affect the observed choice distribution. We first define menus and histories without ties directly from choice behavior. The idea is that menus without ties are characterized by the fact that slightly perturbing their elements has no effect on choice probabilities.[24] We capture such perturbations using convergence in mixture, as defined following Axiom 0.

[22] E.g., in a recent experimental study of temptation and self-control, Toussaert (2016) presents subjects with lotteries over next-period menus to differentiate between so-called random Strotz agents and Gul and Pesendorfer (2001) agents. For related uses of lotteries in lab experiments, see Augenblick, Niederle, and Sprenger (2015).

[23] Lemma 12 verifies that X_t is a separable metric space. Then Mixture Continuity and Finiteness make use of the same convergence notions as defined following Axiom 0.

[24] Lu (2016) and Lu and Saito (2016) use an alternative approach, directly incorporating into the primitive a collection of measurable sets that capture the absence of ties and defining choice probabilities only on measurable subsets of each menu. Their approach requires that ties occur with probability either zero or one, so is not applicable to our setting. Our perturbation-based approach is similar in spirit to Ahn and Sarver (2013).

Definition 4. For any $0 \leq t \leq T$ and $h^{t-1} \in \mathcal{H}_{t-1}$, the set of period-$t$ *menus without ties* conditional on history h^{t-1} is denoted $\mathcal{A}_t^*(h^{t-1})$[25] and consists of all $A_t \in \mathcal{A}_t$ such that for any $p_t \in A_t$ and any sequences $p_t^n \to^m p_t$ and $B_t^n \to^m A_t \smallsetminus \{p_t\}$, we have

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

For $t = 0$, we write $\mathcal{A}_0^* := \mathcal{A}_0^*(h^{t-1})$. The set of period t *histories without ties* is $\mathcal{H}_t^* := \{h^t = (A_0, p_0, \ldots, A_{t-1}, p_{t-1}) \in \mathcal{H}_t : A_k \in \mathcal{A}_k^*(h^{k-1}) \text{ for all } k \leq t\}$.

The following axiom relates choice distributions after nearby histories. To state this formally, we extend convergence in mixture to histories: We say $h^{t,n} \to^m h^t$ if $h^{t,n} = (A_0^n, p_0^n, ..., A_t^n, p_t^n)$ and $h^t = (A_0, p_0, ..., A_t, p_t)$ satisfy $A_k^n \to^m A_k$ and $p_k^n \to^m p_k$ for each k.

Axiom 4 (History Continuity)**.** For all $t \leq T-1$, A_{t+1}, p_{t+1}, and h^t,

$$\rho_{t+1}(p_{t+1}; A_{t+1} | h^t) \in \text{co}\{\lim_n \rho_{t+1}(p_{t+1}; A_{t+1} | h^{t,n}) : h^{t,n} \to^m h^t, h^{t,n} \in \mathcal{H}_t^*\}.$$

In general, if period t histories are slightly altered, we expect subsequent period $t+1$ choice behavior to be adjusted continuously, except when there was tie-breaking in the past. If the agent chose p_t from A_t as a result of tie-breaking, then slightly altering the choice problem can change the set of states at which p_t would be chosen and hence lead to a discontinuous change in the private information revealed by the choice of p_t. The history continuity condition restricts the types of discontinuities ρ_{t+1} can admit, ruling out situations in which choices after some history are completely unrelated to choices after any nearby history. Specifically, the fact that choice behavior after h^t can be expressed as a mixture of behavior after some nearby histories without ties reflects the way in which the agent's tie-breaking procedure may vary with her payoff-relevant private information.

3.4 Representation Theorem

Theorem 1. For any dynamic stochastic choice rule ρ, the following are equivalent:

(i). ρ satisfies Axioms 1–4.

(ii). ρ admits a DREU representation.

The proof of Theorem 1 appears in Appendix B. We now sketch the argument for sufficiency in the two-period setting ($T = 1$). Readers wishing to proceed directly to the analysis of evolving utility and gradual learning may skip ahead to Section 4.

[25] Note that $\mathcal{A}_t^*(h^{t-1}) \not\subseteq \mathcal{A}_t(h^{t-1})$ because the first set contains all menus without ties (we use history h^{t-1} here only to determine where ties could occur) while the second set contains only menus that occur with positive probability after history h^{t-1}—typically very few menus.

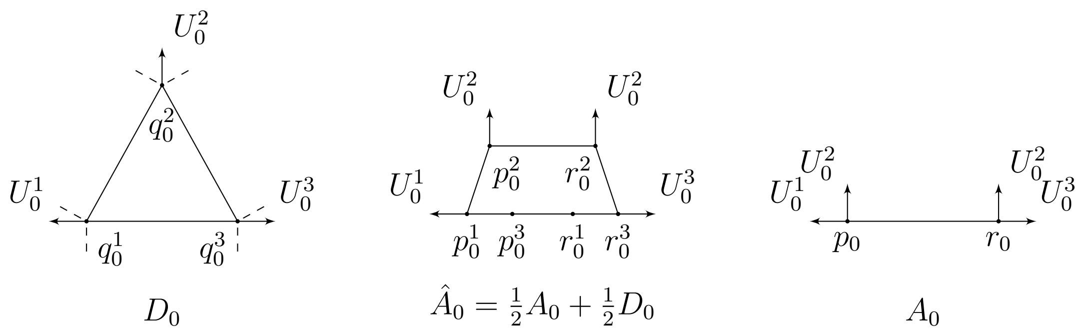

Figure 2: Suppose $S_0 = \{s_0^1, s_0^2, s_0^3\}$ with corresponding utilities U_0^1, U_0^2, U_0^3. Menu D_0 is a separating menu from which q_0^i is chosen precisely in state s_0^i. In menu $A_0 = \{p_0, r_0\}$, p_0 is chosen with probability 1 in state s_0^1; tied with r_0 in s_0^2; and never chosen in s_0^3. In $\hat{A}_0 = \frac{1}{2}A_0 + \frac{1}{2}D_0$, p_0 is replaced with three copies $\{p_0^1, p_0^2, p_0^3\}$: Each p_0^i is chosen in state s_0^i with the same probability with which p_0 is chosen at s_0^i and is never chosen otherwise. Step 3 shows choice probabilities following $(\hat{A}_0, p_0^i)$ are the same as following (D_0, q_0^i). Step 4 shows choice probabilities following (A_0, p_0) are a weighted sum of choice probabilities following $(\hat{A}_0, p_0^i)$, with weights given by $\mu_0(s_0^i|C_0(p_0, A_0))$. Combined with the static representation of $\rho_1(\cdot|D_0, q_0^i)$ (Step 2), this yields $\rho_1(p_1, A_1|A_0, p_0) = \sum_{i=1}^3 \mu_1^{s_0^i}(C_1^i(p_1, A_1))\mu_0(s_0^i|C_0(p_0, A_0))$.

Step 1: Static random expected utility representations. Since each X_t $(t = 0, 1)$ is a separable metric space (Lemma 12), Axiom 3 (History-dependent REU) together with Theorem 0 yields a static REU representation $(\Omega_0, \mathcal{F}_0^*, \mu_0, \mathcal{F}_0, U_0, W_0)$ of ρ_0 and for each $h^0 \in \mathcal{H}_0$, a static REU representation $(\Omega_1^{h^0}, \mathcal{F}_1^{*h^0}, \mu_1^{h^0}, \mathcal{F}_1^{h^0}, U_1^{h^0}, W_1^{h^0})$ of $\rho_1(\cdot|h^0)$. Thus far, there is no relationship between the period-0 and period-1 representations. In the following, we use Axioms 1, 2, and 4 to combine them into a DREU representation, which requires $\rho_1(p_1, A_1|A_0, p_0)$ to be represented as a conditional probability, with respect to a single underlying probability space Ω, of the event $C(p_1, A_1)$ given the event $C(p_0, A_0)$.

Step 2: Period-1 choices conditional on period-0 states. Let S_0 be the finite partition of Ω_0 that generates $\mathcal{F}_0$. We refer to cells $s_0 \in S_0$ as states and let $U_{s_0} = U_0(\omega)$ for any $\omega \in s_0 \in S_0$. For any history (A_0, p_0), define $\mathcal{U}_0(A_0, p_0) := \{U_{s_0} : s_0 \in S_0 \text{ and } p_0 \in M(A_0, U_{s_0})\}$ to be the set of period-0 utilities consistent with the choice of p_0 from A_0. Since $(U_0, \mathcal{F}_0)$ is simple, each U_{s_0} is nonconstant and induces a different preference, so by standard arguments (Lemma 13 in the appendix) we can find a menu $D_0 = \{q_0^{s_0} : s_0 \in S_0\}$ that strictly separates all states, i.e., such that for any $s_0^* \in S_0$ we have $\mathcal{U}_0(D_0, q_0^{s_0^*}) = \{U_{s_0^*}\}$. Figure 2 shows an example. For the remainder of this proof sketch, fix such a separating menu D_0 and define $\rho_1^{s_0}(p_1, A_1) := \rho_1(p_1, A_1|D_0, q_0^{s_0})$ for each A_1 and p_1. The representation of $\rho_1(\cdot|D_0, q_0^{s_0})$ obtained in Step 1 then constitutes a static REU representation $(\Omega_1^{s_0}, \mathcal{F}_1^{*s_0}, \mu_1^{s_0}, \mathcal{F}_1^{s_0}, U_1^{s_0}, W_1^{s_0})$ of $\rho_1^{s_0}$.

Step 3: $\rho_1^{s_0}$ is well-defined. We now use Linear History Independence, Contraction History Independence, and History Continuity to show that for any $(B_0, q_0) \in \mathcal{H}_0$ such that $\mathcal{U}_0(B_0, q_0) = \{U_{s_0}\}$, we have $\rho_1(\cdot, A_1|B_0, q_0) = \rho_1^{s_0}(\cdot, A_1)$; that is, $\rho_1^{s_0}$ describes choice behavior after *any* history that is only consistent with state s_0. To see this, assume first that $M(B_0, U_{s_0}) =$

$\{q_0\}$, i.e., q_0 is the unique maximizer of U_{s_0} in B_0. Define $r_0 := \frac{1}{2}q_0 + \frac{1}{2}q_0^{s_0}$, $\tilde{B}_0 := \frac{1}{2}B_0 + \frac{1}{2}\{q_0^{s_0}\}$, and $\tilde{D}_0 := \frac{1}{2}\{q_0\} + \frac{1}{2}D_0$. Then $\mathcal{U}(\tilde{B}_0, r_0) = \mathcal{U}(\tilde{B}_0 \cup \tilde{D}_0, r_0) = \mathcal{U}(\tilde{D}_0, r_0) = \{U_{s_0}\}$. From the static REU representation of ρ_0 and because $M(B_0, U_{s_0}) = \{q_0\}$, it follows that

$$\rho_0(r_0, \tilde{B}_0) = \rho_0(r_0, \tilde{B}_0 \cup \tilde{D}_0) = \rho_0(r_0, \tilde{D}_0) = \mu_0(s_0). \tag{5}$$

But then

$$\begin{aligned}\rho_1(\cdot, A_1|B_0, q_0) &= \rho_1(\cdot, A_1|\tilde{B}_0, r_0) = \rho_1(\cdot, A_1|\tilde{B}_0 \cup \tilde{D}_0, r_0)\\ &= \rho_1(\cdot, A_1|\tilde{D}_0, r_0) = \rho_1(\cdot, A_1|D_0, q_0^{s_0}) = \rho_1^{s_0}(\cdot, A_1),\end{aligned}$$

where the first and fourth equalities follow from Axiom 2 (Linear History Independence), the second and third equalities from Axiom 1 (Contraction History Independence) and (5), and the final equality holds by definition. Finally, using Axiom 4 (History Continuity), we can extend this argument to the case where in state s_0, q_0 is tied with other lotteries in B_0.

Step 4: Splitting histories into states. Now consider a general history $h^0 = (A_0, p_0)$. By mixing with the separating menu D_0, we can decompose $\rho_1(\cdot|h^0)$ into a weighted sum of choice probabilities conditional on each state s_0, where the weight on s_0 is the μ_0-conditional probability of s_0 given history h^0. Concretely, let $\hat{A}_0 := \frac{1}{2}A_0 + \frac{1}{2}D_0$, and for any $s_0 \in S_0$, let $p_0^{s_0} := \frac{1}{2}p_0 + \frac{1}{2}q_0^{s_0}$. This is depicted in Figure 2. Note that by construction of D_0 and the representation of ρ_0, we have that $\rho_0(p_0^{s_0}, \hat{A}_0) = \mu_0(C_0(p_0, A_0)|s_0)\mu_0(s_0)$, where $C_0(p_0, A_0)$ is the event in Ω_0 that p_0 is chosen from A_0. Moreover, whenever $\rho_0(p_0^{s_0}, \hat{A}_0) > 0$, then $\mathcal{U}_0(\hat{A}_0, p_0^{s_0}) = \{U_{s_0}\}$, so Step 3 together with the representation of $\rho_1^{s_0}$ implies $\rho_1(p_1, A_1|\hat{A}_0, p_0^{s_0}) = \rho_1^{s_0}(p_1, A_1) = \mu_1(C_1^{s_0}(p_1, A_1))$, where $C_1^{s_0}(p_1, A_1)$ is the event in $\Omega_1^{s_0}$ that p_1 is chosen from A_1. Then

$$\begin{aligned}\rho_1(p_1, A_1|A_0, p_0) &= \rho_1(p_1, A_1|\hat{A}_0, \frac{1}{2}\{p_0\} + \frac{1}{2}D_0) = \sum_{s_0 \in S_0} \rho_1(p_1, A_1|\hat{A}_0, p_0^{s_0}) \ \frac{\rho_0(p_0^{s_0}, \hat{A}_0)}{\sum_{s_0' \in S_0} \rho_0(p_0^{s_0'}, \hat{A}_0)}\\ &= \sum_{s_0 \in S_0} \mu_1^{s_0}(C_1^{s_0}(p_1, A_1)) \frac{\mu_0(C_0(p_0, A_0)|s_0)\mu_0(s_0)}{\sum_{s_0' \in S_0} \mu_0(C_0(p_0, A_0)|s_0')\mu_0(s_0')} = \sum_{s_0 \in S_0} \mu_1^{s_0}(C_1^{s_0}(p_1, A_1))\mu_0(s_0|C_0(p_0, A_0)).\end{aligned} \tag{6}$$

Indeed, the first equality follows from Linear History Independence, the second equality from the definition of ρ_1 conditional on a set of histories, the third from the observations of the preceding paragraph, and the fourth from Bayes' rule.

Step 5: Completing the proof. Now define $\Omega = \bigcup_{s_0 \in S_0} s_0 \times \Omega_1^{s_0}$. In the natural way, the partitions S_0 of Ω_0 and $S_1^{s_0}$ of $\Omega_1^{s_0}$ induce a finitely generated filtration on Ω, and the random utilities and tie-breakers on Ω_0 and $\Omega_1^{s_0}$ induce processes of utilities and tiebreakers on Ω.[26]

[26]Specifically, let $\mathcal{F}_0$ be generated by the partition $\{s_0 \times \Omega_1^{s_0} : s_0 \in S_0\}$ and $\mathcal{F}_1$ by the partition $\{s_0 \times s_1 : s_0 \in S_0, s_1 \in S_1^{s_0}\}$. For any $(\omega_0, \omega_1) \in s_0 \times \Omega_1^{s_0}$, let $U_0(\omega_0, \omega_1) = U_0(\omega_0)$, $U_1(\omega_0, \omega_1) = U_1^{s_0}(\omega_1)$ and $W_0(\omega_0, \omega_1) = W_0(\omega_0)$, $W_1(\omega_0, \omega_1) = W_1^{s_0}(\omega_1)$.

Define μ on Ω by $\mu(E_0 \times E_1) = \mu_0(E_0) \times \mu_1^{s_0}(E_1)$ for any measurable $E_0 \subseteq s_0$, $E_1 \subseteq \Omega_1^{s_0}$. From the construction of Ω and (6), it is then easy to see that $\rho_1(p_1, A_1|h^0) = \mu(C(p_1, A_1)|C(p_0, A_0))$, where $C(p_t, A_t)$ denotes the event in Ω that p_t is chosen from A_t. Thus, ρ admits a DREU representation, as required.

4 Evolving Utility vs. Gradual Learning

4.1 History-dependent revealed preference

In the following subsections, we characterize evolving utility and gradual learning. Both models impose additional restrictions on the agent's realized utilities $U_t(\omega)$. However, in our setting, the link between the agent's stochastic and history-dependent choice behavior and her underlying state-dependent utilities is less straightforward than under deterministic choice.[27] To make this link, we identify a collection of incomplete and history-dependent revealed preference relations. For each history h^t and any $q_t, r_t \in \Delta(X_t)$, $q_t \succsim_{h^t} r_t$ reveals that the agent prefers q_t to r_t in any state of the world ω that gives rise to history h^t; that is, $U_t(\omega)(q_t) \geq U_t(\omega)(r_t)$ for all $\omega \in C(h^t)$.

To see the idea, consider any history $h^0 = (A_0, p_0)$ and suppose that

$$\rho_0(\frac{1}{2}p_0 + \frac{1}{2}r_0; \frac{1}{2}A_0 + \frac{1}{2}\{q_0, r_0\}) = 0. \tag{7}$$

Then $U_0(\omega)(q_0) \geq U_0(\omega)(r_0)$ for all $\omega \in C(h^0)$. Indeed, for an expected-utility maximizer, it is optimal to choose $\frac{1}{2}p_0 + \frac{1}{2}r_0$ from menu $\frac{1}{2}A_0 + \frac{1}{2}\{q_0, r_0\}$ if and only if it is optimal to choose p_0 from A_0 *and* to choose r_0 from $\{q_0, r_0\}$.[28] Thus, if $\frac{1}{2}p_0 + \frac{1}{2}r_0$ is never chosen from $\frac{1}{2}A_0 + \frac{1}{2}\{q_0, r_0\}$, this reveals that the agent prefers q_0 to r_0 whenever she would select p_0 from A_0. Conversely, if $U_0(\omega)(q_0) \geq U_0(\omega)(r_0)$ for all $\omega \in C(h^0)$, then (7) continues to hold as long as q_0 and r_0 are perturbed appropriately to eliminate potential ties.

More generally, this suggests the following definition:

Definition 5. For each $t \leq T-1$ and $h^t = (h^{t-1}, A_t, p_t) \in \mathcal{H}_t$ relation $\succsim_{h^t}$ on $\Delta(X_t)$ is defined as follows: For any $q_t, r_t \in \Delta(X_t)$, we have $q_t \succsim_{h^t} r_t$ if there exist $q_t^n \to^m q_t$ and $r_t^n \to^m r_t$ such that

$$\rho_t(\frac{1}{2}p_t + \frac{1}{2}r_t^n; \frac{1}{2}A_t + \frac{1}{2}\{q_t^n, r_t^n\}|h^{t-1}) = 0 \text{ for all } n.$$

[27] It is also less straightforward than in the dynamic logit model characterized by Fudenberg and Strzalecki (2015) where (because of the i.i.d. nature of shocks) the deterministic component of the agent's utility function is identified, as under static logit, by choice with probability greater than a half.

[28] This observation is related to a common preference elicitation method in experimental work. To elicit a subject's ranking over a number of options in an incentive compatible manner, the subject is asked to indicate choices from multiple menus; a lottery then determines which menu (and corresponding choice) is implemented.

Let $\sim_{h^t}$ and $\succ_{h^t}$ respectively denote the symmetric and asymmetric component of $\succsim_{h^t}$.[29]

We show in Appendix C that when ρ admits a DREU representation, then $q_t \succsim_{h^t} r_t$ if and only if $U_t(\omega)(q_t) \geq U_t(\omega)(r_t)$ for all $\omega \in C(h^t)$.[30] Thus, $\succsim_{h^t}$ captures the desired notion of revealed preference.

4.2 Evolving Utility

To characterize evolving utility, the following three axioms employ $\succsim_{h^t}$ to translate conditions from the deterministic menu-choice literature to our setting. First, Separability (e.g., Fishburn 1970, Theorem 11.1) ensures that utility in every state of the world has an additively separable form $U_t(z_t, A_{t+1}) = u_t(z_t) + V_t(A_{t+1})$:

Axiom 5 (Separability)**.** For all $t \leq T-1$, h^t, z_t, x_t and A_{t+1}, B_{t+1}, we have $\frac{1}{2}(z_t, A_{t+1}) + \frac{1}{2}(x_t, B_{t+1}) \sim_{h^t} \frac{1}{2}(x_t, A_{t+1}) + \frac{1}{2}(z_t, B_{t+1})$.

The next axiom translates conditions from Dekel, Lipman, and Rustichini (2001) that ensure that $V_t(A_{t+1})$ captures the option value contained in menu A_{t+1}, i.e., that $V_t(A_{t+1}) = \mathbb{E}[\max_{p_{t+1} \in A_{t+1}} \hat{U}_{t+1}(p_{t+1}) \mid \mathcal{F}_t]$ for some random utility function $\hat{U}_{t+1}$. Part (i) is Kreps's (1979) preference for flexibility axiom; it says that the agent always weakly prefers bigger menus. Part (ii) ensures that the agent cannot affect the filtration. Part (iii) ensures that $\succsim_{h^t}$ is continuous and part (iv) that it induces a nontrivial preference over continuation menus.[31]

Axiom 6 (DLR Menu Preference)**.** For all $t \leq T-1$ and h^t, the following hold:[32]

(i). Monotonicity: For any z_t and $A_{t+1} \subseteq B_{t+1}$, we have $(z_t, B_{t+1}) \succsim_{h^t} (z_t, A_{t+1})$.

(ii). Indifference to Timing: For any z_t, A_{t+1}, B_{t+1}, and $\alpha \in (0,1)$, we have $(z_t, \alpha A_{t+1} + (1-\alpha)B_{t+1}) \sim_{h^t} \alpha(z_t, A_{t+1}) + (1-\alpha)(z_t, B_{t+1})$.

(iii). Continuity: $\succsim_{h^t}$ is continuous.[33]

(iv). Menu Nondegeneracy: There exist B_{t+1}, A_{t+1} such that $(z_t, B_{t+1}) \succ_{h^t} (z_t, A_{t+1})$ for all z_t.

[29] That is, $q_t \sim_{h^t} r_t$ if $q_t \succsim_{h^t} r_t$ and $r_t \succsim_{h^t} q_t$. And $q_t \succ_{h^t} r_t$ if $q_t \succsim_{h^t} r_t$ and $r_t \not\succsim_{h^t} q_t$.

[30] The "if" direction also makes use of Axioms 5 and 6 below. See Corollary C.1.

[31] $\succsim_{h^t}$ also satisfies a version of the DLR finiteness axiom (Axiom DLR 6 in Ahn and Sarver 2013). However, in the presence of Sophistication, this property is inherited from the finiteness axiom on ρ (Axiom 3 (v)), so we do not need to impose it as a separate condition.

[32] In the following, we identify any $(z_t, A_{t+1}) \in X_t$ with the Dirac lottery $\delta_{(z_t, A_{t+1})} \in \Delta(X_t)$.

[33] That is, for all $p_t \in \Delta(X_t)$, the upper and lower contour sets $\{q_t : q_t \succsim_{h^t} p_t\}$ and $\{q_t : p_t \succsim_{h^t} q_t\}$ are closed in $\Delta(X_t)$ endowed with the topology of weak convergence (recall that by by Lemma 12, X_t is a separable metric space). Alternatively, it is enough to require that, for any z_t and A_{t+1}, both $\{B_{t+1} : (z_t, A_{t+1}) \succsim_{h^t} (z_t, B_{t+1})\}$ and $\{B_{t+1} : (z_t, B_{t+1}) \succsim_{h^t} (z_t, A_{t+1})\}$ are closed in $\mathcal{A}_{t+1}$.

The final axiom adapts the sophistication axiom due to Ahn and Sarver (2013). We require that at any history h^t, the agent values a menu B_{t+1} strictly more than its subset A_{t+1} if and only if she in fact chooses something from $B_{t+1} \setminus A_{t+1}$ with strictly positive probability following h^t. This axiom ensures that the agent correctly anticipates her future utility distribution, that is, $\hat{U}_{t+1} = U_{t+1}$.

Axiom 7 (Sophistication)**.** For all $t \leq T-1$, $h^t \in \mathcal{H}_t$, and $A_{t+1} \subseteq B_{t+1} \in \mathcal{A}^*_{t+1}(h^t)$, the following are equivalent:

(i). $\rho_{t+1}(p_{t+1}; B_{t+1}|h^t) > 0$ for some $p_{t+1} \in B_{t+1} \setminus A_{t+1}$

(ii). $(z_t, B_{t+1}) \succ_{h^t} (z_t, A_{t+1})$ for all z_t.

Theorem 2. Suppose that ρ admits a DREU representation. The following are equivalent:

(i). ρ satisfies Axioms 5–7.

(ii). ρ admits an evolving utility representation.

Proof. See Appendix C. ■

4.3 Gradual Learning

Gradual learning is a specialization of evolving utility where the agent's consumption preference is time-invariant but unknown to her and she is learning about it over time. The additional behavioral content of this assumption is captured by restrictions on the agent's preference $\succsim_{h^t}$ over streams of consumption lotteries.

Fix $t \leq T-1$. Given a sequence $\ell_t, \dots, \ell_T \in \Delta(Z)$ of consumption lotteries, let the *stream* $(\ell_t, \dots, \ell_T) \in \Delta(X_t)$ be the period-t lottery that at every period $\tau \geq t$ yields consumption according to ℓ_τ. Formally, for any consumption lottery $\ell \in \Delta(Z)$ and menu $A_{t+1} \in \mathcal{A}_{t+1}$, define $(\ell, A_{t+1}) \in \Delta(X_t)$ to be the period-t lottery that yields current consumption according to ℓ and yields continuation menu A_{t+1} for sure; i.e., $(\ell, A_{t+1}) := \sum_{z_t \in Z} \ell(z_t)\delta_{(z_t, A_{t+1})}$. Then $(\ell_t, \dots, \ell_T) := (\ell_t, A_{t+1}) \in \Delta(X_t)$, where the sequence of menus $A_{t+1}, \dots, A_T$ is defined recursively from period T backwards by $A_T := \{\ell_T\} \in \mathcal{A}_T$ and $A_s := \{(\ell_s, A_{s+1})\} \in \mathcal{A}_s$ for all $s = t+1, \dots, T-1$. We write $(\ell_t, \dots, \ell_\tau, m, \dots, m)$ if $\ell_{\tau+1} = \ldots = \ell_T = m$ for some $m \in \Delta(Z)$ and $\tau \geq t$, and we do not specify the number of m-entries when there is no risk of confusion.

The key axiom capturing learning is the following:

Axiom 8 (Stationary Consumption Preference)**.** For all $t \leq T-1$, $\ell, m, n \in \Delta(Z)$, and h^t, $(\ell, n, \dots, n) \succ_{h^t} (m, n, \dots, n)$ if and only if $(n, \ell, n, \dots, n) \succ_{h^t} (n, m, n, \dots, n)$.

Axiom 8 implies that at any history h^t, the agent's felicity today and her expected felicity tomorrow induce the same preference over consumption lotteries. The following example illustrates the connection with learning.

Example 4. Suppose the agent faces a choice between two providers of some service, e.g., two hairdressers or dentists. Based on her current information, she believes provider ℓ to be better than m (n denotes no consumption), so will select the former if choosing between walk-in appointments today. If her desired appointment date is next week, then the agent may in general prefer to delay her decision, because she may be able to acquire more information in the meantime. However, Axiom 8 says that if she is forced to decide today (say, because advance booking is required), then she must again prefer to commit to ℓ. This is because if the agent currently believes ℓ to be better than m, then by the martingale property of beliefs she should expect her information next week to still favor ℓ on average. ▲

Given Axiom 8, the agent's felicity today and her expected felicity tomorrow can be normalized to be the same. The next axiom ensures that the agent's time preference is deterministic and time-invariant. Suppose that for some consumption lotteries ℓ and m there is a weight α that makes the agent indifferent between getting (ℓ today and m tomorrow) and the lottery $\alpha\ell + (1-\alpha)m$ in both periods. Provided the agent is not indifferent between ℓ and m today, this weight together with the fact that today's felicity equals tomorrow's expected felicity identifies the agent's discount factor. The axiom asserts that this weight, and hence the agent's discount factor, is independent of today's state and time period. We say that $\ell, m \in \Delta(Z)$ are *h^t-nonindifferent* if $(\ell, n, \ldots, n) \not\sim_{h^t} (m, n, \ldots, n)$ for some $n \in \Delta(Z)$.

Axiom 9 (Constant Intertemporal Tradeoff)**.** For all $t, \tau \leq T-1$, if ℓ, m are h^t-nonindifferent and $\hat{\ell}, \hat{m}$ are g^τ-nonindifferent, then for all $\alpha \in [0,1]$ and $n \in \Delta Z$:

$$(\ell, m, n, \ldots, n) \sim_{h^t} (\alpha\ell + (1-\alpha)m, \alpha\ell + (1-\alpha)m, n, \ldots, n)$$
$$\Longleftrightarrow$$
$$(\hat{\ell}, \hat{m}, n, \ldots, n) \sim_{g^\tau} (\alpha\hat{\ell} + (1-\alpha)\hat{m}, \alpha\hat{\ell} + (1-\alpha)\hat{m}, n, \ldots, n).$$

Axiom 9 has no bite if the agent is indifferent between all consumption lotteries at h^t. To rule this out, we impose the following condition:

Condition 1 (Consumption Nondegeneracy)**.** For all $t \leq T-1$ and h^t, there exist h^t-nonindifferent $\ell, m \in \Delta(Z)$.

Theorem 3. Suppose that ρ admits an evolving utility representation and Condition 1 is satisfied. The following are equivalent:

(i). ρ satisfies Axioms 8 and 9.

(ii). ρ admits a gradual learning representation.

The proof is in Appendix D. The argument for sufficiency proceeds in three steps. Consider an evolving utility representation $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t, u_t))$ of ρ, where we can perform appropriate normalizations to ensure that $\sum_{z\in Z} u_t(\omega)(z) = 0$ for all ω and t and that the discount factor is 1. Fix any ω and $t \leq T-1$. We first show that Axiom 8 implies that $u_t(\omega)$ and $\hat{u}_t(\omega) := \mathbb{E}[u_{t+1} \mid \mathcal{F}_t(\omega)]$ represent the same preference over consumption lotteries. Thus, there exists a (possibly state and time-dependent) $\delta_t(\omega)$ such that $\hat{u}_t(\omega) = \delta_t(\omega)u_t(\omega)$. Next, note that if $u_t(\omega)(\ell) \neq u_t(\omega)(m)$, then the unique weight α that makes the agent indifferent between (ℓ today and m tomorrow) and ($\alpha\ell + (1-\alpha)m$ both today and tomorrow) is $\frac{1}{1+\delta_t(\omega)}$. Hence, Axiom 9 together with Condition 1 implies that $\delta_t(\omega) \equiv \delta > 0$ is state and time-invariant. Finally, the above shows that the process $(\delta^{-t}u_t)$ is a martingale, so that $\delta^{-t}u_t(\omega) = \mathbb{E}[\delta^{-T}u_T \mid \mathcal{F}_t(\omega)]$. Thus, replacing U_t with $\delta^{-t}U_t$ and u_t with $\delta^{-t}u_t$ yields a gradual learning representation of ρ, where $\tilde{u} = \mathbb{E}[\delta^{-T}U_T \mid \mathcal{F}_T]$. Finally, we note that δ will be strictly less than 1 if and only if ρ additionally satisfies the following impatience axiom: for all $t \leq T-1$, h^t, and $\ell, m, n \in \Delta(Z)$, if $(\ell, n, \dots, n) \succ_{h^t} (m, n \dots, n)$, then $(\ell, m, n, \dots, n) \succ_{h^t} (m, \ell, n \dots, n)$.

A natural generalization of gradual learning is to replace the discount factor δ in (2) with a random variable $\delta : \Omega \to \mathbb{R}_{++}$ that is measurable with respect to time 0 private information $\mathcal{F}_0$. This captures the idea of a population of agents with heterogeneous discount factors, each of whom is learning over time about their fixed but unknown felicity. An analogous characterization can be obtained in this case: The only difference is that instead of imposing Axiom 9 on arbitrary histories h^t and g^τ, we require that h^t be a subhistory of g^τ.

5 Properties of the Representations

5.1 Uniqueness

The following proposition, which we prove in Supplementary Appendix H, summarizes the uniqueness properties of DREU, evolving utility, and gradual learning.

Proposition 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 $p_t \in A_t \in \mathcal{A}_t$.

If $(\mathcal{D}, (u_t), \delta)$ is an evolving utility representation of ρ, then $(\hat{\mathcal{D}}, (\hat{u}_t), \hat{\delta})$ is an evolving utility representation of ρ if and only if (i)-(iii) hold and additionally

(iv). $\alpha_t(\omega) = \alpha_0(\omega)(\frac{\hat{\delta}}{\delta})^t$ for all $\omega \in \Omega$ and $t = 0, \ldots, T$;

(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 $\gamma_t(\omega) := \beta_t(\omega) - \delta\mathbb{E}[\beta_{t+1}|\mathcal{F}_t(\omega)]$ if $t \leq T-1$.

If $(\mathcal{D}, (u_t), \delta)$ is a gradual learning representation of ρ that satisfies Condition 1, then $(\hat{\mathcal{D}}, (\hat{u}_t), \hat{\delta})$ is a gradual learning representation of ρ if and only if (i)-(v) hold and additionally

(vi). $\delta = \hat{\delta}$

(vii). $\beta_t(\omega) = \frac{1-\delta^{T-t+1}}{1-\delta}\mathbb{E}[\beta_T|\mathcal{F}_t(\omega)]$ for all ω and t.

Points (i) and (ii) of Proposition 1 show that in DREU, the agent's choices uniquely determine her underlying stochastic process of *ordinal* payoff-relevant private information, while point (iii) shows that the (ordinal) distribution of tie-breakers is pinned down for choices featuring ties. This is the period-by-period dynamic analog of known identification results for static REU representations (Proposition 4 in Ahn and Sarver (2013)). Intuitively, applying a different positive affine transformation of the vNM utility in each state does not change the optimal choice, and likewise a measure-preserving relabeling of states does not change the choice distribution.

Point (iv) shows that evolving utility implies strictly sharper identification than DREU of the agent's *cardinal* private information: In particular, the random scaling factor used to transform $\hat{\delta}^t\hat{U}_t$ into $\delta^t U_t$ is given by $\alpha_0(\omega)$, and hence is the same in all periods and measurable with respect to period-0 private information. This allows for meaningful intertemporal comparisons such as "in state ω, the additional period-t utility for p_t over q_t is greater than the additional discounted period-$(t+1)$ utility for p_{t+1} over q_{t+1}" and cross-state comparisons such as "the additional period-t utility for p_t over q_t is greater in state ω than in state $\omega' \in \mathcal{F}_0(\omega)$."[34] In the heterogeneous population interpretation, this means that behavior doesn't change upon multiplying the felicity of a particular agent-type by the same positive number in each period. These numbers can differ across types, except for types that share the same period 0 private information. Thus, the model allows for general intrapersonal intertemporal comparisons of utility, but only for limited interpersonal comparisons. This generalizes the main identification

[34] The reason the discount factor is not identified in this model is similar to the lack of identification of subjective probability when utility is state dependent: multiplying δ by $\lambda > 0$ and dividing u_t by λ leads to the same preferences and hence choice probabilities.

result (Theorem 2) in Ahn and Sarver (2013): In a two-period setting without consumption in period 0, they obtain $U_1(\omega) = \alpha \hat{U}_1(\hat{\omega}) + \beta_1(\omega)$, where α is constant since they do not allow for period-0 private information.

Finally, gradual learning, unlike evolving utility, allows for unique identification of the discount factor (point (vi)) and entails even sharper identification of cardinal private information (point (vii)).[35]

5.2 Choice and Taste Persistence

As discussed in Example 1, consumption persistence is a widely documented phenomenon, for instance in the marketing literature on brand choice and history dependence is one of its possible explanations. This section introduces two formalizations of this notion and characterizes the corresponding felicity processes u_t in the evolving utility representation.

Our first notion of consumption persistence captures the idea that (absent ties) the agent is more likely to choose consumption lottery p today if she chose p yesterday than if she chose q yesterday, provided today's menu does not include any new consumption options relative to yesterday's menu. To state this formally we focus on atemporal menus, i.e., menus that do not feature any intertemporal tradeoffs, so that the agent's choices are governed solely by her preference over current consumption. Formally, for $t \leq T-1$ menu A_t is *atemporal* if for any p_t, $q_t \in A_t$, we have $p_t^A = q_t^A$. We write $A_t^Z := \{p_t^Z : p_t \in A_t\} \subseteq \Delta(Z)$. All period T menus A_T are atemporal with $A_T^Z = A_T$.

Definition 6. ρ features *consumption persistence* if for any histories $h^t = (h^{t-1}, A_t, p_t), g^t = (h^{t-1}, A_t, q_t)$ and $p_{t+1} \in A_{t+1} \in \mathcal{A}^*_{t+1}(h^t) \cap \mathcal{A}^*_{t+1}(g^t)$ with A_t and A_{t+1} atemporal, $A_{t+1}^Z \subseteq A_t^Z$, and $p_t^Z = p_{t+1}^Z$, we have $\rho_{t+1}(p_{t+1}; A_{t+1} | h^{t-1}, A_t, p_t) \geq \rho_{t+1}(p_{t+1}; A_{t+1} | h^{t-1}, A_t, q_t)$.

To obtain a characterization of consumption persistence, we impose the assumption that there are two consumption lotteries $\overline{\ell}$ and $\underline{\ell}$ that are ranked the same way at all histories. This condition is innocuous if, for example, the outcome space includes a monetary dimension.

Condition 2 (Uniformly Ranked Pair)**.** There exist $\overline{\ell}, \underline{\ell} \in \Delta(Z)$ such that for all t, h^t, and $m \in \Delta(Z)$, we have $(\overline{\ell}, m, \ldots, m) \succ_{h^t} (\underline{\ell}, m, \ldots, m)$.

If ρ admits an evolving utility representation, consumption persistence is equivalent to a particular form of taste persistence: for any felicity u and any convex set D containing u, today's felicity is more likely to be (equivalent to) a felicity in D if yesterday's felicity was u than if yesterday's felicity was any other u'. Formally, given any set $D \subseteq \mathbb{R}^Z$ of felicities, let $[D] := \{w \in \mathbb{R}^Z : w \approx v \text{ for some } v \in D\}$.

[35] The discount factor is unique in other special cases of evolving utility, for example if each alternative z consists of wealth and a consumption bundle and the utility of wealth is separable and state-independent.

Proposition 2. Suppose ρ admits an evolving utility representation $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t, u_t))$ and Condition 2 holds. Then the following are equivalent:

(i). ρ features consumption persistence

(ii). for any $u, u' \in \mathbb{R}^Z$, convex $D \subseteq \mathbb{R}^Z$ with $u \in D$, and h^{t-1} with $\mu(\{u_t \approx u\}|C(h^{t-1}))$, $\mu(\{u_t \approx u'\}|C(h^{t-1})) > 0$, we have

$$\mu(\{u_{t+1} \in [D]\}|C(h^{t-1}) \cap \{u_t \approx u\}) \geq \mu(\{u_{t+1} \in [D]\}|C(h^{t-1}) \cap \{u_t \approx u'\}).$$

Proof. See Supplementary Appendix I. ■

An alternative notion, which is neither implied by nor implies consumption persistence, is consumption inertia. This says that if an agent chose a particular consumption lottery yesterday, then (absent ties) she will continue to choose it with positive probability today, as long as today's menu does not include any new consumption options.

Definition 7. ρ features *consumption inertia* if for any $h^t = (h^{t-1}, A_t, p_t) \in \mathcal{H}_t$ and $p_{t+1} \in A_{t+1} \in \mathcal{A}^*_{t+1}(h^t)$ with A_t and A_{t+1} atemporal, $A^Z_{t+1} \subseteq A^Z_t$, and $p^Z_t = p^Z_{t+1}$, we have

$$\rho_t(p_{t+1}; A_{t+1}|h^{t-1}, A_t, p_t) > 0.$$

In the presence of Condition 1 (Consumption Nondegeneracy) from Section 4.3, consumption inertia is equivalent to an alternative notion of taste persistence which requires that if the agent's period t consumption preference is given by u, her period $t+1$ consumption preference remains u with positive probability.

Proposition 3. Suppose that ρ admits an evolving utility representation $(\Omega, \mathcal{F}^*, \mu, (\mathcal{F}_t, U_t, W_t, u_t))$ and Condition 1 is satisfied. Then the following are equivalent:

(i). ρ features consumption inertia

(ii). for any $u \in \mathbb{R}^Z$ and h^{t-1} with $\mu(\{u_t \approx u\}|C(h^{t-1})) > 0$, we have

$$\mu(\{u_{t+1} \approx u\}|C(h^{t-1}) \cap \{u_t \approx u\}) > 0.$$

Proof. See Supplementary Appendix I. ■

The next section illustrates the different implications of the two notions by considering an evolving utility representation in which the felicity u_t follows a finite Markov chain.

5.2.1 Application: Markov Chain

Let $\mathcal{U} = \{u^1, ..., u^m\}$ denote a finite set of possible felicities, where $u^i \not\approx u^j$ for any $i \neq j$ and there exist $\overline{\ell}, \underline{\ell} \in \Delta(Z)$ such that $u^i((\overline{\ell})) > u^i(\underline{\ell})$ for all i. Let M be an irreducible transition matrix, where M_{ij} denotes the probability that period $t+1$ utility is u^j conditional on period t utility being u^i. The initial distribution ξ is assumed to have full support, but need not be the stationary distribution. Any such Markov chain $(\mathcal{U}, M, \xi)$ generates a Markov evolving utility representation.

Corollary 1. Suppose that ρ has a Markov evolving utility representation $(\mathcal{U}, M, \xi)$.

(i). Assume that for any i, j, k, l with $i \notin \{j, k, l\}$, we have $u^i \not\in [\text{co}\{u^j, u^k, u^l\}]$. Then ρ features consumption persistence if and only if the Markov chain is a renewal process, i.e., there exists $\alpha \in [0, 1)$ and $\nu \in \Delta(\mathcal{U})$ such that $M_{ii} = \alpha + (1-\alpha)\nu(u^i)$ and $M_{ij} = (1-\alpha)\nu(u^j)$ for all $i \neq j$.

(ii). ρ features consumption inertia if and only if $M_{ii} > 0$ for every i.

Proof. See Supplementary Appendix I. ■

The assumption that $u^i \not\in [\text{co}\{u^j, u^k, u^l\}]$ in point (i) is a regularity condition on the structure of $\mathcal{U}$ that is generically satisfied if the outcome space is rich enough relative to the number of utility functions. The following example does not satisfy the condition. It features consumption persistence under a non-renewal process.

Example 5 (Random walk over a line)**.** Consider an evolving utility representation where the felicity process is of the form $u_t = w + \alpha(x_t)v$, where $w, v \in \mathbb{R}^Z$ are fixed felicities, $\alpha : \mathbb{Z} \to \mathbb{R}$ is a strictly increasing function, and x_t follows a random walk over $\mathbb{Z}$: it remains at its current value with probability p, increases by one with probability $\frac{1-p}{2}$, and decreases by one with probability $\frac{1-p}{2}$. This agent displays consumption persistence if and only if $p \geq \frac{1}{3}$ and consumption inertia if and only if $p > 0$. ▲

6 Comparison with Dynamic Discrete Choice

In this section, we relate our analysis to the dynamic discrete choice (DDC) literature. Section 6.1 demonstrates that the i.i.d. DDC model that is very widely used in this literature is incompatible with evolving utility, because it violates a key feature of Bayesian rationality, positive option value. Section 6.2 shows that by utilizing menu variation, our analysis yields identification results that are complementary to those in the DDC literature.

6.1 Evolving Utility vs. i.i.d. DDC

Parametric specifications of our evolving utility model are featured in the DDC literature (e.g. Pakes, 1986). By far more widely used, however, is the i.i.d. DDC model, which is the following alternative special case of DREU, see e.g., Miller (1984), Rust (1989), Hendel and Nevo (2006), Kennan and Walker (2011), Sweeting (2013), and Gowrisankaran and Rysman (2012). The DDC literature typically defines choices only on deterministic decision trees; in what follows we study this restriction of our domain. Let $Y_T := Z$ and $\mathcal{A}^d_T := \mathcal{K}(Y_T)$ and recursively for $t \leq T-1$ define $Y_t := Z \times \mathcal{A}^d_{t+1}$ and $\mathcal{A}^d_t := \mathcal{K}(Y_t)$.[36]

Definition 8. The *i.i.d. DDC model* is a restriction of DREU to deterministic decision trees that additionally satisfies the Bellman equation

$$U_t(z_t, A_{t+1}) = v_t(z_t) + \delta \mathbb{E}\left[\max_{y_{t+1} \in A_{t+1}} U_{t+1}(y_{t+1})\right] + \epsilon_t^{(z_t, A_{t+1})}, \tag{8}$$

where the functions $v_t : Z \to \mathbb{R}$ are deterministic; the discount factor $\delta \in (0,1)$; and $\epsilon_t : \Omega \to \mathbb{R}^{Y_t}$ are vectors of zero-mean payoff shocks such that $\epsilon_t^{(z_t, A_{t+1})}$ and $\epsilon_\tau^{(x_\tau, B_{\tau+1})}$ are independently and identically distributed nondegenerate random variables for all (z_t, A_{t+1}) and $(x_\tau, B_{\tau+1})$.

In the i.i.d. DDC model, utilities $U_t(z_t, A_{t+1})$ feature a deterministic component that is additively separable into instantaneous felicity $v_t(z_t)$ and discounted expected continuation value $\delta \mathbb{E}\left[\max_{y_{t+1} \in A_{t+1}} U_{t+1}(y_{t+1})\right]$; to generate randomness, the model adds on a shock $\epsilon_t^{(z_t, A_{t+1})}$ whose distribution is i.i.d. across time and across options (z_t, A_{t+1}). Clearly, the i.i.d. DDC model is not equivalent to evolving utility, because unlike the latter it does not allow for serially correlated utilities and hence does not give rise to history dependent choice behavior. More strongly, however, we now show that the models are in fact incompatible: While evolving utility entails a positive option value, the i.i.d. DDC model has the opposite implication.

A first manifestation of this difference is that the i.i.d. DDC agent sometimes chooses to commit to strictly smaller menus. Let $A_t := \{(z_t, A_{t+1}), (z_t, B_{t+1})\}$ where $A_{t+1} \subsetneq B_{t+1}$. From Axiom 6 it follows that in the evolving utility model, absent ties, $\rho_t\left((z_t, A_{t+1}), A_t | h_t\right) = 0$. By contrast, in the i.i.d. DDC model, (8) implies that as long as the distribution of ϵ-shocks has large enough support,[37] the agent will choose (z_t, A_{t+1}) from A_t with strictly positive probability; in particular, this happens whenever the realization of $\epsilon_t^{(z_t, A_{t+1})}$ exceeds

[36] Alternatively, we could study an extension of the DDC model to lotteries. One natural candidate is a linear extension, under which the DDC model is a special case of DREU. Other extensions to lotteries are possible, but they are less satisfactory, as they violate Axiom 0 and lead to counterintuitive comparative statics as pointed out in the static setting by Apesteguia and Ballester (2017). Our results in this section are independent of the extension since they hold on the subdomain $\mathcal{A}^d_t$.

[37] Indeed, the DDC literature typically assumes this distribution to have full support. On deterministic decision trees this is observationally equivalent to a finitely generated distribution with large enough support.

$\epsilon_t^{(z_t,B_{t+1})}$ by more than the expected utility difference of the two menus. Nevertheless, since $\mathbb{E}[U_{t+1}(z_t, B_{t+1})] \geq \mathbb{E}[U_{t+1}(z_t, A_{t+1})]$, the probability that the i.i.d. DDC agent chooses (z_t, A_{t+1}) is less than 0.5.

More strikingly, there are decision problems for which the i.i.d. DDC agent's behavior displays a negative option value with probability greater than 0.5. Specifically, consider the following problem about the timing of decisions that is illustrated in Figure 3. There are

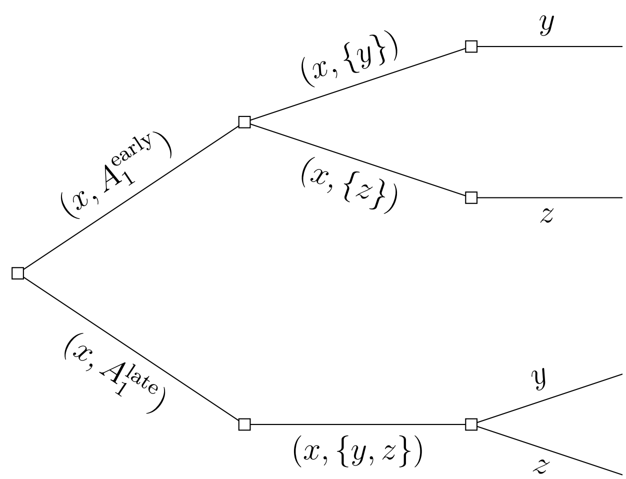

Figure 3: Decision Timing.

three periods $t = 0, 1, 2$. The consumption in period 2 is either y or z, depending on the decision of the agent. The agent can make her decision early, committing in period 1 to receiving y or z the following period; or she can make the decision late, maintaining full flexibility about choosing y or z until the final period. The decision when to choose is made in period 0, and the consumption in periods 0 and 1 is x irrespective of the decision of the agent; for simplicity assume that the utility of x is always zero.[38] Formally, in period 0 the agent faces the menu $A_0 = \{(x, A_1^{\text{early}}), (x, A_1^{\text{late}})\}$ and in period 1 she faces either the menu $A_1^{\text{early}} = \{(x, \{y\}), (x, \{z\})\}$ or the menu $A_1^{\text{late}} = \{(x, \{y, z\})\}$, depending on her period-0 choice.

Note that the agent's decision when to choose does not change the time at which consumptions y or z are received, nor does it affect her consumption in periods 0 or 1, so that there is no penalty to making the decision late. In accordance with positive option value, in the evolving utility model the agent thus chooses to make decisions late with probability 1 (absent ties), because waiting an extra period gives her more information, which enables her to better tailor

[38]Under i.i.d. DDC this means that $v_t(x) = 0$ for all t; under evolving utility that $\mu(\{u_t(x) = 0\}) = 1$. Clearly Proposition 4 does not rely on this normalization.

her choice to the state.[39] To see this, note that

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

By the conditional Jensen inequality and convexity of the max operator, the agent always weakly prefers to decide late. Moreover, this preference 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 $\{\omega', \omega''\}$.

This preference for late decisions does not hold in the i.i.d. DDC model, where we have

$$U_0(x, A_1^{\text{early}}) = \epsilon_0^{(x,A_1^{\text{early}})} + \mathbb{E}[\max\{\delta^2 v_2(y) + \delta\varepsilon_1^{(x,\{y\})}, \delta^2 v_2(z) + \delta\varepsilon_1^{(x,\{z\})}\}]$$
$$U_0(x, A_1^{\text{late}}) = \epsilon_0^{(x,A_1^{\text{late}})} + \mathbb{E}[\max\{\delta^2 v_2(y) + \delta^2\varepsilon_2^y, \delta^2 v_2(z) + \delta^2\varepsilon_2^z\}].$$

The simplest case to analyze is when $v_2(y) = v_2(z)$: In this case, the comparison of the continuation values boils down to the comparison between $\delta\mathbb{E}[\max\{\varepsilon_1^{(x,\{y\})}, \varepsilon_1^{(x,\{z\})}\}]$ and $\delta^2\mathbb{E}[\max\{\varepsilon_2^y, \varepsilon_2^z\}]$. Since the ϵ shocks are i.i.d. and mean zero and $\delta \in (0, 1)$, the former dominates the latter, so that the agent chooses to decide early with probability greater than 0.5. The option of choosing early is attractive because it allows the agent to obtain a positive payoff, namely the maximum of two i.i.d. mean zero shocks, early while deferring the choice delays those payoffs. Proposition 4, which we prove in Supplementary Appendix J, shows that this conclusion holds for any values of $v_2(y)$ and $v_2(z)$.

Proposition 4. In the i.i.d. DDC model $\rho_0((x, A_1^{\text{early}}), A_0) \geq 0.5$.

A special case of this result for logit ϵ shocks was proved by Fudenberg and Strzalecki (2015), by examining the closed-form expressions for continuation values in this setting.[40] Proposition 4 shows that this result does not rely on those specific expressions. Rather, it is a consequence of the mechanical nature of ϵ shocks in *any* i.i.d. DDC model, in particular the fact that shocks to continuation menus are completely detached from their expected continuation value.[41]

Under evolving utility the amount of randomness in choices from menus $\{(z_t, A_{t+1}), (z_t, B_{t+1})\}$ is determined by the randomness in continuation values

[39] A related finding is Theorem 2 of Krishna and Sadowski (2016), who show that one agent prefers to commit to a constant consumption plan more than another agent if and only if her utility process is more autocorrelated, in other words, when she expects to learn less in the future.

[40] Fudenberg and Strzalecki (2015) also introduced a choice-aversion parameter that scales the desire for flexibility and for early decisions. However, in this model the parameter values that imply choice of late decisions with probability higher than 0.5 also imply choice of smaller menus with probability higher than 0.5, thus making violations of positive option value particularly stark in the latter dimension.

[41] Our critique of the mechanical nature of ε shocks is complementary to Apesteguia and Ballester's (2017) critique in the static setting, but the logic of our result is quite different, both formally and conceptually. In particular, in Proposition 4 these mechanical shocks lead to counterintuitive predictions at an *absolute* level, rather than at a comparative level as in their results.

$\mathbb{E}[\max_{y_{t+1}\in A_{t+1}} U_{t+1}(y_{t+1}) \mid \mathcal{F}_t]$; in other words, by how much of the uncertainty about payoffs in $t+1$ is resolved in period t: the more the agent learns in period t, the more random her choices. Thus, the randomness of choices in this model is a reflection of the agent's learning about fundamentals. By contrast, under i.i.d. DDC, the shocks $\epsilon_t^{(z_t,A_{t+1})}$ are i.i.d. across continuation menus, so that randomness here is unrelated to the agent's expectation of fundamentals as captured by the continuation values $\mathbb{E}\left[\max_{y_{t+1}\in A_{t+1}} U_{t+1}(y_{t+1})\right]$. To further see this, consider the modification of equation (8), where the shocks $\epsilon_t^{(z_t,A_{t+1})}$ are i.i.d. over time and across instantaneous consumptions z_t, but satisfy $\epsilon_t^{(z_t,A_{t+1})} = \epsilon_t^{(z_t,B_{t+1})} =: \epsilon_t^{z_t}$ for all continuation menus A_{t+1} and B_{t+1}. This model *is* a special case of evolving utility where the utilities $u_t(z_t) := v_t(z_t) + \epsilon_t^{z_t}$ are independent over time; reflecting the fact that no uncertainty about payoffs in $t+1$ is resolved in period t, choices from the menu $\{(z_t, A_{t+1}), (z_t, B_{t+1})\}$ are deterministic in this case (absent ties).

An advantage that sets the i.i.d. DDC model apart from evolving utility is its statistical convenience: Since shocks do not depend on continuation values their computation does not involve recursive calculations; moreover, by specifying shock distributions with sufficiently large support it is easy to ensure non-degenerate likelihoods (i.e., that all options are chosen with positive probability), whereas under evolving utility some options are necessarily chosen with probability 0. However, as we have highlighted above, this convenience comes at a cost, namely the violation of a key feature of Bayesian rationality, positive option value. Such misspecifications of option value seem particularly problematic in applications where the modeled agents are profit-maximizing firms, and they may potentially lead to biased parameter estimates in these settings.[42] Conceptually, this casts doubt on the typical interpretation of ϵ as "unobserved utility shocks." Another interpretation of ϵ in the DDC literature is that they capture "mistakes" or some small deviations from perfect rationality. However, Proposition 4 shows that the implied deviations are not small as they occur with probability greater than a half; moreover, this interpretation is at odds with the fact that in (8) the ϵ shocks enter into the agent's expected continuation value.

6.2 Identification

There is an extensive literature on identification of DDC models. That literature typically features a state variable with two components: the first one is jointly observed by both the agent and the analyst and the second one is private to the agent. Identification results mostly keep the menu fixed in every period and exploit variation in the jointly observable state variables.[43]

[42] The quantitative importance of such biases is an empirical question, which is beyond the scope of this paper.

[43] For example, exclusion restrictions on terminal states and renewal states are often used (Magnac and Thesmar, 2002); another related set of results is due to Norets and Tang (2013) who use exogenous variation in the transition probabilities of observed variables.

By contrast, our analysis has abstracted away from jointly observed state variables, but our uniqueness results in Section 5.1 rely on the assumption that the analyst observes choices from different menus.[44] Identification results in static models sometimes do utilize menu variation (e.g., Berry, Levinsohn, and Pakes 1995); however, we are not aware of similar results in the dynamic setting. As a result, the two sets of results are mostly complementary. In the rest of this section we highlight the most relevant points of contact, but an exhaustive comparison is beyond the scope of this paper.

Manski (1993) and Rust (1994) show that in a DDC model it is not possible to distinguish a myopic agent ($\delta = 0$) from a patient agent ($\delta > 0$). On the other hand, in the evolving utility model these two cases can be distinguished based on menu variation. More generally, the reasons why δ is not identified are different in the two models: δ is not identified in the DDC model even under the assumption that the function v is time invariant because each action is uniquely associated with a probability distribution over continuation menus and the continuation value can be absorbed into v without changing behavior. On the other hand, the reason δ is not pinned down under evolving utility has to do with time-dependence of the function u_t; see footnote 34. As Magnac and Thesmar (2002) show, the discount factor can be identified in DDC models using additional assumptions on how the utility function depends on the jointly observable variable. As we mentioned in Section 5.1, similar restrictions lead to the identification of δ under evolving utility (see footnote 35). Moreover, another special case when the discount factor can be identified is the gradual learning model.

Many of the results about the identification of the utility function v assume a known conditionally independent distribution of ϵ, see e.g., Magnac and Thesmar (2002).[45] Although the per-period utilities are non-parametrically not identified, certain differences in value functions are. Our approach partially identifies the distribution of u (up to positive affine transformations), which corresponds to jointly identifying v and the distribution of ϵ. This is similar to the partial identification results of Norets and Tang (2013) who also relax the known distribution of ϵ assumption; however, they maintain the conditional independence assumption, whereas our uniqueness result holds also under any possible pattern of serial correlation.[46]

[44] More precisely, in the DDC literature the analyst may observe choices from different menus, but menus are determined by the jointly observable variable which is also an argument of the utility function, preventing a clean separation.

[45] Rust (1994) discusses identification of v in a deterministic choice model without unobservable shocks.

[46] Kasahara and Shimotsu (2009) obtain identification results for finite mixtures of conditionally independent DDC models and Hu and Shum (2012) generalize these result to other forms of serial correlation. However, those papers only identify the mixing probabilities and the choice probabilities conditional on unobserved heterogeneity; they are not after the identification of structural parameters.

7 Extension: Consumption Dependence

Our analysis thus far has focused on isolating a form of history dependence where choices in period t depend on histories $h^t = (A_0, p_0, \ldots, A_{t-1}, p_{t-1})$ of past choices purely due to the fact that h^t partially reveals the agent's serially correlated private information. For this analysis the sequence $(z_0, \ldots, z_t)$ of agent's past consumptions was immaterial, because the fact that $z_k \in \operatorname{supp} p_k^Z$ was realized does not reveal any additional private information to the analyst. In many settings, however, an additional channel through which the agent's past choices can *directly* affect her current choices is that her past *consumption* may change her current preferences. Two prominent examples of this phenomenon, which we call *consumption dependence*, are habit formation (e.g., Becker and Murphy, 1988), where consuming a certain good in the past may make the agent like it more in the present; and active learning/experimentation, where the agent's consumption provides information to her about some payoff-relevant state of the world, as modeled for instance by the multi-armed bandit literature (e.g., Gittins and Jones, 1972; Robbins, 1952).

The present section, in conjunction with Supplementary Appendix K which contains all details, shows that our main insights extend to settings with consumption dependence. To this end, we enrich our primitive: A history $\mathbb{h}^{t-1} = (A_0, p_0, z_0, \ldots, A_{t-1}, p_{t-1}, z_{t-1})$ now summarizes not only that at each period $k \leq t-1$ the agent faced menu A_k and chose p_k, but also that the agent's realized consumption was $z_k \in \operatorname{supp} p_k^Z$. Conditional on $\mathbb{h}^{t-1}$, the analyst observes the frequency $\rho_t(p_t, A_t | \mathbb{h}^{t-1})$ with which the agent chooses p_t from any menu A_t such that $(z_{t-1}, A_t) \in \operatorname{supp} p_{t-1}$.

Theorem 5 in Appendix K shows that natural adaptations of Axioms 1–4 to this setting are equivalent to ρ admitting a *consumption-dependent DREU (CDREU)* representation: At each time t, the agent's choices maximize her vNM utility $U_{s_t} \in \mathbb{R}^{X_t}$, which is determined by a subjective state s_t drawn from a finite state space S_t. There is an initial distribution $\mu_0 \in \Delta(S_0)$, and at each t, today's state s_t and consumption z_t jointly determine the distribution $\mu_{t+1}^{s_t, z_t} \in \Delta(S_{t+1})$ over tomorrow's states. The full formal specification of the representation, including its tie-breaking rule, is in Appendix K. The adaptations of Axioms 1 and 2 still impose history independence of ρ_t across histories $\mathbb{h}^{t-1}$ and $\mathbb{g}^{t-1}$ that are contraction equivalent or linearly equivalent. The sole difference is that in order for such $\mathbb{h}^{t-1}$ and $\mathbb{g}^{t-1}$ to reveal the same private information, we now additionally require that they entail the same sequences $(z_0, \ldots, z_{t-1})$ of realized consumptions; otherwise $\mathbb{h}^{t-1}$ and $\mathbb{g}^{t-1}$ may correspond to different distributions of subjective states due to the consumption dependence of the transition distributions $\mu_{k+1}^{s_k, z_k}$.

As before, we also characterize two important special cases of CDREU that feature dynamic sophistication. *Consumption-dependent evolving utility* requires U_{s_t} to be given by a Bellman equation of the form

$$U_{s_t}(z_t, A_{t+1}) = u_{s_t}(z_t) + \delta \mathbb{E}\left[\max_{p_{t+1}\in A_{t+1}} U_{s_{t+1}}(p_{t+1}) \mid s_t, z_t\right], \tag{9}$$

where the expectation operator is with respect to the distribution $\mu_{t+1}^{s_t,z_t}$ of states that will arise after consuming z_t. Theorem 6 shows that consumption-dependent evolving utility is obtained from CDREU by additionally imposing analogs of the DLR Menu Preference and Sophistication axioms used to characterize evolving utility. Unlike with evolving utility, (9) does not imply any analog of Separability, as current consumption z_t can influence preferences over continuation menus A_{t+1} through its effect on the distribution over tomorrow's utilities $U_{s_{t+1}}$. Consumption-dependent evolving utility can accommodate a number of behavioral phenomena where past consumption directly affects current utility. Example 6 below illustrates this for the case of habit formation;[47] related phenomena include preference for variety (e.g., McAlister, 1982; Rustichini and Siconolfi, 2014), memorable consumption (Gilboa, Postlewaite, and Samuelson, 2016), and endogenous discounting (e.g., Becker and Mulligan, 1997; Uzawa, 1968).[48] In contrast with most existing models of these phenomena, our formulation allows the effect of past consumption on today's utility to be stochastic—arguably a realistic feature in many contexts.

Example 6 (Habit formation)**.** Suppose $\mathcal{V} \subseteq \mathbb{R}^Z$ is a finite set of felicities. There is an initial distribution $\pi_0 \in \Delta(\mathcal{V})$ and a map $\pi : \mathcal{V} \times Z \to \Delta(\mathcal{V})$, capturing the stochastic transition from today's felicity and consumption to tomorrow's felicity. At each period t and current felicity $v_t \in \mathcal{V}$, the agent maximizes

$$U_t^{v_t}(z_t, A_{t+1}) = v_t(z_t) + \delta \int \max_{p_{t+1}} U_{t+1}^{v_{t+1}}(p_{t+1})\, d\pi(v_t, z_t)(v_{t+1})$$

for $t \leq T-1$ and $U_T^{v_T}(z_T) = v_T(z_T)$. To illustrate how this representation can capture a stochastic form of habit formation, suppose for simplicity that $Z = \{0,1\}$ and $\mathcal{V} = \{v^0, v^1\}$, where $v^1(1) = v^0(0) = 1$ and $v^1(0) = v^0(1) = 0$, and let $\pi_{ijk} := \pi(v^i, j)(v^k)$ for $i,j,k \in \{0,1\}$. Then the agent displays habit formation if $\pi_{111} > \pi_{101}, \pi_{011} > \pi_{001}$, so that she is more likely to prefer 1 to 0 today if she preferred 1 yesterday and/or consumed 1 yesterday. ▲

Another important case of consumption-dependent evolving utility is *active learning*, where

$$u_{s_t} = \mathbb{E}\left[u_{s_{t+1}} | s_t, z_t\right] \forall z_t. \tag{10}$$

[47]Several papers have characterized versions of habit formation focusing on deterministic choice, e.g., Gul and Pesendorfer (2007), Rozen (2010). To the best of our knowledge, the non-axiomatic work by Gilboa and Pazgal (2001) is the only stochastic choice model.

[48]Higashi, Hyogo, and Takeoka (2014) characterize a stochastic version of endogenous discounting, but have as their primitive deterministic ex-ante preferences over infinite-horizon decision problems.

Capturing the fact that felicity is fixed but unknown to the agent who learns about it over time, (10) requires that her expectation (prior to consuming any particular z_t) of tomorrow's felicity equal today's felicity (irrespective of the particular z_t). Unlike with gradual learning, today's consumption z_t can have an effect on tomorrow's felicity by affecting tomorrow's information; but unlike with general consumption-dependent evolving utility, this effect is purely informational, and hence z_t does not affect the agent's *expectation* of tomorrow's felicity. As with gradual learning, this additional discipline allows unique identification of the discount factor. Theorem 7 shows that active learning is obtained from consumption-dependent evolving utility by additionally imposing analogs of Axioms 8 and 9 used to characterize gradual learning, as well as a weak form of Separability. The latter requires that if from tomorrow on the agent is committed to a particular stream of consumption lotteries $(\ell_{t+1}, \dots, \ell_T)$, then her preference over today's consumption z_t does not depend on $(\ell_{t+1}, \dots, \ell_T)$. This captures the idea that consumption lottery streams, unlike general continuation menus A_{t+1}, only entail degenerate future choices; hence, today's consumption does not have any informational value in this case, so that the agent evaluates today's consumption myopically, based solely on today's felicity u_{s_t}.

Example 7 below shows that active learning nests the standard independent multi-armed bandit model, as well as models that allow for correlation of arms (e.g., Easley and Kiefer 1988, Aghion, Bolton, Harris, and Jullien 1991).[49] (The active learning model is more general than the one from Example 7 as it allows for very general signal structures, not necessarily i.i.d. conditional on the state.)

Example 7 (Experimentation)**.** Now period-t states correspond to period-t beliefs about a state of the world $\theta \in \Theta$, which captures uncertainty about the true underlying felicity $\tilde{u}_\theta \in \mathbb{R}^Z$. There is a prior $\pi_0 \in \Delta(\Theta)$ and consuming any particular z produces a signal about Θ whose distribution is i.i.d. over time conditional on z and θ. (This specification is general enough to allow the agent to learn about the utility of item x from consuming item z; imposing a further product structure on Θ leads to the usual independent bandit case.) The signal structure induces a map $\pi : \Delta(\Theta) \times Z \to \Delta(\Delta(\Theta))$ from prior beliefs $\nu \in \Delta(\Theta)$ and consumptions z to distributions $\pi(\nu, z) \in \Delta(\Delta(\Theta))$ over posterior beliefs. At each period t and belief ν_t, the agent maximizes

$$U_t^{\nu_t}(z_t, A_{t+1}) = \int \tilde{u}_\theta(z_t)\, d\nu_t(\theta) + \delta \int \max_{p_{t+1} \in A_{t+1}} U_{t+1}^{\nu_{t+1}}(p_{t+1})\, d\pi(\nu_t, z_t)(\nu_{t+1})$$

for $t \leq T-1$ and $U_T^{\nu_t}(z_T) = \int \tilde{u}_\theta(z_t)\, d\nu_T(\theta)$. Note that by the martingale property of beliefs we have $\nu = \int \nu'\, d\pi(\nu, z)(\nu')$ for all ν and z, so that $u_t^{\nu_t} := \int \tilde{u}_\theta\, d\nu_t(\theta) = \int\int \tilde{u}_\theta\, d\nu_{t+1}(\theta)\, d\pi(\nu_t, z_t)(\nu_{t+1}) =: \mathbb{E}[u_{t+1}^{\nu_{t+1}} | \nu_t, z_t]$ for all ν_t, z_t, as required by (10). ▲

[49] Hyogo (2007), Cooke (2016) and Piermont, Takeoka, and Teper (2016) axiomatize related models, taking as their primitive deterministic ex ante preferences over decision problems.

Finally, Heckman (1981) highlights the importance of distinguishing between (what we term) history dependence and consumption dependence, so as to avoid spuriously attributing a causal role to past consumption when observed behavior could instead be explained through serially correlated private information, such as persistent taste heterogeneity. The following condition allows us to make this distinction:

Axiom 10 (Consumption Independence)**.** For all $t \leq T$, if $\mathbb{h}^{t-1} = (A_0, p_0, z_0, \ldots, A_{t-1}, p_{t-1}, z_{t-1})$ and $\mathrm{g}^{t-1} = (A_0, p_0, z'_0, \ldots, A_{t-1}, p_{t-1}, z'_{t-1})$, then $\rho_t(\cdot|\mathbb{h}^{t-1}) = \rho_t(\cdot|\mathrm{g}^{t-1})$.

Consumption independence states that sequences $(z_0, \ldots, z_{t-1})$ of realized consumptions are in fact immaterial for observed choice behavior. When ρ admits a CDREU representation, it is easy to see that it admits a DREU representation if and only if consumption independence is satisfied.[50] Thus, this condition captures precisely the observed choice behavior that can be explained through serially correlated private information alone.

8 Discussion

8.1 Related Literature

An extensive literature studies axiomatic characterizations of random utility models in the *static* setting (Barberá and Pattanaik, 1986; Block and Marschak, 1960; Falmagne, 1978; Luce, 1959; McFadden and Richter, 1990). Our approach incorporates as its static building block the elegant axiomatization of Gul and Pesendorfer (2006) and Ahn and Sarver (2013). A technical contribution of our paper is the extension of their result to an infinite outcome space, which is needed since the space of continuation problems in the dynamic model is infinite. Lu (2016) studies a model with an objective state space where choice is between Anscombe-Aumann acts; by imposing an assumption that utility is state-independent, he traces all randomness of choice to random arrival of signals. This is similar in spirit to our gradual learning representation; however, we do not rely on these strong independence assumptions: our state space is subjective and utility can be state-dependent. A recent paper by Lu and Saito (2016) studies period-0 random choice between consumption lottery streams and attributes the source of randomness of choice to the stochastic discount factor.[51]

The axiomatic literature on *dynamic* random utility, and more generally dynamic stochastic choice, is very sparse. Our choice domain is as in Kreps and Porteus (1978); however, while

[50] Reducing consumption-dependent evolving utility (respectively, active learning) to evolving utility (respectively, gradual learning) additionally requires Separability.

[51] Another recent contribution by Apesteguia, Ballester, and Lu (2017) considers a setting in which choice options are linearly ordered.

they study deterministic choice in each period, we focus on random choice in each period. To the best of our knowledge, Fudenberg and Strzalecki (2015) is the first axiomatic study of stochastic choice in general decision trees, but they focus on the special parametric case of logit utility shocks that are i.i.d. over time, while we characterize a fully non-parametric dynamic random utility model and allow for serially correlated utilities. Because of their i.i.d. assumption, their representation does not give rise to history dependent choice behavior and cannot accommodate phenomena such as learning, choice persistence, and consumption dependence; likewise, challenges such as limited observability do not arise in their setting. In addition, their model features very different attitudes toward option value than our evolving utility model, as we discuss in Section 6.[52] A recent paper by Ke (2017) characterizes a dynamic version of the Luce model, where randomness of choices is caused by mistakes and there is no serially correlated private information. In contrast to the evolving utility model, his model again does not feature positive option value, as larger menus might induce more mistakes.

The literature on menu choice (Dekel, Lipman, and Rustichini, 2001; Dekel, Lipman, Rustichini, and Sarver, 2007; Dillenberger, Lleras, Sadowski, and Takeoka, 2014; Kreps, 1979) considers an agent's deterministic preference over menus (or decision trees) at a hypothetical ex-ante stage in which the agent does not receive any information but anticipates receiving information later. An important difference of our approach is that we study the agent's behavior in actual decision trees, allowing information to arrive in each period and therefore focusing on stochastic choice. We discuss the comparison in more detail in Section 2.2.3. Related papers are Krishna and Sadowski (2014) and Krishna and Sadowski (2016) that study ex-ante preferences over infinite-horizon decision trees and characterize *stationary* versions of our evolving utility representation (making them unsuited to study gradual learning). Another related paper by Ahn and Sarver (2013) studies both ex-ante preference over menus and ex-post stochastic choice from menus; they show how to connect the analysis of Gul and Pesendorfer (2006) and of Dekel, Lipman, and Rustichini (2001) to obtain better identification properties. Their sophistication axiom plays a key role in our characterization of evolving utility.

Finally, as surveyed in Section 6, an extensive empirical literature uses specifications of the i.i.d. DDC model. As mentioned, the specific parametric case of i.i.d. logit utility shocks was axiomatized by Fudenberg and Strzalecki (2015). Our DREU representation nests the i.i.d. DDC model, while we show that i.i.d. DDC is incompatible with evolving utility because it entails a negative option value. We also relate our uniqueness results to the identification results in that literature.

[52] On more limited domains, Gul, Natenzon, and Pesendorfer (2014) study an agent who receives an outcome only once at the end of a decision tree, and characterize a generalization of the Luce model. Pennesi (2017) characterizes a version of the logit model where the analyst observes a history-independent sequence of stochastic choice data over consumption streams. There is also non-axiomatic work studying special cases of our representation where the agent makes a one-time consumption choice at a stopping time, e.g., Fudenberg, Strack, and Strzalecki (2016).

8.2 Conclusion

This paper provides the first analysis of the fully general and non-parametric model of dynamic random utility. When utilities are serially correlated, a key new feature relative to the static and i.i.d. model is that choices appear history dependent, a pervasive phenomenon in economic applications. We axiomatically characterize the implied dynamic stochastic choice behavior, imposing discipline on the form of history dependence that can arise under arbitrary serial correlation.

Stochastic choice data in dynamic domains lets us distinguish important models that coincide in the static setting. In particular, choices that arise from learning rather than from more general taste shocks display a form of stationary consumption preference, capturing the martingale property of beliefs. Moreover, by distinguishing between past choices and realized consumption, we can separate history dependence due to serially correlated utilities from models of habit formation and experimentation, where past consumption directly affects the agent's utility process, and characterize when phenomena such as consumption persistence can be explained through the former channel alone.

Our analysis has implications for the dynamic discrete choice literature. By utilizing menu variation, we provide identification results that are complementary to those in the DDC literature. Moreover, unlike the widely used i.i.d. DDC model, our evolving utility representation specifies payoff shocks in a manner that implies a positive option value. This may motivate developing tractable parametric specifications of evolving utility for use in applications.

We also provide several methodological contributions that we believe may prove useful for future work on stochastic choice: A solution to the limited observability problem that arises from the fact that in dynamic settings past choices typically restrict future opportunity sets; an extension to infinite outcome spaces of Gul and Pesendorfer's (2006) and Ahn and Sarver's (2013) characterization of static random expected utility; and a way to infer, and impose additional structure on, the agent's preference conditional on any particular realization of her private information. The latter revealed preference approach could potentially be used to study other interesting special cases of DREU, such as dynamic models of random temptation or mistakes. Finally, some techniques developed in this paper naturally carry over from the multi-period to the multi-agent setting, and in ongoing work we exploit this to study strategic interactions under correlated private information.

Appendix: Main Proofs

The appendix is structured as follows:

- Section A defines equivalent versions of DREU, evolving utility, and gradual learning, as well as other important terminology that is used throughout the appendix.
- Sections B–D prove Theorems 1–3.
- Section E collects together several lemmas that are used throughout Sections B–D.

The supplementary appendix contains the following additional material:

- Section F proves Theorem 0 (the static REU representation result for arbitrary separable metric spaces of outcomes), which is used in the proof of Theorem 1.
- Section G proves Proposition 5 from Section A.
- Sections H, I, and J collect together proofs for Sections 5.1, 5.2, and 6, respectively.
- Finally, Section K contains formal definitions and representation theorems for the consumption dependent representations from Section 7.

A joint file, including both the main text and supplementary appendix, is available at: https://drive.google.com/file/d/0B-372Fn5SRUAM3BYVmNrR2diR1E/view

A Equivalent Representations

Instead of working with probabilities over the grand state space Ω, our proofs of Theorems 1–3 will employ equivalent versions of our representations, called S-based representations, that look at one-step-ahead conditionals.[53] Section A.1 defines S-based representations. Section A.2 establishes the equivalence between DREU, evolving utility, and gradual learning representations and their respective S-based analogs. Section A.3 introduces important terminology regarding the relationship between states and histories that will be used throughout the proofs of Theorems 1–3.

A.1 S-based Representations

For any $X \in \{X_0, \ldots, X_T\}$, $A \in K(\Delta(X))$, $p \in \Delta(X)$, let $N(A,p) := \{U \in \mathbb{R}^X : p \in M(A,U)\}$ and $N^+(A,p) := \{U \in \mathbb{R}^X : \{p\} = M(A,U)\}$.

Definition 9. A *random expected utility (REU) form* on $X \in \{X_0, \ldots, X_T\}$ is a tuple $(S, \mu, \{U_s, \tau_s\}_{s \in S})$ where

(i). S is a finite state space and μ is a probability measure on S

(ii). for each $s \in S$, $U_s \in \mathbb{R}^X$ is a nonconstant utility over X.

(iii). for each $s \in S$, the tie-breaking rule τ_s is a finitely-additive probability measure on the Borel σ-algebra on $\mathbb{R}^X$ and is *proper*, i.e., $\tau_s(N^+(A,p)) = \tau_s(N(A,p))$ for all A, p.

[53] These are dynamic analogs of the static GP representations in Ahn and Sarver (2013).

Given any REU form $(S, \mu, \{U_s, \tau_s\}_{s\in S})$ on X_i and any $s \in S$, $A_i \in \mathcal{A}_i$, and $p_i \in \Delta(X_i)$, define

$$\tau_s(p_i, A_i) := \tau_s(\{w \in \mathbb{R}^{X_i} : p_i \in M(M(A_i, U_s), w)\}).$$

Definition 10. An *S-based DREU representation* of ρ consists of tuples $(S_0, \mu_0, \{U_{s_0}, \tau_{s_0}\}_{s_0\in S_0})$, $(S_t, \{\mu_t^{s_{t-1}}\}_{s_{t-1}\in S_{t-1}}, \{U_{s_t}, \tau_{s_t}\}_{s_t\in S_t})_{1\le t\le T}$ such that for all $t = 0, \dots, T$, we have:

DREU1: For all $s_{t-1} \in S_{t-1}$, $(S_t, \mu_t^{s_{t-1}}, \{U_{s_t}, \tau_{s_t}\}_{s_t\in S_t})$ is an REU form on X_t such that[54]

(a) $U_{s_t} \not\approx U_{s'_t}$ for any distinct pair $s_t, s'_t \in \operatorname{supp}(\mu_t^{s_{t-1}})$;

(b) $\operatorname{supp}(\mu_t^{s_{t-1}}) \cap \operatorname{supp}(\mu_t^{s'_{t-1}}) = \varnothing$ for any distinct pair s_{t-1}, s'_{t-1};

(c) $\bigcup_{s_{t-1}\in S_{t-1}} \operatorname{supp} \mu_t^{s_{t-1}} = S_t$.

DREU2: For all p_t, A_t, and $h^{t-1} = (A_0, p_0, A_1, p_1, \dots, A_{t-1}, p_{t-1}) \in \mathcal{H}_{t-1}(A_t)$,[55]

$$\rho_t(p_t, A_t|h^{t-1}) = \frac{\sum_{(s_0,\dots,s_t)\in S_0\times\dots\times S_t} \prod_{k=0}^{t} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)}{\sum_{(s_0,\dots,s_{t-1})\in S_0\times\dots\times S_{t-1}} \prod_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)}.$$

An *S-based evolving utility representation* of ρ is an S-based DREU representation such that for all $t = 0, \dots, T$, we additionally have:

EVU: For all $s_t \in S_t$, there exists $u_{s_t} \in \mathbb{R}^Z$ such that for all $z_t \in Z, A_{t+1} \in \mathcal{A}_{t+1}$, we have

$$U_{s_t}(z_t, A_{t+1}) = u_{s_t}(z_t) + V_{s_t}(A_{t+1}),$$

where $V_{s_t}(A_{t+1}) := \sum_{s_{t+1}} \mu_{t+1}^{s_t}(s_{t+1}) \max_{p_{t+1}\in A_{t+1}} U_{s_{t+1}}(p_{t+1})$ for $t \le T-1$ and $V_{s_T} \equiv 0$.

An *S-based gradual learning representation* is an S-based evolving-utility representation such that additionally:

GL: There exists $\delta > 0$ such that for all $t = 0, \dots, T-1$ and $s_t \in S_t$, we have

$$u_{s_t} = \frac{1}{\delta} \sum_{s_{t+1}} \mu_{t+1}^{s_t}(s_{t+1}) u_{s_{t+1}}.$$

A.2 Equivalence Result

Proposition 5. Let ρ be a dynamic stochastic choice rule.

(i). ρ admits a DREU representation if and only if ρ admits an S-based DREU representation.

(ii). ρ admits an evolving utility representation if and only if ρ admits an S-based evolving utility representation.

(iii). ρ admits a gradual learning representation if and only if ρ admits an S-based gradual learning representation.

Proof. See Supplementary Appendix G. ■

[54] For $t = 0$, we abuse notation by letting $\mu_t^{s_{t-1}}$ denote μ_0 for all s_{t-1}.
[55] For $t = 0$, we again abuse notation by letting $\rho_t(\cdot|h^{t-1})$ denote $\rho_0(\cdot)$ for all h^{t-1}.

A.3 Relationship between Histories and States

Throughout the proofs of Theorems 1–3 we will make use of the following terminology concerning the relationship between histories and states. Fix any $t \in \{0, \ldots, T\}$. Suppose that $(S_{t'}, \{\mu_{t'}^{s_{t'-1}}\}_{s_{t'-1} \in S_{t'-1}}, \{U_{s_{t'}}, \tau_{s_{t'}}\}_{s_{t'} \in S_{t'}})$ satisfy DREU1 and DREU2 from Definition 10 for each $t' \leq t$.

Fix any state $s_t^* \in S_t$. We let $\text{pred}(s_t^*)$ denote the unique *predecessor sequence* $(s_0^*, \ldots, s_{t-1}^*) \in S_0 \times \ldots \times S_{t-1}$, given by assumptions DREU1 (b) and (c), such that $s_{k+1}^* \in \text{supp}(\mu_{k+1}^{s_k^*})$ for each $k = 0, ..., t-1$. Given any history $h^t = (A_0, p_0, \ldots, A_t, p_t)$, we say that s_t^* is *consistent* with h^t if $\prod_{k=0}^{t} \tau_{s_k^*}(p_k, A_k) > 0$.

For any $k = 0, \ldots, t$, $s_k \in S_k$, $p_0 \in A_0 \in \mathcal{A}_0$, and $p_{k+1} \in A_{k+1} \in \mathcal{A}_{k+1}$, let

$$\begin{aligned} \mathcal{U}_{s_k}(A_{k+1}, p_{k+1}) &:= \{U_{s_{k+1}} : s_{k+1} \in \text{supp}\, \mu_{k+1}^{s_k} \text{ and } p_{k+1} \in M(A_{k+1}, U_{s_{k+1}})\}; \\ \mathcal{U}_0(A_0, p_0) &:= \{U_{s_0} : s_0 \in S_0 \text{ and } p_0 \in M(A_0, U_{s_0})\}. \end{aligned}$$

A *separating history* for s_t^* is a history $h^t = (B_0, q_0, ..., B_t, q_t)$ such that $\mathcal{U}_{s_{k-1}^*}(B_k, q_k) = \{U_{s_k^*}\}$ for all $k = 0, \ldots, t$ and $h^t \in \mathcal{H}_t^*$, where we abuse notation by letting $\mathcal{U}_{s_{-1}^*}(B_0, q_0)$ denote $\mathcal{U}_0(B_0, q_0)$. Note that separating histories are required to be histories without ties.

We record the following properties:

Lemma 1. Fix any $s_t^* \in S_t$ with $\text{pred}(s_t^*) = (s_0^*, \ldots, s_{t-1}^*)$. Suppose $h^t = (B_0, q_0, \ldots, B_t, q_t)$ satisfies $\mathcal{U}_{s_{k-1}^*}(B_k, q_k) = \{U_{s_k^*}\}$ for all $k = 0, \ldots, t$. Then for all $k = 0, \ldots, t$, s_k^* is the only state in S_k that is consistent with h^k.

Proof. Fix any $\ell = 0, \ldots, t$. First, consider $s'_\ell \in S_\ell \smallsetminus \{s_\ell^*\}$, with $\text{pred}(s'_\ell) = (s'_0, \ldots, s'_{\ell-1})$. Let $k \leq \ell$ be smallest such that $s'_k \neq s_k^*$. Then $s'_k \in \text{supp}\, \mu_k^{s_{k-1}^*}$, so $\mathcal{U}_{s_{k-1}^*}(B_k, q_k) = \{U_{s_k^*}\}$ implies that $q_k \notin M(B_k, U_{s'_k})$. Thus, $\tau_{s'_k}(q_k, B_k) = 0$, whence s'_ℓ is not consistent with h^ℓ.

Next, to show that s_ℓ^* is consistent with h^ℓ, note that $\rho_\ell(q_\ell, B_\ell | h^{\ell-1}) > 0$, so DREU2 implies

$$\sum_{(s_0, \ldots, s_\ell) \in S_0 \times \ldots \times S_\ell} \prod_{k=0}^{\ell} \mu_k^{s_{k-1}}(s_k) \tau_{s_k}(q_k, B_k) > 0. \tag{11}$$

Now, if $(s_0, \ldots, s_{\ell-1}) \neq \text{pred}(s_\ell)$, then $\prod_{k=0}^{\ell} \mu_k^{s_{k-1}}(s_k) = 0$. And if $(s_0, \ldots, s_{\ell-1}) = \text{pred}(s_\ell)$ but $s_\ell \neq s_\ell^*$, then the first paragraph shows $\prod_{k=0}^{\ell} \tau_{s_k}(q_k, B_k) = 0$. Hence, (11) reduces to $\prod_{k=0}^{\ell} \mu_k^{s_{k-1}^*}(s_k^*) \tau_{s_k^*}(q_k, B_k) > 0$, whence s_ℓ^* is consistent with h^ℓ. ■

Lemma 2. Every $s_t^* \in S_t$ admits a separating history.

Proof. Fix any $s_t^* \in S_t$ with $\text{pred}(s_t^*) = (s_0^*, \ldots, s_{t-1}^*)$. By Lemma 13 and DREU1 (a), there exist menus $B_0 = \{q_0(s_0) : s_0 \in S_0\} \in \mathcal{A}_0$ and $B_k(s_{k-1}) = \{p_k(s_k) : s_k \in \text{supp}\, \mu_k^{s_{k-1}}\} \in \mathcal{A}_k$ for each $k = 1, \ldots, t$ and $s_k \in S_k$ such that $\mathcal{U}_0(B_0, q_0(s_0)) = \{U_{s_0}\}$ for all $s_0 \in S_0$ and $\mathcal{U}_{s_{k-1}}(B_k(s_{k-1}), q_k(s_k)) = \{U_{s_k}\}$ for all $s_k \in \text{supp}\, \mu_k^{s_{k-1}}$. Moreover, we can assume that $B_{k+1}(s_k) \in \text{supp}\, q_k(s_k)^A$ for all $k = 0, \ldots, t-1$ and $s_k \in S_k$, by letting each $q_k(s_k)$ put small enough weight on $(z, B_{k+1}(s_k))$ for some $z \in Z$. Then $h^t := (B_0, q_0(s_0^*), \ldots, B_t(s_t^*), q_t(s^*(t))) \in \mathcal{H}_t$. Moreover, since $\mathcal{U}_{s_{k-1}^*}(B_k, q_k(s_k^*)) = \{U_{s_k^*}\}$, Lemma 1 implies that or all for all $k = 0, \ldots, t$, s_k^* is the only state consistent with h^k. Additionally, for all $k = 0, \ldots, t$ and $s_k \in \text{supp}\, \mu_k^{s_{k-1}^*}$, we have $M(B_k(s_{k-1}^*), U_{s_k}) = \{q_k(s_k)\}$ by construction. Hence, by Lemma 14, we have $B_k(s_{k-1}^*) \in \mathcal{A}_k^*(h^{k-1})$. Thus $h^t \in \mathcal{H}_t^*$, so h^t is a separating history for s_t^*. ■

B Proof of Theorem 1

B.1 Proof of Theorem 1: Sufficiency

Suppose ρ satisfies Axioms 1–4. To show that ρ admits a DREU representation, it suffices, by Proposition 5, to construct an S-based DREU representation for ρ.

We proceed by induction on $t \leq T$. First consider $t = 0$. Since ρ_0 satisfies Axiom 3 and X_0 is a separable metric space by Lemma 12, the existence of $(S_0, \mu_0, \{U_{s_0}, \tau_{s_0}\}_{s_0 \in S_0})$ satisfying DREU1 and DREU2 from Definition 10 is immediate from Theorem 4, which extends Gul and Pesendorfer's (2006) and Ahn and Sarver's (2013) characterization result for static S-based REU representations to separable metric spaces and which we prove in Supplementary Appendix F.

Suppose next that $0 \leq t < T$ and that we have constructed $(S_{t'}, \{\mu_{t'}^{s_{t'-1}}\}_{s_{t'-1} \in S_{t'-1}}, \{U_{s_{t'}}, \tau_{s_{t'}}\}_{s_{t'} \in S_{t'}})$ satisfying DREU1 and DREU2 for each $t' \leq t$. We now construct $(S_{t+1}, \{\mu_{t+1}^{s_t}\}_{s_t \in S_t}, \{U_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}})$ satisfying DREU1 and DREU2.

B.1.1 Defining $\rho_{t+1}^{s_t}$ and $(S_{t+1}, \{\mu_{t+1}^{s_t}\}_{s_t \in S_t}, \{U_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}})$:

To this end, we first pick an arbitrary separating history $h^t(s_t)$ for each $s_t \in S_t$ (this exists by Lemma 2) and define

$$\rho_{t+1}^{s_t}(\cdot, A_{t+1}) := \rho_{t+1}(\cdot, A_{t+1}|h^t(s_t))$$

for all $A_{t+1} \in \mathcal{A}_{t+1}$. Note that here $\rho_{t+1}(\cdot, |h^t(s_t))$ is the extended version of $\rho_{t+1}(\cdot|h^t(s_t))$ given in Definition 3; by Axiom 2 and Lemma 15, the specific choice of $\lambda \in (0, 1]$ and $d^{t-1} \in \mathcal{D}_{t-1}$ used in the extension procedure does not matter.

By Axiom 3 and the fact that X_{t+1} is separable metric (Lemma 12), Theorem 4 applied to $\rho_{t+1}^{s_t}$ yields an REU form $(S_{t+1}^{s_t}, \mu_{t+1}^{s_t}, \{U_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}^{s_t}})$ on X_{t+1} such that $U_{s_{t+1}} \not\approx U_{s'_{t+1}}$ for any distinct pair $s_{t+1}, s'_{t+1} \in S_{t+1}^{s_t}$ and such that

$$\rho_{t+1}^{s_t}(p_{t+1}, A_{t+1}) = \sum_{s_{t+1} \in S_{t+1}^{s_t}} \mu_{t+1}^{s_t}(s_{t+1}) \tau_{s_{t+1}}(p_{t+1}, A_{t+1})$$

for all p_{t+1} and A_{t+1}. Without loss, we can assume that $S_{t+1}^{s_t}$ and $S_{t+1}^{s'_t}$ are disjoint whenever $s_t \neq s'_t$. Set $S_{t+1} := \bigcup_{s_t \in S_t} S_{t+1}^{s_t}$ and extend $\mu_{t+1}^{s_t}$ to a probability measure on S_{t+1} by setting $\mu_{t+1}^{s_t}(s_{t+1}) = 0$ for all $s_{t+1} \in S_{t+1} \smallsetminus S_{t+1}^{s_t}$.

By construction, it is immediate that $(S_{t+1}, \{\mu_{t+1}^{s_t}\}_{s_t \in S_t}, \{U_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}})$ thus defined satisfies DREU1 and that

$$\rho_{t+1}^{s_t}(p_{t+1}, A_{t+1}) = \sum_{s_{t+1} \in S_{t+1}} \mu_{t+1}^{s_t}(s_{t+1}) \tau_{s_{t+1}}(p_{t+1}, A_{t+1}) \tag{12}$$

for all p_{t+1} and A_{t+1}. It remains to show that DREU2 is also satisfied.

B.1.2 $\rho_{t+1}^{s_t}$ is well-behaved:

To this end, Lemma 3 below first shows that the definition of $\rho_{t+1}^{s_t}$ is well-behaved, in the sense that for any history h^t that can only arise in state s_t, $\rho_{t+1}^{s_t} = \rho_{t+1}(\cdot|h^t)$.

Lemma 3. Fix any $s_t^* \in S_t$ with $\text{pred}(s_t^*) = (s_0^*, ..., s_{t-1}^*)$. Suppose $h^t = (A_0, p_0, ..., A_t, p_t) \in \mathcal{H}_t$ satisfies $\mathcal{U}_{s_{k-1}^*}(A_k, p_k) = \{U_{s_k^*}\}$ for all $k = 0, 1, \ldots, t$. Then for any $A_{t+1} \in \mathcal{A}_{t+1}$, $\rho_{t+1}(\cdot, A_{t+1}|h^t) = \rho_{t+1}^{s_t^*}(\cdot, A_{t+1})$.

Proof. **Step 1:** Let $\tilde{h}^t = (\tilde{A}_0, \tilde{p}_0, \dots, \tilde{A}_t, \tilde{p}_t)$ denote the separating history for s_t^* used to define $\rho_{t+1}^{s_t^*}$. We first prove the Lemma under the assumption that $h^t \in \mathcal{H}_t^*$, i.e, that h^t is itself a separating history for s_t^*.[56]

Pick $(r_0, ..., r_t) \in \Delta(X_0) \times \ldots \times \Delta(X_t)$ such that $A_{t+1} \in \operatorname{supp} r_t^A$ and for all $k = 0, \dots, t-1$,

$$\operatorname{supp}(r_k^A) \supseteq \{B_{k+1}, \tilde{B}_{k+1}, B_{k+1} \cup \tilde{B}_{k+1}\},$$

where $B_\ell := \frac{1}{3}A_\ell + \frac{1}{3}\{\tilde{p}_\ell\} + \frac{1}{3}\{r_\ell\}$ and $\tilde{B}_\ell := \frac{1}{3}\tilde{A}_\ell + \frac{1}{3}\{p_\ell\} + \frac{1}{3}\{r_\ell\}$ for $\ell = 0, \dots, t$. Define $q_\ell := \frac{1}{3}p_\ell + \frac{1}{3}\tilde{p}_\ell + \frac{1}{3}r_\ell$.

Note that since $h^t, \tilde{h}^t \in \mathcal{H}_t^*$ and $\mathcal{U}_{s_{k-1}^*}(A_k, p_k) = \mathcal{U}_{s_{k-1}^*}(\tilde{A}_k, \tilde{p}_k) = \{U_{s_k^*}\}$, Lemma 14 implies that $M(A_k, U_{s_k^*}) = \{p_k\}$ and $M(\tilde{A}_k, U_{s_k^*}) = \{\tilde{p}_k\}$ for all $k = 0, 1, \dots, t$. By linearity of the U_s, we then also have

$$\begin{aligned}\mathcal{U}_{s_{k-1}^*}(B_k, q_k) = \mathcal{U}_{s_{k-1}^*}(\tilde{B}_k, q_k) = \mathcal{U}_{s_{k-1}^*}(B_k \cup \tilde{B}_k, q_k) = \{U_{s_k^*}\} \text{ and}\\ M(B_k, U_{s_k^*}) = M(\tilde{B}_k, U_{s_k^*}) = M(B_k \cup \tilde{B}_k, U_{s_k^*}) = \{q_k\}.\end{aligned}$$

This implies that for all $k = 0, \dots, t$ and $s_k \in \operatorname{supp} \mu_{k-1}^{s_{k-1}^*}$,

$$\tau_{s_k}(q_k, B_k) = \tau_{s_k}(q_k, \tilde{B}_k) = \tau_{s_k}(q_k, B_k \cup \tilde{B}_k) = \begin{cases} 1 \text{ if } s_k = s_k^* \\ 0 \text{ otherwise} \end{cases}$$

By DREU2 of the inductive hypothesis, it follows that for all $k = 0, \dots, t-1$,

$$\begin{aligned}\mu_t^{s_{t-1}^*}(s_t^*) &= \rho_t(q_t, B_t | B_0, q_0, \dots, B_{t-1}, q_{t-1}) = \rho_t(q_t, \tilde{B}_t | \tilde{B}_0, q_0, \dots, \tilde{B}_{t-1}, q_{t-1})\\ &= \rho_t(q_t, B_t \cup \tilde{B}_t | B_0, q_0, \dots, B_{k-1}, q_{k-1}, B_k \cup \tilde{B}_k, q_k, \dots, B_{t-1} \cup \tilde{B}_{t-1}, q_{t-1})\\ &= \rho_t(q_t, B_t \cup \tilde{B}_t | \tilde{B}_0, q_0, \dots, \tilde{B}_{k-1}, q_{k-1}, B_k \cup \tilde{B}_k, q_k, \dots, B_{t-1} \cup \tilde{B}_{t-1}, q_{t-1}),\end{aligned}$$

whence repeated application of Axiom 1 (Contraction History Independence) yields

$$\begin{aligned}\rho_{t+1}(\cdot, A_{t+1} | B_0, q_0, \dots, B_t, q_t) = \rho_{t+1}(\cdot, A_{t+1} | B_0 \cup \tilde{B}_0, q_0, \dots, B_t \cup \tilde{B}_t, q_t) =\\ \rho_{t+1}(\cdot, A_{t+1} | \tilde{B}_0, q_0, \dots, \tilde{B}_t, q_t).\end{aligned} \tag{13}$$

Moreover, by Axiom 2 (Linear History Independence) and Lemma 15, we have

$$\begin{aligned}\rho_{t+1}(\cdot, A_{t+1} | h^t) = \rho_{t+1}(\cdot, A_{t+1} | B_0, q_0, \dots, B_t, q_t) \text{ and}\\ \rho_{t+1}(\cdot, A_{t+1} | \tilde{h}^t) = \rho_{t+1}(\cdot, A_{t+1} | \tilde{B}_0, q_0, \dots, \tilde{B}_t, q_t).\end{aligned} \tag{14}$$

Combining (13) and (14) we obtain that $\rho_{t+1}(\cdot, A_{t+1} | h^t) = \rho_{t+1}(\cdot, A_{t+1} | \tilde{h}^t) := \rho_{t+1}^{s_t^*}(\cdot, A_{t+1})$. This proves the Lemma for histories $h^t \in \mathcal{H}_t^*$.

Step 2: Now suppose that $h^t \notin \mathcal{H}_t^*$. Take any sequence of histories $h^{t,n} \to^m h^t$ with $h^{t,n} = (A_0^n, p_0^n, ..., A_t^n, p_t^n) \in \mathcal{H}_t^*$ for each n. Note that such a sequence exists by Axiom 4 (History Continuity).

We claim that for all large enough n, $\mathcal{U}_{s_{k-1}^*}(A_k^n, p_k^n) = \{U_{s_k^*}\}$ for all $k = 0, \dots, t$. Suppose for a contradiction that we can find a subsequence $(h^{t,n_\ell})_{\ell=1}^\infty$ for which this claim is violated. Note that for all ℓ, $\rho_k(p_k^{n_\ell}, A_k^{n_\ell} | h^{k-1,n_\ell}) > 0$ for all $k = 0, \dots, t$ (by the fact that h^{t,n_ℓ} is a well-defined history).

[56]Note that $\mathcal{U}_{s_{k-1}^*}(A_k, p_k) = \{U_{s_k^*}\}$ for all $k = 0, 1, \dots, t$ does not by itself imply that h^t is a history without ties.

Hence, DREU2 for $k \leq t$ implies that we can find $s'_{t,n_\ell} \in S_t$ with $\mathrm{pred}(s'_{t,n_\ell}) = (s'_{0,n_\ell}, \ldots, s'_{t-1,n_\ell})$ and $(s'_{0,n_\ell}, \ldots, s'_{t,n_\ell}) \neq (s^*_0, \ldots, s^*_t)$ such that $U_{s'_{k,n_\ell}} \in \mathcal{U}_{s'_{k-1,n_\ell}}(A^{n_\ell}_k, p^{n_\ell}_k)$ for all $k = 0, \ldots, t$. Moreover, since $S_0 \times \ldots \times S_t$ is finite, by choosing the subsequence (h^{t,n_ℓ}) appropriately, we can assume that $(s'_{0,n_\ell}, \ldots, s'_{t,n_\ell}) = (s'_0, \ldots, s'_t) \neq (s^*_0, \ldots, s^*_t)$ for all ℓ. Pick the smallest k such that $s'_k \neq s^*_k$ and pick any $q_k \in A_k$. Since $A^{n_\ell}_k \to^m A_k$ we can find $q^{n_\ell}_k \in A^{n_\ell}_k$ with $q^{n_\ell}_k \to^m q_k$. For all ℓ we have $U_{s'_k} \in \mathcal{U}_{s'_{k-1}}(A^{n_\ell}_k, p^{n_\ell}_k)$, so $U_{s'_k}(p^{n_\ell}_k) \geq U_{s'_k}(q^{n_\ell}_k)$, whence $U_{s'_k}(p_k) \geq U_{s'_k}(q_k)$ by linearity of $U_{s'_k}$. Moreover, by choice of k, $s'_k \in \mathrm{supp}\, \mu^{s'_{k-1}}_{k-1} = \mathrm{supp}\, \mu^{s^*_{k-1}}_{k-1}$. Thus, $U_{s'_k} \in \mathcal{U}_{s^*_{k-1}}(A_k, p_k) = \{U_{s^*_k}\}$. But $s'_k \neq s^*_k$, so by DREU1 (a) of the inductive hypothesis $U_{s'_k} \not\approx U_{s^*_k}$, a contradiction.

By the previous paragraph, for large enough n, $h^{t,n}$ satisfies the assumption of the Lemma. Since $h^{t,n} \in \mathcal{H}^*_t$, Step 1 then shows that $\rho_{t+1}(p_{t+1}, A_{t+1}|h^{t,n}) = \rho^{s^*_t}_{t+1}(p_{t+1}, A_{t+1})$ for all large enough n and all p_{t+1}. By Axiom 4 (History Continuity), this implies that for all p_{t+1}

$$\rho_{t+1}(p_{t+1}, A_{t+1}|h^t) \quad \in \quad \mathrm{co}\{\lim_n \rho_{t+1}(p_{t+1}, A_{t+1}|h^{t,n}) : h^{t,n} \to^m h^t, h^{t,n} \in \mathcal{H}^*_t\} = \{\rho^{s^*_t}_{t+1}(p_{t+1}, A_{t+1})\},$$

which completes the proof. ■

B.1.3 $\rho_{t+1}(\cdot|h^t)$ is a weighted average of $\rho^{s_t}_{t+1}$:

The next lemma shows that $\rho_{t+1}(\cdot|h^t)$ can be expressed as a weighted average of the state-dependent choice distributions $\rho^{s_t}_{t+1}$, where the weight on each $\rho^{s_t}_{t+1}$ corresponds to the probability of s_t conditional on history h^t.

Lemma 4. For any $p_{t+1} \in A_{t+1}$ and $h^t = (A_0, p_0, ..., A_t, p_t) \in \mathcal{H}_t(A_{t+1})$, we have

$$\rho_{t+1}(p_{t+1}, A_{t+1}|h^t) = \frac{\sum_{(s_0,\ldots,s_t) \in S_0 \times \cdots \times S_t} \prod_{k=0}^t \mu^{s_{k-1}}_k(s_k) \tau_{s_k}(A_k, p_k) \rho^{s_t}_{t+1}(p_{t+1}, A_{t+1})}{\sum_{(s_0,\ldots,s_t) \in S_0 \times \cdots \times S_t} \prod_{k=0}^t \mu^{s_{k-1}}_k(s_k) \tau_{s_k}(A_k, p_k)}.$$

Proof. Let $\{s^1_t, ..., s^m_t\}$ denote the set of states in S_t that are consistent with history h^t (as defined in Section A.3). For each j, let $\hat{h}^t(j) = (B^j_0, q^j_0, \ldots, B^j_t, q^j_t)$ be a separating history for state s^j_t. We can assume that for each $k = 1, \ldots, t$, q^j_{k-1} puts small weight on $(z, \frac{1}{2}A_k + \frac{1}{2}B^j_k)$ for some z, so that $h^t(j) := \frac{1}{2}h^t + \frac{1}{2}\hat{h}^t(j) \in \mathcal{H}_t(A_{t+1})$ for all j.

Note first that for all $j = 1, \ldots, m$, we have

$$\rho(h^t(j)) = \prod_{k=0}^t \mu^{s^j_{k-1}}_k(s^j_k) \tau_{s^j_k}(p_k, A_k). \tag{15}$$

Indeed, observe that

$$\begin{aligned} \rho(h^t(j)) &= \prod_{k=0}^t \rho_k(\frac{1}{2}p_k + \frac{1}{2}q^j_k, \frac{1}{2}A_k + \frac{1}{2}B^j_k | \frac{1}{2}h^{k-1} + \frac{1}{2}\hat{h}^{k-1}(j)) \\ &= \sum_{(s_0,\ldots,s_t)} \prod_{k=0}^t \mu^{s_{k-1}}_k(s_k) \tau_{s_k}(\frac{1}{2}p_k + \frac{1}{2}q_k, \frac{1}{2}A_k + \frac{1}{2}B^j_k) \\ &= \prod_{k=0}^t \mu^{s^j_{k-1}}_k(s^j_k) \tau_{s^j_k}(\frac{1}{2}p_k + \frac{1}{2}q^j_k, \frac{1}{2}A_k + \frac{1}{2}B^j_k) = \prod_{k=0}^t \mu^{s^j_{k-1}}_k(s^j_k) \tau_{s^j_k}(p_k, A_k). \end{aligned}$$

The first equality holds by definition. The second equality follows from DREU2 of the inductive hypothesis. For the final two equalities, note that since $\hat{h}^t(j)$ is a separating history for s_t^j, we have for all $k = 0, \dots, t$ that $\mathcal{U}_{s_{k-1}^j}(B_k^j, q_k^j) = \{U_{s_k^j}\}$ with $\{q_k^j\} = M(B_k^j, U_{s_k^j})$ (by Lemma 14). Also, since s_t^j is consistent with h^t, $\tau_{s_k^j}(p_k, A_k) > 0$ for all $k = 0, \dots, t$. This implies that for every $s_k \in \operatorname{supp} \mu_k^{s_{k-1}^j}$, $\tau_{s_k}(\frac{1}{2}p_k + \frac{1}{2}q_k^j, \frac{1}{2}A_k + \frac{1}{2}B_k) > 0$ if and only if $s_k = s_k^j$, yielding the third equality. It also implies that $M(\frac{1}{2}A_k + \frac{1}{2}B_k^j, U_{s_k^j}) = M(\frac{1}{2}A_k + \frac{1}{2}\{q_k^j\}, U_{s_k^j})$, so that $\tau_{s_k^j}(\frac{1}{2}p_k + \frac{1}{2}q_k^j, \frac{1}{2}A_k + \frac{1}{2}B_k^j) = \tau_{s_k^j}(\frac{1}{2}p_k + \frac{1}{2}q_k^j, \frac{1}{2}A_k + \frac{1}{2}\{q_k^j\}) = \tau_{s_k^j}(p_k, A_k)$, yielding the fourth equality.

Now let $H^t := \{h^t(j) : j = 1, \dots, m\} \subseteq \mathcal{H}_t(A_{t+1})$. Note that by repeated application of Axiom 2, we have that

$$\rho_{t+1}(p_{t+1}, , A_{t+1}|h^t) = \rho_{t+1}(p_{t+1}, A_{t+1}|H^t). \tag{16}$$

Moreover, we have that

$$\begin{aligned}
\rho_{t+1}(p_{t+1}, A_{t+1}|H^t) &= \frac{\sum_{j=1}^m \rho(h^t(j))\rho_{t+1}(p_{t+1}, A_{t+1}|h^t(j))}{\sum_{j=1}^m \rho(h^t(j))} \\
&= \frac{\sum_{j=1}^m \prod_{k=0}^t \mu_k^{s_{k-1}^j}(s_k^j)\tau_{s_k^j}(p_k, A_k)\rho_{t+1}(p_{t+1}, A_{t+1}|h^t(j))}{\sum_{j=1}^m \prod_{k=0}^t \mu_k^{s_{k-1}^j}(s_k^j)\tau_{s_k^j}(p_k, A_k)} \\
&= \frac{\sum_j \prod_{k=0}^t \mu_k^{s_{k-1}^j}(s_k^j)\tau_{s_k^j}(p_k, A_k)\rho_{t+1}^{s_t^j}(p_{t+1}|A_{t+1})}{\sum_j \prod_{k=0}^t \mu_k^{s_{k-1}^j}(s_k^j)\tau_{s_k^j}(p_k, A_k)} \\
&= \frac{\sum_{(s_0,\dots,s_t)\in S_0\times\cdots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(A_k, p_k)\rho_{t+1}^{s_t}(p_{t+1}|A_{t+1})}{\sum_{(s_0,\dots,s_t)\in S_0\times\cdots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(A_k, p_k)}.
\end{aligned} \tag{17}$$

Indeed, the first equality holds by definition of choice conditional on a set of histories. The second equality follows from Equation (15). Note next that since $\hat{h}^t(j)$ is a separating history for s_t^j and s_t^j is consistent with h^t, we have that $\mathcal{U}_{s_k^j}(\frac{1}{2}p_k + \frac{1}{2}q_k^j, \frac{1}{2}A_k + \frac{1}{2}B_k^j) = \{U_{s_k^j}\}$ for each k. Hence, Lemma 3 implies that $\rho_{t+1}(p_{t+1}, A_{t+1}|h^t(j)) = \rho_{t+1}^{s_t^j}(p_{t+1}, A_{t+1})$, yielding the third equality. Finally, note that if $(s_0, \dots, s_t) \in S_0 \times \dots S_t$ with $(s_0, \dots, s_t) \neq (s_0^j, \dots, s_t^j)$ for all j, then either $s_t \notin \{s_t^1, \dots, s_t^m\}$, or $s_t = s_j^t$ for some j but $(s_0, \dots, s_{t-1}) \neq \operatorname{pred}(s_t^j)$. In either case, $\prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(A_k, p_k) = 0$, yielding the final equality. Combining (16) and (17), we obtain the desired conclusion. ■

B.1.4 Completing the proof:

Finally, combining Lemma 4 with the representation of $\rho_{t+1}^{s_t}$ in (12) yields that for any $h^t = (A_0, p_0, ..., A_t, p_t) \in \mathcal{H}_t(A_{t+1})$

$$\begin{aligned}
&\rho_{t+1}(p_{t+1}, A_{t+1}|h^t) \\
= &\frac{\sum_{(s_0,\dots,s_t)\in S_0\times\cdots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(A_k, p_k) \sum_{s_{t+1}\in S_{t+1}} \mu_{t+1}^{s_t}(s_{t+1})\tau_{s_{t+1}}(p_{t+1}, A_{t+1})}{\sum_{(s_0,\dots,s_t)\in S_0\times\cdots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(A_k, p_k)} \\
= &\frac{\sum_{(s_0,\dots,s_t,s_{t+1})\in S_0\times\cdots\times S_t\times S_{t+1}} \prod_{k=0}^{t+1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(A_k, p_k)}{\sum_{(s_0,\dots,s_t)\in S_0\times\cdots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(A_k, p_k)}.
\end{aligned}$$

Thus, $(S_{t+1}, \{\mu_{t+1}^{s_t}\}_{s_t \in S_t}, \{U_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}})$ also satisfies requirement DREU2, completing the proof.

B.2 Proof of Theorem 1: Necessity

Suppose ρ admits a DREU representation. By Proposition 5, ρ admits an S-based DREU representation. By Lemma 16, for each t and $h^t \in \mathcal{H}_t$, the (static) stochastic choice rule $\rho_t(\cdot|h^t) : \mathcal{A}_t \to \Delta(\Delta(X_t))$ given by the extended version of ρ from Definition 3 also satisfies DREU2. In other words, $\rho_t(\cdot|h^t)$ admits an S-based REU representation (see Definition 11). Thus, Theorem 4 implies that Axiom 3 holds. It remains to verify that Axioms 1, 2, and 4 are satisfied.

Claim 1. ρ satisfies Axiom 1 (Contraction History Independence).

Proof. Take any $h^{t-1} = (h_{-k}^{t-1}, (A_k, p_k)), \hat{h}^{t-1} = (h_{-k}^{t-1}, (B_k, p_k)) \in \mathcal{H}_{t-1}(A_t)$ such that $B_k \supseteq A_k$ and $\rho_k(p_k; A_k|h^{k-1}) = \rho_k(p_k; B_k|h^{k-1})$. From DREU2 for ρ_k, $\rho_k(p_k; A_k|h^{k-1}) = \rho_k(p_k; B_k|h^{k-1})$ implies that

$$\sum_{(s_0,\dots,s_k)} \prod_{l=0}^{k-1} \mu_l^{s_{l-1}}(s_l)\tau_{s_l}(p_l, A_l)\mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k) = \sum_{(s_0,\dots,s_k)} \prod_{l=0}^{k-1} \mu_l^{s_{l-1}}(s_l)\tau_{s_l}(p_l, A_l)\mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, B_k). \tag{18}$$

Since $B_k \supseteq A_k$ implies $\tau_{s_k}(p_k, A_k) \geq \tau_{s_k}(p_k, B_k)$ for all s_k, the only way for (18) to hold is if $\tau_{s_k}(p_k, A_k) = \tau_{s_k}(p_k, B_k)$ for all s_k consistent with h^k. Thus,

$$\rho_t(p_t; A_t|h^{t-1}) = \frac{\sum_{(s_0,\dots,s_t)\in S_0\times\dots\times S_t} \prod_{l=0}^{t} \mu_l^{s_{l-1}}(s_l)\tau_{s_l}(p_l, A_l)}{\sum_{(s_0,\dots,s_{t-1})\in S_0\times\dots\times S_{t-1}} \prod_{l=0}^{t-1} \mu_l^{s_{l-1}}(s_l)\tau_{s_l}(p_l, A_l)} = \rho_t(p_t; A_t|\hat{h}^{t-1}),$$

as required. ∎

Claim 2. ρ satisfies Axiom 2 (Linear History Independence).

Proof. Take any A_t, $h^{t-1} = (A_0, p_0, \dots, A_{t-1}, p_{t-1}) \in \mathcal{H}_{t-1}(A_t)$, and $H^{t-1} \subseteq \mathcal{H}_{t-1}(A_t)$ of the form $H^{t-1} = \{(h_{-k}^{t-1}, (\lambda A_k + (1-\lambda)B_k, \lambda p_k + (1-\lambda)q_k)) : q_k \in B_k\}$ for some $k < t$, $\lambda \in (0,1)$, and $B_k = \{q_k^j : j = 1,\dots,m\} \in \mathcal{A}_k$. Let $\tilde{A}_k := \lambda A_k + (1-\lambda)B_k$, and for each $j = 1,\dots,m$, let $\tilde{p}_k^j := \lambda p_k + (1-\lambda)q_k^j$ and $\tilde{h}^{t-1}(j) := (h_{-k}^{t-1}, (\tilde{A}_k, \tilde{p}_k^j))$.

By DREU2, for all p_t, we have

$$\rho_t(p_t; A_t|h^{t-1}) = \frac{\sum_{(s_0,\dots,s_t)} \prod_{\ell=0}^{t} \mu_\ell^{s_{\ell-1}}(s_\ell)\tau_{s_\ell}(p_\ell, A_\ell)}{\sum_{(s_0,\dots,s_{t-1})} \prod_{\ell=0}^{t-1} \mu_\ell^{s_{\ell-1}}(s_\ell)\tau_{s_\ell}(p_\ell, A_\ell)}. \tag{19}$$

Moreover, by definition

$$\rho_t(p_t; A_t|H^{t-1}) = \frac{\sum_{j=1}^m \rho(\tilde{h}^{t-1}(j))\rho_t(p_t; A_t|\tilde{h}^{t-1}(j))}{\sum_{j=1}^m \rho(\tilde{h}^{t-1}(j))},$$

where for each $j = 1,\dots,m$, DREU2 yields

$$\rho_t(p_t; A_t|\tilde{h}^{t-1}(j)) = \frac{\sum_{(s_0,\dots,s_t)} \left(\prod_{\ell=0,\dots,t;\ell\neq k} \mu_\ell^{s_{\ell-1}}(s_\ell)\tau_{s_\ell}(p_\ell, A_\ell)\right) \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(\tilde{p}_k^j, \tilde{A}_k)}{\sum_{(s_0,\dots,s_{t-1})} \left(\prod_{\ell=0,\dots,t-1;\ell\neq k} \mu_\ell^{s_{\ell-1}}(s_\ell)\tau_{s_\ell}(p_\ell, A_\ell)\right) \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(\tilde{p}_k^j, \tilde{A}_k)}.$$

and

$$\rho(\tilde{h}^{t-1}(j)) := \prod_{\ell=0,\ldots,t-1;\ell\neq k} \rho_\ell(p_\ell; A_\ell|\tilde{h}^{\ell-1})\rho_k(\tilde{p}_k^j; \tilde{A}_k|\tilde{h}^{k-1})$$

$$= \sum_{(s_0,\ldots,s_{t-1})} \left(\prod_{\ell=0,\ldots,t-1;\ell\neq k} \mu_\ell^{s_{\ell-1}}(s_\ell)\tau_{s_\ell}(p_\ell, A_\ell) \right) \mu_k^{s_{k-1}}(s_k)\tau_{s_k}((\tilde{p}_k^j, \tilde{A}_k)).$$

Combining and rearranging, we obtain

$$\rho_t(p_t; A_t|H^{t-1}) = \frac{\sum_{(s_0,\ldots,s_t)} \left(\prod_{\ell=0,\ldots,t;\ell\neq k} \mu_\ell^{s_{\ell-1}}(s_\ell)\tau_{s_\ell}(A_\ell, p_\ell) \right) \mu_k^{s_{k-1}}(s_k) \sum_{j=1}^m \tau_{s_k}(\tilde{p}_k^j, \tilde{A}_k)}{\sum_{(s_0,\ldots,s_{t-1})} \left(\prod_{\ell=0,\ldots,t-1;\ell\neq k} \mu_\ell^{s_{\ell-1}}(s_\ell)\tau_{s_\ell}(A_\ell, p_\ell) \right) \mu_k^{s_{k-1}}(s_k) \sum_{j=1}^m \tau_{s_k}(\tilde{p}_k^j, \tilde{A}_k)}. \quad (20)$$

But observe that for all s_k,

$$\sum_{j=1}^m \tau_{s_k}(\tilde{p}_k^j, \tilde{A}_k) = \sum_{j=1}^m \tau_{s_k}(\{w \in \mathbb{R}^{X_k} : \tilde{p}_k^j \in M(M(\tilde{A}_k, U_{s_k}), w)\})$$

$$= \sum_{q_k \in B_k} \tau_{s_k}(\{w \in \mathbb{R}^{X_k} : p_k \in M(M(A_k, U_{s_k}), w) \text{ and } q_k \in M(M(B_k, U_{s_k}), w)\}) \quad (21)$$

$$= \tau_{s_k}(\{w \in \mathbb{R}^{X_k} : p_k \in M(M(A_k, U_{s_k}), w)\})$$

$$= \tau_{s_k}(p_k, A_k),$$

where the second equality follows from linearity of the representation, the third equality from the fact that τ_{s_k} is a proper finitely-additive probability measure on $\mathbb{R}^{X_k}$, and the remaining equalities hold by definition. Combining (19), (20), and (21), we obtain $\rho_t(p_t; A_t|h^{t-1}) = \rho_t(p_t; A_t|H^{t-1})$, as required. ■

Claim 3. ρ satisfies Axiom 4 (History Continuity).

Proof. Fix any A_t, $p_t \in A_t$, and $h^{t-1} = (A_0, p_0, ..., A_{t-1}, p_{t-1}) \in \mathcal{H}^{t-1}$. Let $S_{t-1}(h^{t-1}) \subseteq S_{t-1}$ denote the set of period-$(t-1)$ states that are consistent with h^{t-1}. Define $\rho_t^{s_{t-1}}(p_t; A_t) := \sum_{s_t} \mu_t^{s_{t-1}}(s_t)\tau_{s_t}(p_t, A_t)$ for each s_{t-1}. By Lemma 16,

$$\begin{aligned}\rho_t(p_t; A_t|h^{t-1}) &= \frac{\sum_{(s_0,\ldots,s_t)\in S_0\times\cdots\times S_t} \prod_{k=0}^t \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)}{\sum_{(s_0,\ldots,s_{t-1})\in S_0\times\cdots\times S_{t-1}} \prod_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)} \\ &= \frac{\sum_{(s_0,\ldots,s_{t-1})\in S_0\times\cdots\times S_{t-1}} \prod_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k) \sum_{s_t\in S_t} \mu_t^{s_{t-1}}(s_t)\tau_{s_t}(p_t, A_t)}{\sum_{(s_0,\ldots,s_{t-1})\in S_0\times\cdots\times S_{t-1}} \prod_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)}.\end{aligned}$$

Hence, $\rho_t(p_t; A_t|h^{t-1}) \in \text{co}\{\rho_t^{s_{t-1}}(p_t; A_t) : s_{t-1} \in S_{t-1}(h^{t-1})\}$. Fix any $s_{t-1}^* \in S_{t-1}(h^{t-1})$. To prove the claim, it is sufficient to show that

$$\rho_t^{s_{t-1}^*}(p_t; A_t) \in \{\lim_n \rho_t(p_t; A_t|h_n^{t-1}) : h_n^{t-1} \to^m h^{t-1}, h_n^{t-1} \in \mathcal{H}_{t-1}^*\}.$$

To this end, let $\text{pred}(s_{t-1}^*) = (s_0^*, \ldots, s_{t-2}^*)$ and let $\bar{h}^{t-1} = (B_0, q_0, ..., B_{t-1}, q_{t-1}) \in \mathcal{H}_{t-1}^*$ be a separating history for s_{t-1}^*. By Lemma 17, for each $k = 0, \ldots, t-1$, we can find sequences $A_k^n \in \mathcal{A}_k^*(\bar{h}^{k-1})$ and $p_k^n \in A_k^n$ such that $A_k^n \to^m A_k$, $p_k^n \to^m p_k$ and $\mathcal{U}_{s_{k-1}^*}(A_k^n, p_k^n) = \{U_{s_k^*}\}$ for all n and all $k = 0, \dots, t-1$. Working backwards from $k = t-2$, we can inductively replace A^n_k and p^n_k with a mixture putting small weight on (z, A^n_{k+1}) for some z to ensure that $A^n_{k+1} \in \operatorname{supp} p^{n,A}_k$ for all $k \leq t-2$ while maintaining the properties in the previous sentence. Then by construction $h^{t-1}_n := (A^n_0, p^n_0, \dots, A^n_{t-1}, p^n_{t-1}) \in \mathcal{H}^*_{t-1}(A_t)$ and h^{t-1}_n is a separating history for s^*_{t-1}, which by Lemma 16 implies

$$\rho_t(p_t; A_t | h^{t-1}_n) = \frac{\sum_{s_t \in S_t} \left(\prod_{k=0}^{t-1} \mu_k^{s^*_{k-1}}(s^*_k) \tau_{s^*_k}(p_k, A_k) \right) \mu_t^{s^*_{t-1}}(s_t) \tau_{s_t}(p_t, A_t)}{\prod_{k=0}^{t-1} \mu_k^{s^*_{k-1}}(s^*_k) \tau_{s^*_k}(p_k, A_k)}$$
$$= \sum_{s_t} \mu_t^{s^*_{t-1}}(s_t) \tau_{s_t}(p_t, A_t) =: \rho_t^{s^*_{t-1}}(p_t; A_t)$$

for each n. Since $h^{t-1}_n \to^m h^{t-1}$, this verifies the desired claim. ■

C Proof of Theorem 2

C.1 Implications of $\succsim_{h^t}$

We begin with a preliminary lemma that characterizes the implications of the history-dependent revealed preference $\succsim_{h^t}$.

Lemma 5. Suppose that ρ admits an S-based DREU representation. Consider any $t \leq T-1$, $h^t = (A_0, p_0, \dots, A_t, p_t) \in \mathcal{H}_t$, and $q_t, r_t \in \Delta(X_t)$.

(i). If $q_t \succsim_{h^t} r_t$, then $U_{s_t}(q_t) \geq U_{s_t}(r_t)$ for all s_t consistent with h^t.

(ii). Suppose there exist $g, b \in \Delta(X_t)$ such that $U_{s_t}(g) > U_{s_t}(b)$ for all s_t consistent with h^t. If $U_{s_t}(q_t) \geq U_{s_t}(r_t)$ for all s_t consistent with h^t, then $q_t \succsim_{h^t} r_t$.

(iii). If h^t is a separating history for s_t, then $q_t \succsim_{h^t} r_t$ if and only if $U_{s_t}(q_t) \geq U_{s_t}(r_t)$.

Proof. (i): We prove the contrapositive. Suppose there exists s_t consistent with h^t such that $U_{s_t}(q_t) < U_{s_t}(r_t)$. Since s_t is consistent with h^t, we have $\Pi_{k=0}^{t} \mu_k^{s_{k-1}}(s_k) \tau_{s_k}(p_k, A_k) > 0$ for $\text{pred}(s_t) = (s_0, \dots, s_{t-1})$. Moreover, $U_{s_t}(q_t) < U_{s_t}(r_t)$ implies that for any $q^n_t \to^m q_t$ and $r^n_t \to^m r_t$, we have $U_{s_t}(q^n_t) < U_{s_t}(r^n_t)$ for large enough n. But then for all large enough n, we have $\tau_{s_t}(\frac{1}{2}p_t + \frac{1}{2}r^n_t; \frac{1}{2}A_t + \frac{1}{2}\{q^n_t, r^n_t\}) = \tau_{s_t}(p_t, A_t) > 0$, whence by Lemma 16, $\rho_t(\frac{1}{2}p_t + \frac{1}{2}r^n_t; \frac{1}{2}A_t + \frac{1}{2}\{q^n_t, r^n_t\} | h^{t-1}) > 0$. Thus, by definition, $q_t \not\succsim_{h^t} r_t$.

(ii): Let $S_t(h^t)$ denote the set of s_t consistent with h^t. Suppose $U_{s_t}(q_t) \geq U_{s_t}(r_t)$ for all $s_t \in S_t(h^t)$. Then picking $g, b \in \Delta(X_t)$ such that $U_{s_t}(g) > U_{s_t}(b)$ for all $s_t \in S_t(h^t)$ and letting $q^n_t := \frac{n}{n+1} q_t + \frac{1}{n+1} g$ and $r^n_t := \frac{n}{n+1} r_t + \frac{1}{n+1} b$ for all n, we have $q^n_t \to^m q_t$, $r^n_t \to^m r_t$, and $U_{s_t}(q^n_t) > U_{s_t}(r^n_t)$ for all n and $s_t \in S_t(h^t)$. Consider any $(s_0, \dots, s_{t-1}, s_t) \in S_0 \times \dots \times S_{t-1} \times S_t$. Then either $s_t \in S_t(h^t)$, in which case $\tau_{s_t}(\frac{1}{2}p_t + \frac{1}{2}r^n_t; \frac{1}{2}A_t + \frac{1}{2}\{q^n_t, r^n_t\}) = 0$ for all n, so that $\Pi_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k) \tau_{s_k}(p_k, A_k) \mu_t^{s_{t-1}}(s_t) \tau_{s_t}(\frac{1}{2}p_t + \frac{1}{2}r^n_t; \frac{1}{2}A_t + \frac{1}{2}\{q^n_t, r^n_t\}) = 0$. Or $s_t \notin S_t(h^t)$, in which case $\Pi_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k) \tau_{s_k}(p_k, A_k) \mu_t^{s_{t-1}}(s_t) \tau_{s_t}(p_t, A_t) = 0$, in which case again $\Pi_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k) \tau_{s_k}(p_k, A_k) \mu_t^{s_{t-1}}(s_t) \tau_{s_t}(\frac{1}{2}p_t + \frac{1}{2}r^n_t; \frac{1}{2}A_t + \frac{1}{2}\{q^n_t, r^n_t\}) = 0$. By Lemma 16, this implies $\rho_t(\frac{1}{2}p_t + \frac{1}{2}r^n_t; \frac{1}{2}A_t + \frac{1}{2}\{q^n_t, r^n_t\} | h^{t-1}) = 0$ for all n, i.e., $q_t \succsim_{h^t} r_t$.

(iii): Finally, suppose h^t is a separating history for s_t. If $q_t \succsim_{h^t} r_t$, then $U_{s_t}(q_t) \geq U_{s_t}(r_t)$ by part (i). For the converse, note that since U_{s_t} is non-constant, there exist $g, b \in \Delta(X_t)$ such that $U_{s_t}(g) > U_{s_t}(b)$. Since s_t is the only state consistent with h^t (recall Lemma 1), part (ii) implies that if $U_{s_t}(q_t) \geq U_{s_t}(r_t)$ then $q_t \succsim_{h^t} r_t$. ■

C.2 Proof of Theorem 2: Sufficiency

Throughout this section, we assume that ρ admits a DREU representation and satisfies Axioms 5–7. We will show that ρ admits an evolving utility representation. By Proposition 5, it is sufficient to construct an S-based evolving utility representation. Sections C.2.1–C.2.5 accomplish this.

C.2.1 Recursive Construction up to t

The construction proceeds recursively. Suppose that $t \leq T-1$. Assume that we have obtained $(S_{t'}, \{\mu_{t'}^{s_{t'-1}}\}_{s_{t'-1}\in S_{t'-1}}, \{U_{s_{t'}}, \tau_{s_{t'}}\}_{s_{t'}\in S_{t'}})$ for each $t' \leq t$ such that DREU1 and DREU2 hold for all $t' \leq t$ and EVU holds for all $t' \leq t-1$ (see Definition 10 for the statements of these conditions). Note that the base case $t = 0$ is true because of the fact that ρ admits a DREU representation and by Proposition 5 (the requirement that EVU holds for $t' \leq t-1$ is vacuous here). To complete the proof, we will construct $(S_{t+1}, \{\mu_{t+1}^{s_t}\}_{s_t\in S_t}, \{U_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1}\in S_{t+1}})$ such that DREU1 and DREU2 hold for $t' \leq t+1$ and EVU holds for $t' \leq t$.

C.2.2 Properties of U_{s_t}

Using Lemma 5, the following lemma translates Axioms 5 (Separability) and 6 (DLR Menu Preference) into properties of U_{s_t}.

Lemma 6. For any $s_t \in S_t$, there exist functions $u_{s_t} : Z \to \mathbb{R}$ and $V_{s_t} : \mathcal{A}_{t+1} \to \mathbb{R}$ such that

(i). $U_{s_t}(z_t, A_{t+1}) = u_{s_t}(z_t) + V_{s_t}(A_{t+1})$ holds for each (z_t, A_{t+1})

(ii). V_{s_t} is continuous

(iii). V_{s_t} is monotone, i.e., $V_{s_t}(A'_{t+1}) \geq V_{s_t}(A_{t+1})$ for any $A_{t+1} \subseteq A'_{t+1}$

(iv). V_{s_t} is linear, i.e., $V_{s_t}(\alpha A_{t+1} + (1-\alpha)A'_{t+1}) = \alpha V_{s_t}(A_{t+1}) + (1-\alpha)V_s(A'_{t+1})$ for all A_{t+1}, A'_{t+1} and $\alpha \in (0,1)$.

Moreover, there exist C'_{t+1}, $C_{t+1} \in \mathcal{A}_{t+1}$ such that $V_{s_t}(C'_{t+1}) > V_{s_t}(C_{t+1})$ for all $s_t \in S_t$.

Proof. Fix any $s_t \in S_t$ and a separating history h^t for s_t, the existence of which is guaranteed by Lemma 2.

For (i), it suffices, by standard arguments, to show that

$$U_{s_t}(z_t, A_{t+1}) + U_{s_t}(z'_t, A'_{t+1}) = U_{s_t}(z'_t, A_{t+1}) + U_{s_t}(z_t, A'_{t+1})$$

for all z_t, z'_t, and A_{t+1}, A'_{t+1}. Fix any z_t, z'_t, and A_{t+1}, A'_{t+1}. By Axiom 5 (Separability), we have $\frac{1}{2}(z_t, A_{t+1}) + \frac{1}{2}(z'_t, A'_{t+1}) \sim_{h^t} \frac{1}{2}(z'_t, A_{t+1}) + \frac{1}{2}(z_t, A'_{t+1})$, whence Lemma 5 (iii) implies that $\frac{1}{2}U_{s_t}(z_t, A_{t+1}) + \frac{1}{2}U_{s_t}(z'_t, A'_{t+1}) = \frac{1}{2}U_{s_t}(z'_t, A_{t+1}) + \frac{1}{2}U_{s_t}(z_t, A'_{t+1})$. This proves that there exist functions $u_{s_t} : Z \to \mathbb{R}$ and $V_{s_t} : \mathcal{A}_{t+1} \to \mathbb{R}$ such that $U_{s_t}(z_t, A_{t+1}) = u_{s_t}(z_t) + V_{s_t}(A_{t+1})$ for each (z_t, A_{t+1}), as required.

For (ii), note that since $\succsim_{h^t}$ is continuous on $\Delta(X_t)$ by Axiom 6 (iii) (Continuity) and represented by U_{s_t} (by Lemma 5 (iii)), U_{s_t} is continuous. By part (i), this implies that V_{s_t} is continuous on $\mathcal{A}_{t+1}$.

For (iii), suppose $A_{t+1} \subseteq A'_{t+1}$ and fix any z_t. By Axiom 6 (i) (Monotonicity), we have $(z_t, A'_{t+1}) \succsim_{h^t} (z_t, A_{t+1})$. By Lemma 5, this implies that $U_{s_t}(z_t, A'_{t+1}) \geq U_{s_t}(z_t, A_{t+1})$, whence $V_{s_t}(A'_{t+1}) \geq V_{s_t}(A_{t+1})$ by (i).

For (iv), fix any A_{t+1}, A'_{t+1}, z_t, and $\alpha \in (0,1)$. Axiom 6 (ii) (Indifference to Timing) implies $(z_t, \alpha A_{t+1} + (1-\alpha)A'_{t+1}) \sim_{h^t} \alpha(z_t, A_{t+1}) + (1-\alpha)(z_t, A'_{t+1})$, which by Lemma 5 implies $U_{s_t}((z_t, \alpha A_{t+1} +$

$(1-\alpha)A'_{t+1})) = U_{s_t}(\alpha(z_t, A_{t+1}) + (1-\alpha)(z_t, A'_{t+1}))$. By linearity and separability of U_{s_t}, this implies $V_{s_t}(\alpha A_{t+1} + (1-\alpha)A'_{t+1}) = \alpha V_{s_t}(A_{t+1}) + (1-\alpha)V_{s_t}(A'_{t+1})$, as required.

Finally, for the "moreover" part, again consider any s_t^* and separating history h^t for s_t^*. By Axiom 6 (iv) (Non-degenerate Menu Preference), there exist $A'_{t+1}(s_t^*), A_{t+1}(s_t^*)$, and z_t such that $(z_t, A'_{t+1}(s_t^*)) \succ_{h^t} (z_t, A_{t+1}(s_t^*))$. Thus, Lemma 5 (iii) implies $U_{s_t^*}(z_t, A'_{t+1}(s_t^*)) > U_{s_t^*}(z_t, A_{t+1}(s_t^*))$, so $V_{s_t^*}(A'_{t+1}(s_t^*)) > V_{s_t^*}(A_{t+1}(s_t^*))$ by part (i). Now let $C'_{t+1} := \bigcup_{s_t^* \in S_t} \left(A'_{t+1}(s_t^*) \cup A_{t+1}(s_t^*)\right)$ and let $C_{t+1} := \sum_{s_t^* \in S_t} \frac{1}{|S_t|} A_{t+1}(s_t^*)$. Then for all s_t and s'_t, by monotonicity of V_{s_t}, we have $V_{s_t}(C'_{t+1}) \geq V_{s_t}(A_{t+1}(s'_t))$, where by construction this inequality is strict whenever $s_t = s'_t$. By linearity of V_{s_t}, this implies $V_{s_t}(C'_{t+1}) > V_{s_t}(C_{t+1})$. ■

Corollary C.1. *Fix any $t \leq T-1$ and $h^t \in \mathcal{H}^t$. Then $q_t \succsim_{h^t} r_t$ if and only if $U_{s_t}(q_t) \geq U_{s_t}(r_t)$ for all s_t consistent with h^t.*

Proof. The "only if" direction is a restatement of part (i) of Lemma 5. For the "if" direction, let C'_{t+1} and C_{t+1} be as in the "moreover" part of Lemma 6. Pick any $z \in Z$ and let $g_{t+1} := \delta_{(z, C'_{t+1})}$ and $b_{t+1} := \delta_{(z, C_{t+1})}$. Then by Lemma 6, $U_{s_t}(g_{t+1}) > U_{s_t}(b_{t+1})$ for all s_t. Hence, the "if" direction is immediate from part (ii) of Lemma 5. ■

C.2.3 Construction of Random Utility in Period $t+1$

Since ρ admits a DREU representation, it admits an S-based DREU representation by Proposition 5, so in particular we can obtain $(S_{t+1}, \{\mu_{t+1}^{s_t}\}_{s_t \in S_t}, \{\tilde{U}_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}})$ satisfying DREU1 and DREU2 at $t+1$. For any $s_t \in S_t$, define $\rho_{t+1}^{s_t}$ by $\rho_{t+1}^{s_t}(p_{t+1}, A_{t+1}) := \sum_{s_{t+1}} \mu_{t+1}^{s_t}(s_{t+1}) \tau_{s_{t+1}}(p_{t+1}, A_{t+1})$ for all p_{t+1}, A_{t+1}.

C.2.4 Sophistication and Finiteness of Menu Preference

Before completing the representation, we establish two more lemmas. Using Axiom 7 (Sophistication), the first lemma ensures that for each s_t, $\rho_{t+1}^{s_t}$ and the preference over $\mathcal{A}_{t+1}$ induced by V_{s_t} satisfy Axioms 1 and 2 in Ahn and Sarver (2013).

Lemma 7. For any $s_t \in S_t$, separating history h^t for s_t, and $A_{t+1} \subseteq A'_{t+1} \in \mathcal{A}_{t+1}^*(h^t)$, the following are equivalent:

(i). $\rho_{t+1}^{s_t}(A'_{t+1} \smallsetminus A_{t+1}; A'_{t+1}) > 0$.

(ii). $V_{s_t}(A'_{t+1}) > V_{s_t}(A_{t+1})$

Proof. Pick any separating history h^t for s_t. Note that by DREU2 at $t+1$ and Lemma 16, we have $\rho_{t+1}(A'_{t+1} \smallsetminus A_{t+1}; A'_{t+1}|h^t) = \rho_{t+1}^{s_t}(A'_{t+1} \smallsetminus A_{t+1}; A'_{t+1})$. Moreover, by Corollary C.1 and Lemma 6 (i), $V_{s_t}(A'_{t+1}) > V_{s_t}(A_{t+1})$ if and only if $(z_t, A'_{t+1}) \succ_{h^t} (z_t, A_{t+1})$ for all z_t. By Axiom 7, this implies that $V_{s_t}(A'_{t+1}) > V_{s_t}(A_{t+1})$ if and only if $\rho_{t+1}^{s_t}(A'_{t+1} \smallsetminus A_{t+1}; A'_{t+1}) > 0$, as claimed. ■

The next lemma shows that because of Lemma 7, the finiteness of each $\text{supp}\mu_{t+1}^{s_t}$ is enough to ensure that the preference over $\mathcal{A}_{t+1}$ induced by each V_{s_t} satisfies Axiom DLR 6 (Finiteness) introduced by Ahn and Sarver (2013):

Lemma 8. For each $s_t \in S_t$, there is $K_{s_t} > 0$ such that for any A_{t+1}, there is $B_{t+1} \subseteq A_{t+1}$ such that $|B_{t+1}| \leq K_{s_t}$ and $V_{s_t}(A_{t+1}) = V_{s_t}(B_{t+1})$.

Proof. Fix any $s_t \in S_t$ and a separating history h^t for s_t. Let $S_{t+1}(s_t) := \text{supp}\mu_{t+1}^{s_t}$. We will show that $K_{s_t} := |S_{t+1}(s_t)|$ is as required.

Step 1: First consider any $A_{t+1} \in \mathcal{A}^*_{t+1}(h^t)$. Then by Lemma 14, for each $s_{t+1} \in S_{t+1}(s_t)$ we have $|M(A_{t+1}, \tilde{U}_{s_{t+1}})| = 1$. Letting $B_{t+1} := \bigcup_{s_{t+1} \in S_{t+1}(s_t)} M(A_{t+1}, \tilde{U}_{s_{t+1}})$, we then have that $|B_{t+1}| \leq K_{s_t}$ and $\rho^{s_t}_{t+1}(A_{t+1} \smallsetminus B_{t+1} | A_{t+1}) = 0$. By Lemma 7, this implies that $V_{s_t}(A_{t+1}) = V_{s_t}(B_{t+1})$, as required.

Step 2: Next take any $A_{t+1} \notin \mathcal{A}^*_{t+1}(h^t)$. By Lemma 17, we can find a sequence $A^n_{t+1} \to^m A_{t+1}$ with $A^n_{t+1} \in \mathcal{A}^*_{t+1}(h^t)$ for all n. Then by Step 1, we can find $B^n_{t+1} \subseteq A^n_{t+1}$ for all n such that $|B^n_{t+1}| \leq K_{s_t}$ and $V_{s_t}(A^n_{t+1}) = V_{s_t}(B^n_{t+1})$. By definition of $\to^m$, for each $q_{t+1} \in A_{t+1}$, there exists $D_{t+1}(q_{t+1}) \in \mathcal{A}_{t+1}$ and a sequence $\alpha_n(q_{t+1}) \to 0$ such that $B^n_{t+1} \subseteq \bigcup_{q_{t+1} \in A_{t+1}} \alpha_n(q_{t+1}) D_{t+1}(q_{t+1}) + (1 - \alpha_n(q_{t+1}))\{q_{t+1}\}$ for all n. Hence, since $|B^n_{t+1}| \leq K_{s_t}$ for all n, restricting to a subsequence if necessary, there is $B_{t+1} \subseteq A_{t+1}$ such that $B^n_{t+1} \to^m B_{t+1}$ and such that $|B_{t+1}| \leq K_{s_t}$. Finally, by continuity of V_{s_t} (Lemma 6 (ii)), we have $V_{s_t}(B_{t+1}) = V_{s_t}(A_{t+1})$, as required. ■

C.2.5 Completing the representation

Recall that in Section C.2.3, we have obtained $(S_{t+1}, \{\mu^{s_t}_{t+1}\}_{s_t \in S_t}, \{\tilde{U}_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}})$ satisfying DREU1 and DREU2 at $t+1$. We now show that for each $s_{t+1} \in S_{t+1}$ there exist $\alpha_{s_{t+1}} > 0$ and $\beta_{s_{t+1}} \in \mathbb{R}$ such that after replacing $\tilde{U}_{s_{t+1}}$ with $U_{s_{t+1}} := \alpha_{s_{t+1}} \tilde{U}_{s_{t+1}} + \beta_{s_{t+1}}$, we additionally have that EVU holds at time t.

Fix any s_t and let $S_{t+1}(s_t) := \text{supp}\,\mu^{s_t}_{t+1}$. Note that by DREU1 at $t+1$ and since we have defined $\rho^{s_t}_{t+1}$ by $\rho^{s_t}_{t+1}(p_{t+1}, A_{t+1}) := \sum_{s_{t+1} \in S_{t+1}(s_t)} \mu^{s_t}_{t+1}(s_{t+1}) \tau_{s_{t+1}}(p_{t+1}, A_{t+1})$ for all p_{t+1} and A_{t+1}, it follows that $(S_{t+1}(s_t), \mu^{s_t}_{t+1}, \{\tilde{U}_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1} \in S_{t+1}(s_t)})$ is an S-based REU representation of $\rho^{s_t}_{t+1}$ (see Definition 11).

Since all the $U_{s_{t+1}}$ are non-constant and induce different preferences over $\Delta(X_{t+1})$ for distinct $s_{t+1}, s'_{t+1} \in S_{t+1}(s_t)$ and since V_{s_t} is nonconstant by Lemma 6, we can find a finite set $Y \subseteq X_{t+1}$ such that (i) V_{s_t} is non-constant on $\mathcal{A}_{t+1}(Y) := \{B_{t+1} \in \mathcal{A}_{t+1} : \cup_{p_{t+1} \in B_{t+1}} \text{supp}(p_{t+1}) \subseteq Y\}$; (ii) for each $s_{t+1} \in S_{t+1}(s_t)$, $\tilde{U}_{s_{t+1}}$ is non-constant on Y; and (iii) for each distinct pair $s_{t+1}, s'_{t+1} \in S_{t+1}(s_t)$, $\tilde{U}_{s_{t+1}} \not\approx \tilde{U}_{s'_{t+1}}$ on Y.

Observe that by Lemmas 6 and 8, the preference $\succsim_{s_t}$ on $\mathcal{A}_{t+1}(Y)$ induced by V_{s_t} satisfies Axioms DLR 1–6 (Weak Order, Continuity, Independence, Monotonicity, Nontriviality, Finiteness) in Ahn and Sarver (2013) (henceforth AS), so by Corollary S1 in AS $\succsim_{s_t}$ admits a DLR representation (see Definition S1 in AS). Moreover, since $\rho^{s_t}_{t+1}$ admits an S-based REU representation (what AS call a GP representation), so does its restriction to $\mathcal{A}_{t+1}(Y)$. Finally, by Lemma 7, the pair $(\succsim_{s_t}, \rho^{s_t}_{t+1})$ satisfies AS's Axioms 1 and 2 on $\mathcal{A}_{t+1}(Y)$. Thus, by Theorem 1 in AS, we can find a DLR-GP representation of $(\succsim_{s_t}, \rho^{s_t}_{t+1})$ on $\mathcal{A}_{t+1}(Y)$, i.e., an S-based REU representation $(\hat{S}_{t+1}(s_t), \hat{\mu}^{s_t}_{t+1}, \{\hat{U}_{s_{t+1}}, \hat{\tau}_{s_{t+1}}\}_{s_{t+1} \in \hat{S}_{t+1}(s_t)})$ of $\rho^{s_t}_{t+1}$ on $\mathcal{A}_{t+1}(Y)$ such that $\succsim_{s_t}$ restricted to $\mathcal{A}_{t+1}(Y)$ is represented by $\hat{V}_{s_t}$, where $\hat{V}_{s_t}(A_{t+1}) := \sum_{s_{t+1} \in \hat{S}_{t+1}(s_t)} \hat{\mu}^{s_t}_{t+1}(s_{t+1}) \max_{p_{t+1} \in A_{t+1}} \hat{U}_{s_{t+1}}(p_{t+1})$. Since V_{s_t} also represents $\succsim_{s_t}$ restricted to $\mathcal{A}_{t+1}(Y)$, standard arguments yield $\hat{\alpha_{s_t}} > 0$ and $\hat{\beta}_{s_t} \in \mathbb{R}$ such that for all $A_{t+1} \in \mathcal{A}_{t+1}(Y)$, we have $V_{s_t}(A_{t+1}) = \hat{\alpha}_{s_t} \hat{V}_{s_t}(A_{t+1}) + \hat{\beta}_{s_t}$, whence

$$V_{s_t}(A_{t+1}) = \sum_{s_{t+1} \in \hat{S}_{t+1}(s_t)} \hat{\mu}^{s_t}_{t+1}(s_{t+1}) \max_{p_{t+1} \in A_{t+1}} U_{s_{t+1}}(p_{t+1}), \tag{22}$$

where $U_{s_{t+1}} = \hat{\alpha}_{s_t} \hat{U}_{s_{t+1}} + \hat{\beta}_{s_t}$. By the uniqueness properties of S-based REU representations (Proposition 4 in AS), $(\hat{S}_{t+1}(s_t), \hat{\mu}^{s_t}_{t+1}, \{U_{s_{t+1}}, \hat{\tau}_{s_{t+1}}\}_{s_{t+1} \in \hat{S}_{t+1}(s_t)})$ still constitutes an S-based REU representation of $\rho^{s_t}_{t+1}$ on $\mathcal{A}_{t+1}(Y)$. Applying Proposition 4 in AS again, since

$(S_{t+1}(s_t), \mu^{s_t}_{t+1}, \{\tilde{U}_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1}\in S_{t+1}(s_t)})$ also represents $\rho^{s_t}_{t+1}$ on $\mathcal{A}_{t+1}(Y)$, we can assume after relabeling that $S_{t+1}(s_t) = \hat{S}_{t+1}(s_t)$, $\hat{\mu}^{s_t}_{t+1} = \mu^{s_t}_{t+1}$ and that for each $s_{t+1} \in S_{t+1}(s_t)$, there exist constants $\alpha_{s_t+1} > 0$ and $\beta_{s_{t+1}} \in \mathbb{R}$ such that

$$U_{s_{t+1}}(x_{t+1}) = \alpha_{s_{t+1}} \tilde{U}_{s_{t+1}}(x_{t+1}) + \beta_{s_{t+1}} \tag{23}$$

for each $x_{t+1} \in Y \subseteq X_{t+1}$. Since $\tilde{U}_{s_{t+1}}$ is defined on X_{t+1}, we can extend $U_{s_{t+1}}$ to the whole space X_{t+1} by (23). Then $U_{s_{t+1}}$ and $\tilde{U}_{s_{t+1}}$ represent the same preference over $\Delta(X_{t+1})$, so since $(S_{t+1}(s_t), \mu^{s_t}_{t+1}, \{\tilde{U}_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1}\in S_{t+1}(s_t)})$ satisfies DREU1 and DREU2, so does $(S_{t+1}(s_t), \mu^{s_t}_{t+1}, \{U_{s_{t+1}}, \tau_{s_{t+1}}\}_{s_{t+1}\in S_{t+1}(s_t)})$.

It remains to show that (22) holds for all $A_{t+1} \in \mathcal{A}_{t+1}$, so that EVU is satisfied at s_t. To see this, consider any $A_{t+1} \in \mathcal{A}_{t+1}$ and choose a finite set $Y' \subseteq X_{t+1}$ such that $Y \cup \bigcup_{p_{t+1}\in A_{t+1}} \text{supp}(p_{t+1}) \subseteq Y'$. As above, we can again apply Theorem 1 in AS to obtain a DLR-GP representation $(\bar{S}_{t+1}(s_t), \bar{\mu}^{s'_t}_{t+1}, \{\bar{U}_{s_{t+1}}, \bar{\tau}_{s_{t+1}}\}_{s_{t+1}\in \bar{S}_{t+1}(s_t)})$ of the pair $(\succsim_{s_t}, \rho^{s_t}_{t+1})$ restricted to $\mathcal{A}_{t+1}(Y')$. But since this also yields a DLR-GP representation of $(\succsim_{s_t}, \rho^{s_t}_{t+1})$ restricted to $\mathcal{A}_{t+1}(Y)$, by the uniqueness property of DLR-GP representations (Theorem 2 in AS), we can assume that $\bar{S}_{t+1}(s_t) = S_{t+1}(s_t)$, $\bar{\mu}^{s_t}_{t+1} = \mu^{s_t}_{t+1}$ and that there exists $\bar{\alpha}_{s_t} > 0$ and $\bar{\beta}_{s_{t+1}} \in \mathbb{R}$ such that $\bar{U}_{s_{t+1}} = \bar{\alpha}_{s_t} U_{s_{t+1}} + \bar{\beta}_{s_{t+1}}$ for each $s_{t+1} \in S_{t+1}(s_t)$. Since $\succsim_{s_t}$ is represented on $\mathcal{A}_{t+1}(Y')$ by $\bar{V}_{s_t}(B_{t+1}) := \sum_{s_{t+1}\in S_{t+1}(s_t)} \mu^{s_t}_{t+1}(s_{t+1}) \max_{p_{t+1}\in B_{t+1}} \bar{U}_{s_{t+1}}(p_{t+1})$ and since $\bar{\alpha}_{s_t}$ depends only on s_t (and not on s_{t+1}), it follows that $\succsim_{s_t}$ is also represented on $\mathcal{A}_{t+1}(Y')$ by $V'_{s_t}(B_{t+1}) := \sum_{s_{t+1}\in S_{t+1}(s_t)} \mu^{s_t}_{t+1}(s_{t+1}) \max_{p_{t+1}\in B_{t+1}} U_{s_{t+1}}(p_{t+1})$. Thus, the linear functions V_{s_t} and V'_{s_t} represent the same preference on $\mathcal{A}_{t+1}(Y')$ and coincide on $\mathcal{A}_{t+1}(Y)$, so they must also coincide on $\mathcal{A}_{t+1}(Y')$. Thus, (22) holds at A_{t+1}.

This shows that EVU holds at t. Combining this with the inductive hypothesis, it follows that $(S_{t'}, \{\mu^{s_{t'-1}}_{t'}\}_{s_{t'-1}\in S_{t'-1}}, \{U_{s_{t'}}, \tau_{s_{t'}}\}_{s_{t'}\in S_{t'}})$ satisfies DREU1 and DREU2 for all $t' \leq t+1$ and EVU for all $t' \leq t$, as required.

C.3 Proof of Theorem 2: Necessity

Suppose that ρ admits an evolving utility representation. Then by Proposition 5, ρ admits an S-based evolving utility representation $(S_t, \{\mu^{s_{t-1}}_t\}_{s_{t-1}\in S_{t-1}}, \{U_{s_t}, u_{s_t}, \tau_{s_t}\}_{s_t\in S_t})$.

We first show that for every $t \leq T-1$, there exist $g_t, b_t \in \Delta(X_t)$ such that $U_{s_t}(g_t) > U_{s_t}(b_t)$ for all $s_t \in S_t$. By separability of U_{s_t}, it is sufficient to find menus C'_{t+1}, C_{t+1} such that $V_{s_t}(C'_{t+1}) > V_{s_t}(C_{t+1})$ for all s_t. Note first that for any $s_{t+1} \in S_{t+1}$, since $U_{s_{t+1}}$ is nonconstant, we can find $g_{t+1}(s_{t+1}), b_{t+1}(s_{t+1}) \in \Delta(X_{t+1})$ such that $U_{s_{t+1}}(g_{t+1}(s_{t+1})) > U_{s_{t+1}}(b_{t+1}(s_{t+1}))$. Let $C'_{t+1} := \{g_{t+1}(s_{t+1}), b_{t+1}(s_{t+1}) : s_{t+1} \in S_{t+1}\}$, and for every s_t, let $A_{t+1}(s_t) := \{b_{t+1}(s_{t+1})\}$ for some $s_{t+1} \in \text{supp}\mu^{s_t}_{t+1}$. Then $V_{s_t}(C'_{t+1}) \geq V_{s_t}(A_{t+1}(s'_t))$ for all s_t, s'_t, with strict inequality for $s_t = s'_t$. Hence, letting $C_{t+1} := \sum_{s_t\in S_t} \frac{1}{|S_t|} A_{t+1}(s_t)$, linearity implies $V_{s_t}(C'_{t+1}) > V_{s_t}(C_{t+1})$ for all s_t, as required.

By Lemma 5, the previous paragraph implies that for all $t \leq T-1$, h^t and q_t, r_t, we have $q_t \succsim_{h^t} r_t$ if and only if $U_{s_t}(q_t) \geq U_{s_t}(r_t)$ for all s_t consistent with h^t. Axioms 5 (Separability) and 6 (i)–(ii) (Monotonicity and Indifference to Timing) are then straightforward to verify from the representation. Moreover, C'_{t+1} and C_{t+1} from the previous paragraph satisfy $(z_t, C'_{t+1}) \succ_{h^t} (z_t, C_{t+1})$ for all h^t and z_t, implying Axiom 6 (iv) (Nondegeneracy).

To show Axiom 7 (Sophistication), consider any $t \leq T-1$, h^t, z_t, and $A_{t+1} \subseteq A'_{t+1} \in \mathcal{A}^*(h^t)$. Since $A'_{t+1} \in \mathcal{A}^*(h^t)$, Lemma 14 implies that $\rho_{t+1}(A'_{t+1} \smallsetminus A_{t+1}; A'_{t+1}|h^t) > 0$ holds if and only if there exists some s_t consistent with h^t such that $\max_{p_{t+1}\in A'_{t+1}} U_{s_{t+1}}(p_{t+1}) > \max_{p_{t+1}\in A_{t+1}} U_{s_{t+1}}(p_{t+1})$ for some $s_{t+1} \in \text{supp}\mu^{s_t}_{t+1}$, which by the representation is equivalent to $V_{s_t}(A'_{t+1}) > V_{s_t}(A_{t+1})$. By Lemma 5, this is equivalent to $(z_t, A_{t+1}) \not\succsim_{h^t} (z_t, A'_{t+1})$, which by Monotonicity is in turn equivalent to $(z_t, A'_{t+1}) \succ_{h^t} (z_t, A_{t+1})$.

Finally, to show Axiom 6 (iii) (Continuity), note first that for each $s_{T-1} \in S_{T-1}$, $\sum_{s_T \in S_T} \mu_T^{s_{T-1}}(s_T) \max_{p_T \in A_T} U_{s_T}(p_T)$ is continuous in menu A_T. Assuming inductively that for each $k \geq t+1$ and $s_k \in S_k$, $\sum_{s_{k+1} \in S_{k+1}} \mu_{k+1}^{s_k}(s_{k+1}) \max_{p_{k+1} \in A_{k+1}} U_{s_{k+1}}(p_{k+1})$ is continuous in menu A_{k+1}, it also follows that for each t and $s_t \in S_t$, $\sum_{s_{t+1} \in S_{t+1}} \mu_{t+1}^{s_t}(s_{t+1}) \max_{p_{t+1} \in A_{t+1}} U_{s_{t+1}}(p_{t+1})$ is continuous in menu A_{t+1}. Thus for each s_t, $U_{s_t}(p_t)$ is continuous in p_t. Then Continuity follows as, for each p_t, $\{q_t : q_t \succsim_{h^t} p_t\} = \cap_{s_t:\text{consistent with } h^t}\{q_t : U_{s_t}(q_t) \geq U_{s_t}(p_t)\}$ and $\{q_t : p_t \succsim_{h^t} q_t\} = \cap_{s_t:\text{consistent with } h^t}\{q_t : U_{s_t}(p_t) \geq U_{s_t}(q_t)\}$ are closed.

D Proof of Theorem 3

D.1 Proof of Theorem 3: Sufficiency

Suppose that ρ admits an evolving utility representation and that Condition 1 and Axioms 8 (Stationary Consumption Preference) and 9 (Constant Intertemporal Tradeoff) hold. By Proposition 5, ρ admits an S-based evolving utility representation $(S_t, \{\mu_t^{s_{t-1}}\}_{s_{t-1} \in S_{t-1}}, \{U_{s_t}, u_{s_t}, \tau_{s_t}\}_{s_t \in S_t})_{t=0,...,T}$. Up to adding appropriate constants to each utility u_{s_t} and U_{s_t}, we can ensure that $\sum_{z \in Z} u_{s_t}(z) = 0$ for all $t = 0, ..., T$ and $s_t \in S_t$ without affecting that $(S_t, \{\mu_t^{s_{t-1}}\}_{s_{t-1} \in S_{t-1}}, \{U_{s_t}, u_{s_t}, \tau_{s_t}\}_{s_t \in S_t})_{t=0,...,T}$ is an S-based evolving utility representation of ρ. We will show that this representation is in fact an S-based gradual learning representation, i.e., that there exists a discount factor $\delta \in (0,1)$ such that for all $t \leq T-1$ and s_t, we have $u_{s_t} = \frac{1}{\delta} \sum_{s_{t+1}} \mu_{t+1}^{s_t}(s_{t+1}) u_{s_{t+1}}$. By Proposition 5, this implies that ρ admits a gradual learning representation.

Condition 1 implies that each u_{s_t} is nonconstant:

Lemma 9. For each $t = 0, .., T-1$ and $s_t \in S_t$, there exist $\ell, m \in \Delta(Z)$ such that $u_{s_t}(\ell) \neq u_{s_t}(m)$.

Proof. Consider any $t = 0, \dots, T-1$, $s_t \in S_t$ and separating history h^t for s_t. By Condition 1, there exist $\ell, m, n \in \Delta(Z)$ such that $(\ell, n, \dots, n) \not\sim_{h^t} (m, n, \dots, n)$. Then Lemma 5 (iii) implies that $U_{s_t}((\ell, n, \dots, n)) \neq U_{s_t}((m, n, \dots, n))$, whence $u_{s_t}(\ell) \neq u_{s_t}(m)$, as required. ∎

For any $t = 0, \dots, T-1$ and $s_t \in S_t$ and $\ell \in \Delta(Z)$, let

$$\mathbb{E}[u_{t+1}(\ell)|s_t] := \sum_{s_{t+1}} \mu_{t+1}^{s_t}(s_{t+1}) u_{s_{t+1}}(\ell)$$

denote the expected period $t+1$ felicity of ℓ at state s_t. Stationary Consumption Preference implies that u_{s_t} and $\mathbb{E}[u_{t+1}|s_t]$ induce the same preference over $\Delta(Z)$:

Lemma 10. For all $\ell, m \in \Delta(Z)$, $t = 0, ..., T-1$, and $s_t \in S_t$,

$$\mathbb{E}[u_{t+1}(\ell)|s_t] > \mathbb{E}[u_{t+1}(m)|s_t] \iff u_{s_t}(\ell) > u_{s_t}(m).$$

Proof. Fix any $\ell, m, n \in \Delta(Z)$, $t = 0, ..., T-1$, $s_t \in S_t$ and separating history h^t for s_t. Note that $u_{s_t}(\ell) > u_{s_t}(m)$ if and only if $U_{s_t}((\ell, n, \dots, n)) > U_{s_t}((m, n, \dots, n)$, which by Lemma 5 (iii) is in turn equivalent to $(\ell, n, \dots, n) \succ_{h^t} (m, n, \dots, n)$. Likewise, $\mathbb{E}[u_{t+1}(\ell)|s_t] > \mathbb{E}[u_{t+1}(m)|s_t]$ if and only if $U_{s_t}((n, \ell, n, \dots, n)) > U_{s_t}((n, m, n, \dots, n))$, which by Lemma 5 (iii) is equivalent to $(n, \ell, n, \dots, n) \succ_{h^t} (n, m, n, \dots, n)$. Thus, the claim is immediate from Axiom 8. ∎

Given Lemma 10, Constant Intertemporal Tradeoff now allows us to obtain a time-invariant and non-random discount factor $\delta > 0$.

Lemma 11. There exists $\delta \in (0,1)$ such that for all $t = 0, \dots, T-1$ and $s_t \in S_t$, we have $u_{s_t} = \frac{1}{\delta}\mathbb{E}[u_{t+1}|s_t]$.

Proof. Fix any $t, \hat{t} \le T-1$, $s_t \in S_t$, $\hat{s}_{\hat{t}} \in S_{\hat{t}}$, and separating histories h^t for s_t and $\hat{h}^{\hat{t}}$ for $\hat{s}_{\hat{t}}$. By Lemma 10, u_{s_t} and $\mathbb{E}[u_{t+1}|s_t]$ induce the same preference over $\Delta(Z)$, and moreover, u_{s_t} is nonconstant by Lemma 9. Hence, there exist constants $\gamma_{s_t} > 0, \beta_{s_t} \in \mathbb{R}$ such that $u_{s_t} = \gamma_{s_t}\mathbb{E}[u_{t+1}|s_t] + \beta_{s_t}$. Since we have normalized felicities such that $\sum_{z \in Z} u_{s_{t'}}(z) = 0$ for any t' and $s_{t'}$, we must have $\beta_{s_t} = 0$. Similarly, there exists $\hat{\gamma}_{\hat{s}_{\hat{t}}} > 0$ such that $u_{\hat{s}_{\hat{t}}} = \hat{\gamma}_{\hat{s}_{\hat{t}}}\mathbb{E}[u_{t+1}|\hat{s}_{\hat{t}}]$.

Let $\delta_{s_t} := \frac{1}{\gamma_{s_t}}$ and $\hat{\delta}_{\hat{s}_{\hat{t}}} := \frac{1}{\hat{\gamma}_{\hat{s}_{\hat{t}}}}$. We first show that $\delta_{s_t} = \hat{\delta}_{\hat{s}_{\hat{t}}}$. By Condition 1, there exist h^t-nonindifferent $\ell, m \in \Delta(Z)$ and $\hat{h}^{\hat{t}}$-nonindifferent $\hat{\ell}, \hat{m} \in \Delta(Z)$. For any $\alpha \in (0,1)$ and $n \in \Delta(Z)$, Lemma 5 (iii) along with the above implies

$$\begin{aligned}(\alpha\ell + (1-\alpha)m, \alpha\ell + (1-\alpha)m, n, \dots, n) &\sim_{h^t} (\ell, m, n, \dots, n)\\ &\iff\\ (1+\delta_{s_t})(\alpha u_{s_t}(\ell) + (1-\alpha)u_{s_t}(m)) &= u_{s_t}(\ell) + \delta_{s_t}u_{s_t}(m)\\ &\iff\\ \alpha &= \frac{1}{1+\delta_{s_t}},\end{aligned}$$

where the final equivalence holds because $u_{s_t}(\ell) \neq u_{s_t}(m)$ (Lemma 9). Likewise, we have $(\alpha\hat{\ell} + (1-\alpha)\hat{m}, \alpha\hat{\ell} + (1-\alpha)\hat{m}, n, \dots, n) \sim_{\hat{h}^{\hat{t}}} (\hat{\ell}, \hat{m}, n, \dots, n)$ if and only if $\alpha = \frac{1}{1+\hat{\delta}_{\hat{s}_{\hat{t}}}}$. Since by Axiom 9, we have $(\alpha\ell + (1-\alpha)m, \alpha\ell + (1-\alpha)m, n, \dots, n) \sim_{h^t} (\ell, m, n, \dots, n)$ if and only if $(\alpha\hat{\ell} + (1-\alpha)\hat{m}, \alpha\hat{\ell} + (1-\alpha)\hat{m}, n, \dots, n) \sim_{\hat{h}^{\hat{t}}} (\hat{\ell}, \hat{m}, n, \dots, n)$, this implies $\delta_{s_t} = \hat{\delta}_{\hat{s}_{\hat{t}}} =: \delta$.

■

This completes the proof that ρ admits an S-based gradual learning representation.

D.2 Proof of Theorem 3: Necessity

Suppose that ρ admits a gradual learning representation and Condition 1 holds. By Proposition 5, ρ admits an S-based gradual learning representation $(S_t, \{\mu_t^{s_{t-1}}\}_{s_{t-1} \in S_{t-1}}, \{U_{s_t}, u_{s_t}, \tau_{s_t}\}_{s_t \in S_t})_{t=0,\dots,T}$ with discount factor $\delta > 0$.

The same argument as in the proof of the necessity direction of Theorem 2 shows that for all $t \le T-1$, h^t and $q_t, r_t \in \Delta(X_t)$, we have $q_t \succsim_{h^t} r_t$ if and only if $U_{s_t}(q_t) \ge U_{s_t}(r_t)$ for all s_t consistent with h^t.

Given this, Axiom 8 is equivalent to the statement that for all s_t, u_{s_t} and $\mathbb{E}[u_{t+1}|s_t]$ represent the same preference over $\Delta(Z)$. But this is immediate from the fact that for all s_t, we have $u_{s_t} = \frac{1}{\delta}\mathbb{E}[u_{t+1}|s_t]$.

Finally, to establish Axiom 9, consider any $t \le T-1$, h^t, and h^t-nonindifferent $\ell, m \in \Delta(Z)$. By the second paragraph, for any $\alpha \in [0,1]$ and $n \in \Delta(Z)$, we have $(\alpha\ell + (1-\alpha)m, \alpha\ell + (1-\alpha)m, n, \dots, n) \sim_{h^t} (\ell, m, n, \dots, n)$ if and only if $U_{s_t}((\alpha\ell + (1-\alpha)m, \alpha\ell + (1-\alpha)m, n, \dots, n)) = U_{s_t}((\ell, m, n, \dots, n))$ for all s_t consistent with h^t. Since $u_{s_t} = \frac{1}{\delta}\mathbb{E}[u_{t+1}|s_t]$, this is equivalent to

$$(1+\delta)(\alpha(u_{s_t}(\ell) + (1-\alpha)u_{s_t}(m)) = u_{s_t}(\ell) + \delta u_{s_t}(m) \text{ for all } s_t \text{ consistent with } h^t. \tag{24}$$

But since ℓ, m are h^t-nonindifferent, there is some s_t^* consistent with h^t such that $u_{s_t^*}(\ell) \neq u_{s_t^*}(m)$, whence (24) is equivalent to $\alpha = \frac{1}{1+\delta}$. Since this holds for all h^t and h^t-nonindifferent ℓ, m, this establishes Axiom 9.

E Additional Lemmas

Lemma 12. For all $t = 0, \dots, T$, X_t is a separable metric space, where $X_T := Z$ is endowed with the discrete metric and for all $t \leq T-1$, we recursively endow $\Delta(X_{t+1})$ with the induced topology of weak convergence, $\mathcal{A}_{t+1} := \mathcal{K}(\Delta(X_{t+1}))$ with the induced Hausdorff topology, and $X_t := Z \times \mathcal{A}_{t+1}$ with the induced product topology.

Proof. 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. ■

Lemma 13. Let Y be any set (possibly infinite) and let $\{U_s : s \in S\} \subseteq \mathbb{R}^Y$ be a collection of nonconstant vNM utility functions indexed by a finite set S such that $U_s \not\approx U_{s'}$ for any distinct $s, s' \in S$. Then there is a collection of lotteries $\{p^s : s \in S\} \subseteq \Delta(Y)$ such that $U_s(p^s) > U_s(p^{s'})$ for any distinct $s, s' \in S$.

Proof. 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'. ■

Lemma 14. Fix $t = 0, \dots, T$. Suppose $(S_{t'}, \{\mu_{t'}^{s_{t'-1}}\}_{s_{t'-1} \in S_{t'-1}}, \{U_{s_{t'}}, \tau_{s_{t'}}\}_{s_{t'} \in S_{t'}})$ satisfy DREU1 and DREU2 for all $t' \leq t$. Take any $h^{t-1} \in \mathcal{H}_{t-1}$ and let $S(h^{t-1}) \subseteq S_{t-1}$ denote the set of states consistent with h^{t-1}. Then for any $A_t \in \mathcal{A}_t$, the following are equivalent:

(i). $A_t \in \mathcal{A}_t^*(h^{t-1})$

(ii). For each $s_{t-1} \in S(h^{t-1})$ and $s_t \in \operatorname{supp} \mu_t^{s_{t-1}}$, $|M(A_t, U_{s_t})| = 1$.

Proof.
(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 $q_t \in A_t \setminus \{p_t\}$, 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 $\underline{B}_t^n := \{\underline{q}_t^n : q_t \in A_t \setminus \{p_t\}\}$, let $\overline{B}_t^n := \{\overline{q}_t^n : q_t \in A_t \setminus \{p_t\}\}$, and let $B_t^n := \underline{B}_t^n \cup \overline{B}_t^n$. Then $B_t^n \to^m A_t \setminus \{p_t\}$ and $p_t^n \to^m p_t$.

Moreover, since $|M(A_t, U_{s_t})| > 1$, there exists $q_t \in A_t \setminus \{p_t\}$ 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 $s_t' \in S_t \setminus \{s_t\}$, 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 $\mathrm{pred}(s_{t-1}) = (s_0, \ldots, s_{t-2})$, Lemma 16 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) $\Longrightarrow$ (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 16 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. ■

Lemma 15. Suppose that ρ satisfies Axiom 2. Fix $t \geq 1$, $A_t \in \mathcal{A}_t$, $h^{t-1} = (A_0, p_0, \ldots, A_{t-1}, p_{t-1}) \in \mathcal{H}_{t-1}$, and $\lambda = (\lambda_n)_{n=0}^{t-1}$, $\hat{\lambda} = (\hat{\lambda}_n)_{n=0}^{t-1} \in (0,1]^t$. Suppose $d^{t-1} = (\{q_n\}, q_n)_{n=0}^{t-1}$, $\hat{d}^{t-1} = (\{\hat{q}_n\}, \hat{q}_n)_{n=0}^{t-1} \in \mathcal{D}_{t-1}$ satisfy $\lambda h^{t-1} + (1-\lambda)d^{t-1}, \hat{\lambda} h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1} \in \mathcal{H}_{t-1}(A_t)$, where $\lambda h^{t-1} + (1-\lambda)d^{t-1} := (\lambda_n A_n + (1-\lambda_n)\{q_n\}, \lambda_n p_n + (1-\lambda_n)q_n)_{n=0}^{t-1}$ and $\hat{\lambda} h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}$ is defined analogously. Then

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

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

Proof. 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 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 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

$$\begin{aligned} &\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}, \frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1} \in \mathcal{H}_{t-1}(A_t); \\ &\frac{2}{3}B_m + \frac{1}{3}\{\hat{q}_m\}, \{\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m\} \in \operatorname{supp} q^A_{m-1}; \\ &\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \{\frac{1}{2}q_m + \frac{1}{2}\hat{q}_m\} \in \operatorname{supp} \hat{q}^A_{m-1}, \end{aligned} \tag{25}$$

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 $n = 1, \ldots, t-1$

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

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 2 implies

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

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 (25) is satisfied.

Given (25), $\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}, \frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1} \in \mathcal{H}_{t-1}(A_t)$, so the base case of the proof implies

$$\begin{aligned} \rho_t(\cdot; A_t|\lambda h^{t-1} + (1-\lambda)d^{t-1}) &= \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}); \\ \rho_t(\cdot; A_t|\hat{\lambda} h^{t-1} + (1-\hat{\lambda})\hat{d}^{t-1}) &= \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1}). \end{aligned} \tag{26}$$

Also, (25) 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 2

$$\begin{aligned} \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1}) &= \rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1})_{-m}, (\frac{2}{3}B_m + \frac{1}{3}\{\hat{q}_m\}, \frac{2}{3}r_m + \frac{1}{3}\hat{q}_m)); \\ \rho_t(\cdot; A_t|\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1}) &= \rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1})_{-m}, (\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \frac{2}{3}\hat{r}_m + \frac{1}{3}q_m)). \end{aligned} \tag{27}$$

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

$$\begin{aligned}\rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}d^{t-1})_{-m}, (\frac{2}{3}B_m + \frac{1}{3}\{q_m\}, \frac{2}{3}r_m + \frac{1}{3}q_m)) = \\ \rho_t(\cdot; A_t|(\frac{1}{2}h^{t-1} + \frac{1}{2}\hat{d}^{t-1})_{-m}, (\frac{2}{3}\hat{B}_m + \frac{1}{3}\{q_m\}, \frac{2}{3}\hat{r}_m + \frac{1}{3}q_m)).\end{aligned} \tag{28}$$

Combining (26), (27), and (28) 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. ■

Lemma 16. Fix $t = 0, \dots, T$. Suppose $(S_{t'}, \{\mu_{t'}^{s_{t'-1}}\}_{s_{t'-1} \in S_{t'-1}}, \{U_{s_{t'}}, \tau_{s_{t'}}\}_{s_{t'} \in S_{t'}})$ satisfy DREU1 and DREU2 for all $t' \leq t$. Then the extended version of ρ from Definition 3 also satisfies DREU2 for all $t' \leq t$, i.e., for all $p_{t'}, A_{t'}$, and $h^{t'-1} = (A_0, p_0, \dots, A_{t'-1}, p_{t'-1}) \in \mathcal{H}_{t'-1}$,[57] we have

$$\rho_{t'}(p_{t'}, A_{t'}|h^{t'-1}) = \frac{\sum_{(s_0,\dots,s_{t'}) \in S_0 \times \dots \times S_{t'}} \prod_{k=0}^{t'} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)}{\sum_{(s_0,\dots,s_{t'-1}) \in S_0 \times \dots \times S_{t'-1}} \prod_{k=0}^{t-1} \mu_k^{s_{k-1}}(s_k)\tau_{s_k}(p_k, A_k)}.$$

Proof. 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 3 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 $\lambda h^{t'-1} + (1-\lambda)d^{t'-1} \in \mathcal{H}_{t'-1}(A_{t'})$. ■

Lemma 17. Fix $t = 0, \dots, T$. Suppose $(S_{t'}, \{\mu_{t'}^{s_{t'-1}}\}_{s_{t'-1} \in S_{t'-1}}, \{U_{s_{t'}}, \tau_{s_{t'}}\}_{s_{t'} \in S_{t'}})$ satisfy DREU1 and DREU2 for all $t' \leq t$. Fix any $s_{t-1} \in S_{t-1}$, separating history h^{t-1} for s_{t-1}, and $A_t \in \mathcal{A}_t$. Then there exists a sequence $A_t^n \to^m A_t$ such that $A_t^n \in \mathcal{A}_{t+1}^*(h^t)$ for all n. Moreover, given any $s_t^* \in \text{supp}\mu_t^{s_{t-1}}$ and $p_t^* \in M(A_t, U_{s_t^*})$, we can ensure in this construction that there is $p_t^n(s_t^*) \in A_t^n$ with $p_t^n(s_t^*) \to^m p_t^*$ such that $\mathcal{U}_{s_t}(A_t^n, p_t^n(s_t^*)) = \{U_{s_t^*}\}$ for all n.

Proof. Let $S_t(s_{t-1}) := \text{supp}\mu_t^{s_{t-1}}$. 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) $\bigcup_{p_t \in A_t} \text{supp} p_t \subseteq Y_t$. By (i) and (ii) and Lemma 13, 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 $z^{s_t} \in \text{argmax}_{y \in Y} U_{s_t}$ 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

[57] For $t' = 0$, we abuse notation by letting $\rho_{t'}(\cdot|h^{t'-1})$ denote $\rho_0(\cdot)$ for all $h^{t'-1}$.

$A_t^n := \{p_t^n(s_t) : s_t \in S_t(s_{t-1})\} \cup \{r_t^n : r_t \in A_t\}$. 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 14 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. ■

References

Abdulkadiroglu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (forthcoming): "Research design meets market design: Using centralized assignment for impact evaluation," *Econometrica*.

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1991): "Optimal learning by experimentation," *The review of economic studies*, 58(4), 621–654.

Ahn, D. S., and T. Sarver (2013): "Preference for flexibility and random choice," *Econometrica*, 81(1), 341–361.

Aliprantis, C. D., and K. C. Border (2006): *Infinite Dimensional Analysis: a Hitchhiker's Guide*. Springer, Berlin; London.

Angrist, J., P. Hull, P. A. Pathak, and C. Walters (forthcoming): "Leveraging lotteries for school value-added: Testing and estimation," *Quarterly Journal of Economics*.

Anscombe, F. J., and R. J. Aumann (1963): "A definition of subjective probability," *The annals of mathematical statistics*, 34(1), 199–205.

Apesteguia, J., M. Ballester, and J. Lu (2017): "Single-Crossing Random Utility Models," *Econometrica*.

Apesteguia, J., and M. A. Ballester (2017): "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences," *Journal of Political Economy*.

Augenblick, N., M. Niederle, and C. Sprenger (2015): "Working over time: Dynamic inconsistency in real effort tasks," *The Quarterly Journal of Economics*, p. qjv020.

Barberá, S., and P. Pattanaik (1986): "Falmagne and the rationalizability of stochastic choices in terms of random orderings," *Econometrica*, pp. 707–715.

Becker, G. S., and C. B. Mulligan (1997): "The endogenous determination of time preference," *The Quarterly Journal of Economics*, 112(3), 729–758.

Becker, G. S., and K. M. Murphy (1988): "A theory of rational addiction," *Journal of political Economy*, 96(4), 675–700.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, pp. 841–890.

Block, D., and J. Marschak (1960): "Random Orderings And Stochastic Theories of Responses," in *Contributions To Probability And Statistics*, ed. by I. O. et al. Stanford: Stanford University Press.

Cooke, K. (2016): "Preference discovery and experimentation," *Theoretical Economics*.

Dekel, E., B. Lipman, and A. Rustichini (2001): "Representing preferences with a unique subjective state space," *Econometrica*, 69(4), 891–934.

Dekel, E., B. L. Lipman, A. Rustichini, and T. Sarver (2007): "Representing Preferences with a Unique Subjective State Space: A Corrigendum1," *Econometrica*, 75(2), 591–600.

Deming, D. J. (2011): "Better schools, less crime?," *The Quarterly Journal of Economics*, p.

qjr036.

Deming, D. J., J. S. Hastings, T. J. Kane, and D. O. Staiger (2014): "School choice, school quality, and postsecondary attainment," *The American economic review*, 104(3), 991–1013.

Dillenberger, D., J. S. Lleras, P. Sadowski, and N. Takeoka (2014): "A theory of subjective learning," *Journal of Economic Theory*, 153, 287–312.

Easley, D., and N. M. Kiefer (1988): "Controlling a stochastic process with unknown parameters," *Econometrica: Journal of the Econometric Society*, pp. 1045–1064.

Epstein, L. G. (1999): "A definition of uncertainty aversion," *The Review of Economic Studies*, 66(3), 579–608.

Erdem, T., and M. P. Keane (1996): "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets," *Marketing science*, 15(1), 1–20.

Falmagne, J. (1978): "A representation theorem for finite random scale systems," *Journal of Mathematical Psychology*, 18(1), 52–72.

Fishburn, P. (1970): *Utility theory for decision making.*

Fishburn, P. C. (1984): "On Harsanyi's Utilitarian Cardinal Welfare Therem," *Theory and Decision*, 17, 21–28.

Fudenberg, D., P. Strack, and T. Strzalecki (2016): "Speed, Accuracy, and the Optimal Timing of Choices," .

Fudenberg, D., and T. Strzalecki (2015): "Dynamic logit with choice aversion," *Econometrica*, 83(2), 651–691.

Gilboa, I. (1987): "Expected utility with purely subjective non-additive probabilities," *Journal of mathematical Economics*, 16(1), 65–88.

Gilboa, I., and A. Pazgal (2001): "Cumulative discrete choice," *Marketing Letters*, 12(2), 119–130.

Gilboa, I., A. Postlewaite, and L. Samuelson (2016): "Memorable consumption," *Journal of Economic Theory*, 165, 414–455.

Gittins, J. C., and D. M. Jones (1972): *A dynamic allocation index for the sequential design of experiments*. University of Cambridge, Department of Engineering.

Gowrisankaran, G., and M. Rysman (2012): "Dynamics of consumer demand for new durable goods," *Journal of political Economy*, 120(6), 1173–1219.

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," *Econometrica*, 82(5), 1873–1912.

Gul, F., and W. Pesendorfer (2001): "Temptation and self-control," *Econometrica*, 69(6), 1403–1435.

Gul, F., and W. Pesendorfer (2004): "Self-control and the theory of consumption," *Econometrica*, 72(1), 119–158.

Gul, F., and W. Pesendorfer (2006): "Random expected utility," *Econometrica*, 74(1), 121–146.

Gul, F., and W. Pesendorfer (2007): "Harmful addiction," *The Review of Economic Studies*, 74(1), 147–172.

Hausman, J., and D. McFadden (1984): "Specification tests for the multinomial logit model," *Econometrica: Journal of the Econometric Society*, pp. 1219–1240.

Heckman, J. J. (1981): "Heterogeneity and state dependence," in *Studies in labor markets*, pp. 91–140. University of Chicago Press.

Hendel, I., and A. Nevo (2006): "Measuring the implications of sales and consumer inventory behavior," *Econometrica*, 74(6), 1637–1673.

Higashi, Y., K. Hyogo, and N. Takeoka (2014): "Stochastic endogenous time preference," *Journal of Mathematical Economics*, 51, 77–92.

Hu, Y., and M. Shum (2012): "Nonparametric identification of dynamic models with unobserved state variables," *Journal of Econometrics*, 171(1), 32–44.

Hyogo, K. (2007): "A subjective model of experimentation," *Journal of Economic Theory*, 133(1), 316–330.

Kasahara, H., and K. Shimotsu (2009): "Nonparametric identification of finite mixture models of dynamic discrete choices," *Econometrica*, 77(1), 135–175.

Ke, S. (2017): "A Dynamic Model of Mistakes," working paper.

Kennan, J., and J. Walker (2011): "The effect of expected income on individual migration decisions," *Econometrica*, 79(1), 211–251.

Khamsi, M. A., and W. A. Kirk (2011): *An introduction to metric spaces and fixed point theory*, vol. 53. John Wiley & Sons.

Kitamura, Y., and J. Stoye (2016): "Nonparametric analysis of random utility models: testing," .

Kreps, D. (1979): "A representation theorem for" preference for flexibility"," *Econometrica*, pp. 565–577.

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

Krishna, R. V., and P. Sadowski (2014): "Dynamic preference for flexibility," *Econometrica*, 82(2), 655–703.

Krishna, V., and P. Sadowski (2016): "Randomly Evolving Tastes and Delayed Commitment," *mimeo*.

Lu, J. (2016): "Random choice and private information," *Econometrica*, 84(6), 1983–2027.

Lu, J., and K. Saito (2016): "Random intertemporal choice," Discussion paper, mimeo.

Luce, D. (1959): *Individual choice behavior*. John Wiley.

Magnac, T., and D. Thesmar (2002): "Identifying dynamic discrete decision processes," *Econometrica*, 70(2), 801–816.

Manski, C. F. (1993): "Dynamic choice in social settings: Learning from the experiences of others," *Journal of Econometrics*, 58(1-2), 121–136.

McAlister, L. (1982): "A dynamic attribute satiation model of variety-seeking behavior," *Journal of Consumer Research*, 9(2), 141–150.

McFadden, D., and M. Richter (1990): "Stochastic rationality and revealed stochastic preference," *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz*, pp. 161–186.

Miller, R. A. (1984): "Job matching and occupational choice," *The Journal of Political Economy*, pp. 1086–1120.

Noor, J. (2011): "Temptation and Revealed Preference," *Econometrica*, 79(2), 601–644.

Norets, A. (2009): "Inference in dynamic discrete choice models with serially orrelated unobserved state variables," *Econometrica*, 77(5), 1665–1682.

Norets, A., and X. Tang (2013): "Semiparametric Inference in dynamic binary choice models," *The Review of Economic Studies*, p. rdt050.

Pakes, A. (1986): "Patents as options: Some estimates of the value of holding European patent stocks," *Econometrica*, 54, 755–784.

Pennesi, D. (2017): "Intertemporal discrete choice," working paper.

Piermont, E., N. Takeoka, and R. Teper (2016): "Learning the Krepsian state: Exploration through consumption," *Games and Economic Behavior*, 100, 69–94.

Rao, K. P. S. B., and M. B. Rao (2012): *Theory of Charges: A Study of Finitely Additive Measures*. Academic Press.

Robbins, H. (1952): "Some aspects of the sequential design of experiments," *Bulletin of the American Mathematical Society*, 58(5), 527–535.

Rozen, K. (2010): "Foundations of intrinsic habit formation," *Econometrica*, 78(4), 1341–1373.

Rust, J. (1989): "A Dynamic Programming Model of Retirement Behavior," in *The Economics of Aging*, ed. by D. Wise, pp. 359–398. University of Chicago Press: Chicago.

——— (1994): "Structural estimation of Markov decision processes," *Handbook of econometrics*, 4, 3081–3143.

Rustichini, A., and P. Siconolfi (2014): "Dynamic theory of preferences: Habit formation and taste for variety," *Journal of Mathematical Economics*, 55, 55–68.

Savage, L. J. (1972): *The foundations of statistics*. Courier Corporation.

Sweeting, A. (2013): "Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81(5), 1763–1803.

Toussaert, S. (2016): "Eliciting temptation and self-control through menu choices: a lab experiment on curiosity," *mimeo.*

Uzawa, H. (1968): "Time preference, the consumption function, and optimum asset holdings," *Value, capital and growth: papers in honor of Sir John Hicks. The University of Edinburgh Press, Edinburgh*, pp. 485–504.