Preference for Flexibility and Random Choice*

David S. Ahn† Todd Sarver‡

May 14, 2012

Abstract

We study a two-stage model where the agent has preferences over menus as in Dekel, Lipman, and Rustichini (2001) in the first period then makes random choices from menus as in Gul and Pesendorfer (2006) in the second period. Both preference for flexibility in the first period and strictly random choices in the second period can be respectively rationalized by subjective state spaces. Our main result characterizes the representation where the two state spaces align, so the agent correctly anticipates her future choices. The joint representation uniquely identifies probabilities over subjective states and magnitudes of utilities across states. We also characterize when the agent completely overlooks some subjective states that realize at the point of choice.

*We thank Eddie Dekel, Bart Lipman, and Philipp Sadowski for useful comments. We owe special thanks to Haluk Ergin, who contributed crucially to the conception and execution of this research.

†Department of Economics, University of California, Berkeley, 508-1 Evans Hall #3880, Berkeley, CA 94720-3880. Email: dahn@econ.berkeley.edu

‡Northwestern University, Department of Economics, 2001 Sheridan Road, Evanston, IL 60208. Email: tsarver@northwestern.edu.

1 Introduction

Since the work of Kreps (1979), the literature on menu choice interprets a preference for flexibility as indicating anticipation of uncertain future tastes. Suppose the agent is unsure whether she will prefer a salad or a steak for dinner. She may prefer to keep both options available to accommodate this uncertainty. Kreps formally modeled this preference by considering the agent's choices among menus of options, where the chosen menu will be her choice set at a future date. A *preference for flexibility* then corresponds to strictly preferring a restaurant that serves both options to a restaurant that serves only one or the other. The resulting extensive literature focuses almost entirely on preferences over menus, leaving implicit the agent's *actual* choice at the restaurant as part of the interpretation of the utility representation. On the other hand, the actual choice provides some additional information regarding taste uncertainty. For example, if we observe the agent always choosing steak whenever it is available, then perhaps her anticipation of a future taste for salad was incorrect.

This paper addresses the correspondence between anticipated choices and actual choices. Menu choice models largely suppress the second stage of choice from the menu. Without loss of generality, they interpret the implied anticipated choices as perfectly predictive of future decisions. Relaxing this interpretation is potentially fruitful in modeling a variety of behaviors. For example, within Kreps's application of future taste contingencies, an inability to anticipate a subjective state at the time of menu choice might suggest that this state is unforeseen. Within the application of temptation, an inability to anticipate future temptations might suggest a lack of sophistication. Testing a joint prediction on anticipated and actual choices requires two kinds of behavioral data. We use the preferences over menus to infer the decision maker's anticipated taste contingencies and use the random choice from menus to infer the taste contingencies at the time of consumption. Specifically, we incorporate preferences over menus using the framework and representation of Dekel, Lipman, and Rustichini (2001, henceforth DLR), who enriched Kreps's domain of choice from menus of deterministic alternatives to menus of lotteries. We then incorporate the agent's stochastic choice from menus using the framework and random expected utility representation of Gul and Pesendorfer (2006, henceforth GP).

Our main contribution characterizes when the anticipated choices align with the actual choices. The novelty of the axioms is in their restrictions on behavior *across* these domains. Two conditions are important. In the first, if the agent strictly prefers adding an option p to a menu A, i.e., $A \cup \{p\} \succ A$, then she must choose p with strictly positive probability over the options in A. The intuition for the second condition is converse. If p is chosen with strictly positive probability over the options in A, then adding it makes the decision maker strictly better off. However, there is a caveat to this intuition—the selection of p might be due to a tie-breaking procedure. That is, the agent might be indifferent between p and the other elements in A, but selects p after flipping a coin. The second condition controls for tie-breaking: if the selection of p from $A \cup \{p\}$ is not due to tie-breaking, then $A \cup \{p\}$ is strictly better than A alone.

When both conditions are satisfied, the decision maker correctly anticipates all possible payoff contingencies. In the context of menu choice alone, DLR add the following provision to their representation, on page 894:

> By assumption, we are representing an agent who cannot think of all (external) possibilities with an agent who has a coherent view of all payoff possibilities. If the agent does foresee the payoff possibilities, do we really have unforeseen contingencies? We remain agnostic on this point. The key idea is that we have allowed for the possibility of unforeseen contingencies by dropping the assumption of an exogenous state space and characterized the agent's subjective view of what might happen. Whether the agent actually fails to foresee any relevant situations is a different matter. It could be that our representation of the agent is quite literally correct—that is, the agent does in fact foresee the set of future utility possibilities and maximizes as in our representation.

By also considering choice from menus, we can assert a position on the agent's view of the possible taste contingencies. In fact, we can precisely characterize whether the agent foresees the relevant situations and the representation is "literally correct." Conversely, we can also distinguish when the agent fails to foresee a relevant situation.

Our second contribution is the unique identification of beliefs and utilities achieved by our joint representation: the combined data are consistent only with a single probability measure over subjective states and with a single state-dependent expected utility function. This improves the identification within either the DLR model or the GP model alone. In DLR, the belief over subjective states and the magnitudes of utilities are not unique. In contrast, expected utility models with objective states can uniquely identify beliefs and utilities. This identification affords these models a clean separation of beliefs and tastes. Such separation is lacking in the menu choice model. Alternative approaches to identifying beliefs in the DLR model have also been considered. For example, Sadowski (2010) enriches the DLR model with objective states, using beliefs over objective states to calibrate beliefs over subjective states. Schenone (2010) incorporates a monetary dimension to the prize space and assumes utility for money is state invariant in order to identify beliefs. Krishna and Sadowski (2011) use continuation values in a dynamic setting to pin down per-period beliefs in a recursive representation. Our approach is different. We augment the DLR model with random choice data, and the discipline of the choice data delivers the improved identification.

Our final contributions are to the model of menu choice and the model of random utility in themselves. The original representation of GP is not immediately comparable to DLR, as GP use a different state space and beliefs are defined with respect to a non-standard algebra.[1] We format the GP representation in a manner that enables a sensible direct comparison with the DLR representation. Even after suitable formatting, we need the respective state spaces to be finite for analytical tractability. In the supplementary appendix of this paper, we provide new theorems that characterize finite state space representations for both models. While related finite representation results exist for the menu choice model, to our knowledge we are the first to identify appropriate restrictions for the random choice setting.

Throughout the paper, we focus on the original interpretation of menu choice offered by Kreps and refined by DLR. Formally, we assume a monotonicity condition where larger menus are always weakly preferred to their subsets. Monotonicity excludes applications such as temptation or regret, which we hope to explore in future work.

2 The Model

Let Z be a finite set of alternatives containing at least 2 elements. Let $\triangle(Z)$ denote the set of all probability distributions on Z, endowed with the Euclidean metric d. We generally use p, q, r to denote arbitrary lotteries in $\triangle(Z)$. Let $\mathcal{A}$ denote the set of all nonempty, finite subsets of $\triangle(Z)$, endowed with the Hausdorff metric:

$$d_h(A,B) = \max\left\{ \max_{p\in A}\min_{q\in B} d(p,q), \max_{q\in B}\min_{p\in A} d(p,q) \right\}.$$

Elements of $\mathcal{A}$ are called menus, with generic menus denoted A, B, C.

Consider an individual who makes choices in two periods. In period 1, she chooses a menu A. In period 2, she makes a stochastic choice out of the menu A. As a result, her behavior is represented by two primitives corresponding to her choices in the two periods. To represent her preferences over menus in period 1, the first primitive is a binary relation $\succsim$ over $\mathcal{A}$. To represent her stochastic choices in period 2, the second primitive is a function that associates each menu with a probability distribution over its elements. To describe the second primitive formally, we need to introduce some additional notation.

Let $\triangle(\triangle(Z))$ denote the set of all simple probability distributions on $\triangle(Z)$. Let

[1] GP consider the minimal algebra on utility vectors that separates utility vectors that have different maximizers for some menu. As a result, their algebra does not separate a utility vector from its affine transformations.

$\lambda^A \in \triangle(\triangle(Z))$ denote the individual's random choice behavior in period 2 when facing the menu A. The term $\lambda^A(B)$ denotes the probability that the individual chooses a lottery in the set $B \in \mathcal{A}$ when facing menu $A \in \mathcal{A}$. To respect feasibility, only available lotteries can be chosen with strictly positive probability, so $\lambda^A(A) = 1$ for any menu $A \in \mathcal{A}$. Formally:

Definition 1. A *Random Choice Rule (RCR)* is a function $\lambda : \mathcal{A} \to \triangle(\triangle(Z))$ such that $\lambda^A(A) = 1$ for any menu $A \in \mathcal{A}$.

Note that expected-utility functions on $\triangle(Z)$ are equivalent to vectors in $\mathbb{R}^Z$, by associating each expected-utility function with its values for sure outcomes. The following notation will be useful in the sequel. For any menu $A \in \mathcal{A}$ and expected-utility function $u \in \mathbb{R}^Z$, let $M(A, u)$ denote the maximizers of u in A:

$$M(A, u) = \left\{ p \in A : u(p) = \max_{q \in A} u(q) \right\}.$$

We take the pair $(\succsim, \lambda)$ as primitive. In the next two subsections, we formally introduce the DLR representation for $\succsim$ and the GP representation for λ.

2.1 Preference over Menus: The DLR Representation

We will assume that preferences over menus comply with the canonical DLR representation with a finite subjective state space.

Definition 2. A *DLR representation* of $\succsim$ is a triple (S, U, μ) where S is a finite state space, $U : S \times \triangle(Z) \to \mathbb{R}$ is a state-dependent expected-utility function, and μ is a probability distribution on S, such that:

1. $A \succsim B$ if and only if $V(A) \geq V(B)$ where $V : \mathcal{A} \to \mathbb{R}$ is defined by

$$V(A) = \sum_{s \in S} \mu(s) \max_{p \in A} U_s(p). \tag{1}$$

2. *Non-redundancy*: For any two distinct states $s, s' \in S$, U_s and U'_s do not represent the same vNM preference on $\triangle(Z)$.

3. *Minimality*: $\mu(s) > 0$ and U_s is nonconstant for all $s \in S$.

The standard interpretation of Equation (1) is the following. In period 1, when the individual looks ahead to period 2, she anticipates that a subjective state $s \in S$ will realize. She believes that the probability of each state is given by μ. If a particular state $s \in S$ realizes, then her expected-utility function over $\triangle(Z)$ will be U_s. She also anticipates that she will make a choice out of the menu A in period 2 after learning the realized state, selecting an alternative in A maximizing her utility U_s conditional on each state $s \in S$. The value $V(A)$ in Equation (1) is her ex ante expected utility of menu A in period 1, before learning the period 2 state.

2.2 Random Choice Rule: The GP Representation

Taking the interpretation given for the DLR representation one step further, suppose that the individual's subjective model (S, U, μ) about the distribution of her second period tastes is correct. Suppose also that A is a menu such that U_s has a unique maximizer in A for each $s \in S$. Then, for any $p \in A$ the probability that the individual chooses p from A in period 2 is the probability of the event that U_s is maximized by p in A. That is, if we let λ^A denote the stochastic choices of this individual from menu A, then:

$$\lambda^A(p) = \mu(\{s \in S : p \in M(A, U_s)\}).$$

Note that if there exists a state s for which there is more than one maximizer of U_s in A, then the state-dependent utility function does not lead to a unique choice conditional on s. Hence, the DLR representation (S, U, μ) is not enough to determine a unique stochastic choice over the menu A. We will follow Gul and Pesendorfer (2006) and address this problem by assuming that in the case of indifferences, the individual uses a tie-breaking rule to chose a lottery in $M(A, U_s)$. To formalize the notion of a tie-breaker, we need to introduce some additional notation.

Let the set of *normalized (non-constant) expected-utility functions* on $\triangle(Z)$ be

$$\mathcal{U} = \left\{ u \in \mathbb{R}^Z : \sum_{z \in Z} u_z = 0, \sum_{z \in Z} u_z^2 = 1 \right\},$$

and endow $\mathcal{U}$ with its relative topology in $\mathbb{R}^Z$. Let $\mathbf{B}_{\mathcal{U}}$ denote the Borel σ-algebra of $\mathcal{U}$, and let $\triangle^f(\mathcal{U})$ denote the set of all finitely-additive probability measures over $(\mathcal{U}, \mathbf{B}_{\mathcal{U}})$. Given a state s, we will model the tie-breaking conditional on s as a probability distribution $\tau_s \in \triangle^f(\mathcal{U})$.[2] The interpretation is that the individual will choose a maximizer $p \in M(A, U_s)$ with the probability that the tie-breaking distribution draws a utility function $u \in \mathcal{U}$ such that p maximizes the realized tie-breaker u on $M(A, U_s)$.

[2]The modeling choice that τ_s is a probability distribution over normalized expected-utility functions $\mathcal{U}$ instead of all expected utility functions $\mathbb{R}^Z$ is without loss of generality.

Definition 3. Given a finite state space S, a *tie-breaking rule* for S is a function $\tau : S \to \triangle^f(\mathcal{U})$ satisfying the following regularity condition for all $A \in \mathcal{A}$ and $p \in A$:

$$\tau_s\Big(\Big\{u \in \mathcal{U} : u(p) > u(q),\ \forall q \in A \setminus \{p\}\Big\}\Big) = \tau_s\Big(\Big\{u \in \mathcal{U} : u(p) = \max_{q \in A} u(q)\Big\}\Big).$$

The regularity condition in the above definition ensures that the tie-breaking rule does not itself lead to ties with positive probability. The following representation corresponds to a special case of the tie-breaker representation from Gul and Pesendorfer (2006, supplementary material) where the state space is finite.

Definition 4. A *GP representation* of λ is a quadruple (S, U, μ, τ) where S is a finite state space, $U : S \times \triangle(Z) \to \mathbb{R}$ is a state-dependent utility function, μ is a probability distribution on S, and τ is a tie-breaking rule over S such that:

1. For every $A \in \mathcal{A}$ and $p \in A$,

$$\lambda^A(p) = \sum_{s \in S} \mu(s)\tau_s\big(\big\{u \in \mathcal{U} : p \in M(M(A, U_s), u)\big\}\big). \tag{2}$$

2. *Non-redundancy*: For any two distinct states $s, s' \in S$, U_s and U'_s do not represent the same vNM preference on $\triangle(Z)$.

3. *Minimality*: $\mu(s) > 0$ and U_s is nonconstant for all $s \in S$.

Equation (2) formalizes the two-stage maximization procedure described earlier. In the event that the realized utility U_s admits multiple maximizers $M(A, U_s)$, the agent uses a tie-breaker u given by the tie-breaking distribution τ_s. The total probability of choice is thus the joint probability, summed over all possible states, of p surviving both stages: first being optimal for the realized utility U_s, and then being optimal for the realized tie-breaker u.

2.3 Putting the DLR and GP Representations Together

We now introduce our desired joint representation. The decision maker has a prior μ over a subjective state space S of taste contingencies, and U describes the dependence of expected utilities on subjective states. It is essential that the same state space, belief, and utilities can be used to represent both the preference $\succsim$ and the random choice rule λ. In this case, the decision maker perfectly predicts the random choices she will make in the second stage and how much utility these choices will provide. She correspondingly evaluates menus in the first stage according to these correct predictions.

Definition 5. A *DLR-GP representation* of $(\succsim, \lambda)$ is a quadruple (S, U, μ, τ) where (S, U, μ) is a DLR representation of $\succsim$ and (S, U, μ, τ) is a GP representation of λ.

An arbitrary pair of menu preferences and a random choice rule will not generally admit a DLR-GP representation, even if $\succsim$ and λ themselves admit DLR and GP representations, respectively. A single unified DLR-GP representation requires that the behaviors be consistent. Specifically, some restrictions must be placed across the DLR and the GP setting, connecting the inferences regarding anticipation from the menu choice stage with the inferences regarding actualized taste contingencies from random choice stage.

Axiom 1. $A \cup \{p\} \succ A \implies \lambda^{A\cup\{p\}}(p) > 0.$

Axiom 1 requires that if adding the lottery p makes a menu A more appealing, then p must have some chance of being chosen from the menu $A \cup \{p\}$. So, potential flexibility that she will never exercise does not leave the decision maker any better off. This means that availability of options is welfare increasing in only an instrumental sense; availability per se is not desirable.

Now consider the following condition.

Axiom 1* (Consequentialism)**.** $\lambda^A = \lambda^B \implies A \sim B.$

Consequentialism implies that the agent only cares about the distribution of choices that are induced by a menu. For example, options outside the support of λ^A that are never chosen are irrelevant. Some applications of menu choice are at odds with this assumption. For example, in the model of temptation by Gul and Pesendorfer (2001), tempting but unchosen elements can make the agent worse off by forcing her to exert self-control.

Consequentialism relates to classic choice axioms such as independence of irrelevant alternatives or the weak axiom of revealed preference. A standard revealed-preference interpretation of these axioms is that the dominated or unchosen options in a menu cannot alter the preference for or the enjoyment of chosen options. In some environments, such as those with temptations, this interpretation is suspect. Here, consequentialism brings this implicit hedonic inference out in the open: the menu choice stage captures the effect of the overall menu, and consequentialism suppresses any influence outside the direct experience of choice.

Proposition 1. *Suppose $\succsim$ satisfies weak order and DLR monotonicity, i.e., $A \precsim B$ whenever $A \subset B$, and λ satisfies GP monotonicity, i.e., $\lambda^A(p) \geq \lambda^B(p)$ whenever $p \in A \subset B$. Then Axiom 1 and consequentialism are equivalent.*

In particular, since DLR monotonicity and GP monotonicity are necessary for the assumed DLR and GP representations, consequentialism and Axiom 1 are identical in our environment.

The converse of Axiom 1 is that $\lambda^{A\cup\{p\}}(p) > 0$ implies $A \cup \{p\} \succ A$. In words, if there is a chance that an option is chosen from a menu, then removing that option leaves the decision maker strictly worse off. This condition might appear natural at first glance, but in fact is overly restrictive. It overlooks the tie-breaking rule that adjudicates among multiple maximizers. For example, suppose z and z' are copies of the same object, so the decision maker will certainly be indifferent between them. That is, $U_s(z) = U_t(z')$ for all states s, t. However, the tie-breaking rule τ randomizes between them whenever both are optimal elements of the menu. In this case, z' might be chosen from the menu $\{z, z'\}$, but its removal leaves the decision maker no worse off since she is indifferent between the options.

The hypothesis that $\lambda^{A\cup\{p\}}(p) > 0$ must therefore be strengthened to ensure that the selection of p from $A \cup \{p\}$ is not an artifact of tie-breaking. Specifically, we must ensure that p is not redundant in the sense of indifference. Given the implied continuity of the expected utility functions U_s, we can test whether p is redundant by perturbing A and p. If p is strictly preferred to elements of A in some state s, then this preference is maintained in neighborhoods about p and A. So not only is p chosen with positive probability when added to the menu A, but any q close to p is chosen with positive probability when added to any menu B close to A. Then the selection of p is not an artifact of the tie-breaking rule, and its addition to A leaves the decision maker strictly better off.

Axiom 2. *For any A and $p \notin A$, if there exists $\varepsilon > 0$ such that $\lambda^{B\cup\{q\}}(q) > 0$ whenever $d(p, q) < \varepsilon$ and $d_h(A, B) < \varepsilon$, then $A \cup \{p\} \succ A$.*

The following is our main representation result.

Theorem 1. *Suppose $\succsim$ has a DLR representation and λ has a GP representation. Then, the pair $(\succsim, \lambda)$ satisfies Axioms 1 and 2 if and only if it has a DLR-GP representation.*[3]

To illustrate the intuition for Theorem 1, suppose $\succsim$ has a DLR representation (S^1, U^1, μ^1) and λ has a GP representation $(S^2, U^2, \mu^2, \tau^2)$. The first key step in the

[3]Instead of assuming the preference and random choice rule respectively admit DLR and GP representations, we could list the basic axioms that characterize these representations. In the supplementary appendix of this paper, we verify that slight modifications of the original axioms from Dekel, Lipman, and Rustichini (2001) and Gul and Pesendorfer (2006) are equivalent to our versions of the DLR and GP representations.

proof consists of showing that Axiom 1 is equivalent to the following condition: For each $U^1_{s_1}$ in the DLR representation of $\succsim$, there is a corresponding $U^2_{s^2}$ in the GP representation of λ that is a positive affine transformation of $U^1_{s_1}$. Similarly, Axiom 2 is necessary and sufficient for each $U^2_{s^2}$ in the GP representation to have a corresponding positive affine transformation $U^1_{s^1}$ in the DLR representation.[4] Therefore, these axioms together imply that, subject to relabeling of the states, it is without loss of generality to assume the representations share a common state space $S^1 = S^2 = S$ such that U^1_s and U^2_s represent the same expected-utility preference for all $s \in S$.

In order to represent the RCR λ, we construct a DLR-GP representation that uses the probability measure μ^2 and tie-breaking rule τ^2 from the GP representation. Then, considering the latitude permitted in the uniqueness of the DLR representation, we can appropriately adjust the magnitudes of the state-dependent utility functions in the DLR representation so that the corresponding measure is now identical to the measure from the GP representation. Formally, in order to use μ^2 and represent the preference $\succsim$, we need to construct a state-dependent utility function U that satisfies $\mu^2(s)U_s(p) = \mu^1(s)U^1_s(p)$ for all $p \in \triangle(Z)$. In other words, we define $U_s(p) = \frac{\mu^1(s)}{\mu^2(s)}U^1_s(p)$. This implies (S, U, μ^2) is a DLR representation for $\succsim$. Also, since each U_s is an affine transformation of U^1_s which is itself an affine transformation of U^2_s, the tuple (S, U, μ^2, τ^2) is a GP representation for λ. Thus, we have constructed a DLR-GP representation for the pair $(\succsim, \lambda)$. For additional details, see the complete proof in Appendix A.3.

We now discuss the identification of beliefs and utilities in the model. By itself, either the DLR or the GP model leaves some feature of its representation open to a degree of freedom. The DLR representation does not uniquely identify either the belief over subjective states or the magnitudes of utilities across states. The nature of nonuniqueness is similar to the nonuniqueness of the prior in the general Anscombe–Aumann expected-utility model (where utility is permitted to be state-dependent). Preferences over menus cannot distinguish whether a subjective state is important because it is very likely to realize, so $\mu(s)$ is large, or because the utilities of the options are very different in that state, so $\|U_s\|$ is large.

The GP representation affords sharper identification of the belief μ over subjective states. In the random choice model, the belief μ is unique (up to relabeling). On the other hand, the utilities U_s remain identified only up to a state-dependent affine transformation, so there is no restriction on scalings across states. Hence, the magnitudes of utility differences cannot be compared across subjective states. This precludes ex ante welfare analysis, since it is not possible to determine whether adding a sometimes-

[4] In other words, if we identify each subjective state with its corresponding ex post preference, Axiom 1 ensures that the DLR state space can be embedded in the GP state space. Conversely, Axiom 2 ensures that the GP state space can be embedded in the DLR state space.

chosen alternative at the expense of another increases or decreases the agent's expected utility.

By considering both models, we identify a unique belief and a unique set of state-dependent utilities that are consistent with the joint data. As an analogy to our approach, the standard identification of the prior over objective states in the Anscombe–Aumann model privileges the normalization of beliefs where utilities are constant across states. Here, we privilege the normalization of first-stage beliefs that aligns with the statistical choices in the second stage. This normalization is arguably even more compelling since it hinges on a physical benchmark of observed choices, rather than an artificial benchmark such as state independence.

Theorem 2. *Two DLR-GP representations* $(S^1, U^1, \mu^1, \tau^1)$ *and* $(S^2, U^2, \mu^2, \tau^2)$ *represent the same pair* $(\succsim, \lambda)$ *if and only if there exists a bijection* $\pi : S^1 \to S^2$*, a scalar* $\alpha > 0$*, and a function* $\beta : S^1 \to \mathbb{R}$ *such that:*

(a) $U^1_{s^1}(p) = \alpha U^2_{\pi(s^1)}(p) + \beta(s^1)$ *for all* $p \in \triangle(Z)$ *and* $s^1 \in S^1$.

(b) $\mu^1(s^1) = \mu^2(\pi(s^1))$ *for all* $s^1 \in S^1$.

(c) $\tau^1_{s^1}(E) = \tau^2_{\pi(s^1)}(E)$ *for the set* $E = \{u \in \mathcal{U} : p \in M(M(A, U^1_{s^1}), u)\}$ *for every* $s^1 \in S^1$*,* $A \in \mathcal{A}$*, and* $p \in A$.

By considering both stages of choice, Theorem 2 provides strictly sharper identification of the utilities and beliefs than either the DLR or GP representation alone. This is possible since the DLR-GP representation uses the same belief and utility function to represent choice in both stages. The probability is pinned down from the random choice data due to the uniqueness of the belief in the GP model. And once we pin down the belief in the DLR model, the magnitudes of utilities across states are also identified. In contrast, for either the DLR or the GP representation in isolation, the utility U_s can be transformed by a state-dependent scale factor α_s. The scaling is not uniform across states, and hence neither the DLR nor the GP representation can distinguish the differences in utilities *across* states. In the joint representation, the transformation must be uniform across states since α does not depend on s. Thus, a statement such as "the additional utility for z above z' is greater in state s than in state t" is now meaningful. Such comparisons allow for counterfactual comparative statics in *ex post* welfare, after the state is realized. For example, a social planner can assess the welfare implications of removing an option from a menu, even after the agent's taste contingency is realized.

3 Unforeseen Contingencies

As mentioned in the introduction, a leading motivation for studying menu choice in general and preference for flexibility in particular is to generate a decision theoretic model of unforeseen contingencies. One interpretation of the subjective state space in the DLR representation is as the agent's personal view of the relevant uncertainty, summarized as payoff information, in a situation where her understanding of the world is otherwise incomplete. However, as the quotation from DLR suggested, while the representation can accommodate a decision maker who rationally resolves her incomplete understanding, it remains unclear how to test whether her understanding was in fact incomplete.

A fundamental difficulty with such an austere revealed-preference approach is that the agent cannot reveal completely unforeseen contingencies precisely because they are unforeseen to her. The power of the model is substantially improved by equipping the analysis with some additional information. On the other hand, simply assuming an exogenous "correct" state space might seem heavy-handed and runs into the objections that motivated the subjective approach in the first place. Instead, we use actual choices from menus to allow inference about the the complete set of states. This at least allows us to verify whether the agents anticipates all of the relevant taste contingencies.

Definition 6. An *Unforeseen Contingencies representation* of $(\succsim, \lambda)$ is a quintuple (S, T, U, μ, τ) where $S \subseteq T$, $(S, U|_S, \mu|_S)$ is a DLR representation of $\succsim$, and (T, U, μ, τ) is a GP representation of λ.

To interpret, the representation allows for the possibility that the agents overlooks some of the relevant taste contingencies, namely those in $T \setminus S$. However, her likelihood assessments among the foreseen contingencies are consistent with their actual frequencies, since the conditional probability $\mu|_S$ represents her beliefs at the menu choice stage over the anticipated subjective states in S. This representation therefore captures a specific type of unforeseen contingencies—*completely unforeseen contingencies.* The contingencies in S are perfectly foreseen, whereas those in $T \setminus S$ are completely unforeseen.[5]

Theorem 3. *Suppose $\succsim$ has a DLR representation and λ has a GP representation. Then, the pair $(\succsim, \lambda)$ satisfies Axiom 1 if and only if it has an Unforeseen Contingencies representation.*

[5]More realistically, an agent might have a partial sense that her understanding at the menu choice stage is incomplete, which suggests a more subtle relationship between the state spaces in the DLR and GP representations. However, a full discussion of how to model partially unforeseen contingencies is beyond our scope.

In view of Proposition 1, consequentialism and an Unforeseen Contingencies representation are equivalent.

The following result summarizes the uniqueness properties of the Unforeseen Contingencies representation. Since the subjective state spaces are not perfectly aligned, we lose some identification for states in $T \setminus S$.

Theorem 4. *Two Unforeseen Contingencies representations* $(S^1, T^1, U^1, \mu^1, \tau^1)$ *and* $(S^2, T^2, U^2, \mu^2, \tau^2)$ *represent the same pair* $(\succsim, \lambda)$ *if and only if there exists a bijection* $\pi : T^1 \to T^2$*, a scalar* $\bar{\alpha} > 0$*, and functions* $\alpha : T^1 \setminus S^1 \to (0, \infty)$ *and* $\beta : T^1 \to \mathbb{R}$ *such that:*

(a) $U^1_{s^1}(p) = \bar{\alpha} U^2_{\pi(s^1)}(p) + \beta(s^1)$ *for all* $p \in \triangle(Z)$ *and* $s^1 \in S^1$*, and* $\pi(S^1) = S^2$*.*

(b) $U^1_{t^1}(p) = \alpha(t^1) U^2_{\pi(t^1)}(p) + \beta(t^1)$ *for all* $p \in \triangle(Z)$ *and* $t^1 \in T^1 \setminus S^1$*.*

(c) $\mu^1(t^1) = \mu^2(\pi(t^1))$ *for all* $t^1 \in T^1$*.*

(d) $\tau^1_{t^1}(E) = \tau^2_{\pi(t^1)}(E)$ *for the set* $E = \{u \in \mathcal{U} : p \in M(M(A, U^1_{t^1}), u)\}$ *for every* $t^1 \in T^1$*,* $A \in \mathcal{A}$*, and* $p \in A$*.*

The identification of beliefs and utilities on S, the set of anticipated subjective states, is unique. On the set $T \setminus S$ of overlooked subjective states, the probabilities are still uniquely identified by the GP stage of random choice, but the magnitudes of utility differences are unidentified because there is no relevant menu preference with which to compare.

One potential limitation of developing a decision-theoretic foundation for unforeseen contingencies is the observability of the primitive. A standard criticism of any exercise in the elicitation of preferences is the large number of observations required to verify that the posited axioms are satisfied. Our application to unforeseen contingencies arguably sharpens this criticism: the agent should learn about a contingency after it realizes, and hence preferences over menus may not be stationary over time. While it is important to keep this possible issue in mind when interpreting Theorems 3 and 4, this problem seems endemic to any choice-based model of unforeseen contingencies.

On a final note, one could consider the opposite relation, where the anticipated state space S in the first stage is a superset of the state space T realized at the point of choice in the second stage. Such behavior suggests a pure preference for freedom of choice, since the agent may desire additional options that she never actually chooses.[6] It is easy to show that Axiom 2 characterizes this set containment where $S \supseteq T$. We leave additional analysis of the resulting representation as a question for future work.

[6] Barberà, Bossert, and Pattanaik (2004) nicely summarize the literature on preference for freedom of choice.

A Proofs

A.1 A Preliminary Result

The following lemma shows that for any finite set of (nonconstant) expected-utility functions F, there is a finite set of lotteries A such that each utility function in F has a unique maximizer in A, and moreover, two utility functions share the same maximizer in A if and only if they represent the same preference over lotteries. It will be used in the proofs of both the representation and uniqueness results.

Lemma 1. *Fix any finite set F of nonconstant expected-utility functions on $\triangle(Z)$.*[7] *Then, there exists $A \in \mathcal{A}$ such that:*

1. *For each $u \in F$ there is a lottery $p \in A$ such that $u(p) > u(q)$ for all $q \in A \setminus \{p\}$.*
2. *If any two $u, v \in F$ share the same maximizer in A, then they represent the same expected-utility preference: If for some $p \in A$, $u(p) \geq u(q)$ and $v(p) \geq v(q)$ for all $q \in A$, then there exists $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u(r) = \alpha v(r) + \beta$ for all $r \in \triangle(Z)$.*

Proof. Fix any $\delta \in (0, \frac{1}{|Z|})$, let $\bar{p} = (\frac{1}{|Z|}, \ldots, \frac{1}{|Z|})$ denote the uniform distribution, and let $C_\delta(\bar{p}) \equiv \{p \in \triangle(Z) : \|p - \bar{p}\| \leq \delta\}$. Consider first any $u \in \mathcal{U}$. That is, suppose $u \in \mathbb{R}^Z$ satisfies $\sum_{z \in Z} u_z = 0$ and $\|u\| = 1$. By the Cauchy-Schwarz inequality, for any $p \in C_\delta(\bar{p})$,

$$u \cdot p = u \cdot (p - \bar{p}) \leq \|u\| \cdot \|p - \bar{p}\| \leq \delta,$$

with equality if and only if $p - \bar{p} = \alpha u$ for some $\alpha > 0$ and $\|p - \bar{p}\| = \delta$. These two conditions uniquely identify the lottery $p^u = \bar{p} + \delta u$, and therefore $\delta = u(p^u) > u(q)$ for all $q \in C_\delta(\bar{p}) \setminus \{p^u\}$.[8]

Now, consider a finite set F as in the statement of the lemma. For each $u \in F$, there exists a unique $\hat{u} \in \mathcal{U}$ that represents the same expected-utility preference as u. Formally, representing expected-utility functions as vectors in $\mathbb{R}^Z$ and letting $\mathbf{1} = (1, \ldots, 1)$ denote the unit vector, the function $\hat{u}$ is given by

$$\hat{u} = \frac{u - \left(\frac{1}{|Z|} \sum_{z \in Z} u_z\right)\mathbf{1}}{\|u - \left(\frac{1}{|Z|} \sum_{z \in Z} u_z\right)\mathbf{1}\|}. \tag{3}$$

By construction, $\hat{u}$ is well-defined and in $\mathcal{U}$ for any nonconstant $u \in \mathbb{R}^Z$, and since $\hat{u}$ is simply an affine transformation of u, it represents the same expected-utility preference. Therefore, take $p^u = \bar{p} + \delta \hat{u}$. By the preceding arguments and the fact that u and $\hat{u}$ represent the same

[7] That is, F is a finite subset of $\{u \in \mathbb{R}^Z : u_z \neq u_{z'}$ for some $z, z' \in Z\}$.

[8] Note that $p^u \in C_\delta(\bar{p})$. First, since $\sum_{z \in Z} u_z = 0$, we have $\sum_{z \in Z} p^u_z = \sum_{z \in Z} \bar{p}_z = 1$. Second, since $|p^u_z - \bar{p}_z| \leq \|p^u - \bar{p}\| = \delta < \frac{1}{|Z|}$, we have $p^u_z \geq 0$ for all $z \in Z$.

preference over lotteries, $u(p^u) > u(q)$ for all $q \in C_\delta(\bar{p}) \setminus \{p^u\}$. Define p^u in this manner for each $u \in F$, and let $A = \{p^u : u \in F\}$. Then, for any $u \in F$, $u(p^u) > u(q)$ for all $q \in A \setminus \{p^u\}$, proving part 1.

For part 2, suppose that for some $u, v \in F$ and $p \in A$, $u(p) \geq u(q)$ and $v(p) \geq v(q)$ for all $q \in A$. Then, by the definition of A, it must be that $p = p^u = p^v$. By the definition of p^u and p^v, this implies u and v are both positive affine transformations of $p - \bar{p}$. □

A.2 Proof of Proposition 1

First suppose Axiom 1 holds. Let $\lambda^A = \lambda^B$. Let C denote the common support of these two measures. By GP monotonicity, $\lambda^C = \lambda^A = \lambda^B$. Enumerate the finite set $A \setminus C$ as $\{p_1, \ldots, p_n\}$. GP monotonicity implies $\lambda^{C \cup \{p_1\}}(C) \geq \lambda^A(C) = 1$, and hence $\lambda^{C \cup \{p_1\}}(p_1) = 0$. Considering the contraposition of Axiom 1, $C \succsim C \cup \{p_1\}$. By DLR monotonicity, this forces $C \sim C \cup \{p_1\}$. Proceeding inductively, we conclude

$$C \sim C \cup \{p_1\} \sim C \cup \{p_1\} \cup \{p_2\} \sim \cdots \sim C \cup \{p_1, \ldots, p_n\} = A.$$

Similarly, $C \sim B$. Hence $A \sim B$ by transitivity.

Now suppose consequentialism holds. Let $\lambda^{A \cup \{p\}}(p) = 0$. By GP monotonicity, $\lambda^{A \cup \{p\}} = \lambda^A$. But by consequentialism, $A \cup \{p\} \sim \{A\}$, that is, it cannot be that $A \cup \{p\} \succ A$. This demonstrates that Axiom 1 holds contrapositively.

A.3 Proof of Theorem 1

The following result summarizes the key implications of Axioms 1 and 2 that will then be used to prove Theorem 1. We write $U^1_{s^1} \approx U^2_{s^2}$ to indicate that $U^1_{s^1}$ is a positive affine transformation of $U^2_{s^2}$.

Proposition 2. *Suppose $\succsim$ has a DLR representation (S^1, U^1, μ^1) and λ has a GP representation $(S^2, U^2, \mu^2, \tau^2)$.*

1. *The pair $(\succsim, \lambda)$ satisfies Axiom 1 if and only if for every $s^1 \in S^1$ there exists $s^2 \in S^2$ such that $U^1_{s^1} \approx U^2_{s^2}$.*
2. *The pair $(\succsim, \lambda)$ satisfies Axiom 2 if and only if for every $s^2 \in S^2$ there exists $s^1 \in S^1$ such that $U^1_{s^1} \approx U^2_{s^2}$.*

Proof. Part 1 — only if: Suppose $(\succsim, \lambda)$ satisfies Axiom 1. Consider the set of utility functions $F = \{U^1_{s^1} : s^1 \in S^1\} \cup \{U^2_{s^2} : s^2 \in S^2\}$ and take $A \in \mathcal{A}$ as described in Lemma 1. Fix any $s^1 \in S^1$. By part 1 of Lemma 1, there exists $p \in A$ such that $U^1_{s^1}(p) > U^1_{s^1}(q)$ for all $q \in A \setminus \{p\}$. Therefore, the definition of the DLR representation implies $A \succ A \setminus \{p\}$, which by Axiom 1 implies $\lambda^A(p) > 0$. By the definition of the GP representation, this requires that $U^2_{s^2}(p) = \max_{q\in A} U^2_{s^2}(q)$ for some $s^2 \in S^2$. By part 2 of Lemma 1, this implies $U^1_{s^1} \approx U^2_{s^2}$.

Part 1 — if: Suppose for every $s^1 \in S^1$ there exists $s^2 \in S^2$ such that $U^1_{s^1} \approx U^2_{s^2}$. To see that Axiom 1 is satisfied, fix any $A \in \mathcal{A}$ and $p \in \triangle(Z)$ such that $A \cup \{p\} \succ A$. Thus, $U^1_{s^1}(p) > \max_{q\in A} U^1_{s^1}(q)$ for some $s^1 \in S^1$. Take $s^2 \in S^2$ such that $U^1_{s^1} \approx U^2_{s^2}$. Then, $U^2_{s^2}(p) > \max_{q\in A} U^2_{s^2}(q)$. The definition of the GP representation therefore implies $\lambda^{A\cup\{p\}}(p) \geq \mu^2(s^2) > 0$.

Part 2 — only if: Suppose $(\succsim, \lambda)$ satisfies Axiom 2. Consider the set of utility functions $F = \{U^1_{s^1} : s^1 \in S^1\} \cup \{U^2_{s^2} : s^2 \in S^2\}$ and take $A \in \mathcal{A}$ as described in Lemma 1. Fix any $s^2 \in S^2$. By part 1 of Lemma 1, there exists $p \in A$ such that $U^2_{s^2}(p) > U^2_{s^2}(q)$ for all $q \in A \setminus \{p\}$. Since A is finite and $U^2_{s^2}$ is continuous, this implies there exists $\varepsilon > 0$ such that $U^2_{s^2}(r) > U^2_{s^2}(q')$ whenever $d(p, r) < \varepsilon$ and $d(q, q') < \varepsilon$ for $q \in A \setminus \{p\}$.

Fix any r and B such that $d(p, r) < \varepsilon$ and $d_h(A \setminus \{p\}, B) < \varepsilon$. Then, by the definition of the Hausdorff metric, for any $q' \in B$, there exists $q \in A \setminus \{p\}$ such that $d(q, q') < \varepsilon$. Hence, $U^2_{s^2}(r) > U^2_{s^2}(q')$ for all $q' \in B$. The definition of the GP representation then implies $\lambda^{B\cup\{r\}}(r) \geq \mu^2(s^2) > 0$. Since this is true for any r and B satisfying $d(p, r) < \varepsilon$ and $d_h(A\setminus\{p\}, B) < \varepsilon$, Axiom 2 implies $A \succ A\setminus\{p\}$. By the definition of the DLR representation, this requires that $U^1_{s^1}(p) = \max_{q\in A} U^1_{s^1}(q)$ for some $s^1 \in S^1$. By part 2 of Lemma 1, this implies $U^1_{s^1} \approx U^2_{s^2}$.

Part 2 — if: Suppose for every $s^2 \in S^2$ there exists $s^1 \in S^1$ such that $U^1_{s^1} \approx U^2_{s^2}$. To see that Axiom 2 is satisfied, fix any $A \in \mathcal{A}$, $p \notin A$, and $\varepsilon > 0$ such that $\lambda^{B\cup\{q\}}(q) > 0$ whenever $d(p, q) < \varepsilon$ and $d_h(A, B) < \varepsilon$. To show that $A \cup \{p\} \succ A$, it suffices to show there exists $s^1 \in S^1$ such that $U^1_{s^1}(p) > \max_{q\in A} U^1_{s^1}(q)$. Given the assumed relationship between the DLR and GP representations, this can be established by showing there exists $s^2 \in S^2$ such that $U^2_{s^2}(p) > \max_{q\in A} U^2_{s^2}(q)$.

We establish this inequality by showing there is a contradiction if $U^2_{s^2}(p) \leq \max_{q\in A} U^2_{s^2}(q)$ for all $s^2 \in S^2$. Intuitively, if this weak inequality holds for all s^2, then p and A can be perturbed slightly to make the inequality strict for all s^2, contradicting the assumption that $\lambda^{B\cup\{q\}}(q) > 0$ whenever $d(p, q) < \varepsilon$ and $d_h(A, B) < \varepsilon$. Formally, for each $s^2 \in S^2$, let $q^{s^2} \in A$ be such that $U^2_{s^2}(q^{s^2}) = \max_{q\in A} U^2_{s^2}(q)$, and let $r^{s^2} \in \triangle(Z)$ be such that $U^2_{s^2}(r^{s^2}) = \max_{r\in\triangle(Z)} U^2_{s^2}(r)$. Fix any $\alpha \in (0, \varepsilon/\sqrt{|Z|})$, and let $B = A \cup \{\alpha r^{s^2} + (1-\alpha)q^{s^2} : s^2 \in S^2\}$. Since for all $s^2 \in S^2$,

$$\left\|\left(\alpha r^{s^2} + (1-\alpha)q^{s^2}\right) - q^{s^2}\right\| = \alpha\left\|r^{s^2} - q^{s^2}\right\| \leq \alpha\sqrt{|Z|} < \varepsilon,$$

it follows that $d_h(A, B) < \varepsilon$. Let $q = \alpha\bar{p} + (1-\alpha)p$, where $\bar{p} = (\frac{1}{|Z|}, \ldots, \frac{1}{|Z|})$ denotes the uniform distribution. Then, $d(p, q) < \varepsilon$. Since each $U^2_{s^2}$ is nonconstant, we have $U^2_{s^2}(\bar{p}) <$

$U^2_{s^2}(r^{s^2})$ for all $s^2 \in S^2$, and hence

$$\begin{aligned} U^2_{s^2}(q) &= \alpha U^2_{s^2}(\bar{p}) + (1-\alpha) U^2_{s^2}(p) \\ &< \alpha U^2_{s^2}(r^{s^2}) + (1-\alpha) U^2_{s^2}(q^{s^2}) \\ &= U^2_{s^2}(\alpha r^{s^2} + (1-\alpha) q^{s^2}) = \max_{r \in B} U^2_{s^2}(r). \end{aligned}$$

This implies $\lambda^{B \cup \{q\}}(q) = 0$, a contradiction. □

Proof of Theorem 1. Suppose $(\succsim, \lambda)$ has a DLR-GP representation (S, U, μ, τ). By definition, this implies (S, U, μ) is a DLR representation for $\succsim$ and (S, U, μ, τ) is a GP representation for λ. Proposition 2 therefore implies that the pair $(\succsim, \lambda)$ satisfies Axioms 1 and 2.

Conversely, suppose $\succsim$ has a DLR representation (S^1, U^1, μ^1) and λ has a GP representation $(S^2, U^2, \mu^2, \tau^2)$, and suppose the pair $(\succsim, \lambda)$ satisfies Axioms 1 and 2. By Proposition 2, for each $s^1 \in S^1$ there exists $s^2 \in S^2$ such that $U^1_{s^1} \approx U^2_{s^2}$, and for each $s^2 \in S^2$ there exists $s^1 \in S^1$ such that $U^1_{s^1} \approx U^2_{s^2}$. Together with the non-redundancy assumptions in the DLR and GP representations, these conditions are sufficient for there to exist a bijection $\pi : S^1 \to S^2$ such that $U^1_{s^1} \approx U^2_{\pi(s^1)}$ for all $s^1 \in S^2$. Therefore, subject to relabeling, it is without loss of generality to assume that $S^1 = S^2 = S$ for some finite state space S, and $U^1_s \approx U^2_s$ for all $s \in S$.

For each $s \in S$, define $U_s : \triangle(Z) \to \mathbb{R}$ by $U_s(p) = \frac{\mu^1(s)}{\mu^2(s)} U^1_s(p)$ for $p \in \triangle(Z)$. Since $\mu^2(s) > 0$ for all $s \in S$, these functions are well-defined. Also, since $\mu^1(s) > 0$ for all $s \in S$, each U_s is nonconstant. Therefore, U inherits the non-redundancy property from U^1. Finally, let $\mu = \mu^2$ and $\tau = \tau^2$. We claim that (S, U, μ, τ) is a DLR-GP representation for $(\succsim, \lambda)$. The tuple (S, U, μ) is a DLR representation for $\succsim$ since

$$\sum_{s \in S} \mu(s) \max_{p \in A} U_s(p) = \sum_{s \in S} \mu^2(s) \max_{p \in A} \left(\frac{\mu^1(s)}{\mu^2(s)} U^1_s(p) \right) = \sum_{s \in S} \mu^1(s) \max_{p \in A} U^1_s(p).$$

The tuple (S, U, μ, τ) is a GP representation for λ since

$$\begin{aligned} \lambda^A(p) &= \sum_{s \in S} \mu^2(s) \tau^2_s(\{u \in \mathcal{U} : p \in M(M(A, U^2_s), u)\}) \\ &= \sum_{s \in S} \mu(s) \tau_s(\{u \in \mathcal{U} : p \in M(M(A, U_s), u)\}), \end{aligned}$$

where the last equality follows because $U_s \approx U^1_s \approx U^2_s$ implies $M(A, U_s) = M(A, U^2_s)$. This completes the sufficiency part of the proof. □

A.4 Proof of Theorem 2

The next two propositions describe the uniqueness properties of the DLR and GP representations, respectively. Combining the two results yields the asserted uniqueness properties of the DLR-GP representation.

Proposition 3. *Two DLR representations* (S^1, U^1, μ^1) *and* (S^2, U^2, μ^2) *represent the same preference if and only if there exists a bijection* $\pi : S^1 \to S^2$*, a constant* $c > 0$*, and functions* $\alpha : S^1 \to (0, \infty)$ *and* $\beta : S^1 \to \mathbb{R}$ *such that:*

(a) $U^1_{s^1}(p) = \alpha(s^1) U^2_{\pi(s^1)}(p) + \beta(s^1)$ *for all* $p \in \triangle(Z)$ *and* $s^1 \in S^1$.

(b) $\mu^1(s^1) = \frac{c}{\alpha(s^1)} \mu^2(\pi(s^1))$ *for all* $s^1 \in S^1$.[9]

Proposition 4. *Two GP representations* $(S^1, U^1, \mu^1, \tau^1)$ *and* $(S^2, U^2, \mu^2, \tau^2)$ *represent the same RCR if and only if there exists a bijection* $\pi : S^1 \to S^2$ *and functions* $\alpha : S^1 \to (0, \infty)$ *and* $\beta : S^1 \to \mathbb{R}$ *such that:*

(a) $U^1_{s^1}(p) = \alpha(s^1) U^2_{\pi(s^1)}(p) + \beta(s^1)$ *for all* $p \in \triangle(Z)$ *and* $s^1 \in S^1$.

(b) $\mu^1(s^1) = \mu^2(\pi(s^1))$ *for all* $s^1 \in S^1$.

(c) $\tau^1_{s^1}(E) = \tau^2_{\pi(s^1)}(E)$ *for the set* $E = \{u \in \mathcal{U} : p \in M(M(A, U^1_{s^1}), u)\}$ *for every* $s^1 \in S^1$, $A \in \mathcal{A}$*, and* $p \in A$.

Theorem 2 follows directly from Propositions 3 and 4. Briefly, the uniqueness of the probability measure in the DLR-GP representation follows from condition (b) of Proposition 4. Using this fact together with condition (b) of Proposition 3 implies that the state-dependent scalar multiple $\alpha(s_1)$ in condition (a) of Proposition 3 must in fact be constant (state-independent).

A.4.1 Proof of Proposition 3

Fix two DLR representations (S^1, U^1, μ^1) and (S^2, U^2, μ^2). It is easy to see that conditions (a) and (b) imply these represent the same preference.

Conversely, suppose (S^1, U^1, μ^1) and (S^2, U^2, μ^2) represent the same preference. Define the set of normalized expected-utility functions $\mathcal{U}$ as above. For $i = 1, 2$ and any $s^i \in S^i$, since $U^i_{s^i}$ is nonconstant, there exists $a^i_{s^i} > 0$, $b^i_{s^i} \in \mathbb{R}$, and $u^i_{s^i} \in \mathcal{U}$ such that $U^i_{s^i}(p) = a^i_{s^i} u^i_{s^i}(p) + b^i_{s^i}$ for all $p \in \triangle(Z)$ (see Equation (3) for a formal description of the mapping from $U^i_{s^i}$ to $u^i_{s^i}$). Define probability measures η^i for $i = 1, 2$ with finite support on $\mathcal{U}$ by

$$\eta^i(u^i_{s^i}) = \frac{\mu^i(s^i) a^i_{s^i}}{\bar{a}^i}, \qquad s^i \in S^i,$$

[9]Since μ^1 and μ^2 are probability measures, it follows that $c = \sum_{s^1 \in S^1} \alpha(s^1) \mu^1(s^1)$.

where $\bar{a}^i = \sum_{s^i \in S^i} \mu^i(s^i) a^i_{s^i}$. Notice that for $i = 1, 2$:

$$\begin{aligned}\sum_{s^i \in S^i} \mu^i(s^i) \max_{p \in A} U^i_{s^i}(p) &= \sum_{s^i \in S^i} \mu^i(s^i) a^i_{s^i} \max_{p \in A} u^i_{s^i}(p) + \sum_{s^i \in S^i} \mu^i(s^i) b^i_{s^i} \\ &= \bar{a}^i \int_{\mathcal{U}} \max_{p \in A} u(p)\, \eta^i(du) + \sum_{s^i \in S^i} \mu^i(s^i) b^i_{s^i},\end{aligned}$$

and hence

$$A \succsim B \iff \int_{\mathcal{U}} \max_{p \in A} u(p)\, \eta^i(du) \geq \int_{\mathcal{U}} \max_{p \in B} u(p)\, \eta^i(du).$$

Thus, both η^1 and η^2 represent $\succsim$ in the sense of Theorem 1 in the supplementary appendix of this paper, so the uniqueness part of that theorem implies $\eta^1 = \eta^2$. In particular, $\text{supp}(\eta^1) = \text{supp}(\eta^2)$, which (together with the non-redundancy of U^1 and U^2 and the strict positivity of μ^1 and μ^2) implies there exists a bijection $\pi : S^1 \to S^2$ such that $u^1_{s^1} = u^2_{\pi(s^1)}$ for all $s^1 \in S^1$. Thus, letting $c = \bar{a}^1/\bar{a}^2$ and $\alpha(s^1) = a^1_{s^1}/a^2_{\pi(s^1)}$ for $s^1 \in S^1$ yields condition (b):

$$\mu^1(s^1) = \frac{\bar{a}^1}{a^1_{s^1}} \eta^1(u^1_{s^1}) = \frac{\bar{a}^1}{a^1_{s^1}} \eta^2(u^2_{\pi(s^1)}) = \frac{\bar{a}^1}{a^1_{s^1}} \frac{a^2_{\pi(s^1)}}{\bar{a}^2} \mu^2(\pi(s^1)) = \frac{c}{\alpha(s^1)} \mu^2(\pi(s^1)).$$

Letting $\beta(s^1) = b^1_{s^1} - (a^1_{s^1}/a^2_{\pi(s^1)}) b^2_{\pi(s^1)}$ for $s^1 \in S^1$ yields condition (a):

$$\begin{aligned}\alpha(s^1) U^2_{\pi(s^1)}(p) + \beta(s^1) &= \frac{a^1_{s^1}}{a^2_{\pi(s^1)}} \left[a^2_{\pi(s^1)} u^2_{\pi(s^1)}(p) + b^2_{\pi(s^1)} \right] + \left[b^1_{s^1} - \frac{a^1_{s^1}}{a^2_{\pi(s^1)}} b^2_{\pi(s^1)} \right] \\ &= a^1_{s^1} u^1_{s^1}(p) + b^1_{s^1} = U^1_{s^1}(p).\end{aligned}$$

A.4.2 Proof of Proposition 4

Fix two GP representations $(S^1, U^1, \mu^1, \tau^1)$ and $(S^2, U^2, \mu^2, \tau^2)$. To see that conditions (a)–(c) imply these represent the same RCR, first note that (a) implies $M(A, U^1_{s^1}) = M(A, U^2_{\pi(s^1)})$ for every $A \in \mathcal{A}$ and $s^1 \in S^1$. Therefore, by (b) and (c), for any $A \in \mathcal{A}$ and $p \in A$,

$$\begin{aligned}&\sum_{s^1 \in S^1} \mu^1(s^1) \tau^1_{s^1}(\{u \in \mathcal{U} : p \in M(M(A, U^1_{s^1}), u)\}) \\ &\qquad = \sum_{s^1 \in S^1} \mu^2(\pi(s^1)) \tau^2_{\pi(s^1)}(\{u \in \mathcal{U} : p \in M(M(A, U^1_{s^1}), u)\}) \\ &\qquad = \sum_{s^1 \in S^1} \mu^2(\pi(s^1)) \tau^2_{\pi(s^1)}(\{u \in \mathcal{U} : p \in M(M(A, U^2_{\pi(s^1)}), u)\}) \\ &\qquad = \sum_{s^2 \in S^2} \mu^2(s^2) \tau^2_{s^2}(\{u \in \mathcal{U} : p \in M(M(A, U^2_{s^2}), u)\}),\end{aligned}$$

and hence both represent the same RCR.

Conversely, suppose $(S^1, U^1, \mu^1, \tau^1)$ and $(S^2, U^2, \mu^2, \tau^2)$ represent the same RCR λ. To simplify notation, we write $U^1_{s^1} \approx U^2_{s^2}$ to indicate that $U^1_{s^1}$ is a positive affine transformation of $U^2_{s^2}$. To establish condition (a), we first show that for every $s^1 \in S^1$ there exists $s^2 \in S^2$ such that $U^1_{s^1} \approx U^2_{s^2}$. Consider the set of utility functions $F = \{U^1_{s^1} : s^1 \in S^1\} \cup \{U^2_{s^2} : s^2 \in S^2\}$ and take $A \in \mathcal{A}$ as described in Lemma 1. Fix any $s^1 \in S^1$. By part 1 of Lemma 1, there exists $p \in A$ such that $U^1_{s^1}(p) > U^1_{s^1}(q)$ for all $q \in A \setminus \{p\}$. Therefore, $M(A, U^1_{s^1}) = \{p\}$ and hence the definition of the GP representation implies $\lambda^A(p) \geq \mu^1(s^1) > 0$. Since $(S^2, U^2, \mu^2, \tau^2)$ represents the same RCR λ, it must be that $p \in M(A, U^2_{s^2})$ for some $s^2 \in S^2$. By part 2 of Lemma 1, this implies $U^1_{s^1} \approx U^2_{s^2}$. Since s^1 was arbitrary, we have established that for every $s^1 \in S^1$ there exists $s^2 \in S^2$ such that $U^1_{s^1} \approx U^2_{s^2}$. An identical argument shows that for every $s^2 \in S^2$ there exists $s^1 \in S^1$ such that $U^1_{s^1} \approx U^2_{s^2}$. Given the non-redundancy requirement in the GP representation, this is sufficient to imply there exists a bijection $\pi : S^1 \to S^2$ and functions $\alpha : S^1 \to (0, \infty)$ and $\beta : S^1 \to \mathbb{R}$ such that $U^1_{s^1}(p) = \alpha(s^1) U^2_{\pi(s^1)}(p) + \beta(s^1)$ for all $p \in \triangle(Z)$ and $s^1 \in S^1$.

To establish condition (b), take A as above and fix any $s^1 \in S^1$. As previously argued, part 1 of Lemma 1 implies there exists $p \in A$ such that $M(A, U^1_{s^1}) = M(A, U^2_{\pi(s^1)}) = \{p\}$. Moreover, since $U^1_{\bar{s}^1}$ represents a different expected utility preference that $U^1_{s^1}$ for all $\bar{s}^1 \neq s^1$, part 2 of Lemma 1 implies $p \notin M(A, U^1_{\bar{s}^1}) = M(A, U^2_{\pi(\bar{s}^1)})$ for all $\bar{s}^1 \neq s^1$. Therefore, $\lambda^A(p) = \mu^1(s^1)$ and $\lambda^A(p) = \mu^2(\pi(s^1))$. Since $s^1 \in S^1$ was arbitrary, conclude $\mu^1(s^1) = \mu^2(\pi(s^1))$ for all $s^1 \in S^1$.

To establish condition (c), again take A as above, and fix any $s^1 \in S^1$, $B \in \mathcal{A}$, and $p \in B$. As above, Lemma 1 implies there exists $q \in A$ such that $M(A, U^1_{s^1}) = M(A, U^2_{\pi(s^1)}) = \{q\}$ and $q \notin M(A, U^1_{\bar{s}^1}) = M(A, U^2_{\pi(\bar{s}^1)})$ for $\bar{s}^1 \neq s^1$. For $\alpha \in (0,1)$, define $C_\alpha = (\alpha B + (1-\alpha)\{q\}) \cup (A \setminus \{q\})$. By continuity, for $\alpha \in (0,1)$ sufficiently small, we have

$$M(C_\alpha, U^1_{s^1}) = M(C_\alpha, U^2_{\pi(s^1)}) = M(\alpha B + (1-\alpha)\{q\}, U^1_{s^1})$$

and

$$\alpha p + (1-\alpha) q \notin M(C_\alpha, U^1_{\bar{s}^1}) = M(C_\alpha, U^2_{\pi(\bar{s}^1)}), \quad \forall \bar{s}^1 \neq s^1.$$

Therefore,

$$\begin{aligned}
&\lambda^{C_\alpha}(\alpha p + (1-\alpha) q) \\
&\quad = \sum_{\bar{s}^1 \in S^1} \mu^1(\bar{s}^1) \tau^1_{\bar{s}^1}(\{u \in \mathcal{U} : \alpha p + (1-\alpha) q \in M(M(C_\alpha, U^1_{\bar{s}^1}), u)\}) \\
&\quad = \mu^1(s^1) \tau^1_{s^1}(\{u \in \mathcal{U} : \alpha p + (1-\alpha) q \in M(M(C_\alpha, U^1_{s^1}), u)\}) \\
&\quad = \mu^1(s^1) \tau^1_{s^1}(\{u \in \mathcal{U} : \alpha p + (1-\alpha) q \in M(M(\alpha B + (1-\alpha)\{q\}, U^1_{s^1}), u)\}) \\
&\quad = \mu^1(s^1) \tau^1_{s^1}(\{u \in \mathcal{U} : p \in M(M(B, U^1_{s^1}), u)\}).
\end{aligned}$$

Similarly,

$$\lambda^{C_\alpha}(\alpha p + (1-\alpha) q) = \mu^2(\pi(s^1)) \tau^2_{\pi(s^1)}(\{u \in \mathcal{U} : p \in M(M(B, U^1_{s^1}), u)\}).$$

Since $\mu^1(s^1) = \mu^2(\pi(s^1))$ by our previous arguments, this implies

$$\tau^1_{s^1}(\{u \in \mathcal{U} : p \in M(M(B, U^1_{s^1}), u)\}) = \tau^2_{\pi(s^1)}(\{u \in \mathcal{U} : p \in M(M(B, U^1_{s^1}), u)\}).$$

Since $s^1 \in S^1$, $B \in \mathcal{A}$, and $p \in B$ were arbitrary, this completes the proof.

A.5 Proof of Theorem 3

The arguments here parallel those used in the proof of Theorem 1. The necessity of Axiom 1 follows directly from Part 1 of Proposition 2.

Conversely, suppose $\succsim$ has a DLR representation (S^1, U^1, μ^1) and λ has a GP representation $(S^2, U^2, \mu^2, \tau^2)$, and suppose the pair $(\succsim, \lambda)$ satisfies Axiom 1. By Part 1 of Proposition 2, for each $s^1 \in S^1$, there exist $s^2 \in S^2$ such that $U^1_{s^1} \approx U^2_{s^2}$. Together with the non-redundancy assumption in the DLR representation, this is sufficient for there to exist an injection $\pi : S^1 \to S^2$ such that $U^1_{s^1} \approx U^2_{\pi(s^1)}$ for all $s^1 \in S^2$. Therefore, subject to relabeling, it is without loss of generality to assume that $S^1 \subset S^2$ and $U^1_{s^1} \approx U^2_{s^1}$ for all $s^1 \in S^1$.

Let $S = S^1$ and $T = S^2$. For each $t \in T \setminus S$, define $U_t : \triangle(Z) \to \mathbb{R}$ by $U_t(p) = U^2_t(p)$ for $p \in \triangle(Z)$. For each $s \in S$, define $U_s : \triangle(Z) \to \mathbb{R}$ by $U_s(p) = \frac{\mu^1(s)}{\mu^2(s)} U^1_s(p)$. Since $\mu^2(s) > 0$ for all $s \in S$, these functions are well-defined. Also, since $\mu^1(s) > 0$ for all $s \in S$, each U_s is nonconstant. Therefore, $U_t \approx U^2_t$ for all $t \in T$, so U inherits the non-redundancy property from U^2. Finally, let $\mu = \mu^2$ and $\tau = \tau^2$. We claim that (S, T, U, μ, τ) is an Unforeseen Contingencies representation for $(\succsim, \lambda)$. The tuple $(S, U|_S, \mu|_S)$ is a DLR representation for $\succsim$ since

$$\sum_{s \in S} \frac{\mu(s)}{\mu(S)} \max_{p \in A} U_s(p) = \sum_{s \in S} \frac{\mu^2(s)}{\mu^2(S)} \max_{p \in A} \left(\frac{\mu^1(s)}{\mu^2(s)} U^1_s(p) \right) = \frac{1}{\mu^2(S)} \sum_{s \in S} \mu^1(s) \max_{p \in A} U^1_s(p).$$

The tuple (T, U, μ, τ) is a GP representation for λ since

$$\begin{aligned}
\lambda^A(p) &= \sum_{t \in T} \mu^2(t) \tau^2_t(\{u \in \mathcal{U} : p \in M(M(A, U^2_t), u)\}) \\
&= \sum_{t \in T} \mu(t) \tau_t(\{u \in \mathcal{U} : p \in M(M(A, U_t), u)\}),
\end{aligned}$$

where the last equality follows because $U_t \approx U^2_t$ implies $M(A, U_t) = M(A, U^2_t)$. This completes the sufficiency part of the proof.

A.6 Proof of Theorem 4

Conditions (b)–(d) follow directly from Proposition 4. Having established the uniqueness of the probability measure, condition (a) then follows from Proposition 3.

References

Barberà, S., W. Bossert, and P. K. Pattanaik (2004): "Ranking Sets of Objects," in *Handbook of Utility Theory: Volume 2 Extensions*, ed. by S. Barberà, P. J. Hammond and C. Seidl. Dordrecht: Kluwer Academic Publishers, 893–977.

Dekel, E., B. Lipman, and A. Rustichini (2001): "Representing Preferences with a Unique Subjective State Space," *Econometrica*, 69, 891–934.

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

Gul, F., and W. Pesendorfer (2006): "Random Expected Utility," *Econometrica*, 74, 121–146.

Kreps, D. M. (1979): "A Representation Theorem for 'Preference for Flexibility'," *Econometrica*, 47, 565–577.

Krishna, R. Vijay and Philipp Sadowski (2011): "Dynamic Preference for Flexibility." Working paper, University of North Carolina and Duke University.

Sadowski, P. (2010): "Contingent Preference for Flexibility: Eliciting Beliefs from Behavior," Working paper, Duke University.

Schenone, P. (2010): "Identifying Subjective Beliefs in Subjective State Space Models," Working paper, Northwestern University.