A Repeated-Choice Foundation for Stochastic Choice*

Jay Lu† Kota Saito‡

June 2025

Abstract

We characterize when an agent's observed stochastic choice can be represented as a limit frequency of optimal choices over time. In our model, an agent repeatedly chooses risky prospects aware that future preferences will evolve according to a subjective ergodic utility process. We show parameters of the model can be uniquely identified and provide a complete characterization. As part of our axiomatization, we extend the representation theorems of Gul and Pesendorfer (2006) and Dekel et al. (2001, 2007) to an infinite-dimensional space with countably-additive probability measures. This is accomplished through a unified methodology based on Lipschitz continuous utilities.

* This paper subsumes parts of the working paper "Repeated Choice: A Theory of Stochastic Intertemporal Preferences". We thank Andy Atkeson, Simon Board, Larry Epstein, Mira Frick, Faruk Gul, Ryota Iijima, Asen Kochov, Bart Lipman, Rosa Matzkin, Tomasz Strzalecki, John Rust, Yi Xin and seminar audiences at various conferences and institutions. We especially thank the late Kim Border for many helpful discussions. Financial support from the NSF under awards SES-1558757 (Saito), SES-1919263 (Saito) and SES-1919275 (Lu) are gratefully acknowledged.

† Department of Economics, UCLA.

‡ Division of the Humanities and Social Sciences, Caltech.

1 Introduction

The psychological origins of stochastic choice models imply an interpretation where choice probabilities correspond to the long-run frequencies of an agent making choices from a decision problem repeatedly.[1] Much of the literature however, has taken these choice frequencies as given without explicitly modeling the underlying processes generating them. In this paper, we provide the first theoretical foundation for this repeated-choice interpretation of stochastic choice. We introduce and characterize a model where an agent repeatedly chooses from a menu aware that future preferences will evolve according to subjective utility process, and the limit frequency of these optimal choices correspond exactly to stochastic choice.

Our setup is as follows. Following Kreps and Porteus (1978) and Gul and Pesendorfer (2004), we consider an infinite-horizon framework where every period, an agent faces a menu (i.e., a choice set) consisting of risky prospects over consumption today and continuation menus tomorrow. We focus on *repeated* menus where regardless of what is chosen or which outcome is realized, the agent will always face the same menu again after some finite time. This guarantees that the agent will choose from the same menu infinitely many times, generating a time series of choices. We can then interpret stochastic choice as the long-run frequency of this sequence of choices from the repeated menu.

In our representation, the agent's utility every period t depends on some state s_t that evolves according to an ergodic Markov process. For example, the state could be the agent's mood every period which affects his risk-aversion and degree of impatience. In state s_t at time t, the utility of a pair (c, z) of current consumption c and future continuation menu z is recursively given by

$$u_t(c, z) = \phi_{s_t}\left(c, \mathbb{E}_{s_t}\left[\max_{p \in z} u_{t+1}(p)\right]\right). \tag{1}$$

Here, ϕ is a stochastic aggregator that specifies the agent's intertemporal tradeoffs between current consumption and future continuation value. The agent knows the

[1] Early works include Thurstone (1927), Luce (1959), Block and Marschak (1960), and Falmagne (1978). The adoption of these models to study unobserved population heterogeneity led to their widespread use in economics. For an overview of this history, see McFadden (2001).

Markov process and is fully sophisticated, taking expectations anticipating that he will be choosing from z again next period. The probability that an option p is chosen from a repeated menu z is given by

$$\rho_z(p) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} 1\left\{u_{it+1}(p) \geq u_{it+1}(q) \text{ for all } q \in z\right\}, \tag{2}$$

where $1\{\cdot\}$ is the indicator function and $u_t(p)$ is the expected utility of a lottery p according to (1). We call this an *ergodic representation* of stochastic choice. Despite the generality of the model and the fact that we only consider a restricted domain of repeated menus, an analyst observing the agent's choices can fully identify the agent's utility process from stochastic choice over repeated binary menus.

Our main result is a complete characterization of our model. First, we show that the set of repeated menus is in fact dense in the set of all menus. This allows us to extend stochastic choice over repeated menus to over all finite (but not necessarily repeated) menus. Next, note that since the Markov process in our representation is ergodic, it has a stationary distribution. By Birkhoff's ergodic theorem, we can rewrite our representation as a random utility with respect to this stationary distribution.

Our first axiom consists of the conditions from Gul and Pesendorfer (2006) adapted to our setup. This ensures we have a random expected utility representation. Our second axiom is a linearity condition that ensures that continuation menus are evaluated according to the representation in Dekel et al. (2001, 2007). We introduce two new stationarity axioms. The first (*A-Stationarity*) is a weaker analog of the stationarity axiom of Koopmans (1960) for stochastic choice. It allows us to construct a recursive and stationary Markov utility process. The second (*D-Stationarity*) is a condition stating that deterministic stationarity is robust to small perturbations. It ensures ergodicity of the utility process. A final application of Birkhoff's ergodic theorem ensures that we can rewrite our random utility as (2) (for details, see the discussion after Theorem (2)).

Our axiomatization requires extending the representation theorems from Gul and Pesendorfer (2006) and Dekel et al. (2001, 2007) to an infinite-dimensional space with countably-additive probability measures. This is a known technical challenge in the

literature due to the lack of compactness in the infinite-dimensional setting.[2] Our main technical contribution is providing a unified methodology to achieve both (see Theorems 3 and 4 in Appendix C). We focus on the set of Lipschitz continuous utilities with common bound, which is compact via the Arzela-Ascoli theorem. Note that this is not only important for the representation but also crucial for identification as identification would not be possible without such restrictions. Importantly, the use Lipschitz continuity is natural in many settings. This is true in Dekel et al. (2001, 2007) and in other popular models (for an application to ambiguity, see Lu (2021)). We hope our methodology will be useful for future research in other applications.

1.1 Related Literature

There are many papers that study stochastic choice under expected utility. Gul and Pesendorfer (2006), Ahn and Sarver (2013), Lu (2016), and Lu and Saito (2018) study static models of stochastic choice, while Fudenberg and Strzalecki (2015) and Frick et al. (2020) study dynamic random choice.[3] Our paper is most closely related to the latter. The main differences are in motivation and mathematical modeling. Given their motivation to study history dependency in populations, stochastic choice in their model is conditional on past menus, choices, and consumption realizations. On the technical side, they extend Gul and Pesendorfer (2006) to a finitely additive measure via the finiteness condition of Ahn and Sarver (2013), while our extension is to a countably-additive measure via Lipschitz continuity. Our infinite-horizon setup also allows us to study ergodic properties of the utility process.

There is also a large literature on menu preferences under expected utility, beginning with Dekel et al. (2001) and Gul and Pesendorfer (2001). The former was extended to an objective state space by Dillenberger et al. (2014). Gul and Pesendorfer (2004) extended this to a dynamic setting by proposing an infinite-horizon consumption setup, which we adopt in this paper. Other papers that use this framework

[2] Krishna and Sadowski (2014) provides an outline of the technical issues of extending Dekel et al. (2001, 2007). See Ma (2018) and Frick et al. (2020) for other extensions of Gul and Pesendorfer (2006).

[3] See also Duraj (2018), which extends Frick et al. (2020) to a setting with an objective state space, and Ke (2018) which studies expected utility in a Luce model.

include Higashi et al. (2009) and Krishna and Sadowski (2021). The first considers a random discounting model in which the agent anticipates the stochasticity of his future discount factor. The second extends the additive linear representation into an infinite-dimensional space. While their extension is finitely additive, our extension is countably additive which preserves uniqueness of the representation. Krishna and Sadowski (2014) and Dillenberger et al. (2023) augment the dynamic setup with an informational structure (see the latter for a review of this literature).

2 Repeated Menus

Time $T = \{1, 2, \dots\}$ is discrete and $M = [0, m]$ is a closed interval representing consumption (e.g., money). The agent is faced with an infinite-horizon consumption problem, that is, a menu of choice options in which each option corresponds to a lottery over consumption today and a continuation menu tomorrow. We will refer to these simply as *menu*s and denote them by $z \in Z$. From Gul and Pesendorfer (2004), we know that Z is homeomorphic to $\mathcal{K}(\Delta(M \times Z))$, where $\Delta(\cdot)$ denotes the set of probability measures and $\mathcal{K}(\cdot)$ denotes the set of nonempty compact subsets. Thus, we will associate Z with $\mathcal{K}(\Delta(M \times Z))$ without loss of generality.

Let $X = M \times Z$ denote the set of possible *outcomes*. For $x \in X$, we sometimes also let x denote the degenerate lottery $\delta_x \in \Delta X$ and p denote the singleton menu $\{p\} \in Z$.

The main focus of our study will be on menus that repeat themselves after a fixed number of periods. The following example illustrates what we mean by such "repeated" menus.

Example 1. Consider an analyst interested in eliciting an agent's risk aversion which may be stochastic over time. Every period, the agent is offered a choice between a safe option b and a risky option r from the menu $z = \{b, r\}$. For instance, the safe option yields \$5 now and the menu $z \in Z$ again next period. The risky option yields either \$10 or \$0 with equal probability now and the menu $z \in Z$ next period. Note that regardless of what the agent chooses and which outcome is realized, he will always

face the menu z again next period.

The menu in the example above is repeated every period. More generally, we consider menus that repeat (for sure) after a fixed number of periods. Formally, for $z \in Z$, let $R_0(z) = \{z\}$ and recursively define $R_t(z) := \mathcal{K}(\Delta(M \times R_{t-1}(z)))$ for $t \in T$. Thus, $R_t(z) \subset Z$ are the subset of menus that yield z for sure after t periods.

Definition 1. A menu z is *t-period* if $z \in R_t(z)$. The menu z is *repeated* if it is t-period for some $t > 0$.

Let $R \subset Z$ denote the set of repeated menus. When a menu is t-period, the agent chooses from the menu at periods $1, 1+t, 1+2t$ and so forth. Since this is repeated ad infinitum, it generates an infinite sequence of choices from the same menu. We can thus calculate the long-run frequency of these choices, which is the common individual interpretation of stochastic choice.[4] We now model this explicitly. Let $Z^f \subset Z$ denote the set of finite menus and let $Z^* = R \cap Z^f$ denote the set of finite repeated menus.

Definition 2. A *stochastic choice* is a mapping $\rho : Z^* \to \Delta(\Delta X)$ such that for every $z \in Z^*$, ρ_z is a probability distribution on z.

Given a finite repeated menu $z \in Z^*$, $\rho_z(p)$ designates the probability of choosing option p from the menu z. To simplify notation, we sometimes let $\rho(z, y) = \rho_{z \cup y}(z)$ for $z, y \in Z^*$. Thus, $\rho(p, q) = \rho_{\{p,q\}}(p)$ denotes the frequency of choosing p over q.

A couple of remarks are in order. First, in a repeated menu, repetition is completely independent of the agent's choices. As a result, the analyst need not worry about selection biases interfering with the data collection process. In fact, for repeated menus, the analyst can simply record the agent's choices whenever the menu appears; there is no need to know the actual history of choices that led to it.[5] Second, even though repeated menus form a small subset of menus, they are rich enough to approximate any menu, i.e. R is dense in Z (see Corollary 1). Moreover, stochastic choice over repeated menus is sufficient to fully identify our model (see Theorem 1).

[4] See Luce (1959) and Luce and Suppes (1965) for more detailed descriptions of the individual interpretation of stochastic choice.

[5] This would still hold if we extended our domain to include menus that repeat with some probability, which would accommodate some consumption-savings applications.

This obviates the need for considering richer data sets such as the entire time path of choices. Both remarks imply that the burden on data collection is much reduced in our setting.

Finally, a remark on ties in the stochastic choice.[6] As in Lu (2016), we deal with ties in allowing for some probabilities to be unspecified. This is analogous to how under deterministic choice, indifference characterizes exactly when the model is silent about which option the agent will choose. We model this as non-measurability and let ρ denote the corresponding outer measure without loss of generality.[7]

3 An Ergodic Representation

3.1 Utility Process

In our model, the agent's utility at every period is stochastic and depends on the realization of some state $s \in S$ (unobserved by the analyst). For instance, the state could be the agent's mood each day, which affects how risk-averse or how patient he is on that day. It could also reflect the realization of some private news arriving every period which affects the agent's utility that period.

The state evolves according to a Markov process $(s_t)_{t \in T}$ with transition probabilities $P : S \to \Delta S$ and a stationary distribution $\pi \in \Delta S$. The Markov process is known to the agent but unknown to the analyst. We assume P satisfies the continuity condition that $P_s \geq \delta\pi$ for some $\delta > 0$.[8] This ensures that the Markov process has full support with respect to its stationary distribution and guarantees ergodicity.

We now describe the agent's utility. Let U denote the set of utility functions $u : X \to [0, 1]$ normalized such that $u(\underline{x}) = 0$ and $u(\bar{x}) = 1$, where $\underline{x}$ and $\bar{x}$ respectively

[6] To see why ties are needed in our model, note that evaluations of continuation menus are deterministic when utilities are iid. As a result, ties can occur with probability one.

[7] Formally, let $\mathcal{F}$ be a σ-algebra on ΔX. Given any $z \in Z^*$, let ρ_z be the outer measure with respect to the σ-algebra generated by $\mathcal{F} \cup \{z\}$. See Lu (2016) for details.

[8] The classic Doeblin's condition states that $P_s^n \geq \delta\lambda$ for some $n \geq 1$ and probability measure λ where P^n is the n-iteration of the Markov process. Our condition obtains if we set $n = 1$ and $\lambda = \pi$.

correspond to consuming 0 and m forever.[9] For any measurable $u \in U$, we let

$$u(p) := \int_X u(x)\, dp$$

denote the expected utility of $p \in \Delta X$.

Every period $t \in T$, a state $s_t \in S$ is realized that determines the agent's utility $u_{s_t} \in U$ and his expectation about next period's state $s_{t+1} \in S$ (according to P_{s_t}). The agent is fully sophisticated and has correct beliefs. Following Kreps and Porteus (1978), we model utilities recursively as aggregator functions that tradeoff current consumption with future continuation values.

Definition 3. $\phi = (\phi_s)_{s \in S}$ is a *stochastic aggregator* if every $\phi_s : M \times [0,1] \to [0,1]$ is Lipschitz continuous (with bound N) and strictly increasing in the second argument.

We now define a utility process as follows.

Definition 4. A stochastic process $(u_t)_{t \in T}$ on U is a *utility process* if there exists a Markov process P on S and a stochastic aggregator ϕ such that a.s.

$$u_t(c, z) = \phi_{s_t}\left(c, \int_S \sup_{p \in z} u_{t+1}(p)\, dP_{s_t}\right). \tag{3}$$

We say the utility process is *generated by* (P, ϕ) if P and ϕ are the corresponding Markov process and stochastic aggregator. If $s_t = s$ for some $s \in S$, we sometimes use the notation u_s or u_{s_t}, instead of u_t.

The following lemma shows that the utility process is an ergodic Markov process on the space of utilities

Lemma 1. *A utility process is an ergodic Markov process.*

Proof. See Appendix A.2. □

To see why, note that if $u_s = u_{s'}$, then equation (3) imposes that the expectations in the second argument of the aggregator must be the same. This implies that the distributions of the next period's utility must also be the same. Ergodicity then follows from the ergodicity of P.

[9] We use the range $[0,1]$ for convenience but any compact interval will suffice.

3.2 Representation

We are now ready to formally define the representation. Recall π is the stationary distribution of P. We say a utility process is *regular* if $u_s(p) = u_s(q)$ with π-probability zero or one for all $p, q \in \Delta X$, i.e. ties either never occur or occur for sure.

Definition 5. ρ has an *ergodic representation* if there exists a regular utility process generated by (P, ϕ) such that for any t-period $z \in Z^*$, a.s.

$$\rho_z(p) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} 1\left\{u_{it+1}(p) \geq u_{it+1}(q) \text{ for all } q \in z\right\}.$$

As usual, if a stochastic choice has an ergodic representation corresponding to (P, ϕ), then we say it is represented by (P, ϕ). In an ergodic representation, the probability of choosing an option is exactly the long-run frequency of choosing it in an infinite sequence of choices. Every period, the option is chosen only if it is ranked the highest according to the utility process.

Ergodicity of the utility process has two important implications. First, by Birkhoff's ergodic theorem, it means there is an equivalent representation of an ergodic representation via the unique stationary distribution of the utility process, i.e.

$$\rho_z(p) = \pi\left\{s \in S : u_s(p) \geq u_s(q) \text{ for all } q \in z\right\}.$$

Second, it means the time at which the analyst starts collecting choice data is unimportant; all long-run frequencies will eventually converge to this stationarity distribution. Since repeated menus ensure there is no selection bias from menus, this greatly simplifies the data collection process for the analyst.

Two important special cases of our model are as follows.

Example 2 (Standard Utility)**.** A utility process is *standard* if there is a random vNM utility $\hat{u}_s$ over consumption and a random discount factor $\beta_s \in (0, 1]$ such that a.s.

$$\phi_s(c, v) = (1 - \beta_s)\,\hat{u}_s(c) + \beta_s v.$$

Standard utility is additively separable and linear in continuation values. It also cor-

responds to the classic Bellman equation in dynamic discrete choice.[10] Now, consider a 1-period repeated menu z and for every $p \in z$, let $\hat{p} \in \Delta M$ denote its marginal distribution over consumption. Since continuation values for 1-period menus cancel out under standard utility, $u_s(p) \geq u_s(q)$ if and only if $\hat{u}_s(\hat{p}) \geq \hat{u}_s(\hat{q})$. As a result, we have

$$\rho_z(p) = \pi\{s \in S : \hat{u}_s(\hat{p}) \geq \hat{u}_s(\hat{q}) \text{ for all } q \in z\}.$$

This is the random expected utility model of Gul and Pesendorfer (2006).

Example 3 (Stochastic Epstein-Zin)**.** A utility process is *stochastic Epstein-Zin* if there are $\alpha_s > 0$, $\beta_s \in (0,1)$ and $\gamma_s > 0$ such that a.s.

$$\phi_s(c,v) = \left((1-\beta_s)c^{\gamma_s} + \beta_s v^{\frac{\gamma_s}{\alpha_s}}\right)^{\frac{\alpha_s}{\gamma_s}}$$

As in Epstein and Zin (1989), the parameters (α, β, γ) correspond to (stochastic) risk aversion, discounting and the elasticity of intertemporal substitution respectively. When $\gamma = \alpha$, this reduces to standard utility (random expected CRRA utility). A special case is when the state is exactly the parameters $s = (\alpha, \beta, \gamma)$ but in general, the state space can be much richer as it encodes not only intertemporal preferences (i.e. the stochastic aggregator) but also expectations about tomorrow's state.

When the utility process is standard, the analyst does not need to model repetition explicitly as the long-run choice frequency is independent of future continuation menus. When utilities are not standard however as in the case of stochastic Epstein-Zin in Example 3, the repetition structure of menus matters.

We end this section with a uniqueness result. The analyst can completely identify

[10] In that literature, utilities satisfy

$$u_{s_t}(c,z) = (1-\beta)(w(c) + \varepsilon_{s_t}(c,z)) + \beta\mathbb{E}_t\left[\sup_{p\in z} u_{s_{t+1}}(p)\right]$$

where shocks ε are iid across both consumption c and menus z. When the menu is fixed (e.g. z is 1-repeated), letting $\hat{u}_s(c) = w(c) + \varepsilon_s(c)$ corresponds to our standard utility. One caveat is that $\hat{u}_s$ is bounded and Lipschitz continuous which would not technically hold if ε is extreme-value distributed. In most applications however, the set of options is finite in which case this is satisfied without loss.

the agent's utility process in an ergodic representation from stochastic choice over repeated binary menus.[11]

Theorem 1. *Let ρ and τ be represented by (P, ϕ) and (Q, φ) respectively. Then the following are equivalent:*

(1) $\rho(p, q) = \tau(p, q)$ *for all* $\{p, q\} \in Z^*$

(2) (P, ϕ) *and* (Q, φ) *generate the same utility process.*

Proof. See Appendix B. □

For some intuition as to how the utility process can be identified purely from long-run frequencies, consider two utility processes where one is iid and the other is persistent yet both share the same stationary distribution. Since utilities also encode information regarding the agent's expectation about tomorrow's utilities, he would exhibit a greater preference for flexibility in the iid case than in the persistent case. The analyst can thus use attitudes toward continuation menus to distinguish between the two utility processes.

4 Characterization

We now provide an axiomatic characterization of our model. First, we show how repeated menus can be used to approximate any menu. This allows us to extend our primitive to the set of all (finite) menus.

4.1 Repeated Menu Approximations

We first show how repeated menus can be used to approximate any menu $z \in Z$. First, let $r_{y,t}(z)$ denote the menu obtained from replicating z for the first t periods

[11] This does not mean that the Markov process on S is uniquely identified as one can easily relabel or add redundant states. Nevertheless, if we focus on a "minimal" state space such that no two states have the same utility, then unique identification holds.

and ending with the menu $y \in Z$ for sure. We define this inductively as follows. Let $r_{y,0}(z) = y$ and subsequently define

$$r_{y,t}(z) := \{p_{y,t} : p \in z\}$$

where $p_{y,t} \in \Delta X$ is the lottery induced by $r_{y,t-1}$.[12] Lemma 7 in Appendix A.3 shows that this is well-defined. Given any menu $z \in Z$, we can now define what it means to construct a repeated menu that approximates z up to t periods and then repeats.

Definition 6. Given $z \in Z$, let z^t be t-period such that $z^t = r_{z^t,t}(z)$.

The following lemma shows this is well-defined. Moreover, given any menu $z \in Z$, we can use its t-period repeated version to approximate it as we increase the number of periods between each repetition.

Lemma 2. *For every $z \in Z$, z^t exists and $z^t \to z$ as $t \to \infty$.*

Proof. See Appendix A.3. □

An immediate corollary is that we can use repeated menus to approximate any finite menu.

Corollary 1. *Z^* is dense in Z^f.*

Proof. Fix some finite menu $z \in Z^f$ so from Lemma 2, we can find $z^t \in R$ such that $z^t \to z$. Since $z^t = r_{z^t,t}(z)$ and z is finite, z^t is also finite by definition. Thus, $z^t \in R \cap Z^f = Z^*$ as desired. □

4.2 Axiomatic Characterization

The results in the previous section allow us to extend the observed stochastic choice over repeated finite menus to all finite menus as follows. From Corollary 1, for any finite menu $z \in Z^f$, we can find a sequence of finite repeated menus $z^t \in Z^*$ such that $z^t \to z$. Assuming continuity of the stochastic choice mapping (see Axiom 1.4 below), we can thus define stochastic choice over all finite menus by setting $\rho_z = \lim_t \rho_{z^t}$.

[12] Formally, $p_{y,t}(A \times B) = p\left(A \times r_{y,t-1}^{-1}(B)\right)$ for any measurable $A \times B$.

Going forward, let ρ denote the continuous extension of stochastic choice from Z^* to Z^f. We model ties the same way (see the discussion on ties in Section 2) and let $Z^\circ \subset Z^f$ denote the set of finite menus that contain no ties.

We are now ready to state our axioms. The first set consists of conditions for random expected utility. For any $p, q \in \Delta X$ and $a \in [0,1]$, let $ap + (1-a)\,q$ denote the ex-ante mixture between the two lotteries. Let $\text{ext}\,(z)$ denote the extreme options of some finite menu z. Also recall that $\bar{x}$ and $\underline{x}$ are the consumption streams that yield the best outcome (i.e., m) and worst outcome (i.e., 0) respectively forever.

Axiom 1.1 (Monotonicity). *$z \subset y$ implies $\rho_z\,(p) \geq \rho_y\,(p)$.*

Axiom 1.2 (Linearity). *$\rho_z\,(p) = \rho_{\alpha z + (1-\alpha) q}\,(\alpha p + (1-\alpha)\,q)$ for $\alpha > 0$.*

Axiom 1.3 (Extremeness). *$\rho_z\,(ext\,(z)) = 1$.*

Axiom 1.4 (Continuity). *$\rho : Z^\circ \to \Delta\,(\Delta X)$ is continuous.*

Axiom 1.5 (Best-Worst). *$\rho\,(\underline{x}, \bar{x}) = 0$ and $\rho\,(\bar{x}, x) = \rho\,(x, \underline{x}) = 1$ for all $x \in X$.*

Axiom 1.6 (L-continuity). *There exists $N > 0$ such that $|x_1 - x_2| \leq \frac{\alpha}{N}$ implies*

$$\rho\left(\alpha\delta_{\bar{x}} + (1-\alpha)\,\delta_{x_1}, \alpha\delta_{\underline{x}} + (1-\alpha)\,\delta_{x_2}\right) = 1.$$

Axioms 1.1-1.4 are direct from Gul and Pesendorfer (2006). Axiom 1.5 ensures that $\bar{x}$ and $\underline{x}$ truly are the best and worst outcomes. Axiom 1.6 is the stochastic version of the Lipschitz continuity axiom from Dekel et al. (2007). It guarantees that utilities are sufficiently well-behaved in that they are Lipschitz continuous with respect to some common bound N. This is important and ensures the representation is unique.[13] To understand L-continuity intuitively, note that when $N = \infty$, then $x_1 = x_2 = x$ for some x and the axiom reduces to $\rho\left(\alpha\delta_{\bar{x}} + (1-\alpha)\,\delta_x, \alpha\delta_{\underline{x}} + (1-\alpha)\,\delta_x\right) = 1$, which holds by Best-Worst and Linearity. L-continuity requires that this remains true for large enough but finite N.[14] Taken together, Axiom 1 characterizes a random expected Lipschitz utility with best and worst outcomes.

[13] When the outcome space is infinite-dimensional, allowing for all possible vNM utilities would be too permissive and not permit unique identification.

[14] Note that if the condition is satisfied for N, then it must also be satisfied for all $N' \geq N$ so testing the axiom involves finding a large enough N such that the condition holds.

The above conditions are all ex-ante and impose no restrictions on continuation menus. We now define a notion of ex-post mixing. For $\lambda \in [0,1]$, $c, d \in M$ and $z, y \in Z$, define ex-post mixing as

$$\alpha (c, z) \oplus (1-\alpha)(d, y) := (\alpha c + (1-\alpha) d, \alpha z + (1-\alpha) y).$$

The first mixture $\alpha c + (1-\alpha) d$ corresponds to the standard mixing of real numbers (e.g. monetary values) while the second mixture $\alpha z + (1-\alpha) y$ corresponds to Minkowski (ex-ante) mixing of menus. The Minkowski version of ex-post mixing is defined in the usual way

$$\alpha z \oplus (1-\alpha) y := \{\alpha p \oplus (1-\alpha) q : p \in z, q \in y\}$$

whenever the menus z and y consist of degenerate lotteries.

For any $c \in M$, let Z_c^f be the set of finite menus consisting of degenerate lotteries that yield consumption c for sure today, i.e. $\delta_{(c,z)}$. Linearity with respect to mixing of continuation menus would require that for any menu $z \in Z_c^f$, degenerate q and $\alpha > 0$

$$\rho_z (p) = \rho_{\alpha z \oplus (1-\alpha) q} (\alpha p \oplus (1-\alpha) q).$$

This is exactly the ex-post counterpart of Linearity (Axiom 1.2). Under stochastic choice however, this is not sufficient to ensure utilities are linear.[15] Continuation Linearity strengthens linearity to hold even with respect to ex-ante mixing.[16]

Axiom 2 (Continuation Linearity)**.** *For $p \in z \in Z_c^f$, degenerate q and $\alpha, \lambda > 0$,*

$$\rho_z (p) = \rho_{\lambda z + (1-\lambda)(\alpha z \oplus (1-\alpha) q)} (\lambda p + (1-\lambda)(\alpha p \oplus (1-\alpha) q)).$$

Finally, we introduce our two stationarity axioms. Classic Koopmans (1960) stationarity states that choices remain unchanged if everything is delayed by a period. A natural extension of this to stochastic choice would be requiring choice frequencies

[15] Lin (2020) provides examples of random non-linear utilities that violate the linearity axiom.

[16] A weaker axiom would only impose this in cases where today's consumption is the same. This would lead to a model where beliefs about tomorrow's state P_s could also depend on the consumption each period, accommodating stochastic consumption-dependent preferences due to habit formation or experimentation. See Turansick (2023) for related work in a different domain.

to remain unchanged if everything is delayed by a period, i.e.

$$\rho(z, y) = \rho\left(\delta_{(c,z)}, \delta_{(c,y)}\right). \tag{4}$$

While appealing, this condition is too strong in our setup.[17] For example, consider a standard utility process and suppose the state is iid, i.e., $P_s = \pi$ for all $s \in S$. In this case, choice between c or d today would depend on the realization of the state s, i.e. $\hat{u}_s(c)$ versus $\hat{u}_s(d)$. This would be stochastic. On the other hand, choice between c or d delayed by a period would depend on the *expectation* of their utilities, i.e. $\mathbb{E}[\hat{u}_s(c)]$ versus $\mathbb{E}[\hat{u}_s(d)]$. This would be deterministic. As a result, condition (4) would be violated.[18]

Nevertheless, there is a weaker version where stationarity holds on "average". For $\alpha \in [0, 1]$, define the lottery p_α that yields the best prize with probability α and the worst prize otherwise, i.e.

$$p_\alpha := \alpha\delta_{\bar{x}} + (1 - \alpha)\,\delta_{\underline{x}}$$

Now, for any $\alpha \in [0, 1]$, $\rho(z, p_\alpha)$ is the demand for options in menu z given outside option p_α. We can thus interpret

$$v_z := \int_0^1 \rho(z, p_\alpha)\, d\alpha$$

as the "average" demand for z. A-Stationarity below ensures that average demand remains unchanged if all consumptions are delayed by one period.

Axiom 3 (A-Stationarity)**.** *For any $z \in Z^f$ and $c \in M$,*

$$\int_0^1 \rho(z, p_\alpha)\, d\alpha = \int_0^1 \rho\left(\delta_{(c,z)}, \delta_{(c,p_\alpha)}\right) d\alpha.$$

This axiom can be interpreted as a stationarity condition on the *surplus* of menus. Recall from McFadden (1978, 1981) that the consumer surplus of a menu z is given by

[17] See Lu and Saito (2018) for a model where this condition is satisfied.

[18] Interestingly in this example, the discount factor for consumption at period t is given by

$$\beta_{s_1}\mathbb{E}\left[\beta_{s_2}\mathbb{E}\left[\cdots\mathbb{E}\left[\beta_{s_t}\right]\right]\right] = \beta_{s_1}\delta^{t-1}$$

where $\delta := \mathbb{E}[\beta]$. This corresponds to a model of random quasi-hyperbolic discounting where present bias occurs if $\beta_{s_1} < \delta$ and future bias occurs if $\beta_{s_1} > \delta$.

$\int_S \max_{p\in z} u_s(p)\, d\pi$. This is in fact equivalent to the average demand v_z which measures consumer surplus by integrating the demand function with respect to price.[19] The surplus delayed by one period is given by $\int_S \int_S \max_{p\in z} u_{s'}(p)\, dP_s d\pi$. If the Markov process is stationary, i.e. $\pi = \int_S P_s d\pi$, then these two surpluses must be the same.

Finally, consider stationarity for deterministic choices, i.e. if $\rho(z, y) = 1$, then

$$\rho\left(\delta_{(c,z)}, \delta_{(c,y)}\right) = 1$$

for any two menus z and y.[20] The final stationarity axiom require this to be "robust" in the following sense. For any menu $z \in Z$, let

$$p_z := v_z \delta_{\bar{x}} + (1 - v_z)\, \delta_{\underline{x}}$$

denote its *probability-equivalent*. Note that the agent is ex-ante indifferent between the menu and its probability-equivalent. D-stationarity below states that deterministic stationarity is robust even if we perturb the menus z and y slightly by mixing them with their probability-equivalents p_y and p_z respectively.

Axiom 4 (D-stationarity)**.** *There exists $\bar{\varepsilon} > 0$ such that if $\rho(z, y) = 1$, then for all $\varepsilon \leq \bar{\varepsilon}$*

$$\rho\left(\delta_{(c,(1-\varepsilon)z+\varepsilon p_y)}, \delta_{(c,(1-\varepsilon)y+\varepsilon p_z)}\right) = 1$$

D-stationarity not only implies deterministic stationarity but that it also holds in a robust way; this ensures the utility process satisfies Doeblin's condition and is thus ergodic. We now state our representation theorem.

Theorem 2. *ρ satisfies Axioms 1–4 if and only if it has an ergodic representation.*

Proof. See Appendix D □

We now provide an outline for the proof of Theorem 2. The first step is the construction of a random expected utility representation where the probability measure is countably additive and continuation menus are evaluated according to the additive

[19] This can be seen via standard integration by parts. See Lu (2016) for a similar methodology.

[20] This is similar to the menu stationarity axiom of Higashi et al. (2009) except we only require implication in one direction.

linear utility function of Dekel et al. (2001, 2007). For this step, we extend Dekel et al. (2001, 2007) and Gul and Pesendorfer (2006) with countably-additive probability measures to an infinite-dimensional space. Both extensions are known challenges in the literature as the set of utilities over an infinite-dimensional space (without any restrictions) may not be compact.[21]

We employ a unified methodology that achieves both. We first focus on finite-dimensional settings. Then, we apply Kolmogorov's extension theorem followed by the Tietze extension theorem by focusing on the set of Lipschitz continuous utilities with common bound; this forms a nice compact set according to the Arzela-Ascoli theorem. Theorems 3 and 4 in Appendix C are the formal statements. Note that this is not only important for the representation but also crucial for identification in both settings (Theorem 1). In fact, without such a restriction on the set of utilities, identification would not be possible.

Once we have a random expected utility representation where continuation menus are evaluated according to the additive linear functional form, the next step is to show that the random utilities are derived from the stationary distribution of an ergodic utility process. This is where the last two stationarity axioms come into play. First, by using A-Stationarity, we show that the random utility is recursive. This allows us to construct a Markov utility process with a stationary distribution that coincides exactly with the distribution of the random utility from the representation. Next, D-stationarity ensures that this Markov utility process is ergodic. In the final step, the representation is obtained by an application of Birkhoff's ergodic theorem.

[21] For instance, the unit ball is compact in finite-dimensional space but not in infinite-dimensional space. See the discussion after Theorem 3 in Krishna and Sadowski (2014) for more details.

Appendix

A Preliminaries

A.1 Lipschitz Continuous Utilities

In this section, let X denote a compact metric space, $C(X)$ denote the set of continuous functions and $L_N(X)$ denote the set of Lipschitz functions with Lipschitz bound N. We endow $C(X)$ with the topology of uniform convergence. Fix $\bar{x}, \underline{x} \in X$ and define

$$U_N := \{u \in L_N(X) : 0 = u(\underline{x}) \leq u(x) \leq u(\bar{x}) = 1 \text{ for all } x \in X\}. \tag{5}$$

For each $u \in C(X)$ and $p \in \Delta X$, let $u(p) = \int_X u dp$ denote its expectation. The following result shows that the set of utilities we consider is compact. It is crucial for both characterization and identification, and highlights the role of Lipschitz functions.

Lemma 3. *U_N is compact in $C(X)$.*

Proof. First, we show that $L_N(X)$ is equicontinuous. Fix $x \in X$ and $\varepsilon > 0$ and consider $y \in X$ such that $|x - y| < \frac{1}{N}\varepsilon$. Thus, for all $u \in L_N(X)$, $|u(x) - u(y)| \leq N|x - y| < \varepsilon$. Since this holds for all $x \in X$, U_N is equicontinuous. Note that since $0 \leq |u| \leq 1$ for all $u \in U_N$, U_N is pointwise bounded.

Next, we show that U_N is closed. Consider $u_k \in U_N$ such that $u_k \to u$. Since u_k is bounded,

$$u(x) - u(y) = \lim_k (u_k(x) - u_k(y)) \leq \lim_k N|x - y| = N|x - y|$$

so $u \in L_N(X)$. Moreover, since $0 = u_k(\underline{x}) \leq u_k(x) \leq u_k(\bar{x}) = 1$ for all k, $0 = u(\underline{x}) \leq u(x) \leq u(\bar{x}) = 1$. This shows $u \in U_N$ so U_N is closed. By the Arzela-Ascoli Theorem (Theorem 4.43 of Folland (2013)), U_N is compact in $C(X)$. □

The following lemma is useful for characterizing distributions on U_N.

Lemma 4. *For $\mu, \nu \in \Delta U_N$, $\mu = \nu$ iff*

$$\int_{U_N} \max\{u(p), \alpha\} \, d\mu = \int_{U_N} \max\{u(p), \alpha\} \, d\nu$$

for all $\alpha \in [0,1]$ and $p \in \Delta X$.

Proof. Note that one direction is trivial so we will prove the other direction. Let G denote the set of all functions $g : U_N \to \mathbb{R}$ such that

$$g(u) = \sum_{i=1}^{n} a_i e^{r_i u(p_i)}$$

for some n, $a_i \in \mathbb{R}$, $r_i \geq 0$ and $p_i \in \Delta X$. By Theorem 1.57 of Müller and Stoyan (2002), we have

$$\int_{U_N} e^{r_i u(p_i)} d\mu = \int_{U_N} e^{r_i u(p_i)} d\nu$$
$$\int_{U_N} g(u) \, d\mu = \int_{U_N} g(u) \, d\nu$$

for all $g \in G$.

We now show G is dense in $C(U_N)$ via the algebraic version of the Stone-Weierstrass theorem. Note that G is a vector space and includes constants as $e^{0u(p)} = 1$. To see that G is closed under multiplication, note that

$$e^{r_1 u(p_1)} e^{r_2 u(p_2)} = e^{r_1 u(p_1) + r_2 u(p_2)} = e^{(r_1+r_2) u\left(\frac{r_1}{r_1+r_2} p_1 + \frac{r_2}{r_1+r_2} p_2\right)} \in G$$

if $r_1 + r_2 > 0$ and $e^{r_1 u(p_1)} e^{r_2 u(p_2)} = 1 \in G$ otherwise. To see that G separates points in U_N, note that if there are $u, v \in U_N$ such that $u \neq v$, then $u(x) \neq v(x)$ for some $x \in X$ so $e^{u(x)} \neq e^{v(x)}$. Since U_N is compact by Lemma 3, G is a subalgebra, contains the constant function and separates points in U_N, it is dense in $C(U_N)$ by Theorem 9.13 of Aliprantis and Border (2006) (henceforth, AB). Since U_N is compact, $\mu = \nu$ by AB Theorem 15.1. □

Next, we have the usual density result for support functions in $C(U_N)$. It is modified from Lemma 11 from Dekel et al. (2001) to accommodate Lipschitz continuous functions.

Lemma 5. *Let G denote the set of all functions $g : U_N \to \mathbb{R}$ such that*

$$g(u) = r \sup_{p \in z} u(p) - r \sup_{q \in y} u(q)$$

for some $r \geq 0$ and finite $z, y \subset \Delta X$. Then G is dense in $C(U_N)$.

Proof. For ease of notation, let

$$\sigma_z(u) := \sup_{p \in z} u(p)$$

denote the support function given $z \subset \Delta X$. We will show the result via the lattice version of the Stone-Weierstrass theorem. First, we prove that G is a linear subspace. First, note that $0 \in G$ trivially by setting $r = 0$. Moreover, if $g \in G$, then $-g \in G$ so $\lambda g \in G$ for all $\lambda \in \mathbb{R}$. Now, consider $g_1, g_2 \in G$ and assume $r_1 + r_2 > 0$. We thus have

$$\begin{aligned} g_1 + g_2 &= r_1 \sigma_{z_1} - r_1 \sigma_{y_1} + r_2 \sigma_{z_2} - r_2 \sigma_{y_2} \\ &= (r_1 + r_2) \left[(a \sigma_{z_1} + (1-a) \sigma_{z_2}) - (a \sigma_{y_1} + (1-a) \sigma_{y_2}) \right] \\ &= (r_1 + r_2) \left(\sigma_{a z_1 + (1-a) z_2} - \sigma_{a y_1 + (1-a) y_2} \right) \in G \end{aligned}$$

where $a := \frac{r_1}{r_1 + r_2}$. If $r_1 + r_2 = 0$, then $g_1 + g_2 = 0 \in G$ trivially. This proves that G is a linear subspace of $C(U_N)$.

To see that G includes constants, note that $\sigma_{\bar{x}}(u) = 1$ and $\sigma_{\underline{x}}(u) = 0$. That G is a lattice follows from the same arguments as in Lemma 11 of Dekel et al. (2001). Finally, to see that G separates points in U_N, note that if there are $u, v \in U_N$ such that $u \neq v$, then $u(x) \neq v(x)$ for some $x \in X$ so $\sigma_z(u) \neq \sigma_z(v)$ for $z = \{x\}$. Since U_N is compact by Lemma 3, G is a vector lattice, contains the constant function and separates points in U_N, it is dense in $C(U_N)$ by AB Theorem 9.12. □

Finally, the following lemma is modified from Dekel et al. (2007) and characterizes Lipschitz continuous functions on a countable dense subset of X. It will be useful when we prove the extensions of the representation theorems of Dekel et al. (2001, 2007) and Gul and Pesendorfer (2006) in Section C. Since X is a compact metric space and thus separable, let $X^* \subset X$ be a countable dense subset of X and without

loss of generality, assume $\underline{x}, \bar{x} \in X^*$.

Lemma 6. *Suppose $u : X^* \to \mathbb{R}$ is such that $u(\bar{x}) = 1$ and $u(\underline{x}) = 0$. Then $u \in L_N(X^*)$ iff for all $x_1, x_2 \in X^*$ and $\alpha \in [0,1]$, $|x_1 - x_2| \leq \frac{\alpha}{N}$ implies*

$$\alpha u(\bar{x}) + (1-\alpha) u(x_1) \geq \alpha u(\underline{x}) + (1-\alpha) u(x_2).$$

Proof. First, suppose $u \in L_N(X^*)$. Note that if $\alpha = 1$, then the result is trivial so assume $\alpha < 1$. Let $|x_1 - x_2| \leq \frac{\alpha}{N}$ so

$$u(x_2) - u(x_1) \leq N|x_1 - x_2| \leq \frac{N}{1-\alpha}|x_1 - x_2| \leq \frac{\alpha}{1-\alpha}$$

Rearranging yields the result as desired.

Now, suppose the condition holds. Fix some $\bar{\alpha} < 1$ and consider $x_1, x_2 \in X^*$. Set $\alpha := |x_1 - x_2| N$ and first suppose $\alpha \leq \bar{\alpha}$. Now, the condition implies that

$$u(x_2) - u(x_1) \leq \frac{\alpha}{1-\alpha} = \frac{N}{1-\alpha}|x_1 - x_2| \leq \frac{N}{1-\bar{\alpha}}|x_1 - x_2|$$

Next, suppose $\alpha > \bar{\alpha}$. Since X is a convex metric space, we can find $y_i := \left(1 - \frac{i}{n}\right) x_1 + \frac{i}{n} x_2 \in X$ for $i \in \{0, 1, \ldots, n\}$ such that $|y_{i+1} - y_i| = \frac{1}{n}|x_1 - x_2| < \frac{\bar{\alpha}}{N}$. Since X^* is dense in X and the metric mapping is continuous, we can choose n large enough such that for each $\varepsilon > 0$, we can find $y_i^* \in X^*$ such that $|y_i - y_i^*| \leq \varepsilon$ and $\left|y_{i+1}^* - y_i^*\right| < \frac{\bar{\alpha}}{N}$ for all i. Applying the argument above, we have

$$\begin{aligned}
u\left(y_{i+1}^*\right) - u(y_i^*) &\leq \frac{N}{1-\bar{\alpha}}\left|y_{i+1}^* - y_i^*\right| \\
&= \frac{N}{1-\bar{\alpha}}\left(|y_{i+1} - y_i| + \left|y_{i+1}^* - y_{i+1}\right| + |y_i^* - y_i|\right) \\
&\leq \frac{N}{1-\bar{\alpha}}\left(|y_{i+1} - y_i| + 2\varepsilon\right) \\
&= \frac{N}{1-\bar{\alpha}}\left(\frac{1}{n}|x_1 - x_2| + 2\varepsilon\right)
\end{aligned}$$

Since we can set $y_0^* = y_0 = x_1$ and $y_n^* = y_n = x_2$, this implies that

$$u(x_2) - u(x_1) \leq \sum_{1 \leq i \leq n}\left|u(y_i^*) - u\left(y_{i-1}^*\right)\right| \leq \frac{N}{1-\bar{\alpha}}\left(|x_1 - x_2| + 2n\varepsilon\right)$$

Taking $\varepsilon \to 0$ yields $u(x_2) - u(x_1) \leq \frac{N}{1-\bar{\alpha}}|x_1 - x_2|$. Since $\frac{N}{1-\bar{\alpha}} \to N$ as $\bar{\alpha} \to 0$, this

means that $|u(x_2) - u(x_1)| \leq N|x_1 - x_2|$ so $u \in L_N(X^*)$ as desired. $\square$

A.2 Proof of Lemma 1

Fix two states $s_1, s_2 \in S$ and suppose $u_{s_1} = u_{s_2}$. We will show that the induced probability distributions over next period utilities are the same in both states. Define the mapping $\xi : S \to U$ such that $\xi(s) = u_s$. We will show that $\nu_i := P_{s_i} \circ \xi^{-1} \in \Delta U$ are the same for $i \in \{1, 2\}$. Consider the menu $z = \left\{p, \alpha\delta_{\bar{x}} + (1-\alpha)\delta_{\underline{x}}\right\}$. Since $u_{s_1}(c, z) = u_{s_2}(c, z)$ and ϕ is strictly increasing in the second argument,

$$\begin{aligned}\int_U \max\{u(p), \alpha\} d\nu_1 &= \int_S \sup_{p \in z} u_{s'}(p) dP_{s_1} \\ &= \int_S \sup_{p \in z} u_{s'}(p) dP_{s_2} = \int_U \max\{u(p), \alpha\} d\nu_2\end{aligned}$$

for any $\alpha \in [0, 1]$. Thus, $\nu_1 = \nu_2$ by Lemma 4.

We can now define a transition kernel K on U such that $K(u) = P_s \circ \xi^{-1}$ where $u = u_s$ for any $s \in S$. Let $\mu = \pi \circ \xi^{-1} \in \Delta U$ and note that for any measurable $B \subset U$,

$$\begin{aligned}\int_U K_u(B) d\mu = \int_S K_{u_s}(B) d\pi &= \int_S P_s\left(\xi^{-1}(B)\right) d\pi \\ &= \pi\left(\xi^{-1}(B)\right) = \mu(B)\end{aligned}$$

where the third equality follows from the fact that π is a stationary distribution of P. This shows the utility process is a stationary Markov process. Moreover, by the continuity condition of the Markov process P, for any measurable B, we also have μ-a.s.

$$K_u(B) = P_s\left(\xi^{-1}(B)\right) \geq \delta\pi\left(\xi^{-1}(B)\right) = \delta\mu(B)$$

so the utility process also satisfies Doeblin's condition and is ergodic.

A.3 Proof of Lemma 2

Recall that $r_{y,t}(z)$ is the menu that follows $z \in Z$ for t periods and then ends with $y \in Z$ for sure. We first show that this is well-defined and continuous.

Lemma 7. *For any $y \in Z$, $r_{y,t} : Z \to Z$ is well-defined and continuous.*

Proof. We will prove this by induction on t. Clearly this is true for $t = 0$ as $r_{y,0}(z) = y$ for all $z \in Z$. Now, suppose that $r_{y,t-1}$ is well-defined and continuous. This means that for any $p \in z$ and measurable $A \times B$,

$$p_{y,t}(A \times B) = p\left(A \times r_{y,t-1}^{-1}(B)\right)$$

is well-defined. Next, we show that $p_{y,t}$ is continuous in $p \in \Delta X$. Consider $p^n \to p$ and let $u : X \to \mathbb{R}$ be continuous and bounded. Since $r_{y,t-1}$ is continuous,

$$\int_X u(c,z)\, dp_{y,t}^n = \int_X u(c, r_{y,t-1}(z))\, dp^n \to \int_X u(c, r_{y,t-1}(z))\, dp = \int_X u(c,z)\, dp_{y,t}$$

so $p_{y,t}^n \to p_{y,t}$ by AB Theorem 15.3 as desired. Lemma 1(i) from Gul and Pesendorfer (2004) then ensures that $r_{y,t}$ is continuous. By induction, $r_{y,t}$ is well-defined and continuous. □

We now prove Lemma 2. First, we show that $r_{y,t}(z)$ is continuous in y. Again, we will prove this by induction. Note that since $r_{y,0}(z) = y$, this is true for $t = 0$. Now, suppose this holds for $r_{y,t-1}(z)$. If we consider $y_n \to y$, then for any continuous bounded $u : X \to \mathbb{R}$,

$$\int_X u(c,z)\, dp_{y_n,t} = \int_X u(c, r_{y_n,t-1}(z))\, dp \to \int_X u(c, r_{y,t-1}(z))\, dp = \int_X u(c,z)\, dp_{y,t}$$

so $p_{y_n,t} \to p_{y,t}$ by AB Theorem 15.3. Lemma 1(i) from Gul and Pesendorfer (2004) then ensures that $r_{y,t}$ is continuous in y as desired.

Following Gul and Pesendorfer (2004), let $Z_0 := \{0\}$, $X_{t+1} = M \times Z_t$ and $Z_{t+1} := \mathcal{K}(\Delta(X_{t+1}))$. From Theorem A1 of Gul and Pesendorfer (2004), we have the homeomorphism

$$Z \cong \{z \in Z_1 \times Z_2 \times \cdots : z_t = r_{0,t}(z_{t+1})\} \tag{6}$$

where z_t denote t-th argument of z for any $t \in T$ and $r_{0,t}(z) \in Z_t$ denote the t-truncated menu that follows z for t-periods. Note that each Z_t is compact so Z is compact in the product topology.

Since Z is compact, by the Brouwer fixed point theorem (AB Corollary 17.56), there exists some $y \in Z$ such that $y = r_{y,t}(z)$. Note that since y yields y for sure after t periods by definition, $y \in R_t(y)$ and y is t-period. To show that y is well-defined, suppose there exists some other $y' \in Z$ such that $y' = r_{y',t-1}(z)$. Note that y' and y agree on everything up to period t, and since they are both t-period, they agree on all periods. Thus, $y = y'$ from (6). This proves that $z^t = y$ is well-defined.

Finally, we show that $z^t \to z$. Since Z is a compact metric space, assume z^t converges to some $z^* \in Z$ without loss. Note that $z_t^* = z_t$ for all $t \in T$ so $z^* = z$ from (6) as desired. This concludes the proof.

B Proof of Theorem 1

From Lemma 1, the utility process is ergodic and let μ and ν denote the stationary utility distributions for ρ and τ respectively. Suppose ρ and τ agree on all binary repeated menus. We will show that for any $p \in \Delta X$, $u(p)$ has the same distribution under μ and ν. Let $p_\alpha = \alpha\delta_{\bar{x}} + (1-\alpha)\delta_{\underline{x}}$ so $u(p_\alpha) = \alpha$ for all $u \in U$. Let $z = \{p, p_\alpha\}$ and recall $z^t = \{p^t, p_\alpha^t\} \in Z^*$ is the t-period menu that follows z for t periods and then repeats. From Corollary 1, $p^t \to p$ and $p_\alpha^t \to p_\alpha$.

First, suppose $u(p) = \alpha$ μ-a.s.. If $u(p) > \alpha$, then $\lim_t 1_{\{u(p^t) \geq u(p_\alpha^t)\}}(u) = 1$ as otherwise, $u(p^t) < u(p_\alpha^t)$ for some subsequence of t so $u(p) \leq \alpha$ yielding a contradiction. On the other hand, if $u(p) < \alpha$, then $\lim_t 1_{\{u(p^t) \geq u(p_\alpha^t)\}}(u) = 0$ as otherwise $u(p^t) \geq u(p_\alpha^t)$ for some subsequence of t so $u(p) \geq \alpha$ again yielding a contradiction. This implies that μ-a.s.

$$\lim_t 1_{\{u(p^t) \geq u(p_\alpha^t)\}}(u) = 1_{\{u(p) \geq \alpha\}}(u)$$

Thus,

$$\begin{aligned}
\lim_t \rho\left(p^t, p_\alpha^t\right) &= \lim_t \mu\left\{u \in U : u\left(p^t\right) \geq u\left(p_\alpha^t\right)\right\} \\
&= \lim_t \int_U 1_{\{u(p^t) \geq u(p_\alpha^t)\}} d\mu = \int_U 1_{\{u(p) \geq \alpha\}} d\mu \\
&= \mu\left\{u \in U : u(p) \geq \alpha\right\}
\end{aligned}$$

where the first equality follows from Birkhoff's ergodic theorem and the third equality from dominated convergence. Since $\rho\left(p^t, p_\alpha^t\right) = \tau\left(p^t, p_\alpha^t\right)$, this implies that

$$\mu\left\{u \in U : u\left(p\right) \geq \alpha\right\} = \nu\left\{u \in U : u\left(p\right) \geq \alpha\right\}. \tag{7}$$

Since μ is regular, the only other case to consider is when $u\left(p\right) = \alpha^*$ μ-a.s. for some α^*. Applying (7) for all $\alpha \neq \alpha^*$, it must be that $u\left(p\right) = \alpha^*$ ν-a.s.. We have thus shown that $u\left(p\right)$ must have the same distribution under μ and ν, so by Lemma 4, $\mu = \nu$. Since each $u \in U$ determines the transition kernel on U, this means that the Markov utility process induced by μ and ν are the same. The converse is trivial.

C Extensions of DLRS and GP

In this section, we present extensions of the main representation theorems of Dekel et al. (2001, 2007) (henceforth DLRS) and Gul and Pesendorfer (2006) (henceforth GP). We use a unifying methodology to extend their results to countably-additive probability measures in infinite-dimensional settings. In both cases, we achieve this by using Lipschitz continuous utilities with a common bound. The plan of attack consists of two steps. First, we first focus on finite-dimensional settings and apply Kolmogorov's extension theorem to extend the representations to Lipschitz continuous utilities on a countable dense subset. We then use the Tietze extension theorem to extend Lipschitz continuity from the countable dense set to the entire set. The reason for doing this in two steps, rather than using Kolmogorov's extension directly in one go, is to ensure almost sure Lipschitz continuity in the final representation.[22] This is analogous to the method for proving the existence of Brownian motion and ensuring that paths are uniformly continuous almost surely.

In this section, let X be a compact metric space and U_N be the set of Lipschitz continuous utilities with common bound N as defined in (5). Let $X^* \subset X$ be a countable dense subset of X and without loss of generality, assume $\underline{x}, \bar{x} \in X^*$. Let

[22] Other papers that also employ Kolmogorov's extension in this manner include Lu and Saito (2018), who do not address continuity, and Frick et al. (2020), who obtain a measure with finite support (ignoring ties).

$Z = \mathcal{K}(\Delta X)$ denote the set of non-empty compact subsets of ΔX.[23]

C.1 Extension of DLRS

Consider a binary relation $\succeq$ on Z and the following conditions (we omit universal quantifiers unless there is a risk of confusion).

Condition 1.1 (Weak Order)**.** $\succeq$ is complete and transitive.

Condition 1.2 (Flexibility)**.** $z \subset y$ implies $z \preceq y$.

Condition 1.3 (Independence)**.** $z \succeq y$ implies $\alpha z + (1-\alpha) w \succeq \alpha y + (1-\alpha) w$.

Condition 1.4 (Continuity)**.** $\succeq$ is continuous.

Condition 1.5 (Best-Worst)**.** $\bar{x} \succeq \{x, \bar{x}\}$ and $x \succeq \{x, \underline{x}\}$ for all $x \in X$.

Condition 1.6 (L-Continuity)**.** There exists $N > 0$ such that $|x_1 - x_2| \leq \frac{\alpha}{N}$ implies

$$(1-\alpha)\,\delta_{x_1} + \alpha\delta_{\bar{x}} \sim \left\{(1-\alpha)\,\delta_{x_1} + \alpha\delta_{\bar{x}}, (1-\alpha)\,\delta_{x_2} + \alpha\delta_{\underline{x}}\right\}.$$

Conditions 1.1-1.4 are directly from DLRS. Condition 1.5 ensures that $\bar{x}$ and $\underline{x}$ are the best and worst options, which ensures utilities can be bounded. Finally, note that Condition 1.6 (L-continuity) is different than the Lipschitz continuity axiom from Dekel et al. (2007). There, the axiom is used to ensure that the representation is Lipschitz continuous and is in fact unnecessary when preferences satisfy flexibility (Condition 1.2). Here, the condition is needed to ensure that the utility distribution in the representation has support only on Lipschitz continuous utilities.

Theorem 3 (DLRS extension)**.** *$\succeq$ satisfies Conditions 1.1–1.6 iff there is a countably additive probability measure ν on U_N such that $\succeq$ is represented by*

$$v(z) = \int_{U_N} \sup_{p \in z} u(p)\, d\nu.$$

[23] Note that while DLRS formally considers all non-empty subsets of ΔX, it is without loss to focus on compact sets.

We now prove Theorem 3. To see the necessity of the conditions, note that Conditions 1.1–1.4 follow from the same arguments as in DLRS. It is easy to see Condition 1.5 from the representation. Condition 1.6 follows from Lemma 6.

We now show sufficiency. We first construct a representation on a limited domain.

Lemma 8. *There exists a probability measure ν on U_N such that for all finite $W \subset X^*$, $\succeq$ on $\mathcal{K}(\Delta W)$ is represented by*

$$\int_{U_N} \sup_{p \in z} u(p)\, d\nu.$$

Proof. The proof consists of three steps:

Step 1: First, we construct a DLRS representation for any finite menu $W \subset X^*$ where $\underline{x}, \bar{x} \in W$. Note that from DLRS, Conditions 1.1–1.4 imply that there exists a probability measure μ_W on $\mathbb{R}^W$ such that $\int_{\mathbb{R}^W} \sup_{p \in z} u(p)\, d\mu_W$ represents $\succeq$ on $\mathcal{K}(\Delta W)$. By Condition 1.5, we have $u(\underline{x}) \leq u(x) \leq u(\bar{x})$ μ_W-a.s. for all $x \in W$. Thus, we can assume that $u(\bar{x}) = 1$ and $u(\underline{x}) = 0$ without loss of generality. With this normalization of utilities, the DLRS representation is unique. Let $U(X^*)$ denote the set of all utilities $u : X^* \to \mathbb{R}$ such that $0 = u(\underline{x}) \leq u(x) \leq u(\bar{x}) = 1$ for all $x \in X^*$. Since all such μ_W are consistent, we can apply Kolmogorov's extension to obtain a countably additive probability measure μ on $U(X^*)$ such that

$$\int_{U(X^*)} \sup_{p \in z} u(p)\, d\mu$$

represents $\succeq$ on $\mathcal{K}(\Delta W)$ for all finite $W \subset X^*$.

Step 2: Next, we show that μ has support only on Lipschitz continuous utilities over X^*, i.e. $\mu(L_N(X^*)) = 1$. Take N from Condition 1.6 and for every $x_1, x_2 \in X^*$ and $\alpha \in [0,1]$, let $U_\alpha^{x_1,x_2}$ be the set of $u \in U(X^*)$ such that

$$|x_1 - x_2| \leq \frac{\alpha}{N} \Longrightarrow \alpha + (1-\alpha)\, u(x_1) \geq (1-\alpha)\, u(x_2)$$

By Condition 1.6, $\mu(U_\alpha^{x_1,x_2}) = 1$. Let I^* be some countable dense subset of $[0,1]$ and $1 \in I^*$. Since I^* and X^* are countable and μ is countably additive,

$$\mu\left(\bigcap_{x_1,x_2 \in X^*, \alpha \in I^*} U_\alpha^{x_1,x_2}\right) = 1.$$

We will now show that

$$\bigcap_{x_1,x_2\in X^*,\alpha\in I^*} U_\alpha^{x_1,x_2} \subset L_N(X^*)$$

Consider $u \in \bigcap_{x_1,x_2\in X^*,\alpha\in I^*} U_\alpha^{x_1,x_2}$ and suppose $|x_1 - x_2| \leq \frac{\alpha}{N}$ for some $x_1, x_2 \in X^*$ and $\alpha < 1$. Consider a decreasing sequence $\alpha_k \searrow \alpha$ where $\alpha_k \in I^*$. Since $|x_1 - x_2| \leq \frac{\alpha_k}{N}$, $\alpha_k + (1-\alpha_k)\, u(x_1) \geq (1-\alpha_k)\, u(x_2)$ for all k so $\alpha + (1-\alpha)\, u(x_1) \geq (1-\alpha)\, u(x_2)$. By Lemma 6, this means that $u \in L_N(X^*)$ as desired. This implies that $\mu(L_N(X^*)) = 1$.

Step 3: Finally, we extend this to bounded Lipschitz continuous utilities over all X. By the Lipschitz version of the Tietze extension theorem (see McShane (1934)), we can extend μ on $L_N(X^*)$ to a probability measure ν on $L_N(X)$. Define

$$U_N(x) := \{u \in L_N(X) \,:\, 0 = u(\underline{x}) \leq u(x) \leq u(\bar{x}) = 1\}$$

Now, since $\nu(U_N(x)) = 1$ for all $x \in X^*$, by the countable additivity of ν,

$$\nu\left(\bigcap_{x\in X^*} U_N(x)\right) = 1.$$

We will now show that

$$\bigcap_{x\in X^*} U_N(x) \subset U_N$$

Consider $u \in \bigcap_{x\in X^*} U_N(x)$ so $0 \leq u(x) \leq 1$ for all $x \in X^*$. Since u is continuous and X^* is dense in X, this implies $0 \leq u(x) \leq 1$ for all $x \in X$ so $u \in U_N$ as desired. This implies that $\nu(U_N) = 1$ and concludes the proof. □

We now conclude the proof of Theorem 3 by extending the representation to all of $Z = \mathcal{K}(\Delta X)$. First, note the following useful lemma.

Lemma 9. *For every $p \in \Delta X$, there exists a sequence $p_n \to p$ where each p_n has some finite support in X^*.*

Proof. Since X^* is dense and Dirac measures are extreme points in ΔX, this follows from the Krein-Milman theorem (AB Theorem 15.10): □

For any $z \in Z$, define

$$v(z) := \int_{U_N} \sup_{p\in z} u(p)\, d\nu$$

We show that v is continuous. Note that $z_n \to z$ implies $\sup_{p \in z_n} u(p) \to \sup_{p \in z} u(p)$ for all $u \in U_N$. By dominated convergence, $v(z_n) \to v(z)$ so v is continuous.

Now, consider a generic $z \in Z$. For any $p \in \Delta X$, by Lemma 9, we can find p_n with finite support in X^* such that $p_n \to p$. Let $z_n := \{p_n : p \in z\}$ so $z_n \to z$ and $z_n \in \mathcal{K}(\Delta W_n)$ for some finite $W_n \subset X^*$. Define $\alpha_n := v(z_n) \in [0,1]$ and note that

$$z_n \sim \alpha_n \delta_{\bar{x}} + (1 - \alpha_n) \delta_{\underline{x}}$$

Assume that α_n converges to some α without loss of generality. Since v is continuous, $v(z) = \alpha$. Since $z_n \to z$ and $\alpha_n \to \alpha$, by Condition 1.4,

$$z \sim \alpha \delta_{\bar{x}} + (1 - \alpha) \delta_{\underline{x}} = v(z) \delta_{\bar{x}} + (1 - v(z)) \delta_{\underline{x}}$$

Thus, $z \succeq y$ iff

$$v(z) \delta_{\bar{x}} + (1 - v(z)) \delta_{\underline{x}} \succeq v(y) \delta_{\bar{x}} + (1 - v(y)) \delta_{\underline{x}}$$

iff $v(z) \geq v(y)$. Thus, v represents $\succeq$ on Z.

C.2 Extension of GP

We now apply the same methodology to extend GP. Consider a stochastic choice function ρ on finite menus $Z^f \subset Z$, i.e. ρ_z is a probability measure over z for every $z \in Z^f$. Without ties, this construction would be identical that of the DLRS extension. We follow Lu (2016) in allowing for ties in that they can occur with probability one. Let $Z^\circ \subset Z^f$ denote the set of finite menus that contain no ties and say a probability measure on U_N is *regular* if $u(p) = u(q)$ occurs with probability zero or one for all $p, q \in \Delta X$. Consider the following conditions (again, we omit universal quantifiers unless there is a risk of confusion).

Condition 2.1 (Monotonicity)**.** $z \subset y$ implies $\rho_z(p) \geq \rho_y(p)$.

Condition 2.2 (Linearity)**.** $\rho_z(p) = \rho_{\alpha z + (1-\alpha) q}(\alpha p + (1 - \alpha) q)$ for $\alpha > 0$.

Condition 2.3 (Extremeness)**.** $\rho_z(\text{ext}(z)) = 1$.

Condition 2.4 (Continuity)**.** $\rho : Z^\circ \to \Delta(\Delta X)$ is continuous.

Condition 2.5 (Best-Worst)**.** $\rho(\underline{x}, \bar{x}) = 0$ and $\rho(\bar{x}, x) = \rho(x, \underline{x}) = 1$ for all $x \in X$.

Condition 2.6 (L-continuity)**.** There exists $N > 0$ such that $|x_1 - x_2| \leq \frac{\alpha}{N}$ implies

$$\rho\left(\alpha\delta_{\bar{x}} + (1-\alpha)\,\delta_{x_1}, \alpha\delta_{\underline{x}} + (1-\alpha)\,\delta_{x_2}\right) = 1.$$

Conditions 2.1-2.4 are directly from GP with the continuity condition adjusted for ties. Condition 2.5, like Condition 1.5, ensures that $\bar{x}$ and $\underline{x}$ are the best and worst options and ensures utilities can be bounded. Condition 2.6 serves the same role as Condition 1.6 in ensuring that utilities are Lipschitz continuous.

Theorem 4 (GP extension)**.** *ρ satisfies Conditions 2.1–2.6 iff there exists a countably additive regular probability measure μ on U_N such that for any $z \in Z^f$,*

$$\rho_z(p) = \mu\{u \in U_N : u(p) \geq u(q) \text{ for all } q \in z\}.$$

We now prove Theorem 4. To see the necessity of the conditions, note that Conditions 2.1–2.3 follow from the same arguments as in GP while Condition 2.4 follows from the same argument as in Lu (2016). It is easy to see Condition 2.5 from the representation. Condition 2.6 follows from Lemma 6.

We now show sufficiency. The methodology parallels the way in which we extended DLRS in the previous section. In fact, Lemma 10 below corresponds exactly to Lemma 8 above.

Lemma 10. *There exists a probability measure μ on U_N such that for all finite $W \subset X^*$ and finite $z \subset \Delta W$,*

$$\rho_z(p) = \mu\{u \in U_N : u(p) \geq u(q) \text{ for all } q \in z\}.$$

Proof. As in the proof for Lemma 8, the proof consists of three steps:

Step 1: First, we construct a GP representation for any finite menu $W \subset X^*$ where $\bar{x}, \underline{x} \in W$. Note that from GP and Lu (2016), Condition 2.1–2.4 imply that there exists a regular probability measure π_W on $\mathbb{R}^W$ such that for any finite $z \subset \Delta W$, ρ_z has a GP representation with π_W. Since the GP representation is unique with the

appropriate normalization, we can use the same arguments in Step 1 of Lemma 8 to apply Kolmogorov's extension and obtain a countably additive and regular probability measure π on $U(X^*)$ (the set of normalized utilities) such that for for all finite $W \subset X^*$ and finite $z \subset \Delta W$,

$$\rho_z(p) = \pi\{u \in U(X^*) : u(p) \geq u(q) \text{ for all } q \in z\}.$$

Step 2: Next, we show that π has support only on Lipschitz continuous utilities over X^*, i.e. $\pi(L_N(X^*)) = 1$. This follows the exact same arguments as in Step 2 of Lemma 8 above.

Step 3: Finally, we extend this to bounded Lipschitz continuous utilities over all X. By the same argument as in Step 3 of Lemma 8, we can extend π on $L_N(X^*)$ to a probability measure μ on $L_N(X)$. Moreover, by the same arguments as in Step 3 of Lemma 8, $\mu(U_N) = 1$ which concludes the proof. □

We now conclude the proof of Theorem 4 by extending the representation to all of Z^f. The first step is to extend it to menus without ties, i.e. Z°. Consider a menu $z = \{p_1, \ldots, p_m\} \in Z^\circ$ and fix some small $\underline{\alpha} \in (0,1)$. By Lemma 9, for each p_i, we can find a sequence of options p_i^n with finite support in X^* such that $p_i^n \to p_i$ and

$$|p_i - p_i^n| \leq \frac{\underline{\alpha}}{2N(n-1)}$$

Let $z^n := \{p_1^n, \ldots, p_m^n\}$ so $z^n \to z$. Define

$$\tilde{p}_i^n := \frac{1}{n}\left(\alpha_i^n \delta_{\bar{x}} + (1-\alpha_i^n)\delta_{\underline{x}}\right) + \left(1 - \frac{1}{n}\right) p_i^n$$

for some $\alpha_i^n \in (0,1)$. Note that $\tilde{p}_i^n$ all have finite support in X^* and let $\tilde{z}^n := \{\tilde{p}_1^n, \ldots, \tilde{p}_m^n\}$. When are the options in $\tilde{z}^n$ tied? By Lemma 10, $\tilde{p}_i^n$ and $\tilde{p}_j^n$ are tied iff μ-a.s.

$$\begin{aligned} u(\tilde{p}_i^n) &= u\left(\tilde{p}_j^n\right) \\ \frac{1}{n}\alpha_i^n + \left(1 - \frac{1}{n}\right) u(p_i^n) &= \frac{1}{n}\alpha_j^n + \left(1 - \frac{1}{n}\right) u\left(p_j^n\right) \\ u(p_i^n) - u\left(p_j^n\right) &= \frac{\alpha_j^n - \alpha_i^n}{n-1} \end{aligned}$$

Since m is finite, we can choose $0 = \alpha_1^n < \underline{\alpha} < \alpha_i^n$ for all $i > 1$ such that $\tilde{z}^n$ has no ties, i.e. $\tilde{z}^n \in Z^\circ$. From Lemma 10 and the fact that μ is regular on $U(X^*)$,

$$\begin{aligned}
\rho_{\tilde{z}^n}(\tilde{p}_1^n) &= \mu\{u(\tilde{p}_1^n) > u(\tilde{p}_i^n) \text{ for all } i > 1\} \\
&= \mu\left\{u(p_1^n) > u(p_i^n) + \frac{\alpha_i^n}{n-1} \text{ for all } i > 1\right\} \\
&\leq \mu\left\{u(p_1) > u(p_i) + \frac{\alpha_i^n - \underline{\alpha}}{n-1} \text{ for all } i > 1\right\}
\end{aligned}$$

where the inequality follows from the fact that $u \in U_N$. Since $\tilde{p}_i^n \to p_i$, $\tilde{z}^n \to z$ so by Condition 2.4,

$$\begin{aligned}
\rho_z(p_1) &= \lim_n \rho_{\tilde{z}^n}(\tilde{p}_1^n) \\
&\leq \lim_n \mu\left\{u(p_1) > u(p_i) + \frac{\alpha_i^n - \underline{\alpha}}{n-1} \text{ for all } i > 1\right\} \\
&\leq \mu\{u(p_1) > u(p_i) \text{ for all } i > 1\}
\end{aligned}$$

where the last inequality follows from the fact that μ is countably additive. Note that this is true for any $p_i \in z$. Now, suppose there is some $p_k \in z$ such that the above inequality is strict. This implies that

$$1 = \sum_i \rho_z(p_i) < \sum_i \mu\{u(p_i) > u(p_j) \text{ for all } i \neq j\} \leq 1$$

yielding a contradiction. We thus have extended the GP representation to all menus without ties, i.e. Z°.

The final step of the proof is to extend the representation to all menus including those with ties. We will show that if any p and q are tied, then $u(p) = u(q)$ μ-a.s.. First, suppose p and q are tied and both are not tied with $\underline{x}$ so from the GP representation, $\rho(\underline{x}, p) = 0$. Let $p_\varepsilon := (1-\varepsilon)p + \varepsilon\delta_{\underline{x}}$ so $\rho(p_\varepsilon, p) = 0$ by Condition 2.2. Since p and q are tied, $\rho(p_\varepsilon, q) = 0$ by Lemma A.2 of Lu (2016). Since p_ε and q are not tied, from the representation

$$\begin{aligned}
1 = \rho(q, p_\varepsilon) &= \mu\{u(q) > (1-\varepsilon)u(p)\} \\
&= \mu\{u(q) - u(p) > -\varepsilon u(p)\}
\end{aligned}$$

Taking $\varepsilon \to 0$ yields $u(q) \geq u(p)$ μ-a.s.. By symmetric reasoning, $u(p) \geq u(q)$ μ-a.s. so $u(p) = u(q)$ μ-a.s. On the other hand, if p is tied with $\underline{x}$, then p cannot be tied with $\varepsilon\delta_{\bar{x}} + (1-\varepsilon)\delta_{\underline{x}}$. From the representation,

$$1 = \rho\left(\varepsilon\delta_{\bar{x}} + (1-\varepsilon)\delta_{\underline{x}}, p\right) = \mu\{\varepsilon > u(p)\}$$

Taking $\varepsilon \to 0$ yields $u(p) = 0 = u(q)$ μ-a.s.. Applying Lemma A.2 from Lu (2016), we have that for any $z \in Z^f$,

$$\rho_z(p) = \mu\{u \in U_N : u(p) \geq u(q) \text{ for all } q \in z\}.$$

D Proof of Theorem 2

Sufficiency. Note that Axioms 1.1-1.6 correspond exactly to Conditions 2.1–2.6 in Theorem 4. Thus, there exists a regular countably additive probability measure μ on U_N s.t. for any $z \in Z^f$,

$$\rho_z(p) = \mu\{u \in U_N : u(p) \geq u(q) \text{ for all } q \in z\}.$$

Consider $z = \left\{\delta_{(c,z_1)}, \delta_{(c,z_2)}\right\}$ so

$$\alpha z \oplus (1-\alpha)\delta_{(\hat{c},y)} = \left\{\delta_{(d,\alpha z_1+(1-\alpha)y)}, \delta_{(d,\alpha z_2+(1-\alpha)y)}\right\}$$

where $d = \alpha c + (1-\alpha)\hat{c}$. Now,

$$\begin{aligned}\frac{1}{2}z + \frac{1}{2}(\alpha z \oplus (1-\alpha)r) = \Big\{&\frac{1}{2}\delta_{(c,z_1)} + \frac{1}{2}\delta_{(d,\alpha z_1+(1-\alpha)y)}, \frac{1}{2}\delta_{(c,z_2)} + \frac{1}{2}\delta_{(d,\alpha z_1+(1-\alpha)y)},\\ &\frac{1}{2}\delta_{(c,z_1)} + \frac{1}{2}\delta_{(d,\alpha z_2+(1-\alpha)y)}, \frac{1}{2}\delta_{(c,z_2)} + \frac{1}{2}\delta_{(d,\alpha z_2+(1-\alpha)y)}\Big\}\end{aligned}$$

Applying Axiom 2 and using the notation $u_c(z) = u\left(\delta_{(c,z)}\right)$, we have

$$\begin{aligned}\mu\{u_c(z_1) \geq u_c(z_2)\} &= \rho_z\left(\delta_{(c,z_1)}\right) = \rho_{\frac{1}{2}z+\frac{1}{2}(\alpha z\oplus(1-\alpha)r)}\left(\frac{1}{2}\delta_{(c,z_1)} + \frac{1}{2}\delta_{(d,\alpha z_1+(1-\alpha)y)}\right)\\ &= \mu\{u_c(z_1) \geq u_c(z_2), u_d(\alpha z_1 + (1-\alpha)y) \geq u_d(\alpha z_1 + (1-\alpha)y)\}\end{aligned}$$

This implies that μ-a.s.

$$u_c(z_1) \geq u_c(z_2) \Rightarrow u_d(\alpha z_1 + (1-\alpha) y) \geq u_d(\alpha z_1 + (1-\alpha) y) \tag{8}$$

Since Z, M and $[0,1]$ are all separable and any $u \in U_N$ is continuous, by the countable additivity of μ, we have that this holds μ-a.s. for all $c, d \in M$, $z_1, z_2, y \in Z$ and $\alpha \in [0,1]$. Moreover, we can define a preference relation $\succeq_u$ on Z for each $u \in U_N$ such that $z \succeq_u y$ if and only if $u_c(z) \geq u_c(y)$. Note that this is well-defined as it does not depend on $c \in M$ by (8) above.

We now show that $\succeq_u$ satisfies Conditions 1.1–1.6 from Theorem 3 μ-a.s. Note that Condition 1.1 is trivial and Condition 1.3 follows from (8) above. To see Condition 1.2, note that from Axiom 4, for any z, y, if $y \subset z$, then

$$1 = \rho(z, y) = \rho\left(\delta_{(c,z)}, \delta_{(c,y)}\right) = \mu\{u_c(z) \geq u_c(y)\}.$$

Since μ is countably additive, $u \in U_N$ is continuous and Z is separable, Condition 1.2 follows. Note that Condition 1.4 follows from the continuity of $u \in U_N$. Finally, by applying Axiom 4 to Axioms 1.5 and 1.6, we obtain Conditions 1.5 and 1.6 respectively by the same argument as before. We can thus apply Theorem 3 so that $\succeq_u$ is represented by

$$v_u(z) := \int_{U_N} \sup_{p \in z} \hat{u}(p)\, dK_u$$

where K_u is a countably additive probability measure on U_N. Since $u_c(\cdot)$ and $v_u(\cdot)$ represent the same preference, there must exist some $\phi_u : M \times [0,1] \to [0,1]$ strictly increasing in the second argument such that μ-a.s.

$$u_c(z) = \phi_u(c, v_u(z)).$$

We now show μ is an invariant measure using A-Stationarity (Axiom 3).

Lemma 11. *μ is an invariant measure of K.*

Proof. Define the measure μ^* on U_N such that for every measurable $B \subset U_N$,

$$\mu^*(B) := \int_{U_N} K_u(B)\, d\mu.$$

We will show that $\mu^* = \mu$ proving that μ is an invariant measure of K. Consider $z \in Z^f$ and note that $\rho(z, p_\alpha) = \mu\left\{\sup_{p\in z} u(p) \geq \alpha\right\}$ which implies that

$$\int_0^1 \rho(z, p_\alpha)\, d\alpha = \int_{U_N} \sup_{p\in z} u(p)\, d\mu.$$

On the other hand, note that

$$\begin{aligned}\rho((c,z),(c,p_\alpha)) &= \mu\{\phi_u(c, v_u(z)) \geq \phi_u(c, v_u(p_\alpha))\} \\ &= \mu\{v_u(z) \geq v_u(p_\alpha)\} \\ &= \mu\{v_u(z) \geq \alpha\}\end{aligned}$$

which implies that

$$\int_0^1 \rho\left(\delta_{(c,z)}, \delta_{(c,p_\alpha)}\right) d\alpha = \int_{U_N} v_u(z)\, d\mu.$$

Axiom 3 thus implies that

$$\begin{aligned}\int_{U_N} \sup_{p\in z} u(p)\, d\mu = \int_0^1 \rho(z, p_\alpha)\, d\alpha &= \int_0^1 \rho\left(\delta_{(c,z)}, \delta_{(c,p_\alpha)}\right) d\alpha \\ &= \int_{U_N} v_u(z)\, d\mu = \int_{U_N}\left(\int_{U_N} \sup_{p\in z} \hat{u}(p)\, dK_u\right) d\mu \\ &= \int_{U_N} \sup_{p\in z} u(p)\, d\mu^*.\end{aligned}$$

Letting $z = \{p, p_b\}$, we have

$$\int_{U_N} \max\{u(p), b\}\, d\mu = \int_{U_N} \max\{u(p), b\}\, d\mu^*$$

so $\mu = \mu^*$ by Lemma 4. □

We now formally construct the state space. Let $U^1 \subset U_N$ be set of u such that there exist some K_u and ϕ_u where

$$u(c, z) = \phi_u\left(c, \int_{U_N} \sup_{p\in z} \hat{u}(p)\, dK_u\right)$$

From the arguments above, $\mu(U^1) = 1$. Now define recursively

$$U^{n+1} := \{u \in U^n : K_u(U^n) = 1\}$$
$$U^* := \bigcap_{n=1}^{\infty} U^n$$

We will show that $\mu(U^*) = 1$. First, we show that $\mu(U^n) = 1$ for all n by induction. Suppose $\mu(U^n) = 1$ so by Lemma 11,

$$1 = \mu(U^n) = \int_{U_N} K_u(U^n)\, d\mu$$

This implies that $K_u(U^n) = 1$ μ-a.s. so $\mu(U^{n+1}) = 1$. Since $\mu(U^1) = 1$, this means that $\mu(U^n) = 1$ for all n. Since $U^{n+1} \subset U^n$, by Proposition 3.6 of Çınlar (2011), $\mu(U^*) = \lim_n \mu(U^n) = 1$. By Lemma 11 again, we have

$$1 = \mu(U^*) = \int_{U_N} K_u(U^*)\, d\mu$$

so $K_u(U^*) = 1$ μ-a.s. Set the state space $S := U^*$ and define a Markov process P with transition kernel $P_s := K_u$ for all $s = u \in U^*$. Note that P has invariant distribution μ and for any $z \in Z^f$,

$$\rho_z(p) = \mu(\{s \in S : u_s(p) \geq u_s(q) \text{ for all } q \in z\})$$

where

$$u_s(c, z) = \phi_u\left(c, \int_S \sup_{p \in z} u_{\hat{s}}(p)\, dP_s\right).$$

Finally, we show that P is ergodic via Doeblin's continuity. Consider any positive $f \in C(U_N)$. By Lemma 5, we can find $g_k \in G$ such that $g_k \to f$ uniformly. If we let $f_k := \max\{g_k, 0\}$, then $f_k \to f$ as $f \geq 0$ and note that

$$\begin{aligned} f_k(u) &= \max\{r_k \sigma_{z_k}(u) - r_k \sigma_{y_k}(u), 0\} \\ &= r_k \sigma_{z_k \cup y_k}(u) - r_k \sigma_{y_k}(u) \end{aligned}$$

where $\sigma_z(u) := \sup_{p \in z} u(p)$ is the support function. Since $\rho(z_k \cup y_k, y_k) = 1$, by

Axiom 4, there exists some $\varepsilon > 0$ s.t. μ-a.s.

$$\begin{aligned}
\int_S \sigma_{(1-\varepsilon)z_k \cup y_k + \varepsilon p_{y_k}}(u)\, dP_s &\geq \int_S \sigma_{(1-\varepsilon)y_k + \varepsilon p_{z_k \cup y_k}}(u)\, dP_s \\
(1-\varepsilon)\int_S \sigma_{z_k \cup y_k}(u)\, dP_s + \varepsilon v_{y_k} &\geq (1-\varepsilon)\int_S \sigma_{y_k}(u)\, dP_s + \varepsilon v_{z_k \cup y_k} \\
\int_S \left(\sigma_{z_k \cup y_k}(u) - \sigma_{y_k}(u)\right) dP_s &\geq \frac{\varepsilon}{1-\varepsilon}\left(v_{z_k \cup y_k} - v_{y_k}\right) \\
&= \frac{\varepsilon}{1-\varepsilon}\int_S \left(\sigma_{z_k \cup y_k}(u) - \sigma_{y_k}(u)\right) d\mu \\
\int_S f_k(u)\, dP_s &\geq \frac{\varepsilon}{1-\varepsilon}\int_S f_k(u)\, d\mu \\
\int_S f(u)\, dP_s &\geq \frac{\varepsilon}{1-\varepsilon}\int_S f(u)\, d\mu
\end{aligned}$$

where the last inequality follows from dominated convergence. Since μ is countably additive and $C(U_N)$ is separable by AB Lemma 3.99, we have μ-a.s. for all positive $f \in C(U_N)$

$$\int_S f(u)\, dP_s \geq \delta \int_S f(u)\, d\mu$$

where $\delta := \frac{\varepsilon}{1-\varepsilon}$. By AB Theorem 13.9, for any measurable $E \subset S$, there are positive $f_j \in C(U_N)$ such that $f_j \to 1_E$ P_s-a.s. By the dominated convergence theorem, μ-.a.s.

$$\begin{aligned}
P_s(E) &= \int_S \lim_j f_j(u)\, dP_s = \lim_j \int_S f_j(u)\, dP_s \\
&\geq \lim_j \delta \int_S f_j(u)\, d\mu = \delta\mu(E)
\end{aligned}$$

This implies P satisfies Doeblin's condition and is thus uniformly ergodic (Theorem 16.2.3 of Meyn and Tweedie (2012)). By Birkhoff's ergodic theorem (AB Corollary 20.8), μ-a.s.

$$\rho_z(p) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} 1\left\{u_{it+1}(p) \geq u_{it+1}(q) \text{ for all } q \in z\right\}$$

for all t-period $z \in Z^f$ as desired. This concludes the sufficiency proof.

Necessity. By Lemma 1, we can consider the ergodic utility process $u_t = u_{s_t}$ with

stationary distribution μ. Thus, for any t-period $z \in Z^f$

$$\begin{aligned}\rho_z(p) &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} 1\left\{u_{it+1}(p) \geq u_{it+1}(q) \text{ for all } q \in z\right\} \\ &= \mu\left(\left\{u \in U_N : u(p) \geq u(q) \text{ for all } q \in z\right\}\right)\end{aligned}$$

where the second equation follows from Birkhoff's ergodic theorem (AB Corollary 20.8). Note that even though this holds for repeated menus, continuity of ρ, which follows from the same argument as in Lu (2016), allows us to extend this to all finite menus $z \in Z^f$. Axioms 1.1-1.6 then follows immediately from Theorem 4.

For Axiom 2, consider $p, p' \in z \in Z_c^f$ where $p = \delta_{(c,y)}$ and $p' = \delta_{(c,y')}$. Thus, $u(p) \geq u(p')$ iff

$$\begin{aligned}u(c, v_u(y)) &\geq u(c, v_u(y')) \\ v_u(y) &\geq v_u(y')\end{aligned}$$

where

$$v_u(y) := \int_{U_N} \sup_{r \in y} \hat{u}(r)\, dK_u$$

Now, for any degenerate $q = \delta_{(d,w)}$, $\alpha p \oplus (1-\alpha) q = \delta_{(\alpha c + (1-\alpha)d, \alpha z + (1-\alpha)w)}$. Thus, $u(\alpha p \oplus (1-\alpha) q) \geq u(\alpha p' \oplus (1-\alpha) q)$ iff

$$\begin{aligned}u(\alpha c + (1-\alpha) d, v_u(\alpha y + (1-\alpha) w)) &\geq u(\alpha c + (1-\alpha) d, v_u(\alpha y' + (1-\alpha) w)) \\ v_u(\alpha y + (1-\alpha) w) &\geq v_u(\alpha y' + (1-\alpha) w) \\ v_u(y) &\geq v_u(y')\end{aligned}$$

This implies that $u(p) \geq u(p')$ iff $u(\alpha p \oplus (1-\alpha) q) \geq u(\alpha p' \oplus (1-\alpha) q)$ for all $p, p' \in z$. The axiom follows immediately from the linearity of u and the Minkowski mixture operation.

For Axiom 3, note that by the same arguments as in Lemma 11,

$$\int_0^1 \rho(z, p_\alpha)\, d\alpha = \int_{U_N} \sup_{p \in z} \hat{u}(p)\, d\mu$$

and

$$\int_0^1 \rho\left(\delta_{(c,z)}, \delta_{(c,p_\alpha)}\right) d\alpha = \int_{U_N} v_u(z)\, d\mu = \int_{U_N} \left(\int_{U_N} \sup_{p\in z} \hat{u}(p)\, dK_u\right) d\mu$$
$$= \int_{U_N} \sup_{p\in z} \hat{u}(p)\, d\mu$$

where the last equation follows from the fact that μ is the stationary distribution.

Finally, for Axiom 4, note that from Lemma 1, we know there exists some $\delta > 0$ such that μ-a.s.

$$\int_{U_N} f(u)\, dK_u \geq \delta \int_{U_N} f(u)\, d\mu$$

for all measurable $f \geq 0$. Suppose $\rho(z, y) = 1$ so μ-a.s.

$$f(u) := \sup_{p\in z} u(p) - \sup_{p\in y} u(p) \geq 0$$

Moreover, $v_z \geq v_y$. Define $\bar{\varepsilon}$ such that $\delta = \frac{\bar{\varepsilon}}{1-\bar{\varepsilon}}$ and note that for any $\varepsilon \leq \bar{\varepsilon}$, $(1-\varepsilon)\delta - \varepsilon \geq 0$. We thus have μ-a.s.

$$\begin{aligned}
v_u((1-\varepsilon)z + \varepsilon p_y) - v_u((1-\varepsilon)y + \varepsilon p_z) &= (1-\varepsilon)(v_u(z) - v_u(y)) + \varepsilon(v_u(p_y) - v_u(p_z)) \\
&= (1-\varepsilon)\int_{U_N} f(u)\, dK_u + \varepsilon(v_y - v_z) \\
&\geq (1-\varepsilon)\delta \int_{U_N} f(u)\, d\mu - \varepsilon(v_z - v_y) \\
&= (1-\varepsilon)\delta(v_z - v_y) - \varepsilon(v_z - v_y) \\
&= ((1-\varepsilon)\delta - \varepsilon)(v_z - v_y) \geq 0
\end{aligned}$$

so the axiom follows. This concludes the necessity proof.

References

Ahn, D. and T. Sarver (2013): "Preference for Flexibility and Random Choice," *Econometrica*, 81, 341–361.

Aliprantis, C. and K. Border (2006): *Infinite dimensional analysis*, Springer.

Block, H. and J. Marschak (1960): "Random Orderings and Stochastic Theories of Response," in *Contributions to Probability and Statistics*, ed. by I. Olkin, Stanford University Press, 97–132.

Çinlar, E. (2011): *Probability and stochastics*, vol. 261, Springer Science & Business Media.

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

Dekel, E., B. L. Lipman, A. Rustichini, and T. Sarver (2007): "Representing Preferences with a Unique Subjective State Space: A Corrigendum 1," *Econometrica*, 75, 591–600.

Dillenberger, D., V. Krishna, and P. Sadowski (2023): "Subjective Information Choice Processes," *Theoretical Economics*, 18, 529–559.

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

Duraj, J. (2018): "Dynamic Random Subjective Expected Utility," Working paper.

Epstein, L. G. and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Falmagne, J. (1978): "A Representation Theorem for Finite Random Scale Systems," *Journal of Mathematical Psychology*, 18, 52–72.

Folland, G. B. (2013): *Real analysis: modern techniques and their applications*, John Wiley & Sons.

Frick, M., R. Iijima, and T. Strzalecki (2020): "Dynamic random utility," *Econometrica*, 88, 2281–2328.

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

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

——— (2004): "Self-control and the theory of consumption," *Econometrica*, 72, 119–158.

——— (2006): "Random expected utility," *Econometrica*, 74, 121–146.

Higashi, Y., K. Hyogo, and N. Takeoka (2009): "Subjective random discounting and intertemporal choice," *Journal of Economic Theory*, 144, 1015–1053.

Ke, S. (2018): "Rational expectation of mistakes and a measure of error-proneness," *Theoretical Economics*, 13, 527–552.

Koopmans, T. C. (1960): "Stationary ordinal utility and impatience," *Econometrica*, 28, 287–309.

Kreps, D. M. and E. L. Porteus (1978): "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica*, 185–200.

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

——— (2021): "Randomly Evolving Tastes and Delayed Commitment," *Journal of Mathematical Economics*, 92, 81–94.

Lin, Y.-H. (2020): "Random non-expected utility: non-uniqueness," Working paper.

Lu, J. (2016): "Random Choice and Private Information," *Econometrica*, 84, 1983–2027.

——— (2021): "Random Ambiguity," *Theoretical Economics*, 16, 539–570.

Lu, J. and K. Saito (2018): "Random intertemporal choice," *Journal of Economic Theory*, 177, 780–815.

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

Luce, R. and P. Suppes (1965): "Preference, Utility, and Subjective Probability," in *Handbook of Mathematical Psychology*, ed. by R. Luce, R. Bush, and E. Galanter, Wiley.

Ma, W. (2018): "Random expected utility theory with a continuum of prizes," *Annals of Operations Research*.

McFadden, D. (1978): "Modeling the choice of residential location," *Transportation Research Record*.

——— (1981): "Econometric models of probabilistic choice," *Structural analysis of discrete data with econometric applications*, 198272.

——— (2001): "Economic choices," *American Economic Review*, 91, 351–378.

McShane, E. J. (1934): "Extension of range of functions," *Bulletin of the American Mathematical Society*, 40, 837–842.

Meyn, S. P. and R. L. Tweedie (2012): *Markov chains and stochastic stability*, Springer Science & Business Media.

Müller, A. and D. Stoyan (2002): *Comparison Methods for Stochastic Models and Risks*, John Wiley and Sons, Inc.

Thurstone, L. (1927): "A law of comparative judgment," *Psychological Review*.

Turansick, C. (2023): "Random Utility, Repeated Choice, and Consumption Dependence," Working paper.