A Theory of Preference Discovery

Jason Delaney*[1], Sarah Jacobson†[2], and Thorsten Moenig‡[3]

[1]*School of Business, Georgia Gwinnett College, 1000 University Center Lane, Lawrenceville, GA 30043, USA*
[2]*Department of Economics, Williams College, 24 Hopkins Hall Dr., Williamstown, MA 01267, USA*
[3]*Department of Risk, Actuarial Science & Legal Studies, Temple University, 1801 Liacouras Walk, Philadelphia, PA 19122, USA*

January 2025

Abstract

Is the assumption that people automatically know their own preferences innocuous? We present a theory that explores the implications of having to discover one's preferences. We show that if tastes must be learned through experience, preferences for some goods will be learned over time, but preferences for other goods will never be learned. This is because sampling a new item has an opportunity cost. Learning is less likely for people who are impatient, risk averse, low income, or short-lived, and for consumption items that are rare, are expensive, must be bought in large quantities, or are initially judged negatively relative to other items. Choices will eventually stabilize, but they need not stabilize at true preferences. A pessimistic bias about untried goods

*`jdelaney@ggc.edu`; ORCID 0000-0001-7980-1705

†`sarah.a.jacobson@williams.edu`; Corresponding author; ORCID 0000-0002-2052-5627

‡`moenig@temple.edu`; ORCID 0000-0002-7740-5570

should increase with time. Agents will make choice reversals during the learning process. Welfare loss from suboptimal choices will decline over time but need not approach zero. Overall, our results imply that undiscovered preferences could confound interpretation of choice data of all kinds and could have significant welfare and policy implications.

Keywords: discovered preferences, preference stability, learning, behavioral economics
JEL codes: D81, D83, D01

1 Introduction

> "... this time she found a little bottle on it, ("which certainly was not here before," said Alice,) and round the neck of the bottle was a paper label, with the words "DRINK ME," beautifully printed on it in large letters.
> It was all very well to say "Drink me," but the wise little Alice was not going to do *that* in a hurry."
>
> *Alice's Adventures in Wonderland*
> Lewis Carroll

What we thought was our dream car, after three months, needs a fill-up so often that that annoyance is not counterbalanced by pleasure at the car's sleek profile. After choosing a treatment plan, a patient realizes the side effects just aren't worth it. Regretting a choice we make is an everyday occurrence, from a disappointing lunch order to a tragically mismatched partnership. Despite the universal experience of this gap between our choices and our preferences, economists typically assume that people know their own preferences and that their choices reflect this knowledge, or at least that deviations from that perfection are simply uniform noise. Could ignorance of our own preferences have more complicated ramifications? Delaney et al (2020) use an experiment to test hypotheses about preference discovery and welfare loss. They find that participants learn most but not all of their preferences, and this leads to persistent and economically significant welfare loss, with empirical regularities in where those gaps are the largest.

These results highlight a need for a theoretical account of how and when we learn our preferences. In this paper, we develop a model that formalizes the process of discovering our preferences, with a focus on the extensive margin (which preferences are learned), to understand when and why these losses might occur. The model produces stable but suboptimal choice patterns in the long run, with systematic pessimism about untried goods and persistent welfare loss in patterns congruent with Delaney et al (2020).

Neoclassical microeconomics presents analytically tractable models that yield efficiency from relatively simple assumptions and mechanisms: agents with a set of known preferences participating in voluntary exchange can achieve the best of all possible worlds.

These models are simplistic, and so in both theory and experiment, in economics and psychology, scholars have tested the resilience of these models to the introduction of more realism. In the psychology literature, models often entirely eschew the very idea of stable preferences (Ariely et al, 2003; Lichtenstein and Slovic, 2006), making "what one likes" an elusive concept. Kahneman et al (1997), Scitovsky (1976) argued that people are bad at predicting their satisfaction from a prospective choice. Becker (1996) argued the opposite, while Lerner et al (2015) highlight the importance of emotion in decision-making. Empirical evidence from the psychology literature is mixed, although it reveals some notable variation in people's ability to choose according to underlying tastes. Kahneman and Snell (1990) note that people seem fairly good at predicting utility in familiar choices with immediate feedback. Loewenstein and Adler (1995) find people fail to predict changes in their own tastes. Wilson and Gilbert (2005) review extensive evidence showing systematic errors in forecasting happiness. In economics, the idea of preference discovery was introduced by Plott (1996). The small economic literature that sheds light on the phenomenon has focused primarily on the intensive margin (the updating process) rather than the extensive margin (the extent to which learning covers the set of feasible choices). Ferreira and Gravel (2024) offer a general framework for theories governing the sampling and feedback process. Feedback has been a central feature in related work, beginning with Plott (1996). With feedback, choices tend to become more consistent with neoclassical predictions. van de Kuilen and Wakker (2006), for example, find that with feedback, Allais violations decrease, while Weber (2003) finds learning improves with feedback in strategic games. The fact that feedback matters suggests that preferences may indeed need to be learned through experience.

Indirect evidence of preference discovery processes comes from studies of choice inconsistency ("preference reversals" in the literature, as in Cox and Grether, 1996; or as we prefer to call them, choice reversals). Choice reversals are sometimes treated as indications that subjects fail to understand the elicitation method (Charness et al, 2013) and "choice consistency" is often implicitly taken as indicative of higher accuracy in capturing subjects' underlying preferences or of better translation of preferences into decision-making (Rigby et al, 2015; Bruine de Bruin et al, 2007). However, as we show in our model, a choice reversal is a natural consequence of preference learning. Noussair et al (2004) find that people can converge to an induced value with repeated choices, and this convergence with repetition has also been observed in the contexts of willingness-to-pay vs. willingness-to-accept, non-dominant bidding behavior, and strategic games (Coursey et al, 1987; Shogren et al, 1994, 2001; List, 2003), which also implies some learning, though these settings conflate preference learning with learning about the institution (Braga and Starmer, 2005).

This paper theoretically models preference learning. Our contribution is to develop a theory that integrates preference discovery, as described in Plott (1996), into a neoclassical microeconomic framework, focusing on the extensive margin: what items an agent will and will not learn her tastes for. We maintain the assumption of stable underlying preferences but allow for a need to learn them through experience.

Our model is of a sophisticated agent: she knows for each consumption item whether she has already learned her tastes for it, and she maximizes a discounted stream of expected net benefits, so she will intentionally sample some goods to learn how much she likes them. However, learning has an opportunity cost, and since the benefits of learning are finite, learning will not be complete. We show that the agent will exhibit choice reversals as she learns her preferences. We also demonstrate that she will learn her preferences over time for many goods, thus reducing her welfare loss from bad choices. However, she may never fully learn her preferences. Learning failure is more likely for people who are impatient, are risk averse, have low income, or have a short lifespan, and for goods that are initially undervalued relative to other goods, expensive, rare, or that require a large minimum quantity to be consumed in order to be learned. We also show that a more diffuse expectation about the good's parameters could make the good more or less likely to be learned, depending on the implied likelihood that the good is better than the outside options. We discuss that as the agent lives and learns, she is expected to become more pessimistic, since optimistic errors in prior beliefs are more likely to be corrected than pessimistic errors.

The remainder of the paper proceeds as follows. First, we lay out the model structure. Next, we derive results. We then discuss features that are not formalized in our model but that our model allows us to conjecture about. Finally, we conclude.

2 A Model of Preference Discovery

We begin by building a simple model of decision-making for an agent named Alice. Throughout the model, we make many unrealistic simplifying assumptions. These are intended to make the learning process relatively trivial. For example, as we describe shortly, we assume a pathologically primitive utility function so there is very little to learn. We do this because we are interested in the cases in which Alice fails to learn her preferences for something at the extensive margin, and our assumptions make learning unrealistically easy; any failures we highlight in our simple model will generally be made worse by more complex, realistic assumptions. That is, we give preference discovery its best shot so we can highlight its failures.

All proofs are in Appendix A.

2.1 Alice's Tastes

Alice has tastes over $N \in \mathbb{N}$ goods, $i = 1, \ldots, N$. Alice makes a consumption choice in each of the $T \in \mathbb{N}$ time periods, $t = 0, \ldots, T$, in her life: she chooses a bundle from the subset of goods that are available at that time. We use x_i to denote a quantity of good i, and x_i^t as the quantity of good i consumed at time t. We use fruits as our examples of consumption items; thus, in each period, imagine that Alice is confronted with a random basket of fruits from which she can choose her consumption bundle.

We use the term "goods" quite generally, as some might be "bads" and they may represent goods, services, experiences, or attributes. We limit our consideration to deterministic goods: within a type of good, units are undifferentiated and identical in quality.

We assume that Alice has an underlying preference ordering $\succsim$ over bundles $x = (x_1, \ldots, x_N)$ (where each $x_i \geq 0$) of these goods, and that this ordering obeys the standard assumptions of rational preferences, though we will return shortly to monotonicity and convexity.

Axiom 1. Rational Preferences.
Preferences are continuous, reflexive, complete, and transitive.

□

We can therefore represent Alice's tastes with a utility function $u(.)$.[1] Alice knows the form of $u(.)$, but may not know its precise shape. In partic-

[1]We use a utility function for convenience; our conceptual points about preference

ular, we assume she knows the functional form of her utility function but not necessarily its parameters. Her utility is determined by consumption levels as well as $N_1 \geq N$ parameters that can be arranged in a vector β. We denote the true parameters of $u(.)$ by $\hat{\beta} \in \mathbb{R}^{N_1}$, so that her true utility is $u(x; \hat{\beta})$.

We assume that Alice's true utility function determines the utility she realizes from consumption, and we assume that this true utility function and its parameter vector $\hat{\beta}$ are time-invariant:

Axiom 2. Stability of True Preferences.
At any time $t \geq 0$, the agent's realized utility from consuming a bundle of goods x is $u(x; \hat{\beta})$.

□

However, at any time t, Alice may not know all of her true parameter values. Instead, she has beliefs about these parameters. These beliefs are not point estimates because she is sophisticated enough to know when she has not yet learned her tastes: her beliefs are probability distribution functions over possible values. Therefore, we represent Alice's time-t beliefs about her preferences with a (N_1-dimensional) random variable, denoted by β^t. This random variable has a continuous sample space, which is a subset of $\mathbb{R}^{N_1}$. We let $\mathbb{B}$ denote the set of all random variables that assign a positive probability to possible preference vectors in the neighborhood of the true preferences $\hat{\beta}$. That is,

$$\mathbb{B} = \left\{ \beta \mid \forall \epsilon \in \mathbb{R}^{N_1} \text{ with } \epsilon > 0 : \mathbb{P}\left(\beta \in (\hat{\beta} - \epsilon, \hat{\beta} + \epsilon)\right) > 0 \right\} . \tag{1}$$

The random variable β is characterized by a N_1-dimensional probability density function (p.d.f.) $f^{(\beta)}(b) : \mathbb{R}^{N_1} \rightarrow \mathbb{R}_0^+$. We use $b \in \mathbb{R}^{N_1}$ to denote potential outcomes of the random variable β, that is, potential parameter vectors. Thus, Alice's expected utility from consuming bundle x at time t—given her current preference beliefs in the form of the random variable β^t—is

$$Eu(x; \beta^t) = \int_{\mathbb{R}^{N_1}} f^{(\beta^t)}(b) \cdot u(x; b) db \, .$$

The p.d.f. of Alice's true preferences $\hat{\beta}$ is $\hat{f}(.) = \Delta(\hat{\beta})$, where Δ denotes the Dirac delta function, so that $\hat{f}(b)$ has infinite weight for $b = \hat{\beta}$—such that $\mathbb{P}(b = \hat{\beta}) = 1$—and $\hat{f}(b) = 0$ for all other b.

learning can also be made using just preference rankings, as we did in an earlier version of this paper, titled "Discovered Preferences for Risky and Non-Risky Goods."

Alice's prior beliefs about her preferences before she has had any experience are reflected in the random variable $\beta^0 \in \mathbb{B}$ and described by the joint p.d.f. $f^{(\beta^0)}(b)$. These prior beliefs are exogenous and need not be correct; Wilson and Gilbert (2005) review the evidence that people routinely err in forecasting their utility.

Thus, for each fruit, Alice has true preferences that are exogenous parameter values and she has priors that are exogenously-given probability distribution functions over parameters. While both of these are deterministic, her believed preferences at any time t are, as we will show, not deterministic because the process of encountering fruits (and thus potentially learning her true values) is random.

Next, we assume that Alice's utility function is additively separable:

Axiom 3. Separability of Utility.
For all $i, j \in \{1, \ldots, N\}$ with $i \neq j$, and for all $b \in \mathbb{R}^{N_1}$: $\frac{\partial^2 u(x;b)}{\partial x_i \partial x_j} = 0.$

□

As a result of Axiom 3, Alice has a sub-utility function $u_i(.)$ that determines her utility from each good i, and we can state Alice's utility as:

$$u(x; b) = u_1(x_1; b_1) + \ldots + u_N(x_N; b_N)\,.$$

Now we can define, for $i = 1, \ldots, N$, the real-valued vector b_i as a potential realization of the (possibly multi-dimensional) random variable $\beta_i \in \mathbb{B}_i$ pertaining to the sub-utility function Alice has for good i, and $\mathbb{B}_i$ as the set of all random variables of the same dimension as $\hat{\beta}_i$ that assign a positive probability to preference parameters in the neighborhood of $\hat{\beta}_i$, akin to Equation (1). We can form the overall parameter vector, random variable space, and outcome vector as $\beta = (\beta_1, \ldots, \beta_N) \in \mathbb{B}$, $\mathbb{B} = \mathbb{B}_1 \times \ldots \times \mathbb{B}_N$, and $b = (b_1, \ldots, b_N)$, respectively. In addition, we denote the p.d.f. of the random variable β_i^t by $f_i^{(\beta_i^t)}(.)$.

We assume preferences for each item are (weakly) monotonic, but we allow some goods to give positive and some to give negative marginal utility. We do not restrict Alice's beliefs about a good to a single domain (positive or negative): before she has tried it, she may think that a kumquat is likely to be good but has a chance of being bad. We assume preferences are (weakly) convex, which implies a (weakly) concave utility function for each good.

Axiom 4. Shape of Utility Function.

For each $i \in \{1, \ldots, N\}$, the good-i sub-utility function $u_i(.)$ is twice differentiable. We further assume that it is weakly monotonic and weakly concave; that is, for all $b \in \mathbb{R}^{N_1}$ and all $i \in \{1, \ldots, N\}$:

(i) Monotonicity: Either $\frac{d\,u_i(x_i;b_i)}{d\,x_i} \geq 0$ for all $x_i \geq 0$, or $\frac{d\,u_i(x_i;b_i)}{d\,x_i} \leq 0$ for all $x_i \geq 0$.

(ii) Concavity: $\frac{d^2\,u_i(x_i;b_i)}{(d\,x_i)^2} \leq 0$ for all $x_i \geq 0$.

$\square$

2.2 Alice's World

At discrete times $t = 0, \ldots, T$, Alice has access to a random subset, denoted by G^t, of the universe of goods. It is from the goods in G^t that Alice constructs her consumption bundle at time t. The likelihood that good i is available at time t is time-invariant and independent of the availability of any other good. We denote this probability by $q_i := \mathbb{P}(i \in G^t)$ and we require that $0 < q_i < 1$ for $i = 2, \ldots, N$.

In addition to ordinary goods $i = 2, \ldots, N$, there is also a numeraire good, which we index with $i = 1$. The numeraire good is present at all times, so that $q_1 = 1$. The other special feature of the numeraire good is that Alice knows with certainty that it provides a constant marginal utility of $z > 0$. The numeraire good can be thought of as the option to consume nothing, or as some basic good (like bread) that is always available.

At each time t, Alice is endowed with income y, and that income does not change over time. Money cannot be transferred across time periods. The price per unit of good i is also time-invariant and is denoted by $p_i > 0$.

2.3 Experience and Preference Learning

As noted above, Alice's utility is determined by her true utility function, governed by true parameters $\hat{\beta}$, but Alice may not always know her true parameters. Instead, at time t she chooses according to a utility function parameterized by beliefs β^t, which is a random variable with density function $f^{(\beta^t)}(.)$ and which at $t = 0$ is equal to prior beliefs β^0. Alice learns about her tastes by consuming the goods and updates these parameters accordingly.

We make several assumptions about the preference updating process. First, we assume that there exists a "nibble size" or minimal consumption experience m_i for each good i such that if Alice consumes at least this nibble, she accurately perceives her utility from the good, but if she consumes less, she does not. This is like assuming that if Alice gets an atom of an apple on her tongue, it does not inform her about her taste for apples, but if she eats at least a mouthful she learns her taste for apples fully.[2] Second, we assume that Alice can perceive the separate sub-utilities from each good of which she consumes at least a nibble, rather than only perceiving the utility of the bundle, making the consumption items more like different foods on a plate than like inseparable attributes of a product.

Axiom 5. Experience of Utility.
If she consumes a bundle that includes x_i units of good i, Alice gets utility $u_i(x_i;\hat{\beta}_i)$ from good i in addition to any other utility she earns at the same time. If $x_i \geq m_i$, she accurately perceives that utility $u_i(x_i;\hat{\beta}_i)$. If $x_i < m_i$, she does not perceive how much utility she got from good i nor how it contributes to the utility she got from the overall bundle.

□

Axiom 5 implies that Alice learns the full set of parameters $\hat{\beta}_i$ for good i from a single meaningful consumption experience of this good. Therefore we can assume, without loss of generality, that each $\hat{\beta}_i$ is one-dimensional. This implies that for each good $i \in \{1,\ldots,N\}$, $u_i(.)$ is characterized by a single parameter $\hat{\beta}_i \in \mathbb{R}$, and thus we have $N_1 = N$. For ease of exposition, we make two additional assumptions: First, we normalize utility derived from each good to zero if the good is not consumed, so $u_i(0;b_i) = 0$ for all i and all $b_i \in \mathbb{R}$. And second, we specify that larger parameter values always imply (weakly) larger utility; that is, for each good i, $\frac{\partial u_i(x_i;b_i)}{\partial b_i} \geq 0$.

The requirement that Alice have at least minimal consumption of a good to perceive how she likes it, combined with the existence of a numeraire good that is always available, ensures that the opportunity cost for learning an untried good is non-zero and non-vanishing. The presence of the numeraire good ensures that no good ever appears alone or with only unattractive goods. The numeraire thus ensures that every consumption decision carries

[2]What does it mean for Alice to consume a small amount of a good, not know how much utility she gained, but still in some sense earn that utility? Our interpretation of m_i is that it is finite but so small that the tiny amount of utility it yields is difficult for Alice to precisely quantify. In some cases in which the minimal consumption size is nontrivial, like lumpy items such as houses and spouses, it simply is not feasible to consume a quantity less than the nibble size.

an opportunity cost. Also, if we did not require at least a nibble to learn, then Alice could learn her tastes by purchasing an infinitesimally small quantity of each good when it appears for a negligible cost, so she would always fully learn her preferences, as happens in the theories of Easley and Kiefer (1988) and Aghion et al (1991). We make these assumptions because opportunity cost is intuitively important in extensive margin consumption decisions (whether to consume) like those we study.

Thus, following at-least-minimal consumption of a good, Alice perceives its value to her unerringly. We assume this immediate and perfect assessment because our focus is on cases in which her learning might be incomplete as a result of failure to try goods rather than the dynamics by which learning progresses; we do not study the updating process but rather the case of items that are never sampled, since with most reasonable learning processes, goods that are sampled will eventually be learned, and problems only occur when goods are never sampled.

Axiom 5 implies a unique mapping between utility received from a good and the parameter value(s) for that good. Therefore, Alice should update her beliefs about her preferences based on the utility she experiences in time period t from any previously undiscovered good i of which she consumed at least m_i units.

We disallow spillovers in learning by assuming that consumption of one good is uninformative for learning the parameters associated with other goods, so that tasting an apple does not help learn preferences for oranges; we discuss the relaxation of this assumption in the Discussion.

Axiom 6. Separability of Learning.
Experiencing a good has no effect on the agent's perceived parameters of any other good.

$\square$

Axiom 6 implies that for all $i \neq j$ and for all times s and t, β_i^t and β_j^s vary independently from each other. That is, a change in β_i^t does not lead to a change in β_j^s. As a result, for all $\beta \in \mathbb{B}$:

$$f^{(\beta)}(b) = f_1^{(\beta_1)}(b_1) \cdot \ldots \cdot f_N^{(\beta_N)}(b_N) \text{ for all } b = (b_1, \ldots, b_N) \in \mathbb{R}^N. \tag{2}$$

Moreover, we assume that, once learned, parameters are not forgotten.

Axiom 7. Persistent Memory.

If for some time t, $f_i^{(\beta_i^t)} \equiv \hat{f}_i$, then $f_i^{(\beta_i^s)} \equiv \hat{f}_i$ for all $s \geq t$.

$\square$

Together, Axiom 6 and Axiom 7 ensure that believed parameters for some good i only change with experience with good i. This implies that:

Lemma 1. Updating of Preference Beliefs.
For each good $i \in \{2, \ldots, N\}$:

(a) If $x_i^t < m_i$, then $f^{(\beta_i^{t+1})} \equiv f^{(\beta_i^t)}$.

(b) If $x_i^t \geq m_i$ for any t, then $f^{(\beta_i^s)} \equiv \hat{f}_i$ for all $s \geq t+1$.

(c) For all t, $f_i^{(\beta^t)} \in \left\{ f_i^{(\beta_i^0)}, \hat{f}_i \right\}$.

$\square$

That is, if Alice doesn't have at least a nibble of the good, her believed preferences will not change, and if she does, then her believed preferences will become forever stable at her true preferences. Since her preferences start at her priors and can only change to her true values, her believed preferences will always be her prior or her true value.

2.4 Alice's Optimization Problem

At each time $t \in \{0, \ldots, T\}$, Alice decides how much to consume of each good $i \in G^t$. We denote the time-t consumption bundle as $x^t = (x_1^t, \ldots, x_N^t)$. If Alice existed for only one period, or was fully myopic so that she only considered one time period at a time, she would face the following static expected utility maximization problem:

$$U(f^{(\beta^t)}, G^t) := \max_{x_i^t \text{ for } i \in G^t} Eu(x^t; \beta^t) = \max_{x_i^t \text{ for } i \in G^t} \int_{\mathbb{R}^N} f^{(\beta^t)}(b) \sum_{i \in G^t} u_i(x_i^t; b_i) db \,,$$

subject to

$$\begin{aligned} \sum_{i \in G^t} p_i \cdot x_i^t &\leq y \,, \\ x_i^t &\geq 0 \quad \text{for all } i \in G^t \text{ , and} \\ x_i^t &= 0 \quad \text{for all } i \notin G^t \,. \end{aligned} \tag{3}$$

That is, Alice's myopic choice problem is akin to an optimal atemporal consumption decision with multiple goods and a linear or quasi-linear utility function (due to the constant marginal utility of the numeraire good). For instance, if one available good j (say, jackfruit) has, for all possible consumption quantities, a higher expected marginal sub-utility per dollar than the other available goods, then Alice chooses to consume only that good ($x_j^t = y/p_j$ and $x_i^t = 0$ for all $i \neq j$). If instead the expected marginal utilities per dollar of multiple goods are overlapping for the relevant regions, then the x_i^t values for each of these goods are given by equating the marginal (expected) sub-utilities per dollar of all purchased goods.

Essentially, in the myopic choice problem, Alice will never buy a banana if the maximum marginal sub-utility she expects to get from it (which, given concavity, occurs for the first marginal taste of banana, that is at $x_i = 0$) is not greater than the marginal utility she expects from a bundle of other goods excluding this one; and as in the standard choice problem, the marginal utility of money equals the expected marginal utility of each good that is consumed in positive quantity at its optimized quantity divided by its price.

If Alice is not myopic, she maximizes the present value of her stream of expected utilities, using a per-period discount factor δ. This encapsulates the standard assumption of additive separability of utility across time periods. In most models of intertemporal choice, time periods are linked through the ability to shift money back and forth in time. In this model, time periods are instead linked because a costly consumption investment can yield information that can be used later to optimize consumption.

Axiom 8. Discounted Expected Utility.
When choosing a bundle in time t, Alice maximizes the present value of her stream of expected utility over time.

□

We represent the time-t present value of Alice's expected utility stream, based on optimal intertemporal consumption choices at all times according to Axiom 8, by a value function $V^t(.)$. Her optimization problem can then be stated recursively as:

$$V^t(f^{(\beta^t)}, G^t) = \max_{x_i^t \text{ for } i \in G^t} Eu(x^t; \beta^t) + \delta \cdot E_t \left[V^{t+1}(f^{(\beta^{t+1})}, G^{t+1}) \mid f^{(\beta^t)} \right] , \quad (4)$$

subject to the optimization conditions (3), the parameter updating process specified in Lemma 1, and (for finite T) the terminal condition $V^T(f^{(\beta)}, G) = U(f^{(\beta)}, G)$. $E_t[X]$ denotes the expected value of random variable X based on the information available at time t, which is $f^{(\beta^t)}$. Recall that goods appear probabilistically, so in time t Alice must consider not just the uncertainty she has over her own tastes but also the likelihood that any particular basket of goods G will appear in each future period. At time t, Alice generally does not know her future parameter vector β^{t+1}—or, equivalently, the corresponding p.d.f. $f^{(\beta^{t+1})}$—but she knows that if at time t she samples an unlearned good, its parameters will update. She also does not know what basket G^{t+1} will be available to her, but she knows the likelihood of each possible basket.

Because Alice optimizes her discounted stream of utility, she is willing in each period to forego some current expected utility if in expectation it gives her an increase in discounted future utility that is at least as large as the expected utility foregone now. This increase will come from learning her tastes for a previously-unlearned good. This is only a sacrifice if the unlearned good appears unattractive in a myopic optimization problem. We call this act of sacrificing current expected utility for future expected utility by consuming a new good i *experimental consumption* of good i: choosing $x_i = m_i$ even though $x_i < m_i$ maximizes myopic utility because it is worth it to "experiment." When Alice experimentally consumes good i, she will never choose more than nibble size m_i because that minimizes the expected costs of learning.

Imagine that in time t Alice has not yet learned her taste for mangosteen (good i) and that she can currently afford a nibble of this good.[3] We define, for $i \in G^t$ with $f_i^{(\beta_i^t)} \equiv f_i^{(\beta_i^0)}$ and $m_i \cdot p_i \leq y$:

$$U_i(f^{(\beta^t)}, G^t) := Eu_i(m_i; \beta_i^0) + \max_{x_j^t \text{ for } j \in G^t \setminus \{i\}} \sum_{j \in G^t \setminus \{i\}} Eu_j(x_j^t; \beta_j^t) \ ,$$

subject to

$$\begin{aligned} \sum_{j \in G^t \setminus \{i\}} p_j \cdot x_j^t \quad &\leq y - p_i \cdot m_i \ , \\ x_j^t \quad &\geq 0 \quad \text{for all } j \in G^t \setminus \{i\} \ , \text{ and} \\ x_j^t \quad &= 0 \quad \text{for all } j \notin G^t \ . \end{aligned}$$

$U_i(.)$ is Alice's time-t expected utility from consuming a nibble of good i and allocating the rest of her money myopic-optimally among the remaining

[3] We will explore experimental consumption for one good at a time for ease of exposition; the same concepts would apply if, as is possible, Alice chooses to experimentally consume multiple goods in the same period.

goods: trying just enough mangosteen to learn about it and making a bundle that is otherwise myopically optimizing. The time-t loss of current-period utility from experimental consumption of mangosteen is therefore $U(.)-U_i(.)$. This is only a loss if mangosteen appears unattractive to Alice based on her priors; since Alice has clear incentive to consume it (and thus learn her taste) if it does not, we focus on the case in which it is a loss.

Alice's benefit (valued at time $t+1$) from experimentally consuming good i is:

$$\phi_i^{t+1}(f^{(\beta^t)}) := E_t\left[V^{t+1}(f^{(\beta')}, G^{t+1}) \mid f^{(\beta^t)}\right] - E_t\left[V^{t+1}(f^{(\beta'')}, G^{t+1}) \mid f^{(\beta^t)}\right], \tag{5}$$

where

$$\begin{aligned} \beta' &= (\beta_1^{t+1}, \ldots, \beta_{i-1}^{t+1}, \hat{\beta}_i, \beta_{i+1}^{t+1}, \ldots, \beta_N^{t+1}), \quad \text{and} \\ \beta'' &= (\beta_1^{t+1}, \ldots, \beta_{i-1}^{t+1}, \beta_i^t, \beta_{i+1}^{t+1}, \ldots, \beta_N^{t+1}). \end{aligned}$$

Alice only benefits from experimental consumption of i if she does not yet know her preferences for it (that is, if $f_i^{(\beta_i^t)} \not\equiv \hat{f}_i$), and thus she still holds her prior, that is, $f_i^{(\beta_i^t)} \equiv f_i^{(\beta_i^0)}$. If she does know her preferences for good i, $\phi_i^{t+1} = 0$, by definition. In general, the benefit from experimental consumption will always be non-negative since at worst, Alice can choose not to consume the good in future periods, as the following lemma shows.

Lemma 2. Characteristics of the Benefits of Experimental Consumption.
Ceteris paribus, for all $i \in \{2, \ldots, N\}$ and $t \in \{0, \ldots, T\}$:

(a) $\phi_i^{t+1}(.) \geq 0$.

(b) If $T < \infty$, then for all $\beta \in \mathbb{B}$: $\phi_i^{t+1}(f^{(\beta)})$ is non-increasing as t increases.

(c) If $T = \infty$, then for all $\beta \in \mathbb{B}$: $\phi_i^{t+1}(f^{(\beta)})$ is constant as t increases.

□

Note that this result holds given that all beliefs remain the same. As Alice consumes new goods and learns more, it is possible that ϕ_i could increase, depending on the realizations that occur.

We can now identify the conditions for experimental consumption:

Lemma 3. Conditions for Experimental Consumption.
At time t*, with current preference beliefs* $f^{(\beta^t)}$*, the agent chooses experimental consumption of good* i *if all of the following conditions are met:*

(i) $i \in G^t$.

(ii) $p_i \cdot m_i \leq y$.

(iii) $f_i^{(\beta_i^t)} \not\equiv \hat{f}_i$.

(iv) $U(f^{(\beta^t)}, G^t) - U_i(f^{(\beta^t)}, G^t) < \delta \cdot \phi_i^{t+1}(f^{(\beta^t)})$.

□

The first three conditions state that for Alice to experimentally consume a myopically-unattractive good i, i must be available, she must be able to afford a nibble of it, and she must not have discovered her preferences for it yet. Given these, she will try it if the discounted expected benefit from learning her parameter for the good exceeds the cost of learning: that is, the myopic loss from forgoing other goods that appear more attractive right now is less than the expected discounted stream of benefits from better optimization.

Given experimental consumption of some good i, the quantities chosen of other goods like j will generally not be myopically optimizing: since Alice is spending some money to experimentally taste mangosteen, she will spend less overall on apples and bananas.

We can also observe that if Alice does not choose to consume good i when she encounters that good alone (accompanied by no other good except the numeraire), she will never learn her taste for it unless her preferences for other goods change. The caveat about other tastes not having changed is needed because if Alice's believed preferences for other goods change, good i may suddenly seem more appealing in comparison (may promise a larger possible increase in future utility) and experimental consumption of this good may become worthwhile.

Lemma 4. Minimal Consumption Set.
If Alice has not learned her preferences for good i *prior to time* t*, if* $G^t = G_i = \{1, i\}$*, and if Alice chooses not to consume at least a nibble of good* i *at time* t*, then she will not discover her preferences for good* i *as long as her believed preferences for all other goods remain the same.*

□

3 Theory Results

Now that we have constructed the model components, we can proceed to study the model's implications for preference discovery. Table 1 itemizes the coming propositions.

The first proposition trivially says that if Alice can't afford to consume a good, she'll never learn it. Proposition 2 lays out two sides of a coin: if a good has a prior that makes the good appear better than the always-available outside option, then Alice will eventually try it if she lives long enough, but if it does not, she may or may not try it. The third proposition lists characteristics of the agent and the good that make eventual learning less likely because they influence either the opportunity cost of sampling the good or the present value of the expected future gains from learning it. The preference stability of Proposition 4 is like that noted in studies like Andersen et al (2008) and Dasgupta et al (2017), and relates as well to the decline in choice reversals (Proposition 5b) and welfare loss (6b) over time. However, as noted in Proposition 6c, people can continue to lose welfare forever because there are some goods they will simply never try. We argue less formally that this bias is asymmetric in that where errors persist, they are, on average, negative.

Other models could also yield some of these results. However, the most important point that is particular to our model is that welfare loss can persist, so Proposition 6c is in that way our most important result. Also, Propositions 4 (that preference beliefs–and thus choices–appear to become stable eventually even so) and 5a and 5b (that choice reversals occur but decline over time) are important connections to existing literature. The combined result, that welfare loss can persist even in the absence of choice reversals, is of particular relevance. In addition, the elements of Proposition 3 are important in that they provide a set of testable hypotheses about the learning process that are unique to the need for preferences to be learned. Many of the results in this paper are tested in Delaney et al (2020), with results strongly supporting the theory.

3.1 Preference Learning

Let us first explore which goods Alice will and will not learn her tastes for by any given time and as time approaches infinity. We define $L^t \subseteq \{1, \ldots, N\}$ as the set of all goods for which Alice has learned her preferences prior to time

Table 1: Theory Results

Theory Result	Description
Proposition 1	Unaffordable goods are never learned
Proposition 2a	Goods with priors better than outside option are eventually learned
Proposition 2b	A good will never be tried if the opportunity cost of a meaningful taste outweighs the expected gain from future optimized consumption
Proposition 3	A good is less likely to be learned if:
Proposition 3a	The agent has a smaller income (for normal goods).
Proposition 3b	The good is more expensive.
Proposition 3c	The minimum "nibble" size to learn the good is larger.
Proposition 3d	The good appears less frequently.
Proposition 3e	The agent is more impatient.
Proposition 3f	The agent has a shorter lifetime.
Proposition 3g	The good's prior is low.
Proposition 3h	Given a low prior, the agent's belief is less diffuse.
Proposition 3i	Given a high prior and a risk averse agent, the agent's belief is more diffuse.
Proposition 3j	Other goods seem more attractive.
Proposition 4	Preference parameter beliefs eventually become stable.
Informal result	The average parameter belief error becomes negative (pessimistic) over time.
Proposition 5a	Choice reversals occur.
Proposition 5b	The rate of choice reversals declines to zero over time.
Proposition 6a	Unlearned preferences may cause welfare loss.
Proposition 6b	Welfare loss weakly declines over time.
Proposition 6c	Welfare loss need not approach zero as time passes.

t. That is, $i \in L^t$ if and only if $f_i^{(\beta_i^t)} \equiv \Delta(\hat{\beta}_i)$. Because of our assumptions, $L^0 = \{1\}$ (only the numeraire good has been learned) and $L^{t+1} \supseteq L^t$ for all t. We denote the probability that Alice has learned her preferences for good i by time t as $r_i^t := \mathbb{P}(i \in L^t)$.

Let us define some learning benchmarks. *Full discovery* is the state Alice achieves if she learns her preferences for all goods, so that she has achieved full discovery at time t if $i \in L^t \; \forall i \in \{1, \ldots, N\}$. *Full relevant discovery* at time t means that by t she has learned her tastes for all goods that are truly better (at least for the first bite) than the numeraire good, so $i \in L^t$ for all $i \in \{1, \ldots, N\}$ for which $\left.\frac{d\,u_i\left(x_i; \hat{\beta}_i\right)}{dx_i}\right|_{x_i=0} > z \cdot \frac{p_i}{p_1}$. If Alice achieves full relevant discovery then she may still have some unlearned preferences, but they will not affect her wellbeing since all will be goods she wouldn't optimally consume.

First, it is obvious that Alice will never, even as $t \to \infty$, learn her preferences for any good if a nibble of it is too expensive for her to afford. For example, Alice may never consume the pricey Densuke watermelon and thus may forever cherish a misapprehension of her value for it.

Proposition 1. Unaffordable Goods.
For $i \neq 1$, $i \notin L^T$ if $p_i \cdot m_i > y$.

□

Next, given enough time, Alice will learn the true values of two classes of goods. One class comprises goods for which the current-period expected marginal utility per dollar based on the prior achieves a value above the marginal utility per dollar of the numeraire good: mangoes may look relatively tasty, so they will certainly be tried over time. Other goods, like perhaps (for Alice) the mangosteen, are more prospective: goods with lower expected marginal utility will only be discovered through experimental consumption, and that will only occur if the discounted future expected utility gains from learning her true preferences outweigh the expected current-period utility loss from consuming more of this good than is myopically optimal.

Proposition 2. Goods That Will and Will Not Be Learned.
Consider good $i \in \{2, \ldots, N\}$ such that $p_i \cdot m_i \leq y$.

(a) Good i will eventually be learned if

$$\left.\frac{d\,Eu_i\left(x_i; \beta_i^0\right)}{dx_i}\right|_{x_i=m_i} > z \cdot \frac{p_i}{p_1}\,.$$

That is, for such goods, $r_i^t \to 1$ *as* $t \to \infty$*.*

(b) Good i *will never be learned if both of these conditions are met:*

(i) $\left.\frac{d\,Eu_i\left(x_i;\beta_i^0\right)}{dx_i}\right|_{x_i=m_i} < z \cdot \frac{p_i}{p_1}$ *, and*

(ii) $\max\limits_{G\in\mathbb{G}\,,\,\beta\in\mathbb{B}_i'} \delta \cdot \phi_i^1(f^{(\beta)}) - U(f^{(\beta)}, G) + U_i(f^{(\beta)}, G) < 0,$

where

$$
\begin{aligned}
\mathbb{G} &= \{G \subseteq \{1, \ldots, N\} : 1 \in G\}\,,\ \textit{and} \\
\mathbb{B}_i' &= \{\beta \in \mathbb{B} : \beta_i = \beta_i^0 \textit{ and } \beta_j \in \{\beta_j^0, \hat{\beta}_j\} \textit{ for all } j \neq i\}\,.
\end{aligned}
$$

□

Proposition 2 part (b)(i) states that it is not myopically optimal to consume at least a nibble of good i, and part (b)(ii) states that the myopic utility loss from experimentally consuming good i is larger than the discounted stream of gains from improved information for any allowable set of believed preferences and any realized availability of other goods when this stream is evaluated at the earliest possible time (when it is largest).

A good can meet condition (b)(i) but not (b)(ii). Such goods might or might not be learned, depending on the realized availability of and priors for other goods. For example, Alice might have a relatively low prior for rhubarb and a moderately low (but better than the numeraire) prior for kumquats. If Alice's true taste holds kumquats in even higher regard, then if Alice encounters rhubarb alone before learning her taste for kumquats, her opportunity cost for learning is relatively low and she may taste a nibble of rhubarb. But if she learns her taste for kumquats before encountering rhubarb alone, the potential net benefits of learning will change, and could render experimental consumption of rhubarb unattractive.

To return to our learning benchmarks, Proposition 2 implies that Alice generally need not achieve full discovery of her preferences. Given that we place no restrictions on the priors or true values of the goods, this implies that she need not achieve full relevant discovery, either, since some untried goods could have true values that would make them worth consuming.

We now consider what characteristics of the good itself, the other goods, or the agent foster incomplete learning. The determinants come down to the good's availability, factors that influence the opportunity cost of trying the good when it is available ($U(.) - U_i(.)$) and factors that determine the expected benefit of learning the good's value (ϕ_i^{t+1}).

Proposition 3. Factors That Influence Discovery.
Ceteris paribus, Alice is less likely to learn her preferences for a good $i \in \{2, \ldots, N\}$ by time T under any of the following conditions:

(a) She has less income (smaller y), given that i is a normal good.

(b) The good is more expensive (larger p_i).

(c) A larger nibble is required for a meaningful learning experience (larger m_i).

(d) The good appears less frequently (smaller q_i).

(e) She discounts future consumption more heavily (smaller δ).

(f) She has a shorter lifespan (smaller T).

(g) She has a bad prior perception of the good (that is, her prior p.d.f. $f_i^{(\beta_i^0)}$ is shifted further to the left).

(h) She has more confidence in her belief (i.e., less dispersion in $f_i^{(\beta_i^0)}$), given that she has poor priors for the good that make consumption of at least a nibble an unattractive choice relative to the numeraire.

(i) She has less confidence in her belief (i.e., more dispersion in $f_i^{(\beta_i^0)}$), given that she is risk averse and has a positive average prior, in the sense that the per-dollar marginal utility of good i, parameterized with the mean of $f_i^{(\beta_i^0)}$, exceeds the per-dollar marginal utility of the numeraire—that is, if $\left.\frac{d\,u_i\left(x_i; E[\beta_i^0]\right)}{dx_i}\right|_{x_i=m_i} > z \cdot \frac{p_i}{p_1}$.

(j) Other goods appear more attractive (larger $\hat{\beta}_j$ or $f^{(\beta_j^0)}$ shifted further to the right for some $j \neq i$).

$\square$

Some of these cases coincide with cases pointed out in Thaler and Sunstein (2008) as being ripe for behavioral errors. Specifically, Thaler and Sunstein (2008) note that people will tend to make poor choices "in contexts in which they are inexperienced and poorly informed, and in which feedback is slow or infrequent" (p. 7). The general point of our model is that Alice will make errors when she is inexperienced in the sense of having unlearned preferences if her priors are incorrect. But as time progresses and Alice has the opportunity to learn, she will continue to tend to be inexperienced with rare goods (Proposition 3(d)), or goods she'd buy rarely because they are expensive (Proposition 3(b)), or goods that require more consumption to learn (Proposition 3(c)). In our model, being poorly informed is eventually self-correcting unless Alice is poorly informed in the negative direction (Proposition 3(g)).

While we do not show this formally, it is also intuitive that people who are more risk averse are more likely to have incomplete learning. This is because future rewards from experimental consumption are uncertain, so risk aversion reduces their value.

The personal characteristics associated with never learning her preferences are also associated with populations that are already disadvantaged; this is a concern because, as we show later, undiscovered preferences cause welfare loss, thus burdening these people further.

From our earlier discussion we can also conclude that the parameters of Alice's preference beliefs will stabilize, albeit not necessarily at her true preferences.

Proposition 4. Eventual Preference Belief Stability.
If $T = \infty$, *then* $\mathbb{P}\left(f_i^{(\beta_i^s)} \equiv f_i^{(\beta_i^t)}\ \forall\, s \geq t\right) \to 1$ *as* $t \to \infty$.

□

Recall that studies such as Andersen et al (2008) and Dasgupta et al (2017) find some support for stability of preferences over time; our theory shows those results are not evidence against preference discovery.

We also argue that as Alice learns her preferences over time, she becomes increasingly pessimistic. We do not offer a formal proof of this, but the intuition is as follows.

Alice's priors for some goods make them look better than they are: these are goods i for which the true $\hat{\beta}_i$ lies to the left within the distribution $f_i^{(\beta_i^0)}$. This is optimistic error. There are also goods for which Alice's prior makes them look worse than they are: goods j for which the true $\hat{\beta}_j$ lies to the right within the distribution $f_j^{(\beta_j^0)}$; this is pessimistic error. For goods whose true value and prior probability distribution of parameters are both very low, so that marginal expected utility per dollar is well below the numeraire, neither kind of error will be corrected: Alice will never learn whether rotten mango is less disgusting than rotten guava or vice versa. On the other hand, goods with high priors will see errors of both signs corrected: if ambrosia and nectar both appear delicious but she thinks ambrosia is worse than it is and nectar is better than it is, she'll taste both goods eventually and consequently sort out her true values.

However, for goods nearer to the threshold at which consumption becomes myopically optimal, the sign of the error matters. For a given true parameter value, an optimistic bias will make a good more likely to be tried and learned than will a pessimistic bias, by the logic in Proposition 3(g). As a result, unless error is inversely correlated with true values, goods with a pessimistic error will be less likely to ever be tried, and thus more of these goods will persist unlearned forever. As a result, perception errors for some goods will drop to zero through preference learning, but since this will happen more for optimistic than pessimistic errors, the average tendency of the errors that remain will be to see goods as less attractive than they actually are.

The main story of our results so far is that Alice will sample and learn her taste for many goods, but perhaps never for other goods, including some that are affordable and that she would actually like. In Section 3.2, we study how observers may see evidence of the learning process in action. In Section 3.3, we study how Alice loses welfare because of undiscovered preferences.

3.2 Choice Reversals

Consider now the phenomenon of choice reversals, as discussed in work such as Cox and Grether (1996).[4] In a choice reversal, an agent is observed to make one choice (say, bundle A over bundle B) at one time and then a contradictory choice (B over A) at another time, when all external conditions appear to be identical across the two choice scenarios. Our model allows for these reversals at any finite time, but not as $t \to \infty$.

Proposition 5. Choice Reversals.

[4]Most studies refer to the phenomenon as "preference reversals." As we are maintaining an assumption of stable underlying preferences, we say "choice reversals."

(a) *If at time t there exists a good $i \in \{2, \ldots, N\}$, with $i \notin L^t$ and $p_i m_i < y$, then for any $\hat{\beta} \in \mathbb{B}$ there exists a prior $\beta^0 \in \mathbb{B}$ and a set G with $\mathbb{P}(G_t = G) > 0$, such that $\mathbb{P}(x^{t+1} \neq x^t | G^{t+1} = G^t = G) = 1$. This implies a positive probability of a choice reversal, that is for prior $\beta^0 \in \mathbb{B}$: $\mathbb{P}(x^{t+1} \neq x^t | G^{t+1} = G^t) > 0$.*

(b) *The probability of such a choice reversal approaches 0 as $t \to \infty$.*

□

This result accords with studies that show that reversals decline with repetition, such as Cox and Grether (1996). Note that the eventual absence of choice reversals need not, in this case, imply that Alice is choosing her optimal bundle in succeeding periods. We explore this in the next subsection.

3.3 Welfare Implications

Let us define $U(\hat{f}, G)$ as the maximum true myopic utility Alice can attain with the goods available in set G. Consuming any other bundle x' will give her (weakly) less immediate utility. Let us therefore define Alice's time-t expected welfare loss Δu^t as the expected loss of utility she experiences from not choosing according to her true preferences at time t:

$$\Delta u^t = E\left[U(\hat{f}, G^t) - U(f^{(\beta^t)}, G^t)\right] .$$

Here, the expectation is taken based on the information available to Alice at time 0, that is, her priors $f^{(\beta^0)}$. The uncertainty here stems from the randomness in G^t as well as the randomness in the sets of goods that are available to her over the periods up to time t, which influences her beliefs $f^{(\beta^t)}$.

If Alice behaves according to our model, welfare loss will occur for two reasons. Some accidental loss will occur as Alice chooses according to the preferences she believes she has if those beliefs are incorrect. In addition, Alice may intentionally lower her current expected utility, particularly early in her life, by engaging in experimental consumption to sacrifice current utility in hopes of better optimization in the future. Both of these effects tend to diminish over time as Alice discovers her true preferences for at least some of the goods, although in the case of the former (which is the problematic one of the two), it need not decline to zero. We can thus draw the following conclusions about the agent's welfare loss:

Proposition 6. ***Persistent Welfare Loss.***

Suppose there exists a good $i \in \{2, \ldots, N\}$ *for which* $p_i m_i \leq y$ *and* $\left.\frac{d\,u_i(x_i;\hat{\beta}_i)}{dx_i}\right|_{x_i=0} > z \cdot \frac{p_i}{p_1}$. *Then:*

(a) There exists a prior $\beta^0 \in \mathbb{B}$ *such that for all* $t \geq 0$, $\Delta u^t > 0$.

(b) Under the specification of part (a), Δu^t *is (weakly) decreasing in* t.

(c) There exists a prior $\beta^0 \in \mathbb{B}$ *such that* $\Delta u^t \not\to 0$ *as* $t \to \infty$.

□

It might be argued that this welfare loss derives in part from restrictions we have placed on the learning process, such as the lack of spillover learning, or our specification of belief error. Believed and true preference parameter values may be positively correlated; this would be the case if Alice's beliefs are formed based on information gleaned from consumption of other goods, others' experiences, introspection, or other sensible processes. Such informed guesswork may mitigate but not eliminate the failure to try some goods with true values that would render them part of myopically optimal bundles, and thus would not eliminate the resulting welfare loss, as long as the correlation between beliefs and true values is not perfect.

4 Discussion

Our model shows that people may not fully learn their preferences even under the most congenial circumstances, and that lack of knowledge of one's own preference entails both transitory and persistent welfare loss. Many results from this model have been experimentally validated: specifically, Delaney et al (2020) design an online experiment around the predictions of this model and validate the decline in choice reversals over time, as well as incompleteness of learning in particular for goods that are rare or that have low prior beliefs. In addition, the results corroborate the predictions that agents with lower incomes or shorter lifetimes are more likely to have incomplete learning, that noise has varying effects contingent on the value of the prior, and that pessimistic bias develops over time. Finally, the experiment documents the presence of persistent welfare losses stemming from undiscovered preferences.

With more realistic assumptions, preference discovery would be even less likely, thus making the problems we point out even more egregious. Some such complications include: if multiple consumption experiences are required for the agent to learn her true preferences for a good; if the agent can only observe the aggregate utility from the consumption bundle rather than from each good individually; or if the agent forgets her preferences for a good after learning them. In addition, if the benefits from consuming goods are stochastic rather than deterministic, this could make preferences harder to learn as well, perhaps by adding another parameter to learn or by requiring more experience to learn the preference. On the other hand, if learning is not separable across goods, this might make learning easier by generating learning spillovers across consumption experiences, but also could create more parameters (such as coefficients that govern relationships between goods), thus increasing the dimensionality of the learning problem and making it harder, so that the net effect of non-separability on the learning process is ambiguous.

While goods in our study can be bought in continuous quantities, if choice items are discrete and have large consequences (like houses, jobs, or life partners), learning problems are likely to be worse; the analogy in our model is to goods that have a larger "nibble" (minimum consumption size). Another element that would render learning particularly challenging is an agent's inability to directly assess a good's value even when she "consumes" it, as might be the case for credence goods, donations to charity, and environmental valuation. Indeed, as previously noted, the situations we suggest are most likely to give rise to learning failure correlate to the contexts that Thaler and Sunstein (2008) argue cause people to make bad decisions: cases where the agent is inexperienced and poorly informed, and where she will receive little feedback.

Given the potential welfare loss we have identified, truly sophisticated agents (beyond the sort of self-awareness of diffuse priors Alice has) may, for important decisions, turn to other processes or criteria instead of discounted expected utility maximization based on beliefs. For example, people may reduce a complex housing decision to a simpler problem about their beliefs about the value of an asset appreciating over time. Future research could explore whether people do this and whether it is welfare-enhancing, and could study whether specific heuristics can help the preference learning process avoid the gaps we have identified.

5 Conclusion

Most work in economics implicitly or explicitly assumes that people know what they like. We argue that if self-knowledge is not endowed at birth but rather achieved through experience, as suggested by the discovered preference hypothesis (Plott, 1996), then even the most rational and sophisticated people may fail to learn all of their preferences. At the heart of this failure is the fact that learning has an opportunity cost, and thus complete learning is irrational. In this paper, we develop a theory to explore factors that enable or impede learning for certain people or certain consumption items. We start from a premise that preferences must be learned through experience, and we focus on the extensive margin of learning, that is, which goods are learned, rather than the intensive margin of the updating process. We show that in some cases, tastes for some items will never be learned, and welfare will therefore be lost persistently.

A preference discovery process can explain choice instabilities observed in observational and laboratory studies of behavior, especially in cases of items that are unlikely to have been "consumed" often. Moreover, stable choice behavior does not indicate that agents are choosing according to their true underlying preferences: they may simply have stopped experimenting with new goods because they no longer find it worthwhile to do so.

While Delaney et al (2020) confirms many predictions of this model, the preference discovery process must be studied in more detail and in more settings to shed light on how factors internal and external to the agent affect learning and thus welfare loss. It is possible that an agent can perform mental simulations to achieve some learning without consumption; if so, that would alleviate some of the issues we highlight. On the other hand, we made many assumptions to make learning very easy, and those are unlikely to hold, which would exacerbate learning problems. In addition, research about the scale of welfare loss would further illuminate the value of preference discovery.

The implications of having to learn our own preferences through experience are broad. On the one hand, this model can provide new insights on how to get people to try new things, for example, in the case of a company marketing a product or a government or non-profit promulgating a green or otherwise socially beneficial technology. On the other hand, it shows that cross-sectional choice data from any experimental or observational setting may be contaminated by unstable parameters. Worse, choices that appear stable and rational may not reflect what is actually best for the individual making the decision. A tenet undergirding most economics-based policy advice is that people know what's best for them; but if we have undiscovered preferences, that might not be true.

Acknowledgements

We are grateful for comments from Brit Grosskopf, Muriel Niederle, J. Aislinn Bohren, Angela de Oliveira, Jessica Hoel, Jian Li, Yongsheng Xu, Annemie Maertens, and participants at various seminars and conferences.

Statements and Declarations

Competing Interests: The authors have no competing interests to declare.
Funding: We gratefully acknowledge funding from the Williams College Hellman Fellows Grant.

6 References

References

Aghion P, Bolton P, Harris C, Jullien B (1991) Optimal learning by experimentation. The Review of Economic Studies 58(4):621–654, URL http://www.jstor.org/stable/2297825

Andersen S, Harrison GW, Lau MI, Rutstrom EE (2008) Lost in state space: Are preferences stable? International Economic Review 49(3):1091–1112

Ariely D, Loewenstein G, Prelec D (2003) “coherent arbitrariness”: Stable demand curves without stable preferences. The Quarterly Journal of Economics 118(1):73–105

Becker GS (1996) Accounting for tastes. Harvard University Press

Braga J, Starmer C (2005) Preference anomalies, preference elicitation and the discovered preference hypothesis. Environmental and Resource Economics 32(1):55–89

Bruine de Bruin W, Parker A, Fischhoff B (2007) Individual differences in adult decision-making competence. Journal of Personality and Social Psychology

Charness G, Gneezy U, Imas A (2013) Experimental methods: Eliciting risk preferences. Journal of Economic Behavior & Organization

Coursey DL, Hovis JL, Schulze WD (1987) The disparity between willingness to accept and willingness to pay measures of value. The Quarterly Journal of Economics 102(3):679–690

Cox JC, Grether DM (1996) The preference reversal phenomenon: Response mode, markets and incentives. Economic Theory 7(3):381–405

Dasgupta U, Gangadharan L, Maitra P, Mani S (2017) Searching for preference stability in a state dependent world. Journal of Economic Psychology 62(Supplement C):17 – 32, DOI https://doi.org/10.1016/j.joep.2017.05.001, URL http://www.sciencedirect.com/science/article/pii/S0167487016305840

Delaney J, Jacobson SA, Moenig T (2020) Preference discovery. Experimental Economics 23(3):694–715

Easley D, Kiefer NM (1988) Controlling a stochastic process with unknown parameters. Econometrica 56(5):1045–1064, URL http://www.jstor.org/stable/1911358

Ferreira JV, Gravel N (2024) Revealing preference discovery: a chronological choice framework. Theory and Decision

Kahneman D, Snell J (1990) Predicting utility. In: Hogarth RM (ed) Insights in decision making: A tribute to Hillel J. Einhorn, Chicago and London: University of Chicago Press, pp 295–310

Kahneman D, Wakker PP, Sarin R (1997) Back to Bentham? explorations of experienced utility. The Quarterly Journal of Economics 112(2):375–405

van de Kuilen G, Wakker PP (2006) Learning in the Allais paradox. Journal of Risk and Uncertainty 33(3):155–164

Lerner JS, Li Y, Veldesolo P, Kassam KS (2015) Emotion and decision making. Annual Review of Psychology

Lichtenstein S, Slovic P (2006) The construction of preference. Cambridge University Press

List JA (2003) Does market experience eliminate market anomalies? The Quarterly Journal of Economics 118(1):41

Loewenstein G, Adler D (1995) A bias in the prediction of tastes. The Economic Journal 105(431):pp. 929–937, URL http://www.jstor.org/stable/2235159

Noussair C, Robin S, Ruffieux B (2004) Revealing consumers' willingness-to-pay: A comparison of the BDM mechanism and the Vickrey auction. Journal of Economic Psychology 25(6):725–741

Plott CR (1996) Rational individual behaviour in markets and social choice processes: The discovered preference hypothesis. In: Arrow KJ, et al (eds) The rational foundations of economic behaviour: Proceedings of the IEA Conference held in Turin, Italy, IEA Conference Volume, no. 114. New

York: St. Martin's Press; London: Macmillan Press in association with the International Economic Association, pp 225–250

Rigby D, Burton M, Pluske J (2015) Preference stability and choice consistency in discrete choice experiments. Environmental and Resource Economics

Scitovsky T (1976) The joyless economy: An inquiry into human satisfaction and consumer dissatisfaction. Oxford University Press

Shogren JF, Shin SY, Hayes DJ, Kliebenstein JB (1994) Resolving differences in willingness to pay and willingness to accept. American Economic Review 84(1):255–270

Shogren JF, Cho S, Koo C, List J, Park C, Polo P, Wilhelmi R (2001) Auction mechanisms and the measurement of WTP and WTA. Resource and Energy Economics 23(2):97–109

Thaler RH, Sunstein CR (2008) Nudge: Improving Decisions about Health, Wealth, and Happiness. New Haven and London:

Weber RA (2003) 'learning' with no feedback in a competitive guessing game. Games and Economic Behavior 44(1):134 – 144, DOI http://dx.doi.org/10.1016/S0899-8256(03)00002-2, URL http://www.sciencedirect.com/science/article/pii/S0899825603000022

Wilson TD, Gilbert DT (2005) Affective forecasting: Knowing what to want. Current Directions in Psychological Science 14(3):131–134, DOI 10.1111/j.0963-7214.2005.00355.x, URL http://cdp.sagepub.com/content/14/3/131.abstract

A Appendix: Proofs

Proof of Lemma 1

(a) According to Axiom 5, if $x_i^t < m_i$, Alice has no reason to update her preferences for good i at that time. Axiom 6 ensures that there is no possible experience with any other goods that would lead Alice to update $f_i^{(\beta_i^t)}$. As a result, $f_i^{(\beta_i^{t+1})} \equiv f_i^{(\beta_i^t)}$.

(b) Axiom 5 implies that Alice updates her preferences to the true $\hat{\beta}_i$ upon her meaningful consumption experience at time t, so that $f_i^{(\beta_i^{t+1})} \equiv \hat{f}_i$. Then, by Axiom 7, she will maintain these true preferences into perpetuity.

(c) This follows directly from parts (a) and (b) of this lemma: preference belief for good i starts at $f_i^{(\beta_i^0)}$ and can only change to $\hat{f}_i$, if at all.

□

Proof of Lemma 2

We first observe that learning $\hat{\beta}_i$ provides potentially increased expected utility to the agent for all future periods. For $k \geq 1$, we denote the difference in expected utility from period $(t+k)$ consumption based on whether or not the agent learned $\hat{\beta}_i$ in period t by $\alpha_i^{t,t+k}(f^{(\beta^t)})$. Recalling that $f_i^{(\beta_i^t)}(.)$ can only either be $f_i^{(\beta_i^0)}$ or $\hat{f}_i$, we can write:

$$\begin{aligned} \alpha_i^{t,t+k}(f^{(\beta^t)}) \quad := \quad & E_t\left[\max_{x_j \text{ for } j\in G^{t+k}} Eu(x;\beta^{t+k}) \,\middle|\, f_i^{(\beta_i^{t+1})} \equiv \hat{f}_i\right] \\ & -E_t\left[\max_{x_j \text{ for } j\in G^{t+k}} Eu(x;\beta^{t+k}) \,\middle|\, f_i^{(\beta_i^{t+1})} \equiv f_i^{(\beta_i^t)}\right], \end{aligned} \tag{6}$$

whereby all optimization problems are subject to the usual constraints (see Equation (3) and Lemma 1). Of course, at time t, Alice does not know her exact value of $\alpha_i^{t,t+k}$ since she does not know $\hat{\beta}_i$. Since—conditional on her current beliefs $f^{(\beta^t)}$—both the random availability of goods and the learning process from time t to time $t+1$ are time-independent, the right-hand side of Equation (6) is independent of t, and the only time value that matters is

k, the number of periods since the learning has occurred. We can therefore use the shortened notation α_i^k in place of $\alpha_i^{t,t+k}$.

Alice cannot do better for herself than to optimize based on her true parameters. Therefore, if she optimizes based on any other parameters, her utility must be less than or equal to the utility she gets when optimizing based on her true parameters. Therefore, $\alpha_i^k \geq 0$.

Moreover, by definition of ϕ_i^t, and for all $\beta \in \mathbb{B}$, with $f := f^{(\beta)}$:

$$\phi_i^t(f) = \begin{cases} \alpha_i^1(f) + \delta\alpha_i^2(f) + \delta^2\alpha_i^3(f) + \ldots + \delta^{T-t-1}\alpha_i^{T-t}(f) & \text{, if } T < \infty \\ \alpha_i^1(f) + \delta\alpha_i^2(f) + \delta^2\alpha_i^3(f) + \ldots & \text{, if } T = \infty \end{cases} .$$

We can therefore conclude that:

(a) $\phi_i^{t+1} \geq 0$ because it is the sum of non-negative numbers.

(b) For $T < \infty$,

$$\phi_i^{t+1}(f) - \phi_i^t(f) = -\delta^{T-t-1}\alpha_i^{T-t}(f) \leq 0 .$$

(c) Similarly, for $T = \infty$,

$$\phi_i^{t+1}(f) - \phi_i^t(f) = 0 .$$

□

Proof of Lemma 4

By the conditions of the lemma, we can let $x_i^* < m_i$ denote Alice's optimal time-t consumption choice of good i when the available set of goods is $G_i = \{1, i\}$. We will show that for any $s \geq t$ and any set $G^s \supseteq G_i$, if $f^{(\beta^s)} \equiv f^{(\beta^t)}$, then Alice's optimal time-s consumption bundle includes $x_i^s \leq x_i^*$ units of good i. This implies the statement of the lemma because, if good i is not sufficiently attractive to warrant experimental consumption in time t, and since the value of experimental consumption is non-increasing in time, good i will not become sufficiently attractive to warrant experimental consumption while she maintains her current preference beliefs.

We divide this proof into two parts.

(i) We first show that Alice's dynamically optimal time-s consumption choice of good i, x_i^s, is no greater than x_i^* for $s > t$ if $G^s = G_i = \{1, i\}$.

The solution to Alice's *myopic* optimization problem is independent of time, as it solely depends on the set of available goods as well as the current preference parameters for these goods. Per our assumption, both are identical at times s and t. Therefore, the solution to the myopic choice problem is identical at both times.

Even if it is not myopically optimal, Alice might choose to *experimentally* consume m_i units. For this to happen, according to Lemma 3, it must be true that for $\tau \in \{t, s\}$:

$$U(f^{(\beta^\tau)}, G^\tau) - U_i(f^{(\beta^\tau)}, G^\tau) \leq \delta \cdot \phi_i^{\tau+1}(f^{(\beta^\tau)}) \,.$$

The left-hand side of this inequality is identical for $\tau = t$ and $\tau = s$, while the right-hand side is non-increasing over time (Lemma 2), since $f^{(\beta^s)} \equiv f^{(\beta^t)}$ by our assumption. As a result, if the inequality is not satisfied at time t, it will not be satisfied at time $s > t$.

Therefore, under the given assumptions, the consumption choice of good i at time s cannot exceed the consumption choice of good i at time t given the same preference beliefs and the same minimal choice set.

(ii) Second, if x_i' denotes Alice's optimal choice at time $s \geq t$ if $G^s = G_i = \{1, i\}$, we show that $x_i^s \leq x_i'$ if $G^s \supsetneq G_i = \{1, i\}$. In this case, there exists at least one good $j \in G^s \setminus G_i$. If Alice chooses to consume a positive quantity of good j, she will do so at the possible expense of good i (as well as the numeraire). If strict convexity holds, then this requires Alice to consume no more than x_i' of i. If Alice does not choose a positive quantity of any other good j, then she will choose the same quantity x_i' of good i. Thus, $x_i^s \leq x_i' \leq x_i^*$ by (i).

If convexity is weak but not strict, so that there is constant marginal utility, the results mostly still hold. If good i does not provide the same marginal utility as the numeraire, then both parts will still hold: if the later basket is G_i, then the myopic solution will be identical at both times and experimental consumption will not happen later if it does not happen earlier; and if the later basket is not G_i, the addition of some other good j will not increase and may decrease choice of i. If good i does provide the same marginal utility as the numeraire, then the choice of x_i^t is not unique, and thus we are not guaranteed that a larger x_i^s will be chosen in either case, but this is a pathological case.

$\square$

Proof of Proposition 2

(a) Let $T = \infty$ and define $G_i = \{1, i\}$. If at some time t, the set of available goods is $G^t = G_i$, and if

$$\left.\frac{d\, Eu_i(x_i; \beta_i^0)}{dx_i}\right|_{x_i = m_i} > z \cdot \frac{p_i}{p_1} \,,$$

then Alice will choose to consume at least m_i units of good i, based on the solution to her myopic optimization problem. This is because this minimal consumption set contains only good i and the numeraire, and this inequality says that the marginal utility of good i evaluated at nibble size m_i is at least as high per dollar as the numeraire. Since marginal utility is (at least weakly) diminishing, the marginal utility for consumption up to m_i is at least as large as this, and thus at least m_i units are worth purchasing.

The numeraire does not need to be learned and Alice is already choosing based on her myopic motives to consume enough of i to learn her taste for it, so Alice's myopic choice to consume at least a nibble of i is the same as her dynamically optimal choice. As a result, by Lemma 1, Alice will learn her true preferences for good i at time t.

Lastly, note that the number of goods is finite, and that the probability $\mathbb{P}(G^\tau = G_i)$ that G_i is the available basket in any given round τ lies strictly between 0 and 1 and is time-invariant. Therefore, for any $t > 0$ and any $i \in \{2, \ldots, N\}$:

$$\begin{aligned}
\mathbb{P}(i \in L^t) &= \mathbb{P}(\exists\, s < t \text{ s.t. } G^s = G_i) = 1 - \mathbb{P}(G^s \neq G_i\, \forall\, s < t) \\
&= 1 - (1 - \underbrace{\mathbb{P}(G^\tau = G_i)}_{\in(0,1)})^t \to 1 \text{ as } t \to \infty\,.
\end{aligned}$$

Basket G_i is not the only basket from which Alice might choose to consume at least $x_i = m_i$ of good i, so this probability understates the true likelihood of learning i by time t, but the analytical point is that the probability converges to 1, which would obviously be equally true if other cases give rise to learning as well.

(b) For good i to not be learned even given a life that could be infinitely long, it must appear so unattractive that neither myopic nor experimental consumption seem worthwhile under any circumstance. The first condition of part (b) of this proposition ensures that, from a myopic perspective, Alice always prefers the numeraire good to a nibble (or more) of good i. Thus, the only way she could learn it would be through experimental consumption. The second condition ensures that for any set of preference beliefs β that Alice may have, if she encounters a basket with just this good in the first period ($t = 0$), she will not consume at least m_i of it and thus won't learn it. Since this is true for any possible preference beliefs, then by Lemma 4 if she won't learn it in the first period, she won't learn it in any period regardless of what preference beliefs she has at that later period, because those preference beliefs will be one of the possible preference beliefs for which Alice refuses to learn good i in time 0.

This proves that under the two conditions specified in the proposition, Alice will never consume at least a nibble of good i, so that by Lemma 1(b), she will never learn her preferences for this good.

□

Proof of Proposition 3

To learn her preferences for good i, Alice must satisfy all of these conditions:

(1) $p_i \cdot m_i \leq y$ (affordability), and either

(2) There exists $t \in \{0, \ldots, T\}$ such that $i \in G^t$ and the myopic optimization problem yields $x_i^{t*} \geq m_i$ (myopic consumption), or

(3) There exists $t \in \{0, \ldots, T\}$ such that $i \in G^t$ and $U(f^{(\beta^t)}, G^t) - U_i(f^{(\beta^t)}, G^t) \leq \delta \cdot \phi_i^{t+1}(f^{(\beta^t)})$ (experimental consumption).

Using these conditions, we address each subitem of this Proposition.

(a) Because the budget constraint is tighter relative to our baseline level of y, including in future periods, future consumption is lower, which reduces

ϕ_i^{t+1} and thus the right side of the inequality in (3), making experimental consumption less likely. If i is normal, then a smaller y also means that the optimal myopic choice in the current period is smaller and could fall below the nibble size, so myopic optimization (2) might cease to select enough of this good to learn it; and if it is already below the nibble size, then sampling this good requires a larger utility sacrifice in experimental consumption, increasing the left side of the inequality in (iii) and making experimental consumption less likely. (If i were inferior, the income effect would push in the other direction, rendering the effect of a change in y ambiguous.) These points together are unambiguous for normal goods and thus sufficient to show that a lower y reduces the chance of learning i, but other factors may aggravate the effect of a smaller y. First, a smaller y can make the inequality in (1) fail to hold, so that a nibble of i becomes unaffordable. Second, if there is diminishing marginal utility and the other goods that could be consumed are normal, a lower y would increase the present sacrifice associated with sampling this good, increasing the left side of the inequality in (3), making experimental consumption less likely.

(b) A larger p_i tightens the budget constraint and thus has the same effects as a reduction in y. Moreover, it makes learning good i more costly, so that $U(.) - U_i(.)$ will be larger, increasing the left side of the inequality in (3), making experimental consumption less likely. It also renders a nibble of the good less likely to be affordable, so it could make (1) cease to be met.

(c) An increase in m_i has the same effect on the good's affordability and the cost of learning as an increase in p_i, leading to the same conclusion as (b).

(d) Lowering q_i reduces the chance that $i \in G^t$ for any given t, which means that for a finite T, even if there exists a basket in which Alice would myopically consume (as in (2)) good i, she may not encounter that basket during her life. In addition, since there are fewer future consumption opportunities in which consumption of this good can be optimized, ϕ_i^{t+1} is reduced, which reduces the right side of the inequality in (3) and makes experimental consumption less likely.

(e) A smaller δ reduces the right-hand side of the inequality in (3), making experimental consumption less likely.

(f) A smaller T reduces ϕ_i^{t+1} by restricting the number of future periods in which Alice can benefit from better knowing her preferences. This reduces the right side of the inequality in (3), making experimental consumption less likely.

(g) A lower prior, that is a left-shifted $f_i^{(\beta_i^0)}$, by our assumption that utility is increasing in parameters, means that Alice's expected utility from good i is lower. This makes consumption of any amount of the good less likely to ever be myopically optimal (2) and increases the current-period sacrifice for experimental consumption, which increases the left side of the inequality in (3) and makes experimental consumption less likely.

(h) If Alice has a low prior, i.e. a left-shifted $f_i^{(\beta_i^0)}$ for good i such that it is not myopically optimal to consume a positive quantity of it, then a narrower probability density function will put less probability weight on parameters that would make i attractive enough to be tried. This lowers the good's upside potential, thus reducing ϕ_i^{t+1}, thus reducing the right side of the inequality in (3) and making experimental consumption less likely.

(i) If Alice is extremely confident in her prior preference beliefs such that $f_i^{(\beta_i^0)}$ (or, more precisely, the corresponding random variable β_i^0) has close to zero dispersion, then

$$\left.\frac{d\,Eu_i\,(x_i;\beta_i^0)}{dx_i}\right|_{x_i=m_i} \approx \left.\frac{d\,u_i\,(x_i;E[\beta_i^0])}{dx_i}\right|_{x_i=m_i}. \tag{7}$$

The condition in the proposition gives us that these values are greater than $z \cdot \frac{p_i}{p_1}$. In this case, by Proposition 2(a), good i will eventually be learned.

Let us now increase the dispersion of β_i^0 while keeping its mean constant. As we do so, $\left.\frac{d\,u_i\left(x_i;E[\beta_i^0]\right)}{dx_i}\right|_{x_i=m_i}$ remains the same, whereas $\left.\frac{d\,Eu_i\left(x_i;\beta_i^0\right)}{dx_i}\right|_{x_i=m_i}$ declines because Alice is risk averse and thus loses expected utility as a result of the uncertainty in β_i^0. The more uncertain her beliefs, the lower her expected utility. As a result, with enough uncertainty in her beliefs, the left side of Equation (7) will fall below the right side. If this effect is large enough relative to how "positive" the average prior is, it could make it so that good i will not be learned through myopic consumption, i.e., so that condition (2) does not hold. The level of dispersion of β_i^0 and her believed preferences for the other goods could also make this expected utility so low that the current period utility sacrifice, the left side of the inequality in (3), is so high that experimental consumption does not occur.

Thus, under the given assumptions of risk aversion and a positive prior, a higher level of uncertainty around her beliefs can prevent Alice from learning her preferences for a good.

(j) If $\hat{\beta}_j$ is larger for (at least one) learned good $j \neq i$, or if $f_j^{(\beta_j^0)}$ is more right-shifted for a not-yet-learned good $j \neq i$, then good i appears relatively less attractive. This reduces ϕ_i^{t+1}, since the net gain that could be achieved from consuming i in the future is lower if the utility from consuming counterfactual goods is higher. This is sufficient to show that more attractive other goods make it less likely to learn preferences for a good. Other channels may also be relevant. For example, increased attractiveness of other goods may reduce the optimal myopic choice of i in some periods, which might drop the myopic optimal choice below a nibble and would further increase the sacrifice involved in experimental consumption of good i if it was already not myopically optimal to learn.

$\square$

Proof of Proposition 4

Let $T = \infty$. Suppose that at some time t, Alice's preference beliefs are unstable in the sense that they will change at some later time. For this to be true, there must be a good for which she will learn her preferences at some point in the future, because in our model that is the only way that preference beliefs change. Therefore, there exists some set G and some good $i \notin L^t$, such that under her current preference beliefs β^t, Alice will choose to consume at least a nibble of good i, thus learning her preferences for it, if G appears as the available set of goods.

Let $\tau > t$ denote the first time (since t) that the set $G_i = \{1, i\}$ appears.[5]

[5] If Alice chooses at least a non-trivial consumption amount m_i under set G, she would also choose to consume at least m_i units of good i if set G_i appeared. Therefore we can

It must be true that $L^{\tau+1} \neq L^t$, that is, that Alice will have learned a new good between time t and time τ. This is because either (i) Alice updated her beliefs between time t and τ due to some other consumption experience, so the learned set must expand based on that preference belief change; or (ii) her beliefs have not changed in that time so that $f_i^{(\beta_i^\tau)} \equiv f_i^{(\beta_i^t)}$, in which case Lemma 4 implies that good i will now be learned, that is $i \in L^{\tau+1}$ when we know it was not in L^t. Thus, we have shown that unstable preference beliefs will result in a preference belief change as a good is added to the learned set by (or at) the time Alice encounters the minimal set that includes the good in question.

In each period, the probability that G_i is the available set of goods is non-zero and time-invariant, since there is only a finite number of goods (and thus a finite number of possible sets G) and since the probabilities with which goods appear are constant and independent from each other. Let

$$\rho = \min_{i \in \{2,\ldots,N\}} \mathbb{P}(G_i) > 0$$

denote the probability that the set G_i appears in any given period for the non-numeraire good i whose minimal set is least likely to appear. This need not be the good whose learning triggers the learned set change discussed in the first two paragraphs of this proof, but since that scenario involved either the good i under consideration or some other unknown good learned between t and τ, we can't identify which good and thus which probability to use, so we use that of the good least likely to appear, as that will give the smallest (most conservative) possible probability ρ.

Combining this with what precedes it, if Alice's preference beliefs are currently unstable, there is a positive, time-invariant probability (greater or equal to ρ) in each future period that she will change her preference beliefs in that period, until the first change occurs. Let $T_1 \geq 1$ denote the number of periods it takes for such a change of preference beliefs to occur for the first time. This is a random number because it depends on realized appearances of goods. Similarly, let $T_2 \geq 1$ denote the additional number of periods until the second change of preference beliefs, etc. Note that each T_j measures the number of periods until an event occurs, which happens with probability of at least ρ (which is a constant) each period. Therefore, T_j follows a geometric distribution with a probability parameter of at least ρ.

restrict our attention to the sets G_i without loss of generality for the purpose of this proof.

Since there are only $N - 1 < \infty$ goods to be discovered, and since preference beliefs for each good remain stable once discovered (Lemma 1), there can be at most $N - 1$ preference belief changes in Alice's lifetime. (There are fewer such changes if she discovers multiple goods at the same time, or if some goods are destined to remain forever undiscovered.) Therefore, let $M \leq N - 1$ denote the number of changes to Alice's preference beliefs that will occur over her lifetime under the realized set of baskets G^t for $t \geq 1$. The time of her final change of preference beliefs is thus no greater than $T_1 + \ldots + T_M$. Note that the sum of geometric distributions (with the same parameter) follows a negative binomial distribution and reflects how many periods it takes for (in this case) M events to occur.

Therefore, if T^* denotes a random variable that follows a negative binomial distribution with probability parameter ρ and frequency parameter M, we can conclude that:

$$\mathbb{P}(f_i^{(\beta_i^s)} \equiv f_i^{(\beta_i^t)} \; \forall \, s \geq t) \geq \mathbb{P}(T_1 + \ldots + T_M \leq t) \geq \mathbb{P}(T^* \leq t) \to 1 \text{ as } t \to \infty \, .$$

The first inequality follows from our earlier discussion that preference beliefs will not change after time $T_1 + \ldots + T_M$. Note that the T_j each have an event probability of greater or equal to ρ, while T^* assumes a probability of ρ for each period. Therefore, the sum of the T_j is more likely to be smaller than t compared to the likelihood that T^* is smaller than t, for any $t > 0$. This is reflected in the second inequality. Lastly, the convergence is a property of the negative binomial cumulative distribution function, and since this probability approaches 1, so too must the (larger) probability that preference beliefs, and thus choices, become stable.

Note that this proof does not show that all goods will be learned eventually, because it is not true that all goods have some corresponding set under which the good will be chosen.

□

Proof of Proposition 5

(a) Let $i \in \{2, \ldots, N\}$ denote a good that as of time t is undiscovered ($i \notin L^t$) and affordable ($p_i m_i < y$). Let x'_i denote the quantity of good i that Alice would choose to consume as the solution to her myopic choice problem if the set of available goods was $G^t = G_i = \{1, i\}$.

Consider first the case in which under her true preferences $\hat{\beta}_i$, she would choose $x'_i \neq m_i$. Consider a good-i prior $\beta_i^0 \in \mathbb{B}_i$ that satisfies

$$\left.\frac{d\,Eu_i\,(x_i;\beta_i^0)}{dx_i}\right|_{x_i=m_i} = z \cdot \frac{p_i}{p_1}\,.$$

Under this prior, Alice consumes $x_i^t = m_i$ at time t, and thus learns her true preferences for good i by Lemma 1. That is, $f_i^{(\beta_i^{t+1})} \equiv \Delta(\hat{\beta}_i)$. If at time $t+1$ Alice has the same set of goods available, i.e. $G^{t+1} = G_i$, then she knows her preferences for all available goods. Therefore, there is no opportunity for experiential learning at this time, and therefore her optimal consumption bundle is equal to the solution of her myopic choice problem, which entails $x_i^{t+1} \neq m_i$, by assumption.

Secondly, for the alternative case where under her true preferences $\hat{\beta}_i$ Alice would choose $x'_i = m_i$, consider an (overly optimistic) good-i prior $\beta_i^0 \in \mathbb{B}_i$ such that $x_i^t = y/p_i > m_i$. Following the same logic as in the previous paragraph, we arrive at the same conclusion that $x^{t+1} \neq x^t$ if $G^t = G^{t+1} = G_i$.

In either case, the choice reversal occurs if we have both $G^t = G_i$ and $G^{t+1} = G_i$. By our assumptions in Section 2.2, the two events are independent and have a strictly positive probability of occurring. This proves the first part of the proposition. Therefore, under the given prior, there is a non-zero probability that we observe a choice reversal: despite facing the same basket of available goods at two different times ($G^{t+1} = G^t$), Alice chooses a different bundle ($x^{t+1} \neq x^t$).

We have proved this for particular priors in each case, but it should be evident that many other configurations can also lead to choice reversals.

(b) Once preference beliefs become stable—which Proposition 4 guarantees to happen eventually—Alice will always choose her consumption in order to maximize her myopic expected utility. Since preference beliefs no longer change, this choice is time-invariant, conditional on the available set of goods. In other words, choice reversals no longer occur.

□

Proof of Proposition 6

Let $i \in \{2, \ldots, N\}$ denote a good for which $p_i m_i \leq y$ and $\left.\frac{d\, u_i(x_i;\hat{\beta}_i)}{dx_i}\right|_{x_i=0} > z \cdot \frac{p_i}{p_1}$. Such a good exists based on the assumptions of the proposition.

(a) Choose a prior $\beta_i^0 \in \mathbb{B}_i$ such that both of the following conditions are satisfied:

  (i) $\left.\frac{d\, Eu_i(x_i;\beta_i^0)}{dx_i}\right|_{x_i=0} < z \cdot \frac{p_i}{p_1}$, and

  (ii) $\max\limits_{G \in \mathbb{G}\,,\, \beta \in \mathbb{B}_i'} \delta \cdot \phi_i^1(f^{(\beta)}) - U(f^{(\beta)}, G) + U_i(f^{(\beta)}, G) < 0$, with $\mathbb{G}$ and $\mathbb{B}_i'$ defined in Proposition 2.

  Note that
  $$\left.\frac{d\, Eu_i(x_i;\beta_i^0)}{dx_i}\right|_{x_i=0} \geq \left.\frac{d\, Eu_i(x_i;\beta_i^0)}{dx_i}\right|_{x_i=m_i},$$
  since $u_i(.)$ is concave (by Axiom 4) and since $Eu_i(.)$ is a linear combination of concave functions and thus also concave. Therefore, the marginal expected utility is a (weakly) decreasing function of x_i.

  Given these conditions, by Proposition 2(b), good i will never be learned. Since it is not worth consuming either myopically or experimentally, Alice will always choose to consume $x_i^t = 0$ units of good i.

  However, since the *true* marginal utility of good i exceeds that of the numeraire good, it would be optimal for Alice to consume a positive quantity of the good at least every time the choice set $G_i = \{1, i\}$ appears (and possibly in some other baskets). As a result, whenever $G^t = G_i$, Alice will make a suboptimal consumption choice and thus lose a positive amount of welfare. Since the probability that $G^t = G_i$ is strictly positive (and constant over time), the expected welfare loss is positive for all $t \geq 0$.

(b) Welfare loss results from suboptimal consumption choices due to either (i) lack of knowledge of true preferences, or (ii) experimental consumption for the purpose of learning the true preferences. Both of these effects diminish over time, as more parameters are being discovered. Preference discovery brings the current preference beliefs that Alice uses for her decision making closer to her true preferences, thus reducing both the likelihood and severity of the expected welfare loss in any given period. In addition, over time, experimental consumption happens less often, because fewer parameters will be unknown (and, if time is finite, because the benefits of learning diminish in expectation as time passes, since fewer periods remain), while the cost of learning is time-invariant (again, in expectation), so that goods that remain undiscovered become increasingly less likely to ever be discovered. Therefore, the expected welfare loss Δu^t (weakly) decreasing as t increases.

(c) The example provided in the proof to part (a) of this proposition entails a case in which $\Delta u^t \not\to 0$ as $t \to \infty$.

$\square$