Preference discovery and experimentation

KEVIN COOKE
Department of Economics, Boston University

I provide axiomatic foundations for a model of *taste uncertainty* with endogenous learning through consumption. In this setting, uncertainty is over an unobservable, subjective state space. Preference over lottery–menu pairs is sufficient to identify the state space and the learning process. In this model, the agent is viewed *as if* he learns the utility of an object upon its consumption. This information is used to improve choice from the follow-on menu. This implies a trade-off between consumption value and information leading to *experimentation*. I provide a behavioral definition of experimentation. While the literature focuses on identifying subjective states through a demand for flexibility, I show that experimentation also (partially) identifies taste uncertainty.

KEYWORDS. Taste uncertainty, learning through consumption, experimentation, endogenous learning, subjective state space, decision theory.

JEL CLASSIFICATION. D11, D83.

1. INTRODUCTION

Traditional economic models of uncertainty treat agents as knowing their preferences and having uncertainty over the extrinsic state of the world. However, individuals can also face uncertainty over their own preferences. This *taste uncertainty* may exist despite perfect information about the extrinsic state of the world. Importantly, taste uncertainty is over a subjective state space that is unobservable to the analyst. This paper studies the choice behaviors that can be used to make inferences about an agent's taste uncertainty.

The literature has shown that subjective states can be inferred from preference over menus (Kreps (1979) and Dekel et al. (2001) (henceforth DLR)). When an agent displays a preference for flexibility, strictly preferring a menu to all of its subsets, he can be viewed as having uncertainty over the ranking of objects in the menu. Under this interpretation, it is the anticipated resolution of that uncertainty that drives the demand for flexibility. Typically, the subjective state space is viewed as a set of preference relations. Seemingly, taste uncertainty is a natural application of such models. However, these models

make use of an implicit and exogenous information process. When information is exogenous to the agent's choices, any form of learning through consumption is precluded. I contend that endogenous learning through consumption is an essential ingredient in a model of taste uncertainty. Intuitively, the act of consuming an object of unknown value must provide information. Therefore, a main goal of my analysis is to show how the DLR framework can be adjusted to accommodate learning through consumption. To facilitate this goal, I extend the DLR domain from preference over menus to preference over consumption–menu pairs, (c, A). This extended domain allows several notable behaviors to be captured.

First, I observe that learning through consumption means that the demand for flexibility may depend on consumption. For instance, suppose Alice lives in a city where the Uber ride sharing company has just entered the market. She can commute to work by bus, b, or by Uber, u. She knows her tastes for buses, but is uncertain of whether she will like Uber. Further, suppose that she believes the bus is most likely better than Uber. Then we might expect to observe

$$(b, \{b, u\}) \sim (b, \{b\}) \quad \text{and} \quad (u, \{b, u\}) \succ (u, \{b\}). \tag{1}$$

That is to say, if Alice plans to take the bus today (and gain no taste information), she is indifferent between committing to use the bus tomorrow (perhaps by buying a monthly pass) or being able to choose between the bus and Uber when tomorrow comes. However, if she plans to try Uber today, she may wish to keep the option open of taking Uber again tomorrow, in case she discovers that Uber is awesome! Experiencing Uber induces a demand for flexibility that the bus does not.

Second, flexibility may be demanded even when today's consumption is not included in tomorrow's menu. Suppose Alice is on a trip today where she expects to use Uber's competitor, Lyft, l. If she believes her tastes for all ride share services are correlated, we may well observe

$$(l, \{b, u\}) \succ (l, \{b\}) \succsim (l, \{u\}). \tag{2}$$

Here the larger menu is strictly preferred to committing to either the bus or Uber for tomorrow. This is because Alice believes that experiencing Lyft will inform her about her taste for Uber. This indirect learning results from a correlation in Alice's prior beliefs about her tastes. More generally, every consumption choice of (subjectively) uncertain value provides information to the agent. This information can be used to update beliefs and improve decision making in the future.

Finally, learning through consumption leads to a trade-off between current consumption utility and (endogenously generated) taste information. This can lead to *experimentation*: purposeful action with the goal of resolving uncertainty. Formally, experimentation can be captured by the preferences

$$(b, \{w\}) \succ (u, \{w\}) \quad \text{and} \quad (u, \{b, u\}) \succ (b, \{b, u\}). \tag{3}$$

To understand experimentation in this context, consider the following example: Tomorrow Alice plans to start a new job to which she will walk, w. She still has to get to

work today. She can take either the bus or the new Uber service. We observe her taking the bus to work. This corresponds to $(b, \{w\}) \succ (u, \{w\})$. A standard revealed preference approach would conclude that Alice prefers the bus to the Uber.

However, during the workday she finds out she will be with her current company for a few extra weeks. Knowing that she will face the choice between the bus and the Uber several more times, we see her take the Uber home from work; that is, $(u, \{b, u\}) \succ (b, \{b, u\})$. Now we might conclude Alice is revealing a preference for Uber over the bus. Is this consistent with what we observed before?

Yes, there is an important informational difference between these situations. In the first situation, Alice expected the Uber to provide only current consumption value. However, in the second situation Alice expects to face the bus/Uber choice again in the future. This means taking the Uber provides Alice with taste information that will be useful in her future decisions. This information may be valuable enough to sway Alice's decision, causing her to experiment.

To accommodate these behaviors, I extend the DLR framework to include consumption prior to the choice from the menu. Formally, I study preference over pairs of current consumption lotteries and follow-on menus (of lotteries). A typical choice object is denoted (p, A), where p is a lottery over current consumption and A is a follow-on menu. Each (pure) current consumption alternative is viewed as inducing a positive additive expected utility (PAEU) preference over menus (as in DLR). While in principle consumption can reveal information in a variety of ways, I model the agent *as if* he learns $u(c)$ after consumption of alternative c. In other words, prior to consumption, the agent has a rough idea of his utility function; he knows it lies in set $\mathcal{U}$ and has formed a prior probability, μ, over this set. Now suppose p resolves to yield a current consumption of c. Following consumption of c, the agent expects to learn $u(c)$. Given this information, he will be able to refine his knowledge (and update μ accordingly) to the smaller set $P_c(u) \subset \mathcal{U}$ in which $u' \in P_c(u)$ if and only if $u'(c) = u(c)$. Using this new information, he then selects tomorrow's consumption q from the previously chosen follow-on menu A. Importantly, A is selected prior to consumption; thus the agent is concerned with its ex ante value. This ex ante value is the expected utility of the menu conditional on the endogenous information generated by the choice of current consumption. This process is captured by the *preference discovery* (PD) representation

$$W(p, A) = \mathbb{E}_p \sum_{u \in \mathcal{U}} \left[u(c) + \delta \max_{q \in A} \sum_{u' \in P_c(u)} u'(q) \mu(u' \mid P_c(u)) \right] \mu(u),$$

where $P_c(u) = \{u' \in \mathcal{U} : u'(c) = u(c)\}$.

A notable feature of the above representation is that there is no exogenous set of states. Rather, the agent's state space, $\mathcal{U}$, is subjective and can be identified through choice over consumption–menu pairs. When we think about taste uncertainty, the natural state space to consider is the set of possible utility functions, which is by nature subjective. However, subjective states are unobservable. So not only is a subjective state space natural, but in fact this unobservability precludes the use of an objective state space. To see this, note that in the Savage and Anscombe–Aumann settings, the role

of the objective state space is to allow event-contingent claims to be used to identify an agent's utility and subjective probability. However, in the case of taste uncertainty, there is no longer an incentive compatible way to condition payoffs on particular events because the events are subjective and unverifiable. Therefore, a subjective state space framework, like DLR, is necessary to study taste uncertainty.

After introducing the model and studying some of its implications, I turn to the question of eliciting parameters. Specifically, because full menu choice data are not always available, I focus on providing results that aid in revealing taste uncertainty with limited choice data. First, I show that experimentation can partially identify the agent's taste uncertainty. Thus, in a repeated choice context, observed preference reversals are enough to reveal an agent's taste uncertainty without needing to directly observe an agent's choice of follow-on menu. This means experimentation is a potential behavioral alternative to flexibility that may be more readily observable. The parameters of the PD model cannot be uniquely identified by experimentation behavior alone; however, I show that preferences over a special class of small menus (analogous to financial options) can be used to construct an agent's entire subjective state space.

The paper proceeds as follows. Section 2 discusses the related literature. Section 3 begins by describing the setting, preference domain, and timing assumptions. I then describe the baseline model—a simple extension of the DLR framework to include a prior consumption stage. The baseline model puts very few restrictions on the learning process. The PD model restricts learning to be Bayesian: this model is introduced in Section 4. This section also characterizes a special case of the PD model in which there is no indirect consumption learning (Theorem 1). Section 5 uses this restricted PD representation to introduce the *potential value* concept. Section 6 uses this concept to explore the case of indirect consumption learning. This leads to the main representation result (Theorem 2). Examples of several types of taste uncertainty permitted by the PD model are discussed.

Finally, Section 7 focuses on eliciting the parameters of the PD model, with a particular focus on limited data settings. The axioms imply that an agent with taste uncertainty has an experimentation incentive. A behavioral definition of experimentation is provided. By observing this experimentation behavior, I show that taste uncertainty can be (partially) identified. I then give a graphical interpretation of the *potential value* concept. This points to the final result of the paper: the parameters of the PD representation can be elicited with a "small" set of choice data (Theorem 3). Proofs are provided in the main text when useful; all omitted proofs are collected in Appendix A.3. Additionally, Appendix A.2 discusses the independence of the axioms and their relation to alternate learning processes.

2. Related literature

Subjective state space models were introduced by Kreps (1979). His study of preferences over menus yielded a representation featuring the agent's implicit view of the world. A seminal contribution of that paper was to suggest that it is possible to model uncertainty without needing to assume a particular form for that uncertainty (i.e., an objective

set of states). However, Kreps's state space was not unique. This made the interpretation of his representation difficult. This issue is addressed by DLR. By expanding the domain from Kreps's menus of alternatives to menus of lotteries, they show that under minimal assumptions (weak order, continuity, indifference to randomization (IR), and nontriviality), this additional structure delivers a unique state space consisting of a set of preference relations. When DLR strengthen IR to independence and rule out a preference for commitment (preferring smaller menus), they get a positive additive expected utility (PAEU) representation. The PAEU representation provides the backbone of the model I describe in this paper. My main point of divergence from that model is the addition of the prior consumption period; this allows the information process to be endogenous to consumption. Behaviorally, endogenizing information acquisition leads to a taste for experimentation.

Hyogo (2007) axiomatizes a model of experimentation. Hyogo studies an environment with an abstract set of first stage actions that deliver a subjective signal about the objective state of the world. This information is used to choose an Anscombe–Aumann act from a follow-on menu. Conceptually, this is quite similar to the environment I study. In fact, I begin my analysis by adapting his model to the subjective state space context. Nonetheless, the choice data we utilize is quite different. First, I do not require an objective set of states. This permits the interpretation of my model as one of taste uncertainty. This interpretation is precluded when states are assumed to be directly observed. Second, I restrict his abstract first period actions to be consumption alternatives drawn from the same domain as the elements of the follow-on menu. This restriction tightly links my model to the idea of learning through consumption. Most importantly, I focus on a particular signal process in which the agent's realized consumption utility acts as the signal.

A key point of interest in my paper is the axiomatic characterization of Bayesian updating in a subjective state space framework. Riella (2013) studies a closely related question. In standard objective state space models, imposing dynamic consistency ensures Bayesian updating. However, Riella points out that the standard dynamic consistency axiom relies on objective states. He develops an alternative formulation, which he calls flexibility consistency, that plays the same role in a subjective state space environment. My paper differs from his in several ways. First, he works with exogenous, objective information. Instead, I study a setting where subjective information is endogenously generated through consumption. Additionally, we use different timing assumptions. In his model, an event is observed prior to the selection of the menu. My framework focuses on partitional information where the event is realized only after the menu is selected. This timing assumption is also adopted in Dillenberger et al. (2014) and Krishna and Sadowski (2014). However, both of these papers focus on exogenous information processes and require a set of objective states. Since information is exogenous to the agent's actions, these papers preclude the experimentation behavior that motivates my work.

The most closely related papers to mine are Ke (2013) and Piermont, Takeoka, and Teper (2016). Both of these studies explore the role of consumption in determining future preference. Ke models consumption externalities more generally. For instance, his

framework allows first period consumption to affect future preference through a preference for variety or consumption complementarities. I instead focus specifically on the informational channel. Also, Ke works in a two-period Krepsian world. This leads to the same nonuniqueness as in Kreps (1979).

In contemporaneous work, Piermont et al. (2016) also focus on a setting with taste uncertainty in which information is only generated through consumption. In their model, after an agent tries two alternatives, he learns his ordinal preference between them. This process implies that an anticipated consumption stream induces an ex ante partition of the agent's subjective state space. This generates a demand for flexibility over menus that follow the consumption stream. Observing this demand for flexibility allows the analyst to infer the induced partition and thus identify the agent's anticipated information process. Their main analysis takes place in a recursive menu framework that allows them to identify the agent's anticipated conditional learning process. That is, conditional on a certain signal realization, they can infer which future signals the agent believes possible.

My analysis differs from Piermont et al. (2016) in several ways. Conceptually, our models differ in the type of learning that we study. In my representation, the agent learns the cardinal utility of the object he consumes rather than simply its ordinal ranking relative to previously experienced goods. This means the agent expects to learn about his preferences after a single instance of consumption. However, I do not explore multiple consumption periods, which means I cannot identify the agent's anticipated conditional learning process. Additionally, I offer a behavioral definition of experimentation and show how this relates to the parameters of my model. My technical analysis is also distinct. By focusing on a setting with lotteries and universal best and worst alternatives, I show how the parameters of the model can be elicited with much less choice data than is typical in the literature.

3. The baseline model

In this paper, I study an agent with taste uncertainty. I investigate the choice behavior such an agent exhibits when some of his uncertainty can be resolved through experience. By imposing axioms consistent with experiential learning, I derive a utility representation with the interpretation that the act of consumption allows the decision maker (DM) to learn about his own preferences.

Formally, consider a finite set of consumption alternatives $\mathcal{C}$. Let $\Delta\mathcal{C}$ denote the set of lotteries over $\mathcal{C}$ and let $\mathcal{K}$ be the collection of nonempty, compact subsets of $\Delta\mathcal{C}$. Endow $\Delta\mathcal{C}$ with the standard Euclidean topology. The collection of menus, $\mathcal{K}$, is endowed with the Hausdorff topology[1] and a mixture operation such that for $\alpha \in [0, 1]$,

$$\alpha A + (1 - \alpha)B = \big\{\alpha p + (1 - \alpha)q \mid p \in A, q \in B\big\}.$$

---

[1]The Hausdorff topology refers to the topology induced by the Hausdorff metric, where the Hausdorff metric is

$$d_H(A, B) = \max\Big\{\sup_{a \in A}\inf_{b \in B} d(a, b), \sup_{b \in B}\inf_{a \in A} d(a, b)\Big\}.$$
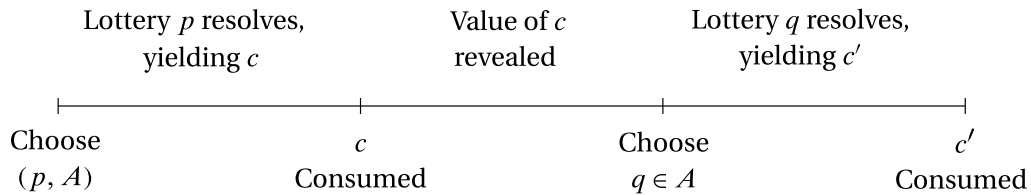
The product space, $\Delta\mathcal{C} \times \mathcal{K}$, has the product topology.

Notice that each $c \in \mathcal{C}$ can be naturally identified with the degenerate lottery $\delta_c \in \Delta\mathcal{C}$ and each lottery $p \in \Delta\mathcal{C}$ can be identified with the singleton menu $\{p\} \in \mathcal{K}$. Where it does not cause confusion, I abuse notation and write c instead of δ_c. Likewise, the set of degenerate lotteries is denoted $\mathcal{C}$ and each singleton menu, $\{p\}$, is identified with its unique element p.

I take a binary relation $\succsim$ on $\Delta\mathcal{C} \times \mathcal{K}$ as primitive. A typical choice object is (p, A). Here p is a lottery over possible consumption alternatives. The realization of this lottery is consumed prior to the agent's choice from the follow-on menu, A. Therefore, different lotteries provide potentially different information to the agent and thus induce different preferences over menus. Let this induced preference, $\succsim_p$, be defined such that

$$A \succsim_p B \quad \text{iff} \quad (p, A) \succsim (p, B).$$

The timing is as follows:

Lottery p resolves, yielding c		Value of c revealed		Lottery q resolves, yielding c'	
Choose (p, A)	c Consumed		Choose $q \in A$		c' Consumed

First, the agent chooses a consumption lottery and a follow-on menu of lotteries. After the realization of the first period lottery, consumption occurs. The agent observes the utility of the consumed object and updates his beliefs accordingly. Then, in an unmodeled second period, he chooses an option from the follow-on menu.

The induced preferences are assumed to be representable in the following way.

DEFINITION 1. Say $(\mathcal{U}, \mu)$ is a *finite positive additive expected utility* (fPAEU) representation of the induced preference $\succsim_c$ if there exists a finite set of expected utility functions, $\mathcal{U}$, and a strictly positive measure, μ, such that the function

$$V_c(A) = \sum_{u \in \mathcal{U}} \max_{a \in A} u(a)\mu(u)$$

satisfies $A \succsim_c B$ if and only if $V_c(A) \geq V_c(B)$ for all $A, B \in \mathcal{K}$.

The above representation was introduced in DLR (2001, 2009). These papers show that the fPAEU representation is equivalent to the induced preference, $\succsim_c$, satisfying weak order, continuity, independence, monotonicity (i.e., a preference for flexibility), and a technical finiteness condition. These conditions are stated formally in the Appendix. The major point of interest is that the set of utility functions, $\mathcal{U}$, acts as a subjective state space. Here the uncertainty can be understood to be over preferences; this provides a natural starting point for a model of taste uncertainty.

Under this interpretation of the fPAEU model, the agent should be viewed as expecting the true state, $u \in \mathcal{U}$, to be revealed prior to choosing from the menu. Therefore,

the demand for flexibility is driven by an unmodeled, exogenous information process. However, as argued in the Introduction, taste uncertainty is resolved in a natural and endogenous way through consumption. To capture this feature, I extend the model to include a prior consumption stage in which (expected) consumption can affect menu preferences. To begin my analysis, I consider a simple extension of the fPAEU framework to incorporate an explicit prior consumption phase. I relegate a full behavioral characterization of this to Appendix A.1, but I note that the associated representation theorem is a slight adaptation of the results of Hyogo (2007) to the subjective state space setting.

DEFINITION 2. A *baseline representation* of $\succsim$ is a tuple $(\{\mathcal{U}_c, \mu_c\}_{c \in \mathcal{C}}, \delta)$ such that $\succsim$ is represented by a function of the form

$$W(p, A) = \mathbb{E}_p\big[V(c) + \delta V_c(A)\big], \tag{4}$$

where for some $c^*, c_* \in \mathcal{C}$ and for all $c, c', c'' \in \mathcal{C}$, it is true that:

(i) $\delta > 0$,

(ii) $V_c(\{c''\}) = V_{c'}(\{c''\}) \equiv V(c'')$,

(iii) $V_c(A) = \sum_{u \in \mathcal{U}_c} \max_{a \in A} u(a) \mu_c(u)$, such that $(\mathcal{U}_c, \mu_c)$ is an fPAEU representation of $\succsim_c$.

(iv) μ_c is a probability measure with support $\mathcal{U}_c$, and

(v) $u(c^*) = 1$, $u(c_*) = 0$, and $u(c') \in [0, 1] \; \forall u \in \mathcal{U}_c$.

The above model is a natural generalization of the fPAEU model to include informative first period consumption. Equation (4) is a standard ex ante expected utility model in which the pair (p, A) is valued according to the (expected) weighted sum of first period consumption utility and the anticipated utility of the follow-on menu. Notice that the value of the follow-on menu depends on first period consumption (the realization of the lottery p). I impose five restrictions on this general form.

The first two restrictions require that the agent (i) values second period consumption and (ii) ranks follow-on singleton menus the same way as he ranks the analogous first period consumption, regardless of his prior consumption decision. This second restriction rules out changing tastes as well as non-informational, intertemporal complementarities. For instance, models of habit formation or a preference for variety would not fit into this framework. Recall that the follow-on menu is selected concurrently with first period consumption. When the menu is singleton, consumption information cannot affect the ex ante valuation of that menu.

Next, (iii) assumes that any degenerate first period consumption induces a menu preference, $\succsim_c$, which has an fPAEU representation. This allows me to carry the subjective state space framework into this extended domain. Then (iv) imposes that μ_c is a probability measure rather than the general positive measure allowed in the fPAEU representation.

Finally, (v) restricts attention to settings that feature a best and worst alternative, which are common across all subjective states. This assumption greatly simplifies exposition; however, it can be relaxed substantially.[2] One benefit of such settings is that they suggest a natural restriction on the state space:

DEFINITION 3. A pair $(\mathcal{U}_c, \mu_c)$ is *regular* if μ_c is a probability measure with support $\mathcal{U}_c$ and

$$\mathcal{U}_c \subseteq \{u : \Delta C \to [0, 1] \mid u(\{\beta c^* + (1 - \beta)c_*\}) = \beta\}.$$

By restricting attention to cases where $u(c^*) = 1$ and $u(c_*) = 0$ for all $u \in \mathcal{U}_c$, the fPAEU representation is uniquely identified. In DLR, $\mathcal{U}_c$ and μ_c are only jointly identified. Under this restriction, a meaningful probability measure can be obtained.

Additionally, under this restriction, each menu is indifferent to some singleton menu containing only a lottery between c^* and c_*. In other words, for each pair (c, A), there exists some $\beta \in [0, 1]$ such that

$$A \sim_c \{\beta c^* + (1 - \beta)c_*\} \equiv \{\hat{\beta}\}.$$

Notice that β is a real number between 0 and 1, while $\hat{\beta}$ is a lottery that places probability β on the best outcome and probability $(1 - \beta)$ on the worst outcome. It may be useful to think of these $\hat{\beta}$s as money values. Under this interpretation, c^* (c_*), can be seen as an amount of money greater (less) than all possible valuations of anything else in C.

4. (ONLY) DIRECT CONSUMPTION LEARNING

In the baseline model, the value of the follow-on menu is dependent on first period consumption. Therefore, it permits learning through consumption. However, the baseline model does not rule out other learning processes. For instance, if $V_c = V_{c'}$ for every $c, c' \in C$, then it is as if the agent receives the same information regardless of his consumption choice. Such a situation should not be interpreted as learning through consumption, but rather as an exogenous information process. Furthermore, the baseline model imposes very little discipline on the relationship across the induced preferences. Thus, no specific learning process is implied.

In this subsection, I place a simple restriction on the learning process that pins down a particularly simple endogenous information process as well as a Bayesian updating rule.

NO INDIRECT LEARNING (*NIL*). For any $\beta \in [0, 1]$ and $c, c' \in C$,

$$c \neq c' \implies \{c', \hat{\beta}\} \sim_c \{c'\} \quad \text{or} \quad \{c', \hat{\beta}\} \sim_c \{\hat{\beta}\}.$$

Notice that if consuming c provides taste information about c', the agent may wish to delay the choice between c' and $\hat{\beta}$ until after consumption of c. This would generate a

---

[2]All results hold as long as the agent strictly prefers c^* to c_* with certainty. See Remark 1 in the Appendix for a behavioral statement of this condition.

preference for flexibility, the agent strictly preferring $\{c', \hat{\beta}\}$ to either of its singleton sub-sets. The NIL axiom rules out this behavior. An agent satisfying NIL will only demand flexibility if c' will be tried prior to choosing from the menu. If this condition holds for all c', then the agent can be thought of as only learning through direct experience. In particular, for each c, the induced preference, $\succsim_c$, will only exhibit a demand for flexibility over menus that contain c or a lottery with c in its support.

Suppose an agent has uncertainty about his tastes for Uber. NIL implies this uncertainty can only be resolved by using Uber. Consumption of other goods or services (for instance, Lyft) would not help the agent make a better determination of his valuation of Uber rides. In the Introduction, three motivating behaviors were discussed. Consumption-dependent flexibility, (1), and experimentation, (3), are compatible with NIL. However, indirect consumption-dependent flexibility, (2), is ruled out by NIL.

Adding this assumption to the baseline model pins down a particularly simple learning process and yields an easily interpretable model. Before showing this, I first define a few terms.

DEFINITION 4. A *preference discovery* (PD) representation of $\succsim$ is a tuple $(\mathcal{U}, \mu, \delta)$, where $(\mathcal{U}, \mu)$ is a regular pair in which $\mathcal{U}$ is a finite set of expected utility functions and δ is a positive constant such that $\succsim$ is represented by

$$W(p, A) = \mathbb{E}_p \sum_{u \in \mathcal{U}} \left[u(c) + \delta \max_{q \in A} \sum_{u' \in P_c(u)} u'(q) \mu\big(u' \mid P_c(u)\big) \right] \mu(u),$$

$$\text{where } P_c(u) = \big\{ u' \in \mathcal{U} : u'(c) = u(c) \big\}.$$

The PD representation is the main model of interest in this paper. While the baseline representation allows for nearly arbitrary learning processes, the PD model focuses on a specific model of learning in which consumption of c informs the agent of $u(c)$.[3] This information is then used to Bayesian update the agent's beliefs about his tastes. This model reduces the parameter space of the baseline model from $(\{\mathcal{U}_c, \mu_c\}_{c \in \mathcal{C}}, \delta)$ to $(\mathcal{U}, \mu, \delta)$, thus offering a more parsimonious view of the agent.

DEFINITION 5. A PD representation, $(\mathcal{U}, \mu, \delta)$, has a *product structure* if for all $u \in \mathcal{U}$,

$$\mu(u) = \prod_{c \in \mathcal{C}} \mu\big(P_c(u)\big),$$

$$\text{where } P_c(u) = \big\{ u' \in \mathcal{U} : u'(c) = u(c) \big\}.$$

Imposing this product structure on the PD representation rules out any correlation across alternatives in the agent's prior belief about his tastes. Therefore, consumption is only able to provide direct information about the consumed alternative. Theorem 1 demonstrates that this structure is equivalent to the behavioral condition, NIL.

---

[3] The requirement that $(\mathcal{U}, \mu)$ is regular is what allows the cardinal utility of c to be interpreted as a meaningful quantity.

THEOREM 1. *A preference, $\succsim$, has a baseline representation and satisfies NIL if and only if it has a PD representation with a product structure. Moreover, a preference, $\succsim$, has at most one PD representation with a product structure.*

This theorem shows that the NIL condition characterizes a specific learning process for the agent.[4] Specifically, in the PD representation, $(\mathcal{U}, \mu, \delta)$, μ can be viewed as the agent's prior belief about his tastes. This prior belief, μ, has support $\mathcal{U}$. This prior is consistent with the fPAEU representation, $(\mathcal{U}_c, \mu_c)$, of each induced menu preference, $\succsim_c$, in the sense that

$$\mu_c\big(P_c(u)\big) = \mu\big(P_c(u)\big) \quad \text{for every } c \in \mathcal{C} \text{ and } u \in \mathcal{U}.$$

So in the PD model, the agent can be interpreted as expecting to learn the utility value of his first period consumption, c, prior to choosing from the follow-on menu. The product structure rules out any other learning. Specifically,

$$\mu\big(P_{c'}(u)\big) = \mu\big(P_{c'}(u) \mid P_c(u)\big) \quad \text{for every } c, c' \in \mathcal{C} \text{ and } u \in \mathcal{U}.$$

That is, given the product structure, learning the value of c cannot affect the agent's beliefs about the value of c'.

5. c-POTENTIAL VALUES

The next step in exploring the PD model is to look for a weakening of NIL that provides the PD structure while allowing the possibility of indirect learning. In service of this goal, this section calls attention to some features of the representation that can be mapped directly into behavior. Specifically, I investigate the structure of the set of utility functions, $\mathcal{U}$. Given some $\mathcal{U}$, it will be useful to consider the set

$$K_{c,c} = \big\{u(c) \in [0, 1] : u \in \mathcal{U}\big\},$$

where $K_{c,c}$ is the set of *potential values* of alternative c.[5] This is the set of the agent's subjectively possible utility realizations of c. Given this definition, I can state the following simple result.

PROPOSITION 1. *If $(\mathcal{U}, \mu, \delta)$ is a PD representation, then*

$$\mathcal{U} \subseteq \prod_{c \in \mathcal{C}} K_{c,c}.$$

Moreover, if $(\mathcal{U}, \mu, \delta)$ has a product structure, then $\mathcal{U} = \prod_{c \in \mathcal{C}} K_{c,c}$.

This proposition highlights the essential role played by potential values in the construction of the PD representation. Notably, these potential values can be identified behaviorally; this is shown below.

---

[4]The example in Section 7.4 explores the limitations of the uniqueness result in Theorem 1.

[5]The reason for the double subscript become clear shortly.

PROPOSITION 2. *Suppose $\succsim$ has a PD representation, $(\mathcal{U}, \mu, \delta)$. Then for any $c \in \mathcal{C}$,*

$$\beta \in K_{c,c} \quad \Longleftrightarrow \quad \{\hat{\beta}, c\} \cup \left\{\frac{1}{2}c + \frac{1}{2}\hat{\beta}'\right\} \succ_c \{\hat{\beta}, c\} \quad \text{for every } \beta' \in (\beta, 1].$$

To understand this proposition, first notice that the agent's only taste uncertainty in the menus on the right is about c. Moreover, the agent expects to consume c (and learn its true value) prior to a choice from the menu. Therefore, if $\beta = \beta'$, the agent would always be indifferent between the two menus, i.e.,

$$\{\hat{\beta}, c\} \cup \left\{\frac{1}{2}c + \frac{1}{2}\hat{\beta}\right\} \sim_c \{\hat{\beta}, c\}.$$

This indifference follows because if consumption of c reveals $u(c) > \beta$, the agent chooses c and otherwise chooses $\hat{\beta}$. The agent never strictly prefers the lottery. Now suppose that for every $u \in \mathcal{U}$, $u(c) \neq \beta$. Then, for any sufficiently small ϵ, the above indifference continues to hold even if the lottery $\frac{1}{2}c + \frac{1}{2}\hat{\beta}$ is slightly improved to $\frac{1}{2}c + \frac{1}{2}\widehat{(\beta + \epsilon)}$. So the only time the right hand side of the proposition can be true *for all $\beta' > \beta$* is when the agent believes he may learn that $u(c) = \beta$. Thus the set of *potential values* of c is exactly characterized by this behavior. This potential value concept focuses on what the agent believes he may learn about $u(c)$ given the consumption of c.

Next, I introduce a generalization of this concept. This is used extensively in the next section to explore the case of indirect learning.

DEFINITION 6. *Fix $c \in \mathcal{C}$ and $p \in \Delta\mathcal{C}$. Then $\beta \in [0, 1]$ is a c-potential value of p if*

$$\{\hat{\beta}, p\} \cup \left\{\frac{1}{2}p + \frac{1}{2}\hat{\beta}'\right\} \succ_c \{\hat{\beta}, p\} \quad \text{for every } \beta' \in (\beta, 1].$$

While potential values only talk about how consumption teaches the agent about the thing that was consumed, the c-potential value concept extends the logic of the potential value to handle cases in which consuming c can inform the agent about his tastes for an arbitrary lottery p. The set of *c-potential values* of p (denoted $K_{c,p}$) should be thought of as the set of updated expected values of p that the agent believes he may face conditional on observing the realized value of c. The next section shows how c-potential values can be used to relax NIL and characterize a version of the PD model in which indirect learning is permitted.

6. INDIRECT CONSUMPTION LEARNING

In Section 4, I characterized a special case of the PD model that rules out indirect consumption learning. However, there are many cases in which this is not desirable. For instance, an agent may believe his tastes for products within the same category are correlated. Thus consuming one product may affect the agent's beliefs about his tastes for related products. Prior to exploring such examples in Section 6.3, I characterize another special case of the PD model in which indirect learning is allowed. The characterization is based on two axioms that jointly weaken NIL and a structural condition that ensures

uniqueness of the representation. This last condition is incompatible with NIL; therefore, I am unable to provide a unique representation result for the general PD model. This is because the two-period setting examined here is not sufficient to pin down the agent's conditional beliefs. That is to say, in terms of the PD model, for a given realization of the $u(c)$ all we can infer about c' is $\mathbb{E}_\mu(u'(c') \mid P_c(u))$ and not some finer information structure. This is more thoroughly discussed in Section 7.4.

6.1 *Weakening NIL*

Axiom A1 (Weak Indirect Learning).   *For any $c \in \mathcal{C}$ and $p \in \Delta\mathcal{C}$, $|K_{c,p}| \leq |K_{c,c}|$.*

Loosely, this axiom says that consumption of c tells the agent more about c than about things that are not c. More formally, A1 ensures that two subjective states that agree on the value of c cannot be distinguished through the consumption of c.

This restriction is what delivers the key feature of the PD representation, that the agent only learns by discovering $u(c)$ upon consumption of c. Since this learning process is so specialized, A1 is ruling out quite a few other potential avenues of learning. For instance, a first period consumption of known constant value cannot be informative about the follow-on menu.

Additionally, A1 is potentially incompatible with separately learning about the constituent parts of a consumption object. Suppose c is a meal consisting of two ingredients that both have an uncertain value. The agent is certain that he dislikes one ingredient and likes the other, but is unsure which is which. If in both possible states, the complete meal has the same aggregate utility, then A1 requires the agent not learn anything. Specifically, anticipated consumption of the two-ingredient meal cannot induce a demand for flexibility in follow-on menus containing one or both of these ingredients.

Furthermore, A1 rules out a wide class of exogenous information processes that could, in principle, provide information regardless of the agent's consumption choices. One notable process that is ruled out by this axiom is learning through introspection. Suppose, that between choosing (c, A) and picking an element from the menu A, the agent plans to spend time thinking about his preferences for things in A. Then even if c is of known consumption value (having a unique potential value), we should expect the agent to have multiple c-potential values for the options he spends time introspecting about. This would violate A1. Adding introspection to a model of preference discovery is an interesting direction for future work. Of particular interest is endogenous introspection, which can vary with (c, A).

Weak indirect learning is implied by NIL. To see this, suppose NIL holds. If $c \neq c'$, then the unique c-potential value of c' would be the ex ante expected value of c'. It is straightforward to show that there is always at least one c-potential value of c; therefore, A1 is satisfied. It is also worth noticing that since the $\hat{\beta}$ lotteries are of certain expected value, every lottery p has a unique $\hat{\beta}$-potential value. So under A1, information can only be revealed if the value of the first period consumption is uncertain.

Axiom A1 places restrictions on what each induced preference may look like. However, it places no restriction across different induced preferences. The following axiom

imposes this type of consistency condition, ensuring that the various induced prefer-
ences can be viewed as Bayesian posteriors of the same prior.

Axiom A2 (Choice Relevant Value). *Take any $c \in C$, $\beta \in [0, 1]$, and any sequence of lotter-
ies, $\{p_n\}$, in which $p_n \to c$. Define $\hat{\beta}'_{p_n}$ such that $\{\hat{\beta}'_{p_n}\} \sim_c \{p_n, \hat{\beta}\}$. Then for any $c' \in C$,*

$$\{p_n, \hat{\beta}\} \succ_{c'} \{\hat{\beta}'_{p_n}\} \quad \forall n \quad \Longrightarrow \quad \beta \text{ is a c-potential value of c.}$$

Let us understand this axiom in parts. First, $\hat{\beta}'_{p_n}$ is defined to be the ex ante ex-
pected value of the menu $\{p_n, \hat{\beta}\}$ conditional on the information process generated by
consuming c. This means $\{p_n, \hat{\beta}\} \succ_{c'} \{\hat{\beta}'_{p_n}\}$ can only be true if c' provides more *choice
relevant* information than c, in the sense that consumption of c' instead of c improves
the expected choice from the menu.

Next, observe that if $\succsim$ has a baseline representation, then A2 implies that consum-
ing c provides as much information about c as does consuming any c'. Formally, A2
implies[6] that for any $\beta, \beta' \in [0, 1]$,

$$\{\hat{\beta}'\} \sim_c \{c, \hat{\beta}\} \quad \Longrightarrow \quad \{\hat{\beta}'\} \succsim_{c'} \{c, \hat{\beta}\} \quad \forall c' \in C.$$

Given this observation, if $\{p_n, \hat{\beta}\} \succ_{c'} \{\hat{\beta}'_{p_n}\}$ is true for all n, then under continuity
(implicit in the baseline representation) we find

$$\{c, \hat{\beta}\} \sim_{c'} \{\hat{\beta}'\} \sim_c \{c, \hat{\beta}\}.$$

It is easy to verify that this implies that the set of c'-potential values of c (denoted $K_{c',c}$)
exactly coincides with $K_{c,c}$. Intuitively, this means that A2 ensures that there is no better
way to learn about c than to try it (loosely, for every c', $|K_{c',c}| \le |K_{c,c}|$). This is subtly
different from A1, which restricts the degree of indirect learning. That is, A1 limits what
the consumption of c can teach the agent about c' (i.e., $|K_{c,c'}| \le |K_{c,c}|$). In other words,
A1 places an upper bound on indirect learning and A2 places a lower bound on direct
learning.

Now notice that as p_n approaches c, each c'-potential value of p_n approaches some
particular c'-potential value of c. Thus, if $\beta \notin K_{c',c}$, there is some sufficiently large n
such that the realized c'-potential value of p_n is greater than β if and only if the realized
c'-potential value of c is greater than β. Since $K_{c',c} = K_{c,c}$, this implies that consumption
of c' instead of c can only affect the agent's (state contingent) choice from the menu if
$\beta \in K_{c,c}$.

When $\succsim$ has a baseline representation and NIL holds, A2 has no additional content.
This result follows since A2 only has content if $K_{c',c} = K_{c,c}$. But under NIL, this can only
happen if there is no taste uncertainty about c (i.e., $K_{c,c}$ is singleton). But in that case, A2
only imposes that $\{c\}$ has the same $\hat{\beta}$ equivalent under each induced preference. This is
a form of stationarity that is implied by the baseline model (see B5 in the Appendix).

---

[6]This implication is proven in Proposition 5 in the Appendix.

So given a baseline representation, A1 and A2 are jointly weaker than NIL. Unfortunately, these axioms do not provide quite enough structure to yield a unique representation theorem. To provide a result for the case of indirect learning, I will restrict attention to settings in which there is a maximally informative alternative (MIA). Consumption of the MIA provides at least as much information about taste as can be gleaned from any other consumption alternative. Notice that the existence of an MIA is incompatible with NIL. This means the resulting representation theorem is not a generalization of Theorem 1, but rather characterize a distinct special case of the general PD model. Behaviorally, settings with an MIA satisfy the following condition.

Axiom A3 (Maximally Informative Alternative). *There exists some $\bar{c} \in \mathcal{C}$ such that for every $c \in \mathcal{C}$ and $A \in \mathcal{K}$,*

$$A \sim_c \{\hat{\beta}\} \quad \implies \quad A \succsim_{\bar{c}} \{\hat{\beta}\}.$$

For each c, every menu has some β-equivalent that can be interpreted as the expected value of that menu given the anticipated information revealed by consumption of c. If c is replaced by a more informative alternative, we expect the valuation of all menus to increase. This axiom assumes there is some $\bar{c} \in \mathcal{C}$ that provides a superset of the information of all other consumption possibilities. At first glance this assumption seems very strong. However, this axiom only imposes an assumption on the information value of c, not its consumption value. Because of this, oftentimes a suitable transformation of the domain yields such a $\bar{c}$.

For instance, suppose $\mathcal{C} = \{\text{steak, chicken, lobster}\}$. Here consuming chicken may provide different information than the other meals. This situation fails to satisfy A3. However, augmenting the set of outcomes to include a sampler platter with all three meals intuitively solves this issue.[7] Though consuming this sampler platter likely provides more information than any of the individual meals, it may be very costly. But that is fine; the maximally informative alternative may be nearly as bad as the worst alternative.

Ultimately, this axiom provides a measuring stick with which to compare all other induced preferences. In particular, the proof of the representation theorem relies on the sequences in A2 to ensure that more informative consumption choices induce fPAEU preferences that are consistent with all less informative choices.[8] However, in general, information is partially ordered. This axiom ensures there is a maximum of this ordering on $\mathcal{C}$. In turn, this allows the PD representation to be uniquely identified using only ex ante preference data over two periods. Axiom A3 could be relaxed if more periods of choice data are observed. In essence, if the agent has enough periods to try all the alternatives in $\mathcal{C}$ before selecting from the menu, it is as if he tried the maximally informative alternative. In this sense, A3 provides a sufficient condition for the unique identification

---

[7]Note that there is some tension between this sampler platter idea and the multi-ingredient meal issue pointed out in the discussion of A1. However, as long as each subjective state is associated with a distinguishable consumption value for the sampler platter, this tension disappears.

[8]Here the information ranking can be thought of in terms of the Blackwell ordering. Formally, this can be shown by adapting Theorem 2 from Dillenberger et al. (2014) to the subjective state space setting.

of the PD representation without having to consider a much larger domain. Piermont et al. (2016) study an infinite horizon problem, which alleviates the need for an analogous assumption in their setting.

With these three axioms in hand, I can now turn to the main result of this paper. But first I need one definition.

DEFINITION 7. A PD representation, $(\mathcal{U}, \mu, \delta)$, is *separating* if there is some $c \in \mathcal{C}$ such that when $u \neq u'$,

$$u, u' \in \mathcal{U} \quad \Longrightarrow \quad u(c) \neq u'(c).$$

This defines a restriction on the PD model such that alternative c separates the states in $\mathcal{U}$ in the sense that $u(c)$ takes a distinct value for each $u \in \mathcal{U}$. In this paper, the maximally informative alternative, $\bar{c}$, plays this role. We can think of $u(\bar{c})$ as a sufficient statistic with respect to the agent's subjective state space. So learning the value of the maximally informative alternative resolves all the uncertainty that is possible to resolve through a single instance of consumption.

Now the main representation theorem.

THEOREM 2. *A preference, $\succsim$, has a baseline representation and satisfies A1, A2, and A3 if and only if it has a unique separating PD representation.*

Theorem 2 shows that relaxing NIL to A1 and A2 removes the product structure restriction from the representation. Because of this, there can now be correlation across alternatives in the agent's beliefs about his taste uncertainty. This allows the possibility of indirect learning. However, to get a unique representation, I assume the existence of a maximally informative alternative. Since an MIA is incompatible with NIL, Theorem 2 is not a generalization of Theorem 1. Instead it characterizes a distinct special case of the PD representation.

6.2 *Proof sketch of Theorem 2*

A full proof is provided in the Appendix, but I discuss the proof of sufficiency of the axioms here in broad strokes. The starting point of the proof is the baseline representation. By assumption, we know that $\succsim$ has a baseline representation. This means that there exists a tuple $(\{\mathcal{U}_c, \mu_c\}_{c \in \mathcal{C}}, \delta)$ such that $\succsim$ is represented by

$$W(p, A) = \mathbb{E}_p\big[V(c) + \delta V_c(A)\big],$$

where for all $c, c', c'' \in \mathcal{C}$, $V_c(A) = \sum_{\mathcal{U}_c} \max_{a \in A} u(a)\mu_c(u)$, $(\mathcal{U}_c, \mu_c)$ is a regular fPAEU of the induced preference $\succsim_c$, and $V_c(\{c''\}) = V_{c'}(\{c''\}) \equiv V(\{c''\})$.

This already has a flavor of learning through consumption. Specifically, the valuation of A, $V_c(A)$, is indexed by c. Thus c is valuable not only for its consumption utility, $V(c)$, but also for its effect on the value of the menu A. However, the nature of the learning process in this representation is arbitrary. To develop a full model of consumption

learning, it is desirable to have an explicit model of the relationship between the various pairs $(\mathcal{U}_c, \mu_c)$.

The key task of this proof is to show that A1–A3 imply that this relationship is Bayesian. Axiom A3 postulates the existence of a maximally informative alternative, $\bar{c}$. With A1, I show that the fPAEU representation of the menu preference induced by this maximally informative alternative, $\succsim_{\bar{c}}$, has the separating form. The remainder of the proof shows that all other induced preferences are "consistent" with this one in the sense that their fPAEU representations can be viewed as coarser partitions of $(\mathcal{U}_{\bar{c}}, \mu_{\bar{c}})$.

The main insight is that comparisons of the form $\{p, \hat{\beta}\} \sim_c \{\hat{\beta}'\}$ tell us about the agent's conditional expected utility for p. As is discussed in Section 7.2 (Theorem 3), this can be used to construct the agent's induced state space for each c. The contrapositive of A2 implies that for a sequence of lotteries p_n that approaches c, there is some sufficiently large n for which the following statement is true:

$$\text{For } \beta \notin K_{c,c}, \quad \{p_n, \hat{\beta}\} \sim_c \{\hat{\beta}'\} \quad \implies \quad \{p_n, \hat{\beta}\} \sim_{c'} \{\hat{\beta}'\} \quad \forall c' \in \mathcal{C}.$$

This condition allows me to relate the conditional expected value of p_n,

$$\mathbb{E}_{\mu_c}\big[u(p_n) \mid u(p_n) > \beta\big],$$

across the menu preferences induced by different consumptions. The meat of the proof argues that this is sufficient to provide the desired consistency. Having shown this, it becomes clear that $(\mathcal{U}_{\bar{c}}, \mu_{\bar{c}}, \delta)$ is a separating PD representation that is numerically identical to the baseline representation. Since $(\mathcal{U}_{\bar{c}}, \mu_{\bar{c}})$ is a regular fPAEU representation, uniqueness follows from the DLR result.

6.3 *Examples of taste uncertainty*

To explore some possible applications of the PD model, I consider two special cases in which additional structure is placed on the domain of consumption objects, $\mathcal{C}$. This highlights two particular types of taste uncertainty that fit into the PD framework. Each is substantiated by a real world example.

Related experience goods   The most straightforward interpretation of the PD model is that of an agent learning the match specific utility of some experience good through consumption. When an agent believes his tastes to be correlated across related experience goods, he can learn through indirect consumption. That is, consuming one product can help the agent learn about his tastes for related products.

For instance, antidepressants are a well known example of experience goods.[9] Treatment success for this class of drugs tends to have idiosyncratic variation across patients. Therefore, a patient's utility for a particular prescription is unknown until he has spent time taking the drug. Moreover, there are several families of drugs that are known to affect similar patients in similar ways. Thus if one drug is unsuccessful, the patient (or

---

[9]This example is adapted from Dickstein (2014). See his article for a thorough description of the market and a structural model that permits indirect consumption learning within product categories.

more probably his doctor) not only learns that his utility for that drug is low, but also updates his beliefs about the value of taking other antidepressants from the same family.

This sort of learning only within (exogenous) product categories is equivalent to the following variant of the No Indirect Learning axiom.

CATEGORY LEARNING (CL). *Suppose C is partitioned into categories. If c and c' are in distinct categories, then for each $\beta \in [0, 1]$, either $\{c', \hat{\beta}\} \sim_c \{c'\}$ or $\{c', \hat{\beta}\} \sim_c \{\hat{\beta}\}$.*

DEFINITION 8. A PD representation, $(\mathcal{U}, \mu, \delta)$ has a *product structure across categories* if for any c and c' in distinct categories and $u \in \mathcal{U}$,

$$\mu\big(P_c(u)\big) = \mu\big(P_c(u) \mid P_{c'}(u)\big).$$

The proof of Theorem 1 can be easily adapted to show that any PD representation of a preference satisfying CL must have a *product structure across categories*. Alternatively, CL can be viewed as a test for whether an agent's behavior is consistent with a particular (partitional) categorization scheme.

Trade-offs across product attributes[10]   Now consider the case where $C \subset \mathbb{R}^N$. In this setting, consumption objects can be viewed as multi-attribute products. An agent may well be certain of his tastes within each particular attribute. Even so, he may be uncertain of his ranking of all objects in C. Specifically, he may be unsure of the trade-off across attributes.

For instance, when searching for apartments, a renter tends to consider many attributes. Among other features, the renter might value apartment size, age, and price. He may even know he strictly prefers larger, newer, and cheaper apartments. However, he might have difficulty ranking a smaller, cheaper apartment against a larger, more expensive one. To be concrete, consider the following (equally new) apartments:

  Apartment A: 900 sq. ft, $900/month rent

  Apartment B: 500 sq. ft, $900/month rent

  Apartment C: 900 sq. ft, $2500/month rent.

The renter may well be certain that apartment A is better than B or C, yet be uncertain of his relative preference for apartments B and C. For instance, he may be uncertain whether he values the apartments according to $u_1 = \frac{\text{size}}{\text{rent}}$ or $u_2 = \frac{\text{size}}{30\sqrt{\text{rent}}}$. This approximately corresponds to

	u_1	u_2
A	1	1
B	0.56	0.56
C	0.36	0.6

[10]Thanks to Ran Spiegler for bringing this type of taste uncertainty to my attention.

If the renter views u_1 and u_2 as equally likely and apartment A is unavailable, we might expect him to lease apartment B. Notice that under the PD framework, this would not help him to resolve his taste uncertainty. While apartment C has a lower expected consumption value, it also has the benefit of helping the renter to resolve his taste uncertainty. Therefore, the renter may well find it optimal to take a short term lease on apartment C and leave open the option of moving to the smaller, cheaper apartment B once he figures out his preferences. In terms of the PD model, this could occur if δ is sufficiently large. Intuitively, this happens if the agent rents for sufficiently many periods to make the information gain outweigh the expected present consumption loss. This would be an example of experimentation. This idea is explored more formally in the next section.

7. Identifying taste uncertainty

7.1 *Partial identification through experimentation*

The motivating idea behind this paper is the observation that (unobservable) taste uncertainty generates an incentive to experiment and that this (observable) experimentation can in turn reveal an agent's taste uncertainty. In this subsection, I explore the question of what exactly experimentation can reveal. In service of this question, I begin by formally defining what I mean by experimentation.

DEFINITION 9. Given $\succsim$, $\langle a, b, A \rangle$ is an *experimentation tuple* if and only if for every $c \in \mathcal{C}$,

$$\big(a, \{c\}\big) \succ \big(b, \{c\}\big) \quad \text{and} \quad (b, A) \succ (a, A).$$

The idea behind this definition is that we can say the agent is experimenting if and only if he is willing to forego current consumption utility to gain (taste) information relevant to his future choice problem, A. When the follow-on menu is a singleton, there is no instrumental role for information. Therefore, the ranking of current consumption alternatives should reveal the agent's true preference ordering. If the agent's ranking of current consumption choices changes when the follow-on menu is a nonsingleton, then we may infer that he is engaging in costly experimentation.[11] That is, he is willing to sacrifice current consumption utility to obtain new information.

This behavioral definition of experimentation allows the analyst to observe experimentation without needing any knowledge of the payoffs associated with individual consumption objects. Moreover, this does not assume any particular learning rule for the agent. I believe this is the first nonparametric definition of experimentation in the literature. The existing literature on experimentation relies on knowledge of the agent's beliefs, updating rule, and/or payoffs to identify when an agent is experimenting.

---

[11]Requiring Definition 9 to hold for every c is unnecessary in the context of the PD model since this consistency is imposed by the baseline representation. However, defining experimentation in this way allows Definition 9 to be viewed as a general definition of experimentation, independent from the assumptions underlying the PD model. Imposing this consistency requirement rules out alternate causes of the preference reversal, such as preference for variety or habit formation.

With this definition in hand, I move on to the question of what we can learn from observed experimentation. This exercise is motivated by the observation that menu choice data (that is, choice between menus) may be difficult to observe in many empirical settings. In such settings, the standard approach of eliciting subjective states by observing the demand for flexibility cannot be applied.

For instance, doctors do not choose the menu of drugs from which they can write prescriptions; this is set by the marketplace. However, we do observe their choices of what drugs to prescribe. Suppose there are two depressed patients with identical symptoms who go visit a benevolent doctor. The doctor expects the first patient to visit regularly, but believes the second patient is unlikely to return for a future visit. In the first case the doctor will be able to adjust the medication in the future, essentially facing a repeated choice problem over drugs. For the other, the prescription is a commitment to only use a single course of treatment. Antidepressants typically have idiosyncratic efficacy across patients and experimentation can be useful to determine the best drug–patient match. If the doctor prescribes different antidepressants to these two patients, we might infer that he is "experimenting" on the first patient. That is, he assigns a drug that has a lower immediate expected payoff, but may be more informative about future treatment options.[12] Notice in this example that the doctor's preference for flexibility is not observed. Whether the follow-on menu was singleton was simply a characteristic of the patient he was treating and not his own choice. Nonetheless, we could observe his experimentation. Thus, when the data do not permit the demand for flexibility to be observed, it may still be possible to detect experimentation.

The following two propositions demonstrate what this experimentation can teach us about taste uncertainty. Both of them focus on a simple type of binary experimentation, which I call the experimentation pair.

DEFINITION 10. Given $\succsim$, $\langle a, b \rangle$ is an *experimentation pair* if and only if for every $c \in \mathcal{C}$,

$$\left(a, \{c\}\right) \succ \left(b, \{c\}\right) \quad \text{and} \quad \left(b, \{a, b\}\right) \succ \left(a, \{a, b\}\right).$$

First, I observe that experimentation pairs occur in the PD model when b is inferior to a in terms of expected consumption value, but promises to provide (sufficiently) more taste information. This can only happen when there is uncertainty with respect to the relative ranking of a and b. This very intuitive property can be formalized as follows.

PROPOSITION 3. *Fix a PD representation* $(\mathcal{U}, \mu, \delta)$. *If* $\langle a, b \rangle$ *is an experimentation pair, then there are* u *and* u' *in* $\mathcal{U}$ *such that* $u(a) > u(b)$ *and* $u'(a) < u'(b)$.

Proposition 3 shows that observed experimentation is a sufficient (but not necessary) condition to infer *relative* taste uncertainty. Specifically, an agent may believe that there is some subjective state in which $u(b) > u(a)$ and still not be willing to experiment in the first period if the immediate loss in expected utility is too large. Thus Proposition 3

---

[12]Perhaps one class of drugs is unlikely to completely control the depression, but produces a range of minor side effects that are differentially informative about what other antidepressants might work.

can be interpreted as showing that relative taste uncertainty can be revealed by experimentation, but only sometimes. However, the next proposition shows that the *absolute* presence of taste uncertainty is *always* revealed by experimentation.

PROPOSITION 4. *Fix a PD representation* $(\mathcal{U}, \mu, \delta)$ *and some* $c \in \mathcal{C}$. *Then the following statements are true:*

(i) *There are* $u, u' \in \mathcal{U}$ *such that* $u(c) \neq u'(c)$ *if and only if there exists* $\beta \in [0, 1]$ *such that* $\langle \hat{\beta}, c \rangle$ *is an experimentation pair.*

(ii) *If* $\langle \hat{\beta}, c \rangle$ *is an experimentation pair for some* β, *then*

$$V(c) \equiv \mathbb{E}_\mu \big[u(c) \big] = \inf \big\{ \beta \in [0, 1] \mid \langle \hat{\beta}, c \rangle \text{ is an experimentation pair} \big\}.$$

This proposition shows that (i) taste uncertainty *always* manifests itself in experimentation and that (ii) when an agent is uncertain about his tastes for c, the expected utility of c can be identified as the lower bound of the set of βs for which $\langle \hat{\beta}, c \rangle$ is an experimentation pair. So while experimentation behavior alone cannot fully identify the parameters of the PD representation, it nonetheless is useful in diagnosing the presence of taste uncertainty and provides a way to elicit expected utilities.

7.2 *Identification with option menus*

In Section 5, it was demonstrated that c-potential values play a crucial role in the construction of the PD representation. A behavioral description of this concept was already developed; however, a graphical representation reveals another important insight. This subsection uses this graphical representation to fully identify the parameters of a regular fPAEU representation with limited choice data.

Recall that every induced menu preference, $\succsim_c$, has a regular fPAEU representation. In this section, I describe a novel approach to eliciting the representation's parameters, $(\mathcal{U}_c, \mu_c)$. The literature (see DLR) typically relies on studying instances where flexibility is demanded to elicit the subjective state space, $\mathcal{U}_c$. When there exists a known best and worst alternative, I observe that this elicitation can be achieved with substantially less choice data. This is motivated by viewing the βs as money values. Under this interpretation, menus of the form $\{p, \hat{\beta}\}$ can be seen as an analogue to financial options.

DEFINITION 11. A menu, $A \in \mathcal{K}$, is an *option* if $A = \{p, \hat{\beta}\}$, where $p \in \Delta \mathcal{C}$ and $\beta \in [0, 1]$.

Consider the option, $\{p, \hat{\beta}\}$. To make the analogy to financial options clear, I refer to the lottery p as the underlying asset and to β as the strike price. There is (potentially) subjective uncertainty about the value of the underlying lottery, p. Alternatively, the strike price yields a state-independent expected payoff. Here p is analogous to a share of stock and β is analogous to a fixed money value. I refer to the β-equivalent of an option as its price. Formally, for some particular initial consumption c, we have the following definition.
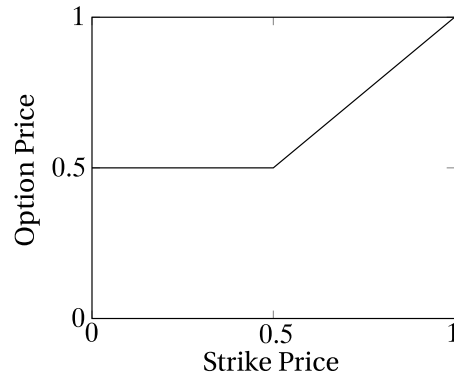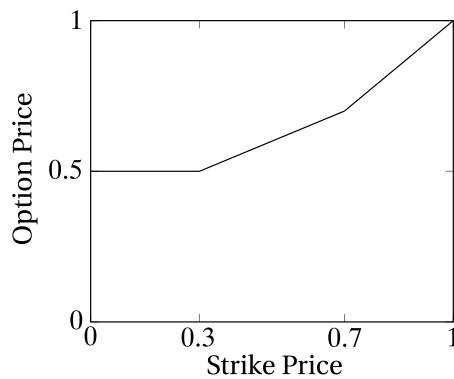
FIGURE 1. No taste uncertainty.

FIGURE 2. Two subjective states.

DEFINITION 12. The value $\beta' \in [0, 1]$ is the *price* of the option $\{p, \hat{\beta}\}$ if $\{p, \hat{\beta}\} \sim_c \{\hat{\beta}'\}$.

It is helpful to think of an option's price, β', as a function of its strike price, β. Such functions are useful for identifying the parameters of the PD representation. Figures 1 and 2 illustrate the essential features of this function under two simple cases.

First, I consider a standard expected utility (EU) maximizing agent who is subjectively certain that the degenerate lottery c has an expected value of 0.5 utils. Behaviorally, this implies whenever $\beta \leq 0.5$, the agent is ex ante certain that he will choose c from the menu $\{c, \hat{\beta}\}$.

Thus, we should expect $\{c, \hat{\beta}\} \sim_c \{c\} \sim_c \{\widehat{0.5}\}$ whenever $\beta \leq 0.5$ and, by similar reasoning, $\{c, \hat{\beta}\} \sim_c \{\hat{\beta}\}$ whenever $\beta > 0.5$. This is illustrated in Figure 1.

Figure 2 illustrates the case of an agent who has subjective uncertainty about the value of c. In particular, he believes that, with equal probability, c is worth either 0.3 or 0.7 utils and that this uncertainty will resolve prior to choosing from the menu. Notice that c is still worth 0.5 utils in expectation. As before, we should expect the price of the option to be 0.5 for low strike prices and equal to the strike price for high strike prices. In particular, $\{c, \hat{\beta}\} \sim_c \{\widehat{0.5}\}$ if $\beta \leq 0.3$ and $\{c, \hat{\beta}\} \sim_c \{\hat{\beta}\}$ whenever $\beta \geq 0.7$. In these cases, the

agent is certain of what he will choose from the option menu even before the value of c is realized. The interesting question is what price the option should take for intermediate values of β. When $\beta \in (0.3, 0.7)$, the agent conditions his choice from the menu on the realized value of c.

After this realization, he determines whether it is better to choose $\hat{\beta}$ or c. Ex ante, the agent believes each utility level of c is equally likely (probability $= \frac{1}{2}$). Therefore, he is (ex ante) indifferent between the option $\{c, \hat{\beta}\}$ and the singleton $\{\frac{1}{2}(\widehat{0.7}) + \frac{1}{2}\hat{\beta}\}$. That is, he believes that with probability 0.5, the state in which the value of c is 0.7 will occur and he will choose c from the menu. The rest of the time he anticipates choosing $\hat{\beta}$. Since we know

$$\{c, \hat{\beta}\} \sim_c \left\{ \frac{1}{2}(\widehat{0.7}) + \frac{1}{2}\hat{\beta} \right\},$$

we can apply the PD representation to fill in the middle part of the function in Figure 2. Specifically, for intermediate values of the strike price (i.e., $\beta \in (0.3, 0.7)$),

$$W\left(c, \{c, \hat{\beta}\}\right) = W\left(c, \left\{ \frac{1}{2}(\widehat{0.7}) + \frac{1}{2}\hat{\beta} \right\}\right),$$

$$V(c) + \delta V_c\left(\{c, \hat{\beta}\}\right) = V(c) + \delta V_c\left(\left\{ \frac{1}{2}(\widehat{0.7}) + \frac{1}{2}\hat{\beta} \right\}\right),$$

$$\underbrace{V_c\left(\{c, \hat{\beta}\}\right)}_{\text{option price}} = 0.35 + \underbrace{0.5}_{\mu\{u|u(c)<\beta\}} * \underbrace{\beta}_{\text{strike price}}.$$

In Figures 1 and 2, notice that the x-axis position of each kink corresponds directly to a c-potential value of c. If instead we consider the option price induced by initial consumption of c' (rather than c), then these kinks will be the c'-potential values of c. If we change the underlying asset from c to p, then the kinks are the c-potential values of p. A function with more kinks corresponds to more potential values and, therefore, more states that the agent expects to become distinguishable prior to choice from the menu.

I have shown that the kinks of the option price curve associated with $\{c, \hat{\beta}\}$ can be interpreted as the c-potential values of the c. Interestingly, we can also interpret the slope of this curve as the agent's subjective cumulative distribution function (CDF) over these c-potential values of c. Specifically, at a strike price of β, the slope of the option curve equals $\mu_c(u \in \mathcal{U}_c \mid u(c) \leq \beta)$. Since the option curve can be derived from behavior, this gives us a way to elicit the $u(c)$ marginals of μ_c. Moreover, the option curves for nondegenerate underlying lotteries allow the full identification of $(\mathcal{U}_c, \mu_c)$. This idea is formalized in Theorem 3.[13]

THEOREM 3. *Consider two induced preferences, $\succsim_a$ and $\succsim_b$, that both have fPAEU representations and share common best and worst alternatives.[14] If, for every option*

---

[13]Lu (2016) proves a result with a similar flavor in the context of an objective state space. He also finds that binary menus are enough to identify his model; however, his technique does not readily generalize in the absence of objective states. Thus the method of proof for Theorem 3 is novel to this paper.

[14]An fPAEU representation, $(\mathcal{U}_c, \mu_c)$, has a known best and worst alternative if there exist c^* and c_* such that for every $u \in \mathcal{U}_c$, $u(c^*) = 1$, and $u(c_*) = 0$, and for all $c' \in \mathcal{C}$, $u(c) \in [0, 1]$. A behavioral characterization

menu, $\{p, \hat{\beta}\}$,

$$\{p, \hat{\beta}\} \sim_a \{\hat{\beta}'\} \quad \Longleftrightarrow \quad \{p, \hat{\beta}\} \sim_b \{\hat{\beta}'\}, \quad then \ \succsim_a = \succsim_b .$$

Theorem 3 shows that if two preferences share a common state-independent best and worst alternative and have distinct fPAEU representations, then they must price some option menu differently. In other words, if the consumption of a and b each induces the same preference over option menus, then they must induce the same preference over all menus. If we restrict attention to a preference known to have a regular fPAEU representation, this theorem implies that preference data on options are sufficient to identify its parameters.

The proof of this theorem is constructive and therefore could be useful in applications. Moreover, it demonstrates how the c-potential value concept aids in model identification. For instance, models that characterize the optimal pricing of experience goods take the distribution of possible utilities as an input. The proof suggests that observing an agent's willingness to pay for option menus would allow a firm to learn this distribution. Of particular interest is the case of related experience goods. There, the construction of the state space reveals the agent's beliefs about the correlation of his utility across goods.

7.3 *Proof of Theorem 3*

The $\hat{\beta}$ lotteries over the common best and worst alternatives provide a numeraire with which to "price" the option menus using each preference. By assumption, $\succsim_a$ and $\succsim_b$ agree on this price for all option menus. Then, by Lemma 1 in the Appendix, both of the induced preferences will have unique regular fPAEU representations, respectively, $(\mathcal{U}_a, \mu_a)$ and $(\mathcal{U}_b, \mu_b)$. The body of this proof constructs a regular pair $(\mathcal{U}, \mu)$ and argues that this must be equal to both $(\mathcal{U}_a, \mu_a)$ and $(\mathcal{U}_b, \mu_b)$.

Use the common price data on options to construct the option price curves for any underlying lottery p. Recall that the kinks in the option price functions correspond to c-potential values of p. Since we are only dealing with the induced portion of the preference, the information process (or, equivalently, the prior consumption, c) can be thought of as given. Therefore, I drop the c and refer to the set of these potential values (for some p) as K_p. Now define the potential state space as

$$\widehat{\mathcal{U}} = \prod_{c \in \mathcal{C}} K_c.$$

Because $\succsim_a$ and $\succsim_b$ have fPAEU representations, we know $\#\widehat{\mathcal{U}} < \infty$. For either preference, the agent's regular subjective state space will be a subset of $\widehat{\mathcal{U}}$. If the agent believed that all his taste uncertainty was uncorrelated across consumption alternatives, we would find that the subjective state space would be this entire set. So the objective is to figure out what correlations the agent believes to be impossible. In other words,

---

of this condition is provided in the Appendix. Moreover, best/worst can be relaxed to $u(c_*) < u(c^*)$ for all $u \in \mathcal{U}_c$. This relaxation is further discussed in a remark at the end of Appendix A.1.

which potential states are given a subjective probability of 0? Choose a lottery q such that there is some $u, u' \in \widehat{\mathcal{U}}$, where $u(q) \neq u'(q)$. If no such q exists, then $\widehat{\mathcal{U}}$ is singleton and the proof is complete.

Now think of q as partitioning $\widehat{\mathcal{U}}$ such that each cell agrees on the utility of q. If q fully separates $\widehat{\mathcal{U}}$, in the sense that all cells are singleton, then we can use the option price curve for q to read off the probabilities of each cell and the proof is once again complete. So the only interesting case is when some cell contains multiple distinct states. Choose two distinct states (u_1 and u_2) in such a cell. Since these states are distinct, they must disagree on the utility of some outcome c. Fix some small $\epsilon > 0$ and define $q' = (1 - \epsilon)q + \epsilon c$. So by construction, $u_1(q') \neq u_2(q')$. Moreover, because there are finitely many states, each cell can be uniformly separated from another in the sense that for some $\delta > 0$, if u and u' are in different cells, $|u(q) - u'(q)| > \delta$. Therefore, for sufficiently small ϵ, q' will generate a strictly finer partition of $\widehat{\mathcal{U}}$ than q did. Once again taking advantage of finiteness, this process can be iterated (a finite number of times) until every cell of the partition consists of only one state. In particular, note that there is some lottery $\bar{q}$ that induces the complete partition of $\widehat{\mathcal{U}}$.

As before, it is straightforward to use the slope of the option price function for $\bar{q}$ to read off the subjective probability of each cell. Let $\mathcal{U}$ be the subset of $\widehat{\mathcal{U}}$ that takes positive probability and let μ be the probability measure on $\mathcal{U}$ constructed as described.

Finally, suppose $(\mathcal{U}_a, \mu_a) \neq (\mathcal{U}_b, \mu_b)$. We already know $\mathcal{U}_a, \mathcal{U}_b \subseteq \widehat{\mathcal{U}}$. So we can simply extend the strictly positive probability measures to the larger set by adding zeros where necessary. However, if the measures differ anywhere, then they generate different option value curves for $\bar{q}$, which violates the assumption that options are priced the same in both preference orderings. Moreover, it is easy to verify that $(\mathcal{U}, \mu)$ must be the fPAEU representation of both preferences. This completes the proof.

7.4 *Eliciting the parameters of the PD model*

A PD model has three main components, $(\mathcal{U}, \mu, \delta)$. The last of these is straightforward to identify from preference data. For instance, when $\delta \leq 1$, it will be exactly equal to the $\beta \in (0, 1]$ such that $(c_*, \{c^*\}) \sim (\hat{\beta}, \{c_*\})$. The two remaining components have a more subtle connection to the observables. Recall that Proposition 1 shows $\mathcal{U} \subset \prod_{c \in \mathcal{C}} K_{c,c} \equiv \widehat{\mathcal{U}}$ and that each $K_{c,c}$ is shown to be observable in Proposition 2. This places an upper bound on the set of utility functions that the agent believes may represent his preferences. Later, Section 7.2 introduced the option value curves. These curves are in principle observable and can be used to infer some of the structure of μ. Specifically, the slope of the option value curve (induced by consumption of c) for $\{c, \hat{\beta}\}$ corresponds to the agent's subjective CDF over c-potential values of c. This is sufficient to identify $\mu(P_c(u))$ for every $u \in \widehat{\mathcal{U}}$. When attention is restricted to PD representations with the product structure, these marginal beliefs are sufficient to identify the model. Moreover, when there is an MIA, $\bar{c}$, $P_{\bar{c}}(u)$ is a singleton for every $u \in \widehat{\mathcal{U}}$. So once again this is sufficient to identify the model.

However, in the general case there are additional degrees of freedom related to the beliefs of the agent conditional on learning some $u(c)$. Given the domain considered

here, there is only enough preference data to infer $\mu\{u \in \widehat{\mathcal{U}} : u(c) = \gamma \wedge u(c') = \gamma'\}$.[15] I cannot, for instance, identify $\mu\{u \in \widehat{\mathcal{U}} : u(c) = \gamma \wedge u(c') = \gamma' \wedge u(c'') = \gamma''\}$. For this reason, Theorem 1 only has a uniqueness result within the product structure class of PD representations. This issue is illustrated in the following example.

Define

$$\mathcal{C} = \{a, b, c^*, c_*\},$$

$$K_{a,a} = \{0, 0.5\},$$

$$K_{b,b} = \{0, 0.5, 1\},$$

$$\mathcal{U} = K_{a,a} \times K_{b,b} \times \{1\} \times \{0\},$$

$$\mu(u) = \frac{1}{6} \quad \forall u \in \mathcal{U},$$

where $K_{a,a}$ is the set of a-potential values of a and $\mathcal{U}$ can be viewed as a set of utility functions on $\mathcal{C}$. Let $u_{i,j}$ denote the utility function in which $u(a) = i$ and $u(b) = j$. Also we can treat μ as a probability measure on $\mathcal{U}$. Notice that K_a and K_b are observable through behavior. Moreover, for any $c \in \mathcal{C}$ and $u \in \mathcal{U}$, $\mu(P_c(u))$ is also observable (this can be determined by the option value curves as seen in the proof of Theorem 3). Now observe that, for any $\delta > 0$, $(\mathcal{U}, \mu, \delta)$ is a regular PD representation with a product structure.

Now consider μ' such that

$$\mu'(u_{0,0}) = \frac{1}{12}, \qquad \mu'(u_{0,0.5}) = \frac{4}{12}, \qquad \mu'(u_{0,1}) = \frac{1}{12},$$

$$\mu'(u_{0.5,0}) = \frac{3}{12}, \qquad \mu'(u_{0.5,0.5}) = 0, \quad \text{and} \quad \mu'(u_{1,1}) = \frac{3}{12}.$$

The PD representation, $(\mathcal{U}, \mu', \delta)$, does not have a product structure. However, it is observably equivalent to $(\mathcal{U}, \mu, \delta)$ on the considered domain (i.e., the two represent the same preference $\succsim$).

8. Concluding remarks

In this paper, I axiomatized a model of taste uncertainty in which an agent learns the cardinal utility of an object upon its consumption. This endogenously generated information produces an experimentation incentive. Using a novel, nonparametric definition of experimentation I show how the agent's taste uncertainty can be inferred from his observed experimentation behavior. This departs from the literature that typically identifies subjective states through an agent's demand for flexibility. In many empirical settings, menu choice data are often unavailable. In such situations, experimentation offers a useful alternative method of eliciting taste uncertainty.

---

[15]Intuitively, this can be determined by looking at how the c-potential values of $(1 - \epsilon)c + \epsilon c'$ are affected by moving from $\epsilon = 0$ to some small positive ϵ. Details of this can be found in the proof of Theorem 2.

Additionally, my model permits indirect learning. Thus consumption of one object can provide taste information about untried alternatives. This indirect learning is a result of correlations in an agent's subjective belief about his tastes. These subjective correlations can be elicited from choice data. To the extent that agents believe their tastes for "similar" products to be correlated, this suggests a method of eliciting a subjective similarity relation. This observation may provide an interesting avenue for future work.

The main interpretation of the preference discovery model is one of complete learning after a single instance of consumption. Adding more periods of choice data would likely allow this to be relaxed so that repeated consumption of a single good can provide increasingly precise taste information. I conjecture that a suitable generalization of the c-potential value concept would prove useful in characterizing this type of partial learning model. Specifically, treating c as a consumption stream would allow a generalized c-potential value to be used to elicit the agent's conditional beliefs.

Appendices

A.1 *Behavioral foundations for the baseline model*

The main text of this paper begins the analysis, taking for granted that the observed preference can be represented by a separable representation based on DLR menu representation. This appendix describes the axiomatic foundations for such a model.

Axiom B1 (Continuous Weak Order). *The relation $\succsim$ is complete, transitive, and, for any pair $(p, A) \in \Delta\mathcal{C} \times \mathcal{K}$, the upper and lower contour sets,*

$$\big\{(p', A') \in \Delta\mathcal{C} \times \mathcal{K} \mid (p', A') \succsim (p, A)\big\} \quad and \quad \big\{(p', A') \in \Delta\mathcal{C} \times \mathcal{K} \mid (p, A) \succsim (p', A')\big\},$$

are closed in the product topology.

Definition 13. The subset $A' \subseteq \text{conv}(A)$ is *critical* for A if for all B with $A' \subseteq \text{conv}(B) \subseteq \text{conv}(A)$, we have $(p, A) \sim (p, B) \; \forall p \in \Delta\mathcal{C}$.

Axiom B2 (Ex post fPAEU). *For each $c \in \mathcal{C}$, the induced menu preference, $\succsim_c$, satisfies the following properties:*

(a) *Monotonicity: $B \subset A \implies A \succsim_c B$.*

(b) *Independence: For any $\alpha \in [0, 1]$ and any $A, A', B \in \mathcal{K}$ (monotonicity),*

$$A \succ_c A' \implies \alpha A + (1 - \alpha)B \succ_c \alpha A' + (1 - \alpha)B.$$

(c) *Finiteness: Every menu, A, has a finite critical subset.*

(d) *Weak Order: $\succsim_c$ is complete and transitive.*

(e) *Continuity: $\{A' \mid A' \succsim_c A\}$ and $\{A' \mid A \succsim_c A'\}$ are closed in the Hausdorff topology.*

Axiom B2 is a set of standard axioms that together are necessary and sufficient for each initial consumption choice, $c \in \mathcal{C}$, to induce a preference over menus, $\succsim_c$, which has an fPAEU representation. This is explored in DLR (2009). Obviously parts (e) and (f) of B2 are implied by B1. Axiom B1 strengthens completeness, transitivity, and continuity to the extended domain considered in this paper rather than only to the domain of menus considered by DLR.

The main text of this paper also makes extensive use of the presence of a state-independent best and worst outcome. Behaviorally this is equivalent to the following condition.

AXIOM B3 (Best/Worst). *There exist c^* and c_* such that for all $c \in \mathcal{C}$ and for all menus $A \in \mathcal{K}$, the following properties hold:*

(a) *Nondegeneracy:* $\{c^*\} \succ_c \{c_*\}$.

(b) *Best always chosen:* $\{c^*\} \sim_c A \cup \{c^*\}$.

(c) *Worst never chosen:* $A \sim_c A \cup \{c_*\}$.

The next three axioms impose some consistency conditions on the types of interaction between the initial consumption and the associated menu preferences.

AXIOM B4 (Rational Expectations). *For $\alpha, \beta, \beta' \in [0, 1]$ and $p, q \in \Delta\mathcal{C}$,*

$$A \sim_p \{\hat{\beta}\} \quad and \quad A \sim_q \{\hat{\beta}'\} \quad \Longrightarrow \quad A \sim_{\alpha p + (1-\alpha)q} \{\alpha\hat{\beta} + (1-\alpha)\hat{\beta}'\}.$$

This axiom is best understood for the case of degenerate consumption lotteries. Let $p = c$ and $q = c'$. Suppose the DM expects the menu A to be worth β given a first period consumption lottery c and worth β' given c'. This axiom requires a first period consumption lottery of $\alpha c + (1 - \alpha)c'$ to lead the DM's expected value of A to be $\alpha\beta + (1 - \alpha)\beta'$. Since I am interested in modeling an agent who learns through the act of consumption, only the (expected) outcome of the consumption lottery should affect the ex ante valuation of the menu.

This axiom rules out any correlation between the first period randomization device and either the agent's taste uncertainty or second stage randomization devices. For instance, if both current and future lotteries were decided by a single (publicly observed) roll of a die, then this axiom would not be satisfied. In this sense, this axiom rules out certain types of information processes.

AXIOM B5 (Weak Separability). *For any $a, b, c, c' \in \mathcal{C}$,*

$$\left(a, \{c\}\right) \succsim \left(a, \{c'\}\right) \quad \Longleftrightarrow \quad \left(c, \{b\}\right) \succsim \left(c', \{b\}\right).$$

Axiom B5 combines two assumptions. First, let $a = b$. Then the above expression reduces to $(b, \{c\}) \succsim (b, \{c'\}) \Longleftrightarrow (c, \{b\}) \succsim (c', \{b\})$. This is a stationarity assumption. It requires that if c is expected to be preferred to c' tomorrow, then it must also be preferred today.

Second, combining the axiom with its implication of stationarity (stated above) shows that B5 also implies $(a, \{c\}) \succsim (a, \{c'\}) \Longrightarrow (b, \{c\}) \succsim (b, \{c'\})$. That is to say, I assume that the induced relations $\succsim_a$ and $\succsim_b$ agree on singleton menus. Notice that it is likely that the consumption of different objects will provide different information that could lead the agent's ranking of menus to be different. However, the induced menu preferences are determined ex ante. The singleton menu is chosen prior to any consumption activity and thus all anticipated information is irrelevant since there is no active choice following the information revelation. Singleton menus will be valued at their ex ante expected value, regardless of the expected information process. A similar axiom is discussed by Hyogo (2007).

AXIOM B6 (Additivity). *(a) For every $c, c', c'' \in \mathcal{C}$ and $p, p', p'' \in \Delta \mathcal{C}$,*

$$(c, \{p'\}) \sim (c', \{p\}) \quad and \quad (c, \{p''\}) \sim (c'', \{p\}) \Longrightarrow (c', \{p''\}) \sim (c'', \{p'\}).$$

(b) For $\alpha \in [0, 1]$ and $p, p', q, q', r, r' \in \Delta \mathcal{C}$,

$$(p, \{p'\}) \succsim (q, \{q'\})$$
$$\Longleftrightarrow \quad (\alpha p + (1 - \alpha)r, \{\alpha p' + (1 - \alpha)r'\}) \succsim (\alpha q + (1 - \alpha)r, \{\alpha q' + (1 - \alpha)r'\}).$$

Part (a) is the well known hexagon condition from Debreu (1960). To understand the hexagon condition, suppose c is better than c', which is better than c'' and similarly for the p's. Now suppose the gain from consuming c instead of c' is exactly offset by switching from p to p' tomorrow and similarly that consuming c instead of c'' is exactly offset by switching from p to p''. Then the hexagon condition requires the gain from consuming c' instead of c'' to be exactly offset by the loss of switching from p' to p'' tomorrow. Part (b) is a weakening of independence such that it only holds when the follow-on menu is restricted to singletons. This weakening (relative to full independence) allows for the possibility that first period consumption can have intertemporal complementarities with nontrivial menus. This is essential if learning is to be valuable to the DM. Additionally, this implies an analogue to the "uniform utility differences" condition introduced in Karni (2004). This analogue is used more explicitly in Hyogo (2007). If I were to expand the domain to $\Delta(\Delta \mathcal{C} \times \mathcal{K})$, then part (a) would be unnecessary and an Anscombe–Aumann-like argument could be used to derive additivity instead.

Together these conditions B1–B6 provide the behavioral foundations for the baseline representation. To formalize this result, I first state a series of lemmas. The first lemma shows that one of the useful features of a domain with a best and worst alternative is to provide a unique regular fPAEU representation. Recall, an fPAEU representation, $(\mathcal{U}_c, \mu_c)$, is regular if $\mathcal{U}_c \subseteq \{u : \Delta \mathcal{C} \to [0, 1] \mid u(\{\beta c^* + (1 - \beta)c_*\}) = \beta\}$ and μ_c is a probability measure with support $\mathcal{U}$.

LEMMA 1. *If $\succsim$ satisfies B2 and B3, then for each induced menu preference, $\succsim_c$, there exists a unique regular fPAEU representation, $(\mathcal{U}_c, \mu_c)$.*

PROOF. It follows from DLR's (2009) Theorem 6 that B2 is sufficient to ensure each induced preference has an fPAEU representation, which I denote $(\mathcal{U}, \mu)$. In this representation, $\mathcal{U}$ is a set of expected utility functions and μ is a positive measure with support $\mathcal{U}$. Axiom B3 implies the existence of universal best and worst alternatives, c^* and c_*. The fPAEU representation, $(\mathcal{U}, \mu)$, can be easily "regularized" as

$$\mathcal{U}_c \equiv \left\{ \frac{u - u(c_*)}{u(c^*)} \,\middle|\, u \in \mathcal{U} \right\} \quad \text{and} \quad \mu_c(u) \equiv \frac{u(c^*)\mu(u)}{\mu(\mathcal{U})}.$$

It is easy to verify that $(\mathcal{U}_c, \mu_c)$ is a regular fPAEU representation of $\succsim_c$. It remains to show that it is the unique one. Notice that the DLR theorem guarantees that $\mathcal{U}_c$ is ordinally unique. That means the set $\mathcal{U}_c$ is identified up to positive affine transformations of its members. However, regularity removes these degrees of freedom by requiring $u(\hat{\beta}) = \beta$. It only remains to show that no probability measure, other than μ_c, will represent $\succsim_c$.

Suppose toward a contradiction that $(\mathcal{U}_c, \mu')$ also is a regular fPAEU representation of $\succsim_c$. Without loss of generality (WLOG), there is some $u' \in \mathcal{U}_c$ such that $\mu_c(u') > \mu'(u')$.

Adapting an argument from DLR, we know there are menus A and B such that

$$\max_{a \in A} u'(a) > \max_{b \in B} u'(b) \quad \text{and} \quad \max_{a \in A} u(a) = \max_{b \in B} u(b) \quad \text{if } u \neq u'.$$

Let $\hat{\beta}_A$ and $\hat{\beta}_B$ be the β-equivalents of A and B, respectively. Since we have $\hat{\beta}_A \sim A$ and $\hat{\beta}_B \sim B$, it must be that

$$\beta_A - \beta_B = \sum_{u \in \mathcal{U}_c} \left[\max_{a \in A} u(a)\mu_c(u) - \max_{b \in B} u(a)\mu_c(u) \right]$$

$$= \sum_{u \in \mathcal{U}_c} \left[\max_{a \in A} u(a)\mu'(u) - \max_{b \in B} u(a)\mu'(u) \right],$$

$$\left(\max_{a \in A} u'(a) - \max_{b \in B} u'(b) \right)\mu_c(u') = \left(\max_{a \in A} u'(a) - \max_{b \in B} u'(b) \right)\mu'(u'),$$

$$\mu_c(u') = \mu'(u').$$

This is a contradiction because we said that $\mu_c(u') > \mu'(u')$. □

The next lemma shows that singleton independence, my B6(b), implies part (b) of the additivity condition used in Hyogo (2007). Moreover, it is straightforward to show that adding B5 implies that if we interchange period consumptions [such that (a, b) becomes (b, a)], this implication will continue to hold.

LEMMA 2. *Suppose $\succsim$ satisfies transitivity and B6(b). Then for every $c, c', c'' \in C$ and $p, p', q, q' \in \Delta C$,*

$$(c, \{p'\}) \sim (c'\{p\}) \quad \text{and} \quad \left(c, \left\{ \frac{1}{2}p' + \frac{1}{2}q \right\} \right) \sim \left(c, \left\{ \frac{1}{2}p + \frac{1}{2}q' \right\} \right)$$

$$\implies \quad (c, \{q'\}) \sim (c'\{q\}).$$

PROOF. By assumption we have

$$(c, \{p'\}) \sim (c', \{p\}) \quad \text{and} \quad \left(c, \left\{\tfrac{1}{2}p' + \tfrac{1}{2}q\right\}\right) \sim \left(c, \left\{\tfrac{1}{2}p + \tfrac{1}{2}q'\right\}\right).$$

Applying singleton independence B6(b), we can mix each side of the first indifference with $\tfrac{1}{2}(c, q)$ to find

$$\left(c, \left\{\tfrac{1}{2}p' + \tfrac{1}{2}q\right\}\right) \sim \left(\tfrac{1}{2}c' + \tfrac{1}{2}c, \left\{\tfrac{1}{2}p + \tfrac{1}{2}q\right\}\right).$$

Then, by the second indifference and transitivity, we find

$$\left(\tfrac{1}{2}c' + \tfrac{1}{2}c, \left\{\tfrac{1}{2}p + \tfrac{1}{2}q'\right\}\right) \sim \left(\tfrac{1}{2}c' + \tfrac{1}{2}c, \left\{\tfrac{1}{2}p + \tfrac{1}{2}q\right\}\right).$$

Another application of B6(b) lets us "unmix" a $\tfrac{1}{2}(c, p)$ from each side, yielding

$$(c, \{q'\}) \sim (c', \{q\}). \qquad \square$$

Using these lemmas, I now state the formal characterization of the baseline representation.

THEOREM 4. *The preference $\succsim$ satisfies B1–B6 if and only if it has a baseline representation, $(\{\mathcal{U}_c, \mu_c\}_{c \in \mathcal{C}}, \delta)$. Moreover, a baseline representation is unique.*

PROOF. Necessity is trivial. To show sufficiency, I adapt Lemma 4 from Hyogo (2007) to the case of a subjective state space. I begin by restricting attention to the restricted domain $\Delta \mathcal{C} \times \Delta \mathcal{C}$. Notice that this space is isomorphic to $[0, 1]^{|\mathcal{C}|} \times [0, 1]^{|\mathcal{C}|}$, and therefore is connected and separable.

So by Debreu (1960), B1 (order and continuity), B3(a) (nondegeneracy), B5 (weak separability), and B6(a) (hexagon condition) imply that there exist continuous and non-constant functions $v : \Delta \mathcal{C} \to \mathbb{R}$ and $w : \Delta \mathcal{C} \to \mathbb{R}$ such that $\succsim$ is represented by

$$\widehat{W}(p, \{q\}) = v(p) + w(q).$$

For each induced preference $\succsim_c$, Lemma 1 pins down a unique pair $(\mathcal{U}_c, \mu_c)$ such that $\succsim_c$ is represented by

$$V_c(A) = \sum_{u \in \mathcal{U}_c} \max_{a \in A} u(a) \mu'(u) \quad \text{for } A \in \mathcal{K}.$$

By B4, when a lottery $p \in \Delta \mathcal{C}$ over prior consumption is chosen, the induced preference $\succsim_p$ can be represented by

$$V_p(A) = \mathbb{E}_p \sum_{u \in \mathcal{U}_c} \max_{a \in A} u(a) \mu'(u) \quad \text{for } A \in \mathcal{K}.$$

When we restrict the domain of V_p to singleton menus, we know that it must be a positive monotonic transformation of w. Moreover, weak separability (Axiom B5) implies

that the representations of all induced preferences must agree on singletons. That is, $V_c(\{c''\}) = V_c(\{c''\})$ for all $c, c', c'' \in \mathcal{C}$. So when they are restricted to singleton menus, the representations are independent of c; thus we can drop the c and just call this function V. Further, the stationarity component of B5 implies that V must also be a positive monotonic transformation of v. Let ϕ and ϕ' be these transformations. Now we can express the representation as follows: constant functions $v : \Delta\mathcal{C} \to \mathbb{R}$ and $w : \Delta\mathcal{C} \to \mathbb{R}$ such that $\succsim$ is represented by

$$\widehat{W}(p, \{q\}) = \phi(V(p)) + \phi'(V(q)).$$

Now I wish to show that ϕ' is affine. Toward that end, Lemma 2 shows that a condition analogous to Karni's (2004) uniform utility differences is implied by B6(b). Following Hyogo (2007), this is sufficient to show that the restricted preference can be represented by

$$\widehat{W}(p, \{q\}) = \phi(V(p)) + V(q).$$

In fact (by an analogue to Lemma 2), B5 and B6(b) together imply a symmetric condition that guarantees that ϕ is also affine. Thus there is a $\delta > 0$ such that

$$\widehat{W}(p, \{q\}) = V(p) + \delta V(q).$$

Further, the uniqueness portion of Debreu's theorem implies that δ is unique. Best/worst and continuous weak order guarantee all menus have a singleton β-equivalent. Thus there is a natural extension of $\widehat{W}$ to $\Delta\mathcal{C} \times \mathcal{K}$. Denote this by

$$W(p, A) = V(p) + \delta V_p(A) = \mathbb{E}_p[v(c) + \delta V_c(A)].$$

This is the baseline representation that we were seeking. Uniqueness follows from Lemma 1 and Debreu.   □

REMARK 1. The results of this paper utilize a state-independent best and worst outcome. In many settings this is an odious assumption. However, it is not necessary. In fact, B3 can be relaxed to the existence of a pair of alternatives that are strictly ranked (the same way) in every subjective state. All of my results can be suitably modified to hold in this setting. This condition can be formalized in terms of the preference relation as follows.

AXIOM B3* (Better/Worse). *There exist c^* and c_* such that for all $c \in \mathcal{C}$ and $A \in \mathcal{K}$, there is a sufficiently small $\epsilon > 0$ such that*

(a) $\{c^*\} \succ_c \{c_*\}$

(b) $\{c^*\} \sim_c \{c^*\} \cup \{(1 - \epsilon)c_* + \epsilon A\}$

(c) $\{(1 - \epsilon)c^* + \epsilon A\} \sim_c \{(1 - \epsilon)c^* + \epsilon A\} \cup \{c_*\}.$

A.2 *Independence of axioms*

In this appendix, I explore the logical independence of the main learning axioms, A1 and A2. To this end, I give examples of learning processes that satisfy one but not the other.

In terms of the PD representation, A1 requires that the signal generated by consumption of c is measurable with respect to P_c, the partition of $\mathcal{U}$ generated by $u(c)$. One implication of A2 is that consuming c should be the best way to learn about $u(c)$.

Consider a model in which consumption produces a signal coarser than P_c. Clearly such a model satisfies A1. In particular, fix some $\epsilon > 0$ and consider the partition P_c^ϵ in which cells are defined by

$$P_c^\epsilon(u) = \big\{ u' \in \mathcal{U} \mid \epsilon \geq \big| u(c) - u'(c) \big| \big\}.$$

If we modify the PD representation such that $P_c^\epsilon(u)$ replaces $P_c(u)$, there will be sets of utility functions, $\mathcal{U}$, such that trying c is no longer the best way to learn $u(c)$. Specifically consider a separating PD representation (with the above modification) with a maximally informative alternative, $\bar{c}$. That is to say $u, u' \in \mathcal{U}$ implies $|u(\bar{c}) - u'(\bar{c})| > \epsilon$. If there are u and u' such that $|u(c) - u'(c)| \leq \epsilon$, then trying c will not be able to separate these states even though $\bar{c}$ could. This sort of "bandwidth" learning would violate A2 while satisfying A1.

For the opposite case, I apply similar logic and consider a learning process by which consumption of c produces a more informative signal than P_c. Any such process will violate A1. The question is whether A2 can continue to hold in this setting. To find the answer, I need only consider the simplest possible case in which every consumption choice produces the same signal (assumed to be a partition finer than P_c). Now A2 is vacuously true. To see this, notice that A2 only has content if consuming some $c' \neq c$ is sometimes more informative about a menu $\{p_n, \hat{\beta}\}$ than consuming c. However, if c and c' generate the same signal (and thus induce identical menu preferences), this can never be relevant.

Finally notice that a preference with a PD representation without the separating qualification necessarily satisfies A1 and A2, but will not have a maximally informative alternative and therefore fails to satisfy A3.

A.3 *Proofs*

This appendix collects the proofs omitted from the main body of the paper. I begin by considering the case of a preference with a baseline representation that satisfies No Indirect Learning (NIL). Recall from Appendix A.1 that a baseline representation specifies a unique set of fPAEU representations $\{\mathcal{U}_c, \mu_c\}_{c \in \mathcal{C}}$ that represent the induced preferences $\{\succsim_c\}_{c \in \mathcal{C}}$. To prove Theorem 1, the definitions

$$\widehat{\mathcal{U}} \equiv \big\{ u \in [0,1]^{|\mathcal{C}|} \mid \text{ for each } c, u(c) = u'(c) \text{ for some } u' \in \mathcal{U}_c \big\},$$

$$P_c(u) \equiv \big\{ u' \in \widehat{\mathcal{U}} \mid u'(c) = u(c) \big\}$$

will be helpful and $\hat{\mu}$ is a probability measure satisfying

$$\hat{\mu}(u) = \prod_{c \in \mathcal{C}} \mu_c\{P_c(u)\} \quad \text{for every } u \in \hat{\mathcal{U}}.$$

Before proceeding to Theorem 1, I provide a lemma that helps illustrate the restriction NIL places on the baseline model.

LEMMA 3. *If $\succsim$ has a baseline representation and satisfies NIL, then the induced preference, $\succsim_c$, can be represented by*

$$V_c(A) = \sum_{u \in \hat{\mathcal{U}}} \max_{q \in A} \left\{ \begin{array}{ll} \sum_{u' \in P_c(u)} u'(q)\hat{\mu}(u' \mid P_c(u)) & \text{if } c \in \text{supp}(q) \\ V(q) & \text{else} \end{array} \right\} \hat{\mu}(u).$$

PROOF. Notice, for every $c' \neq c \in \mathcal{C}$, that NIL implies there is exactly one c-potential value of c'. The proof of Theorem 3 demonstrates that the unique regular fPAEU representation of $\succsim_c$, denoted $(\mathcal{U}_c, \mu_c)$, satisfies $\mathcal{U}_c \subseteq \prod K_{c,c'}$. Here $K_{c,c'}$ refers to the set of c-potential values of c'. Thus distinct subjective states in $\mathcal{U}_c$ can only differ on the valuation of c. Immediately we can see that the agent cannot demand flexibility in menus that do not contain c (or a lottery placing weight on c). Furthermore, because we know that all induced preferences agree on the ex ante value of singletons, we can determine

$$u'' \in \mathcal{U}_c \quad \Longrightarrow \quad u''(c') = V(c') \quad \forall c' \neq c.$$

Noting that V is linear in probabilities, we find that

$$V_c(A) = \sum_{u'' \in \mathcal{U}_c} \max_{q \in A} u''(q)\mu_c(u'') = \sum_{u'' \in \mathcal{U}_c} \max_{q \in A} \left\{ \begin{array}{ll} u''(q) & \text{if } c \in \text{supp}(q) \\ V(q) & \text{else} \end{array} \right\} \mu(u'').$$

Now we can simply identify $u'' \in \mathcal{U}_c$ with the cell $P_c(u'')$. Given this identification, it is straightforward to see the desired numerical equivalence:

$$V_c(A) = \sum_{u \in \hat{\mathcal{U}}} \max_{q \in A} \left\{ \begin{array}{ll} \sum_{u' \in P_c(u)} u'(q)\hat{\mu}(u' \mid P_c(u)) & \text{if } c \in \text{supp}(q) \\ V(q) & \text{else} \end{array} \right\} \hat{\mu}(u). \qquad \square$$

With Lemma 3 in place, the proof Theorem 1 becomes fairly straightforward.

THEOREM 1 (restatement). *A preference, $\succsim$, has a baseline representation and satisfies NIL if and only if it has a PD representation with a product structure. Moreover, a preference, $\succsim$, has at most one PD representation with a product structure.*

PROOF. Necessity is obvious. To show sufficiency, first notice that the baseline representation is numerically equivalent to

$$W(p, A) = \mathbb{E}_p \sum_{u \in \hat{\mathcal{U}}} [u(c) + \delta V_c(A)]\hat{\mu}(u).$$

Now by applying Lemma 3, this can be rewritten as

$$W(p, A) = \mathbb{E}_p \sum_{u \in \widehat{\mathcal{U}}} \left[u(c) + \delta \max_{q \in A} \left\{ \begin{array}{ll} \displaystyle\sum_{u' \in P_c(u)} u'(q)\hat{\mu}\big(u' \mid P_c(u)\big) & \text{if } c \in \text{supp}(q) \\ V(q) & \text{else} \end{array} \right\} \right] \hat{\mu}(u).$$

Taking advantage of the product structure of $(\widehat{\mathcal{U}}, \hat{\mu})$, I observe that for every $u \in \widehat{\mathcal{U}}$,

$$V(q) = \sum_{u' \in P_c(u)} u''(q)\hat{\mu}\big(u'' \mid P_c(u)\big) \quad \text{as long as } c \notin \text{supp}(q).$$

This observation allows me to show that

$$W(p, A) = \mathbb{E}_p \sum_{u \in \widehat{\mathcal{U}}} \left[u(c) + \delta \max_{q \in A} \left\{ \begin{array}{ll} \displaystyle\sum_{u' \in P_c(u)} u'(q)\hat{\mu}\big(u' \mid P_c(u)\big) & \text{if } c \in \text{supp}(q) \\ \displaystyle\sum_{u'' \in P_c(u)} u''(q)\hat{\mu}\big(u'' \mid P_c(u)\big) & \text{else} \end{array} \right\} \right] \hat{\mu}(u),$$

$$W(p, A) = \mathbb{E}_p \sum_{u \in \widehat{\mathcal{U}}} \left[u(c) + \delta \max_{q \in A} \sum_{u' \in P_c(u)} u'(q)\hat{\mu}\big(u' \mid P_c(u)\big) \right] \hat{\mu}(u).$$

Thus the tuple $(\widehat{\mathcal{U}}, \hat{\mu}, \delta)$ is a PD representation of $\succsim$. Given the definitions of $\widehat{\mathcal{U}}$ and $\hat{\mu}$, it is easy to verify that this tuple has the desired product structure.

To establish uniqueness, suppose there is a second PD representation with a product structure, $(\mathcal{U}', \mu', \delta')$, that also represents $\succsim$. Together, the restatements of Propositions 1 and 2 below demonstrate that $\widehat{\mathcal{U}} = \mathcal{U}'$. Suppose $\delta \leq 1$. Find β such that $(\hat{\beta}, c_*) \sim (c_*, c^*)$. Clearly, $\delta = \delta' = \beta$. For $\delta > 1$, find β such that $(c_*, \hat{\beta}) \sim (c^*, c_*)$. Here, $\delta = \delta' = 1/\beta$. Finally, suppose toward a contradiction that $\hat{\mu} \neq \mu'$. Because of the product structure restriction, this can only be true if there is some $c \in \mathcal{C}$ and $u \in \hat{\mathcal{U}}$ for which $\hat{\mu}(P_c(u)) \neq \mu'(P_c(u))$. However, if this were true, then the two representations would imply different subjective CDFs over the c-potential values of c. This is not possible since this would imply different observable behavior that can be seen as changing the slope of the option value curve for $(c, \hat{\beta})$. Hence there is a contradiction and $(\widehat{\mathcal{U}}, \hat{\mu}, \delta)$ must be the only PD representation of $\succsim$ that satisfies the product structure restriction. $\qquad\square$

PROPOSITION 1 (modified restatement). *If $(\mathcal{U}, \mu, \delta)$ is a PD representation and $K_{c,c} \equiv \{u(c) \mid u \in \mathcal{U}\}$, then*

$$\mathcal{U} \subseteq \prod_{c \in \mathcal{C}} K_{c,c}.$$

Moreover, if $(\mathcal{U}, \mu, \delta)$ has a product structure, then $\mathcal{U} = \prod_{c \in \mathcal{C}} K_{c,c}$.

PROOF. The first part of the proposition is trivially true by construction of $K_{c,c}$. The second part of the proposition is more subtle. However, it follows as a consequence of the construction used to prove Theorem 1. If $(\mathcal{U}, \mu, \delta)$ has a product structure, then Theorem 1 implies there is a corresponding baseline representation, $(\{\mathcal{U}_c, \mu_c\}_{c \in \mathcal{C}}, \delta)$. As was shown in the proof of Theorem 1, $\mathcal{U}$ can be expressed as

$$\mathcal{U} = \big\{ u \in [0, 1]^{|\mathcal{C}|} \mid \text{for each } c, u(c) = u'(c) \text{ for some } u' \in \mathcal{U}_c \big\}. \tag{$\star$}$$

Now suppose toward a contradiction that there exists some $u \in \prod K_{c,c} \setminus \mathcal{U}$. For every c, we know by definition of $K_{c,c}$ that there is a $u_c \in \mathcal{U}$ such that $u(c) = u_c(c)$. However, this means that there was some $u'_c \in \mathcal{U}_c$ such that $u'_c(c) = u_c(c) = u(c)$. Since this is true for every $c \in \mathcal{C}$, $(\star)$ implies that $u \in \mathcal{U}$. This is a contradiction and thus proves the result.   □

PROPOSITION 2 (restatement). *Suppose $\succsim$ has a PD representation, $(\mathcal{U}, \mu, \delta)$. Then*

$$\beta \in K_{c,c} \quad \Longleftrightarrow \quad \{\hat{\beta}, c\} \cup \left\{\frac{1}{2}c + \frac{1}{2}\hat{\beta}'\right\} \succ_c \{\hat{\beta}, c\} \quad whenever \; \beta < \beta' \leq 1.$$

PROOF. ($\Longleftarrow$) Let

$$\mathcal{U}_c = \left\{u' \,\middle|\, u' = \sum_{u \in \mathcal{U}} u\mu(u \mid u(c) = \alpha) \text{ for some } \alpha \in [0, 1]\right\} \quad \text{and} \quad \mu_c(u') = \mu(P_c(u')).$$

It is straightforward to see that $(\mathcal{U}_c, \mu_c)$ is a regular fPAEU representation of the induced preference $\succsim_c$. We know that, for every $\beta' \in (\beta, 1]$,

$$\sum_{u \in \mathcal{U}_c} \max\left(u(\hat{\beta}), u(c), u\left(\frac{1}{2}\hat{\beta}' + \frac{1}{2}c\right)\right)\mu(u) > \sum_{u \in \mathcal{U}_c} \max\left(u(\hat{\beta}), u(c)\right)\mu(u).$$

This means that, for some state $u' \in \mathcal{U}_c$, it must be the case that

$$\frac{1}{2}\beta' + \frac{1}{2}u'(c) = u'\left(\frac{1}{2}\hat{\beta}' + \frac{1}{2}c\right) > \max(\beta, u'(c)) \geq \frac{1}{2}\beta + \frac{1}{2}u'(c).$$

This must hold for all β' arbitrarily close to β. This can only be true if $\beta = u'(c)$, and by the construction of $\mathcal{U}_c$, this means that $P_c(u') \subset \mathcal{U}$ is nonempty. Thus $\beta \in K_{c,c}$.

($\Longrightarrow$) This direction is trivial and can be shown by following the above logic in reverse order.   □

In the discussion of A2, I claim that the axiom (together with the baseline representation) implies that "consuming c provides as much information about c's value as does consuming any c'." This is shown below:

PROPOSITION 5. *If $\succsim$ has a baseline representation and satisfies A2, then for any β, $\beta' \in [0, 1]$,*

$$\{\hat{\beta}'\} \sim_c \{c, \hat{\beta}\} \quad \Longrightarrow \quad \{\hat{\beta}'\} \succsim_{c'} \{c, \hat{\beta}\} \quad \forall c' \in \mathcal{C}.$$

PROOF. The supposition that $\succsim$ has a baseline representation ensures that all induced preferences share a commonly agreed upon state-independent best and worst alternative. The contrapositive of A2 implies

$$\beta \notin K_{c,c} \quad \Longrightarrow \quad \{\hat{\beta}'_{p_n}\} \succsim_{c'} \{p_n, \hat{\beta}\} \quad \text{for some } n.$$

In the above expression, p_n is any sequence approaching c and c' is an arbitrary element of $\mathcal{C}$. Clearly the constant sequence $p_n = c$ for all n works. Then, for any $\beta \notin K_{c,c}$, this gives

$$\{c, \hat{\beta}\} \sim_c \{\hat{\beta}'\} \quad \Longrightarrow \quad \{\hat{\beta}'\} \succsim_{c'} \{c, \hat{\beta}\} \quad \forall c' \in \mathcal{C}.$$

As long as $K_{c,c}$ is finite, the continuity of $\succsim_{c'}$ gives the required result. These finiteness and continuity conditions are guaranteed by the fPAEU form. These conditions are formally stated in Appendix A.1.   □

Propositions 3 and 4 demonstrate the implications that can be drawn from experimentation behavior.

PROPOSITION 3 (restatement). *Fix a PD representation, $(\mathcal{U}, \mu, \delta)$. If $\langle a, b \rangle$ is an experimentation pair, then there are u and u' in $\mathcal{U}$ such that $u(a) > u(b)$ and $u'(a) < u'(b)$.*

PROOF. First notice that $\langle a, b \rangle$ is an experimentation pair if and only if

$$\delta\big(V_b(\{a, b\}) - V_a(\{a, b\})\big) > V(a) - V(b) > 0.^{16}$$

Clearly $V(a) > V(b)$ implies that for some $u \in \mathcal{U}$, $u(a) > u(b)$. Suppose toward a contradiction this is true for all $u \in \mathcal{U}$. Then there is no need for taste information since the agent will choose a out of the menu $\{a, b\}$ in every contingency. Thus we would find $V_b(\{a, b\}) = V_a(24\{a, b\})$, but this is a contradiction.   □

PROPOSITION 4 (restatement). *Fix a PD representation, $(\mathcal{U}, \mu, \delta)$, and some $c \in \mathcal{C}$. Then the following statements are true:*

(i) *There are $u, u' \in \mathcal{U}$ such that $u(c) \neq u'(c)$ if and only if there exists $\beta \in [0, 1]$ such that $\langle \hat{\beta}, c \rangle$ is an experimentation pair.*

(ii) *If $\langle \hat{\beta}, c \rangle$ is an experimentation pair for some β, then*

$$V(c) \equiv \mathbb{E}_\mu\big[u(c)\big] = \inf\big\{\beta \in [0, 1] \mid \langle \hat{\beta}, c \rangle \text{ is an experimentation pair}\big\}.$$

PROOF. For part (i), if $\langle \hat{\beta}, c \rangle$ is an experimentation pair (EP), the existence of an appropriate u and u' follows from Proposition 3. Conversely, if there are $u, u' \in \mathcal{U}$ such that $u(c) \neq u'(c)$, then WLOG I can assume $u(c) > V(c) > u'(c)$. Let $\beta = V(c) + \epsilon$. Then $\langle \hat{\beta}, c \rangle$ is an EP if and only if $\delta \cdot \mu[u : u(c) > V(c) + \epsilon] \cdot (\mathbb{E}_\mu[u(c) : u(c) > V(c) + \epsilon] - V(c) - \epsilon) > \epsilon > 0$.

Notice that the finiteness of $\mathcal{U}$ implies that for sufficiently small ϵ, we have both

$$\bar{\mu} \equiv \mu\big[u : u(c) > V(c) + \epsilon\big] = \mu\big[u : u(c) > V(c)\big] > 0 \quad \text{and}$$
$$\bar{V} \equiv \mathbb{E}_\mu\big[u(c) : u(c) > V(c) + \epsilon\big] = \mathbb{E}_\mu\big[u(c) : u(c) > V(c)\big] > V(c).$$

For small ϵ, $\bar{\mu}$ and $\bar{V}$ are independent of ϵ. Now the sufficient condition for $\langle \hat{\beta}, c \rangle$ to be an EP can be rewritten as

$$\bar{\mu} \cdot \big(\bar{V} - V(c)\big) > \frac{(1 + \delta)\epsilon}{\delta} > 0.$$

---

[16]Here $V_c(A)$ is the expected utility of A conditional on consumption c and $V(p)$ is the ex ante expected utility of p. Formally,

$$V_c(A) = \sum_{u \in \mathcal{U}} \max_{q \in A} \sum_{u' \in P_c(u)} u'(q)\mu\big(u' \mid P_c(u)\big)\mu(u) \quad \text{and} \quad V(p) \equiv V_c(\{p\}) = V_{c'}(\{p\}) \quad \forall c, c' \in \mathcal{C}.$$

As long as a sufficiently small positive ϵ is chosen this will be satisfied. This proves part (i). The proof of part (ii) also follows from this argument. By assumption there is some $\langle \hat{\beta}, c \rangle$ that is an EP. Therefore, we know there are WLOG $u, u' \in \mathcal{U}$ such that $u(c) > V(c) > u'(c)$. By repeating the argument above, we can rule out the case that the infimum of $\{\beta \in [0, 1] \mid \langle \hat{\beta}, c \rangle$ is an experimentation pair$\}$ is strictly greater than V(c). The definition of an EP rules out the infimum being strictly less than V(c); thus, we are left to conclude that part (ii) must be true. $\square$

Now I turn to proving the main representation result of the paper.

THEOREM 2 (restatement). *A preference $\succsim$ has a baseline representation and satisfies A1, A2, and A3 if and only if it has a unique separating preference discovery representation.*

PROOF. ($\Longleftarrow$) Necessity of the axioms is straightforward, though it is worth pointing out that A3 is only necessary because of the separating qualification. This qualification ensures there is some $\bar{c}$ that separates the subjective state space, $\mathcal{U}$. This will act as the maximally informative alternative.

($\Longrightarrow$) The existence of a baseline representation means that each induced preference, $\succsim_c$, has an fPAEU representation denoted $(\mathcal{U}_c, \mu_c)$. The strategy of this proof is to show that A1–A3 imply that these various representations are consistent with the representation of the maximally informative alternative $(\mathcal{U}_{\bar{c}}, \mu_{\bar{c}})$.

Step 1. For each c (including $\bar{c}$), $(\mathcal{U}_c, \mu_c)$ is regular and $\mathcal{U}_c$ has a separating form. Specifically, for distinct elements $u, u' \in \mathcal{U}_c$ it is the case that $u(c) \neq u'(c)$.

Fix some $c \in \mathcal{C}$. Regularity is a condition of the baseline representation. Now suppose toward a contradiction that $\mathcal{U}_c$ is not separating. Then there exists $u, u' \in \mathcal{U}_c$, such that $u(c) = u'(c)$ and $u(p) \neq u'(p)$ for some $p \in \Delta\mathcal{C}$. Notice that the set of c-potential values of p can be expressed as

$$K_{c,p} = \{u(p) \mid u \in \mathcal{U}_c\}.$$

Now fix some arbitrarily small $\epsilon > 0$ and perturb p toward c. Call this perturbed lottery $p' = (1 - \epsilon)p + \epsilon c$. Any us that differ on <u>either</u> the value of c or p will also differ on this perturbed lottery. This implies $\#K_{c,p'} > \#K_{c,c}$, which is a violation of A1.

It is now sufficient to show that for each fPAEU representation, $(\mathcal{U}_c, \mu_c)$, the following equalities are true:

$$\mathcal{U}_c = \bigcup_{u \in \mathcal{U}_{\bar{c}}} \left\{ \sum_{u' \in P_c(u)} u'(q)\mu\big(u' \mid P_c(u)\big) \right\} \quad \text{and}$$

$$\mu_c \bigg(\sum_{u' \in P_c(u)} u'(q)\mu\big(u' \mid P_c(u)\big) \bigg) = \mu_{\bar{c}}\big(P_c(u)\big),$$

where $P_c(u) = \{u' \in \mathcal{U}_{\bar{c}} \mid u'(c) = u(c)\}$.

Step 2. For $c \in \mathcal{C}$, $\mu_c(v) = \mu_{\bar{c}}(P_c(v))$ $\forall v \in \mathcal{U}_c$, where $P_c(v) \equiv \{u' \in \mathcal{U}_{\bar{c}} \mid u'(c) = v(c)\}$.

To prove this step it will be useful to work with the (logically equivalent) contrapositive of A2 rather than A2 itself:

AXIOM A2* (Contrapositive of A2). *If $\beta \notin K_{c,c}$ and $p_n \to c$, then for sufficiently large n,* $\{\hat{\beta}'\} \sim_c \{p_n, \hat{\beta}\} \Longrightarrow \{\hat{\beta}'\} \sim_{\bar{c}} \{p_n, \hat{\beta}\}.$

Recall that $\bar{c}$ denotes the maximally informative alternative assumed to exist by A3. The baseline representation requires that both $\mathcal{U}_{\bar{c}}$ and $\mathcal{U}_c$ have only finitely many elements. Let γ_i and λ_i denote the ith highest utility value of c on the sets $\mathcal{U}_{\bar{c}}$ and $\mathcal{U}_c$, respectively. Now define

$$\mu^i = \mu_{\bar{c}}\big(u \in \mathcal{U}_{\bar{c}} : u(c) = \gamma_i\big),$$

$$\eta^i = \mu_c\big(v \in \mathcal{U}_c : v(c) = \lambda_i\big).$$

For example, for $i = 1$, μ^1 is the subjective probability (conditional on trying $\bar{c}$) that $u(c)$ is the highest value that is (subjectively) possible. I proceed by induction over the values of i.

For the base case, I show that $\lambda_1 = \gamma_1$ and $\mu^1 = \eta^1$. Suppose toward a contradiction that $\lambda_1 > \gamma_1$. By finiteness, there exists a $\beta \in (\gamma_1, \lambda_1) \setminus K_{c,c}$. Then by A2*, it must be that

$$\sum_{v \in \mathcal{U}_c} \max\{v(c), \beta\} \mu_c(v) = \sum_{u \in \mathcal{U}_{\bar{c}}} \max\{u(c), \beta\} \mu_{\bar{c}}(u),$$

$$\lambda_1 \eta^1 + \beta\big(1 - \eta^1\big) = \beta$$

$$\implies \quad \eta^1 = 0 \quad \text{(because } \lambda_1 > \beta\text{)}.$$

But given the definition of η^1 and the fact that $\mathcal{U}_c$ is nonempty, this cannot be true. The case where $\lambda_1 < \gamma_1$ is identical. Thus we must have $\gamma_1 = \lambda_1$.

Given this, choose $\beta' \in (\max\{\gamma_2, \lambda_2\}, \lambda_1) \setminus K_{c,c}$. Apply A2* again to find

$$\gamma_1 \mu^1 + \beta'\big(1 - \mu^1\big) = \lambda_1 \eta^1 + \beta'\big(1 - \eta^1\big),$$

$$\lambda_1 \mu^1 + \beta'\big(1 - \mu^1\big) = \lambda_1 \eta^1 + \beta'\big(1 - \eta^1\big) \quad (\lambda_1 = \gamma_1 \text{ from above}),$$

$$\beta'\big(\eta^1 - \mu^1\big) = \lambda_1\big(\eta^1 - \mu^1\big)$$

$$\implies \quad \eta^1 = \mu^1 \quad \text{(because } \beta' < \lambda_1\text{)}.$$

Now assume for $i \leq k-1$, that $\lambda_i = \gamma_i$ and $\mu^i = \eta^i$. Finally for the inductive step, I need to show that this holds for $i = k$.

Suppose toward a contradiction that $\lambda_k < \gamma_k$. Choose $\beta \in (\max\{\lambda_k, \gamma_{k+1}\}, \gamma_k) \setminus K_{c,c}$. Once again apply A2*:

$$\sum_{v \in \mathcal{U}_c} \max\{v(c), \beta\} \mu_c(v) = \sum_{u \in \mathcal{U}_{\bar{c}}} \max\{u(c), \beta\} \mu_{\bar{c}}(v),$$

$$\sum_{i=1}^{k-1} \lambda_i \eta^i + \beta\left(1 - \sum_{i=1}^{k-1} \eta^i\right) = \sum_{i=1}^{k-1} \gamma_i \mu^i + \gamma_k \mu^k + \beta\left(1 - \sum_{i=1}^{k-1} \mu^i\right) - \beta \mu^k,$$

$$(\gamma_k - \beta)\mu^k = 0 \quad \text{(by inductive assumption)},$$

$$\gamma_k = \beta.$$

This is a contradiction.

The case where $\lambda_k > \gamma_k$ is identical. Thus we must have $\lambda_k = \gamma_k$.

WLOG suppose $\gamma_k > \lambda_{k+1} > \min_{u \in \mathcal{U}_{\bar{c}}} u(c)$. Otherwise the proof is trivial. Now choose $\beta' \in (\max\{\gamma_{k+1}, \lambda_{k+1}\}, \gamma_k)$. Another application of A2* yields

$$\sum_{i=1}^{k} \lambda_i \eta^i + \beta'\left(1 - \sum_{i=1}^{k} \eta^i\right) = \sum_{i=1}^{k} \gamma_i \mu^i + \beta'\left(1 - \sum_{i=1}^{k} \mu^i\right),$$

$$(\lambda_k - \beta')\eta^k = (\gamma_k - \beta')\mu^k,$$

$$\eta^k = \mu^k \quad (\text{since } \lambda_k = \gamma_k).$$

This completes Step 2. I have shown that if $v_i \in \mathcal{U}_c$ assigns the ith highest utility value to c on $\mathcal{U}_c$, then

$$\mu^i = \mu_{\bar{c}}\{u \in \mathcal{U}_{\bar{c}} \mid u(c) = v^i(c)\} = \mu_{\bar{c}}\{P_c(v_i)\} = \eta_i = \mu_c\{v \in \mathcal{U}_c \mid v(c) = \lambda_i\} = \mu_c(v_i).$$

The last equality follows since c separates $\mathcal{U}_c$ (by Step 1).

Step 3. For every $b, c \in \mathcal{C}$ and $v \in \mathcal{U}_c$, it must be that

$$v(b) = \sum_{u \in P_c(v)} u(b)\mu_{\bar{c}}(u \mid P_c(v)) = \mathbb{E}_{\mu}[u(b) \mid u(c) = v(c)].$$

From Step 1 we know that each $v \in \mathcal{U}_c$ takes a distinct value for $v(c)$. Index these vs so that v_i takes the ith largest value of $v(c)$. Let $b' = \alpha c + (1 - \alpha)b$. Notice that for α sufficiently close to 1, if the vs were indexed by their valuation of b', the ordering would be the same.

For any $\beta \notin K_{c,c}$, A2* ensures that there is some $\underline{\alpha}$ such that for $\alpha \in (\underline{\alpha}, 1]$ we have

$$\sum_{v \in \mathcal{U}_c} \max\{v(b'), \beta\}\mu_c(v) = \sum_{u \in \mathcal{U}_{\bar{c}}} \max\{u(b'), \beta\}\mu_{\bar{c}}(u).$$

Picking a sufficiently large α ensures

$$\sum_{\{i|\lambda_i > \beta\}} v_i(b')\eta^i + \sum_{\{i|\lambda_i \leq \beta\}} \beta\eta^i = \sum_{\{u \in \mathcal{U}_c \mid u(c) > \beta\}} u(b')\mu_{\bar{c}}(u \mid u(c) > \beta) + \sum_{\{i|\gamma_i \leq \beta\}} \beta\mu^i.$$

Apply Step 2 to find

$$\sum_{\{i|\gamma_i > \beta\}} v_i(b')\mu^i = \sum_{\{u \in \mathcal{U}_c \mid u(c) > \beta\}} u.$$

Since v and u are EU functions, apply linearity to find

$$\sum_{\{i|\gamma_i > \beta\}} [\alpha v_i(c) + (1 - \alpha)v_i(b)]\mu^i$$

$$= \sum_{\{u \in \mathcal{U}_c \mid u(c) > \beta\}} [\alpha u(c) + (1 - \alpha)u(b)]\mu_{\bar{c}}(u \mid u(c) > \beta).$$

One more application of Step 2 gives

$$\sum_{\{i\mid\gamma_i>\beta\}} v_i(b)\mu^i = \sum_{\{u\in\mathcal{U}_c\mid u(c)>\beta\}} u(b)\mu_{\bar{c}}\big(u\mid u(c)>\beta\big).$$

Choose $\beta_1\in(\gamma_2,\gamma_1)\setminus K_{c,c}$. Then we have

$$v_1(b) = \sum_{\{u\in\mathcal{U}_c\mid u(c)=\gamma_1\}} u(b)\mu_{\bar{c}}\big(u\mid u(c)=\gamma_1\big)\quad\text{(by Step 2)}$$

or, equivalently,

$$v_1(b) = \mathbb{E}_\mu\big[u(b)\mid u(c)=v_1(c)\big].$$

Now by choosing $\beta_2\in(\gamma_3,\gamma_2)\setminus K_{c,c}$, we can use similar logic to find

$$v_1(b)\mu^1 + v_2(b)\mu^2 = \mathbb{E}_\mu\big[u(b)\mid u(c)=\gamma_1\big]\mu^1 + \mathbb{E}_\mu\big[u(b)\mid u(c)=\gamma_2\big]\mu^2,$$

$$v_2(b) = \mathbb{E}_\mu\big[u(b)\mid u(c)=\gamma_2\big].$$

Iterating this procedure completes the step. Together Steps 2 and 3 show that all of the induced fPAEU representations take the desired form. Specifically, they can all be viewed as coarsenings of the representations induced by $\bar{c}$.

Step 4. Within the class of separating PD representations, there is a unique tuple that represents $\succsim$.

Notice that for any separating PD representation, $(\mathcal{U},\mu,\delta)$, satisfying A3, it must be that $(\mathcal{U},\mu)$ is the regular fPAEU representation of the menu preference induced by the maximally informative alternative. Therefore, uniqueness of the separating representation can be established by the uniqueness of the regular fPAEU representation (Lemma 1) and the fact that δ can be uniquely elicited from time preference over $\hat{\beta}$s. $\quad\square$

References

Debreu, Gerard (1960), *Topological Methods in Cardinal utility. Stanford University Press*, Stanford. [1335, 1337]

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2001), "Representing preferences with a unique subjective state space." *Econometrica*, 69, 891–934. [1307, 1313]

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2009), "Temptation driven preferences." *The Review of Economic Studies*, 76, 937–971. [1313, 1334, 1336]

Dickstein, Michael J. (2014), "Efficient provision of experience goods: Evidence from antidepressant choice." Unpublished, Stanford University. [1323]

Dillenberger, David, Juan Lleras, Philipp Sadowski,and Norio Takeoka (2014), "Subjective learning." *Journal of Economic Theory*, 153, 287–312. [1311, 1321]

Hyogo, Kazuya (2007), "A subjective model of experimentation." *Journal of Economic Theory*, 133, 316–330. [1311, 1314, 1335, 1336, 1337, 1338]

Karni, Ed (2004), "Additive representations over actions and acts." *Mathematical Social Sciences*, 48, 113–119. [1335, 1338]

Ke, Shaowei (2013), "Consumption experimentation externality." Unpublished, Princeton University. [1311]

Kreps, David M. (1979), "A representation theorem for 'preference for flexibility'." *Econometrica*, 47, 565–577. [1307, 1310, 1312]

Krishna, Vijay and Philipp Sadowski (2014), "Dynamic preference for flexibility." *Econometrica*, 82, 655–703. [1311]

Lu, Jay (2016), "Random choice and private information." *Econometrica*, 84, 1983–2027. [1329]

Piermont, Evan, Norio Takeoka, and Roee Teper (2016), "Learning the Krepsian state space: exploration through consumption." *Games and Economics Behavior*, 100, 69–94. [1311, 1312, 1322]

Riella, Gil (2013), "Preference for flexibility and dynamic consistency." *Journal of Economic Theory*, 148, 2467–2482. [1311]