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]*Temple University, Fox School of Business, Alter Hall 611, 1801 Liacouras Walk, Philadelphia, PA 19122, USA*

March 2017

Abstract

We develop an axiomatic theory that integrates the discovered preference hypothesis into neoclassical microeconomic choice theory, making predictions amenable to empirical tests. Several regularities in economic literatures could be explained by a theory in which preferences must be discovered through experience. These include: choice reversals as seen in various contexts, instability as seen in risky choice, and errors that decline with repetition as seen in contingent valuation. With reasonable assumptions, we show that choices may appear unstable while preferences are being learned, and that unlearned preferences are associated with welfare loss. We also show that even after choices appear to stabilize, agents face the potential for continued welfare loss due to persistent mis-ranking because of selection bias in the feedback and learning process. The transitory welfare loss that occurs during

*jdelaney@ggc.edu
†sarah.a.jacobson@williams.edu; Corresponding author
‡moenig@temple.edu

the learning process decreases over time, with more common goods, and with more income. For large discrete items purchased a small number of times (like houses), this transitory welfare loss may continue the agent's whole life. The long-run welfare loss caused by persistent mis-ranking is primarily determined by initial misperceptions of goods. In extensions, we demonstrate that imperfect memory of learned tastes and stochasticity in the consumption experiences may make preference learning harder, and that learning spillovers across goods and sophisticated agents who know they need to learn their preferences may or may not alleviate welfare loss.

Keywords: discovered preferences, preference stability, learning, risk preferences

JEL codes: D81, D83, D01, D03

1 Introduction

"You do not like them. So you say. Try them! Try them! And you may."

Green Eggs and Ham
Dr. Seuss

Neoclassical microeconomic choice theory is grounded in the assumption that people choose according to a stable ranking of preferences: that these preferences exist and that people know their own preferences and make choices in accordance with them. However, as discussed in Plott (1996), it is possible that people don't know their tastes until they discover them through consumption experience. When preferences are as-yet-undiscovered, people may make choices that are suboptimal. If this is the case, then some results derived from foundational neoclassical assumptions come into question. In this paper, we construct a model of a preference discovery process to explore its implications for choice patterns and welfare.

Consider an encounter with a strange new food. For example, one of us had never eaten celeriac until recently. She didn't know whether she would like it, and *a priori* believed it untasty. However, after trying it she discovered that it suited her tastes very well. Through experience, she was able to arrive at a more accurate assessment of her preferences, and now enjoys much more efficient levels of celeriac consumption. However, because of her initial misperception, she might have missed out on a lifetime of delicious consumption had she not been induced to try it—indeed, she is likely missing out on other delicious fruits and vegetables on a regular basis due to mistaken beliefs and a lack of experience.

The idea of tastes that are not fully known to the decision-maker has received a small amount of attention in economics but much more in psychology. As discussed by Kahneman et al (1997), Scitovsky (1976) argued that people are bad at predicting utility from a prospective choice. Becker (1996) argued the opposite, and indeed, Kahneman and Snell (1990) note that, when experiences are familiar and immediate, people seem fairly good at predicting utility. Many results from psychology and economics support Scitovsky's claim, however. Loewenstein and Adler (1995) find people fail to predict changes in tastes, and Wilson and Gilbert (2005) review extensive evidence showing systematic errors in forecasting happiness.

Psychological models generally do not feature stable, context-independent preferences (Ariely et al, 2003; Lichtenstein and Slovic, 2006). In contrast, nearly all microeconomic empirical and theoretical work, and most macroeconomic work, depends implicitly on an assumption of stable preferences.

A few papers have studied preference discovery from an economic perspective and there is some evidence that choice instability and feedback interact.[1] van de Kuilen and Wakker (2006) find that repeated trials without feedback do not reduce Allais paradox violations, but with feedback, the violations decrease. Weber (2003) finds that repeated plays of a strategic game exhibit more apparent learning when feedback is provided. This is consonant with the importance of feedback in the preference discovery process highlighted in Plott (1996).

Other economic experiments on repeated choice with feedback provide further suggestive evidence of preference discovery. With repeated choice, people can converge to a true induced value (Noussair et al, 2004). Errors and biases also often decline with repeated choice, as observed in the gap between willingness-to-pay and willingness-to-accept, non-dominant bidding behavior, and in strategic games (Coursey et al, 1987; Shogren et al, 1994, 2001; List, 2003).

Our contribution is to develop a theory that fully integrates preference discovery, as described in Plott (1996), into a neoclassical microeconomic framework. We maintain the assumption of stable underlying preferences but allow for a need to learn them through experience. We show that if preferences must be discovered, we will eventually observe choices that appear consistent with stable underlying preferences in most contexts: discovery occurs, choices become more consistent with some stable ranking, but these choices may not necessarily be consistent with a person's underlying stable preferences and thus the agent may lose welfare from suboptimal choices. In other words, incorrect preference rankings and thus choices that do not fully maximize utility can persist for a person's whole life. We show the kinds of items for which tastes are most likely to be undiscovered: those encountered rarely, those initially believed to be distasteful, and those most different from

[1] We distinguish between learning about objective circumstances and about one's own tastes, which Braga and Starmer (2005) refer to as "institutional learning" and "value learning" respectively. Our focus is on value learning, so we assume the agent knows the objective features of all goods. Institutional learning is best separately modeled, e.g., in experimental consumption models (Kihlstrom et al, 1984) or the two-armed bandit problem (Rothschild, 1974).

our expectations about them. If learning spillovers across goods is possible, welfare losses may be reduced, but learning need not occur any faster or more completely. We further show that risk preferences are harder to learn than other kinds of preferences.

Our model generates several testable hypotheses, some of which are consistent with existing experimental evidence and some of which remain to be explored. First, we find that choice instability is more likely to appear in repeated choice among people who are inexperienced (young), poor, or forgetful. In addition, we should see more instability in unfamiliar choice environments: goods that are rare or expensive; large, discrete, infrequently chosen goods (especially high-stakes goods, e.g., houses); or goods that have many varying characteristics such that any given consumption experience is unique. We also expect choice instability to occur in the presence of uncertainty. Finally, we predict that the more experienced someone is, the more likely it is that a new experience will not be voluntarily chosen, but also that the new experience will result in a positive (rather than negative) surprise, as the choice-experience process leads to negatively biased expectations of untried experiences.

This paper proceeds as follows. First, we establish a simple model of preferences with item-specific parameters, in which a finitely-lived agent has intertemporally-stable true parameters as well as estimates (which may change over time) of those parameters. We introduce a model of learning through experience. In the next section, we show that even in a "very long" life, some parameters will remain mis-estimated, and we show what circumstances are most likely to give rise to undiscovered preferences. In the following section, we extend the model in four ways. First, we relax the assumption that preferences, once learned, are remembered forever. Second, we relax the assumption that one can learn nothing about good A by experiencing good B. Third, we allow the agent to be a bit sophisticated by allowing her to know that she doesn't yet know her own tastes. Fourth, we discuss the special case of risk preferences. Finally, we conclude.

2 A Model of Preference Discovery

We model the decision-making of an agent, Alice, who has tastes over $N \in \mathbb{N}$ goods, $i = 1, \ldots, N$. We use the term "goods" quite generally as some might be "bads" and they may represent goods, services, experiences, or attributes.

For ease of exposition, we use food items as examples throughout the paper. In this section, we limit our consideration to *deterministic* goods: goods that are, within a class of goods, undifferentiated and identical in quality.

We assume that Alice has a stable underlying preference ordering $\succsim$ over these goods, and that this ordering obeys the standard assumptions of rational preferences.

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

We can therefore represent Alice's tastes with a utility function. Let her utility from consuming x_i units of good i (for each $i \in \{1, \ldots, N\}$) be $u(x_1, \ldots, x_N \mid \boldsymbol{B})$ with a set of parameters $\boldsymbol{B}$.[2]

Axiom 2. Stability of True Preferences.
At any time $t > 0$, the agent's realized utility from consuming a bundle of goods, $(x_1, \ldots, x_N)$, is $u\left(x_1, \ldots, x_N \mid \boldsymbol{B}\right)$. In particular, both $u(.)$ and $\boldsymbol{B}$ are time-invariant.

Because this utility function represents Alice's stable underlying preferences, the utility function is stable over time. We further assume that preferences for each item are monotonic, but we allow some goods to give positive and some to give negative marginal utility. Finally, we assume that preferences are convex, which implies a concave utility function.

Axiom 3. Shape of Utility Function.
The agent's utility function $u\left(x_1, \ldots, x_N \mid .\right)$ is twice differentiable, either monotonically increasing or monotonically decreasing, and concave with respect to each good. That is, for any $i \in \{1, \ldots, N\}$, for any possible set of parameter values $\tilde{\boldsymbol{B}}$, and for any set of non-negative values x_j for $j \neq i$:

(i) Monotonicity: Either $\frac{\partial u\left(x_1,\ldots,x_N \mid \tilde{\boldsymbol{B}}\right)}{\partial x_i} > 0$ for all $x_i \geq 0$ or $\frac{\partial u\left(x_1,\ldots,x_N \mid \tilde{\boldsymbol{B}}\right)}{\partial x_i} < 0$ for all $x_i \geq 0$.

[2] We use a utility function for convenience; our conceptual points about preference 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."

(ii) *Concavity:* $\frac{\partial^2 u\left(x_1,\ldots,x_N \mid \tilde{\boldsymbol{B}}\right)}{(\partial x_i)^2} < 0$ *for all* $x_i \geq 0$.

Because $u(x_1, \ldots, x_N \mid \boldsymbol{B})$ is her true utility function, Alice is able to correctly perceive the utility she experiences upon consuming a bundle of goods. However, in contrast to the traditional microeconomic theory, we do not require that Alice *knows* her utility function. For simplicity, we assume that she knows the functional form of $u(x_1, ..., x_N)$, but not necessarily its true parameters $\boldsymbol{B}$. Instead, at any time $t \geq 0$, she believes that her utility function is parameterized with $\boldsymbol{B}^t$, and this belief may change over time. As a result, her consumption choices at time t maximize $u\left(x_1, \ldots, x_N \mid \boldsymbol{B}^t\right)$, subject to the availability of goods and to her budget constraint.

Let us assume that at discrete times $t = 0, 1, 2, \ldots$, the agent has access to a random subset—denoted by G^t—of the universe of goods, and 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 $q_i := \mathbb{P}\left(i \in G^t\right)$ and we require that $0 < q_i < 1$. At each time t, Alice is endowed with money m^t that she can spend. For simplicity, money cannot be transferred across time periods. The price per unit of good i $(i = 1, \ldots, N)$ is time-invariant and denoted by $p_i > 0$.

In each period, then, Alice maximizes the utility function she believes she has at that time. Her optimization problem is:

$$\max_{x_i (i \in G_t)} u\left(x_1, \ldots, x_N \mid \boldsymbol{B}^t\right) ,$$

subject to

$$\begin{aligned} \sum_{i \in G^t} p_i \cdot x_i &\leq m^t , \\ x_i &\geq 0 \text{ for all } i \in G^t , \quad \text{and} \\ x_i &= 0 \text{ for all } i \in \{1, \ldots, N\} \setminus G^t . \end{aligned} \tag{1}$$

Axioms 1 and 3, as in the standard choice problem, ensure the existence and uniqueness of a solution to the agent's optimization problem. Specifically, for available goods $j \in G^t$ for which the perceived marginal utility is always (and thus is at consumption level $x_j = 0$) below a certain level, Alice will choose to consume zero units, while she will consume a positive amount x_i of all other available goods $i \in G^t$ to the point where the perceived (decreasing) marginal utilities per dollar of each of these goods are equalized (to a value equal to the marginal utility of money). All perceived "bads" always fall below that threshold, and some goods may as well; the threshold is determined by the functional form of the utility function, Alice's money, the prices of the goods, and which other goods are available.

The agent's loss of welfare (in the form of lost utility)[3] at time t as a result of her undiscovered preferences is then given by

$$\Delta u^t = \max_{x_i (i \in G_t)} u\left(x_1, \ldots, x_N \mid \boldsymbol{B}\right) - u\left(x_1^*, \ldots, x_N^* \mid \boldsymbol{B}\right) \ ,$$

subject to Equation (1), and with $x_1^*, \ldots, x_N^*$ as the solutions to the agent's optimization problem.

Immediately, we note that the expected welfare loss from Alice's information-constrained choices is larger the farther Alice's believed parameters are from their true values, that is, the more inaccurate her guess about her tastes.

Lastly, we make the simplifying assumption that Alice's utility function is *separable*, so that the utility Alice gets from each good does not depend on the amount of other goods she consumes in the same bundle:

Axiom 4. Separability of Utility.
For all goods $i, j \in \{1, \ldots, N\}$ *with* $i \neq j$, $\frac{\partial^2 u(x_1, \ldots, x_N \mid .)}{\partial x_i \partial x_j} = 0$.

As a result of Axiom 4, we can state Alice's utility function as follows:

$$u(x_1, x_2, ..., x_N \mid \boldsymbol{B}) = v_1\left(x_1 \mid \boldsymbol{\beta_1}\right) + v_2\left(x_2 \mid \boldsymbol{\beta_2}\right) + \ldots + v_N\left(x_N \mid \boldsymbol{\beta_N}\right) \ ,$$

where $\boldsymbol{\beta_i}$ is the parameter vector of Alice's value function $v_i\left(x_i \mid .\right)$ for good i. Therefore, $\boldsymbol{B} = \{\boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_N}\}$ is the set of Alice's actual parameter vectors for each good. Similarly, $\boldsymbol{\beta_i^t}$ is Alice's perceived parameter vector for good i at time t, so that $\boldsymbol{B}^t = \{\boldsymbol{\beta_1^t}, \ldots, \boldsymbol{\beta_N^t}\}$. Separability allows Alice to learn her taste for one good at a time even though consumption occurs in bundles. If preferences were not separable, learning of tastes may be more difficult in that it may take longer (though it would not change our key insights).

[3]Since utility is not cardinal, it is often preferable to define welfare losses in terms of compensating or equivalent variation, particularly when comparing utility across agents. However, since we restrict our attention to a single agent at a single point in time, utility loss is equally appropriate for illustrative purposes.

2.1 Learning by Experience

In this section, we model Alice's belief about her parameters $\boldsymbol{B}^t$ as a point estimate belief.[4] For ease of exposition, we use the following notation:

Definition 1.

(a) *W^t denotes the set of "bads" according to the agent's beliefs at time t:*

$$W^t = \left\{ i \in \{1, \ldots, N\} : \frac{\partial u\left(x_1, \ldots, x_N \mid \boldsymbol{B}^t\right)}{\partial x_i} < 0 \right\} .$$

In particular, W^0 is the set of items she considers to be bads before she has consumed any goods (according to her initial beliefs $\boldsymbol{B}^0$). Similarly, W without a time index denotes the set of all "bads" according to the agent's true preferences $\boldsymbol{B}$.

(b) *Z^t denotes the set of items that the agent consumes a positive amount of at time t:*

$$Z^t = \{ i \in \{1, \ldots, N\} : x_i > 0 \text{ at time } t \} .$$

(c) *L^t is the set of goods $i \in \{1, \ldots, N\}$ for which the agent has correctly learned her preference parameters at time t. That is:*

$$L^t = \left\{ i \in \{1, \ldots, N\} : \boldsymbol{\beta}_{\boldsymbol{i}}^t = \boldsymbol{\beta}_{\boldsymbol{i}} \right\} .$$

Recall from our earlier discussion that at any time $t = 0, 1, 2, \ldots,$ the agent chooses not to consume any good that is in W^t. That is:

Lemma 1.

$$Z^t \cap W^t = \varnothing .$$

□

[4] In Section 3.3 we discuss the case in which Alice is uncertain about the parameters. This will allow her to know that she may not have correct parameters, while a model with parameter point estimates treats an estimate identically whether it is guessed or learned by experience.

We further assume that learning the parameters associated with one good is uninformative for learning the parameters associated with other goods. We relax this assumption in Section 3.2.

Axiom 5. ***Separability of Learning.***
Experiencing a good has no effect on the agent's perceived parameters of any other good.

Moreover, we assume that once learned, parameters are not forgotten. Section 3.1 explores what happens when this assumption is relaxed.

Axiom 6. ***Persistent Memory.***
The agent updates her beliefs for a good only as a result of a relevant experience (of this or any other good).

We now can make some observations about properties of the learning process.

Lemma 2.

(a) If $i \notin Z^t$ *for some* $t \geq 0$*, then* $\boldsymbol{\beta}_{\boldsymbol{i}}^{t+1} = \boldsymbol{\beta}_{\boldsymbol{i}}^{t}$.

(b) For any values of time s *and* u *with* $s < u$*:* $W^s \subseteq W^u$.

(c) For any values of time s *and* u *with* $s < u$*:* $L^s \subseteq L^u$.

□

Thus, Alice only updates a good's parameters when that good is actually experienced so that goods not sampled at a given time will not see their parameters updated in that time.[5] Moreover, once a good finds itself in the "basket of unpalatables," it will remain there forever; that state is absorbing. The same is true for the set of learned goods.

Every time Alice experiences a bundle of goods, she notes whether her experience differs from her belief at the time—that is, if she consumes some bundle of goods at time t, she notices if $u(x_1, \ldots, x_N \mid \boldsymbol{B}^t) \neq u(x_1, \ldots, x_N \mid \boldsymbol{B})$.

[5] There are cases in which Alice can mathematically identify parameters for one good she's experienced in the past in combination with other goods at the moment she fully learns the parameters associated with the other goods. We ignore these cases for simplicity.

Ultimately, once she has experienced a good often enough so that she has been able to mathematically identify her true preferences for this good, she will update the associated parameters accordingly, as the following axiom states.

Axiom 7. ***Updating of Learned Preferences:***
The agent updates her beliefs about a good if and when she has experienced it sufficiently to determine her true set of parameters for that good, so that if the last experience necessary for mathematical identification occurs in time t, then $\boldsymbol{\beta}_{\boldsymbol{i}}^{t+1} = \boldsymbol{\beta}_{\boldsymbol{i}}$.

For example, if Alice's value function for good i has M_i parameters, given separable utility, it will take her M_i trials of different quantities of good i alone to fully determine her preferences over this good. Thus, if her value function is $v_i(x_i) = \alpha_i x_i^{\gamma_i}$, she will learn her taste for good i in two trials (each of a different quantity) of the good. However, when she consumes a mixed bundle containing not just good i but also other goods, she will only know the utility of that bundle; if she has not already mathematically correctly identified the parameters of the other goods in the bundle, she must experience their combination enough times to separately identify each parameter. The maximum number of relevant[6] trials it could take to learn parameters for all goods would be $M_1 + M_2 + \ldots + M_N$.

Since goods appear in baskets with positive and independent probabilities, we can then conclude the following.

Lemma 3.

$$\textit{If } i \notin W^0, \ \textit{ then } \lim_{t \to \infty} \mathbb{P}\left(i \in L^t\right) = 1 \,.$$

□

In other words, any good that doesn't start out in Alice's original believed basket of unpalatables will be eventually learned.

We now define some concepts related to Alice's knowledge of her utility function.

Definition 2.

[6] "Relevant" in this context means that the agent consumes at least one good in the trial that she has not fully learned her preferences of.

(a) Full discovery *of preferences is the state in which the agent has a correct estimate of her preference parameters such that* $\boldsymbol{B}^t = \boldsymbol{B}$*. This occurs when* $L^t = \{1, \ldots, N\}$*.*

(b) Full relevant discovery *is the state in which the agent has a correct estimate of her preference parameters for all goods that are not truly bads, and correctly categorizes each choice item* i *as either a good or a bad, such that* $\boldsymbol{\beta}_{\boldsymbol{i}}^t = \boldsymbol{\beta}_{\boldsymbol{i}}$ *for all* i *not in* W*. This occurs when both* $L^t \supseteq \{1, \ldots, N\} \setminus W$ *and* $W^t = W$*.*

(c) Full voluntary discovery *is the state in which the agent has a correct estimate of her preference parameters for all goods that she does not believe to be bads, such that* $\boldsymbol{\beta}_{\boldsymbol{i}}^t = \boldsymbol{\beta}_{\boldsymbol{i}}$ *for all* i *not in* W^t*. This occurs when* $L^t \supseteq \{1, \ldots, N\} \setminus W^t$*.*

Full *relevant* discovery allows Alice to have incorrect estimates of items that are bads for her. For example, she may correctly rank all fruits, and she may correctly assess that ham and beef are bads for her, but she may believe that ham gives more marginal (dis-)utility than beef when the reverse is actually true. This mis-ranking is irrelevant because it could never cause her to choose a bundle that would not maximize her utility. When Alice has full voluntary discovery, however, she has discovered her tastes for all of the goods that she does not believe to be in W^t, but she might be wrong about her classification of some goods as bads. Full discovery implies full relevant discovery, which implies full voluntary discovery.

It therefore follows that:

Proposition 1. *The agent's time-*t *welfare loss* Δu^t *equals* 0 *with probability* 1*, if and only if she has achieved full relevant discovery prior to time* t*.*

□

In other words, Alice will not lose utility from mistaken choices once she achieves full relevant discovery, but she may continue to lose utility when she is not in that state. In particular, the state of full *voluntary* discovery does not ensure zero utility loss from misunderstood preferences.

We can now state the conditions under which Alice will eventually discover her preferences.

Proposition 2. *With probability 1 as* $t \to \infty$*, the agent will achieve:*

(i) Full voluntary discovery.

(ii) Full relevant discovery, if and only if $W^0 \subseteq W$.

(iii) Full discovery, if $W^0 = \varnothing$.

□

Note that for Proposition 2 (iii), the reverse statement is not necessarily true; that is, Alice can achieve full discovery even if $W^0 \neq \varnothing$. This is the case if the agent's preferences for a good $i \in W^0$ are accidentally correct. That is, $\boldsymbol{\beta}_{\boldsymbol{i}}^0 = \boldsymbol{\beta}_{\boldsymbol{i}}$, even though the agent never experiences this good. Formally, $i \in L^t$ for all $t \geq 0$, and therefore it is possible that $L^t = \{1, \ldots, N\}$ with probability 1 as $t \to \infty$; however $W^0 \neq \varnothing$.

Otherwise, if there is an item in W^0, Alice will not experience it and therefore not discover her true preferences for that item, and thus she will not achieve full discovery. This will only impede full relevant discovery if any item is mistakenly in W^0, i.e. if she believes something is a bad when it is actually a good for her. Some goods that are outside of W^0 may belong in W, but she will learn this with experience eventually and demote them into W^t at some time t.

For example, if she believes (correctly) that all meats are in her set W but mis-ranks some initially in her W^0, those mis-rankings will persist. This mis-ranking will not matter for her choices. She will, however, try all fruits if they are all outside of W^0, and as time progresses she will try each one more and more times. Imagine that the fruits that exist are apples, oranges, and bananas. Alice believes that apples are preferable to oranges, and that these are the only foods worth eating—the meats and bananas are all repugnant. Imagine further, however, that according to her true tastes, Alice's most-preferred fruit is bananas, followed by oranges, while the apples and meats are actually noxious to her.

Over the course of Alice's life, she encounters opportunities to eat fruits and meats and chooses bundles containing combinations of apples and oranges. Her mistaken belief about apples is self-correcting: she will take one bite of an apple, spit it out, and forgo apples thenceforward. Thus she will show a choice reversal on apples: later, given the same choice set, she will choose a different bundle. In all later choice opportunities, Alice will eat oranges if they are available and nothing else, and she will understand her tastes for oranges very well (full voluntary discovery). Her tastes will look stable to any econometrician analyzing her choice data after that fateful apple encounter, and the econometrician might believe that her choices maximize her utility. However, a welfare loss occurs every time bananas appear in her choice set, because she chooses not to eat them (preferring to eat oranges or to go hungry) and thus she never learns that she really likes bananas. Because bananas are in her W^0 but not her W, she can never achieve full relevant discovery and thus fails to optimize.[7]

In the long run, then, this selective feedback process implies that there will be a systematic downward bias in our beliefs about untried goods. As Alice tries goods, she will find that she was incorrect in her beliefs about how much she would like each good. In the cases in which her error was positive (she expected she would like it more than she did) this error will lead some goods to fall outside of W^t when in fact they are in W. These goods will eventually be tried and experience will lead to a correction of the error. Where her error was negative, this error will lead some goods to fall within W^t when they are not in fact in W. These errors will lead Alice to never try these goods. In this way, positive errors correct themselves but negative errors do not, and so under full voluntary discovery, Alice will have, on average, an underestimate of how much she would like a given untried good. This is consistent with evidence indicating that people are too cautious when considering large life changes (Levitt, 2016).

We can now state some regularities about the learning process.

Proposition 3.

Under the conditions that (according to Proposition 2) must hold to yield each type of discovery (full, full relevant, and full voluntary), the agent is more likely to achieve that type of discovery by time $t \in (0, \infty)$ ...

(i) As time passes; that is, if t is larger.

(ii) If the agent has more real purchasing power available each period; that is, if m^s is larger for all (or some) s, or if the goods' prices are low.

(iii) For goods that are more likely to appear; that is, if q_i is larger.

□

[7] One can also think of this as though Alice solves her optimization problem with an additional constraint that $x_i^t = 0$ for all $x_i \in W_t$. However, this additional constraint is imposed by her misunderstanding, and thus her behavior is not globally utility-maximizing.

In the short run, then, at finite times t, the situation is worse than in the long run. Alice is prone to inefficient consumption decisions until she reaches full relevant discovery; that is, for all times $s < t$ until the time t when good i enters into the "learned set" L^t, she will construct bundles with too much or too little of good i relative to what would actually maximize her utility, even if she doesn't mis-classify the good as a bad. Thus she loses welfare from suboptimal bundle choices; what's more, her tastes will look unstable because given the exact same budget constraint and available basket (i.e., $m^s = m^t$ and $G^s = G^t$) she will choose a different bundle in time s than in time t. Therefore, in expectation, her welfare loss $\Delta u^{\cdot}$ is larger, at time s than at time $t > s$.

Agents with more real purchasing power will be able to consume more goods and thus may learn their tastes faster. However, if we calculate lifetime welfare lost as $\Delta U = \sum_{s=0}^{t} \Delta u^t$, it is not clear whether that total loss will be less or greater if Alice has more real purchasing power: if she has more money, she will learn her tastes faster and thus lose utility for fewer periods, but her total utility and thus the scale of utility loss will be larger when her tastes are as-yet-unlearned.[8]

Finally, it is not surprising that goods that are rarer are harder to learn tastes for. This extends to goods that are rarely experienced, such as large purchases always chosen in discrete quantities, such as houses and spouses. Similarly, most people make an insurance decision only a small number of times in life; life insurance, for example, is particularly hard to learn about one's tastes for since it is rarely purchased and no-one redeems it more than once. Of course, insurance also involves risk preferences, which may themselves be difficult to learn, as noted in Section 3.4.

Some of the situations that we note are most problematic for learning preferences are the situations called out in Thaler and Sunstein (2008) as cases in which libertarian paternalism might improve outcomes. Such goods may or may not yield large welfare losses: since they appear rarely, even if the incorrect choice is made every time they appear, that occurs rarely. However, many of these are quite consequential.

We further note that one way in which paternalistic policies may improve

[8] We abstract away from the possibility that poverty creates cognitive stress, as in Mani et al (2013), which will make it harder to optimize and thus increase welfare loss, though we note that one way of modeling such stress is to relax the memory axiom, which we do in Section 3.1.

welfare is by forcing Alice to learn her own preferences. This could be literal paternalism (or maternalism), when Alice is a child. If Alice can be induced or forced to sample the various goods enough (including items she perceives as bads), she will reach full preference discovery, where $\boldsymbol{B}^t = \boldsymbol{B}$. When Alice was a child, her parents likely cajoled and coerced her to try many foods, multiple times each, to help her discover her tastes so that she could be more successful at maximizing her utility throughout her life.

If full relevant discovery occurs at some time T, then Alice should make utility-maximizing choices, just as in the traditional model of choice. We can see, though, that under many possible circumstances, it is likely that one or more of the perceived "bads" is actually a good. If this is the case, then choices made under full voluntary discovery will be stable and consistent so long as choice remains voluntary and the agent is not induced to try something perceived as a "bad." These stable, consistent choices nonetheless will not be efficient choices, and our agent will choose suboptimally in perpetuity, losing welfare as a result.

3 Extensions

The learning process presented above is very simple: Alice has a mistaken belief about her taste for a good until she has sufficient direct consumption experience of the good, after which she forever has exactly the correct belief about her taste for that good. In this section, we explore some extensions to make the model more realistic in ways that generate testable hypotheses for human behavior.

3.1 Relaxing Perfect Memory

In Section 2 we assumed that when we are not experiencing a good, our memory of the utility we get from it is perfect (Axiom 6). However, no-one would deem our story unrealistic if we wrote that Alice said, "I forgot how much I liked mangos!" Let us therefore assume that memory is imperfect, and specifically that there is a fundamental prior belief toward which believed preference parameters decay. We call this a "prior" because it is reasonable to think of it as the beliefs Alice first formed, and thus we denote it as $\boldsymbol{\beta}_{\boldsymbol{i}}^0$.

We retain all other axioms and modeling assumptions from Section 2. Moreover, we assume for simplicity that if Alice had learned (and partially forgotten) her preferences for good i (that is, if $i \in L^s$ for some time s prior to today), she will reacquire her true preferences $\boldsymbol{\beta_i}$ with a single additional experience of that good (even if it is consumed in a bundle with other goods).

Let us denote today as time t, and suppose that Alice's most recent experience of good i occurred at time $s < t$, at which point she correctly updated her preferences for this good: $\boldsymbol{\beta_i^{s+1}} = \boldsymbol{\beta_i}$.

To illustrate the forgetting-and-remembering scenario, we assume a decay process applies to each of the M_i parameters $\boldsymbol{\beta_{im}^t}$ of the agent's value function for good i. For example, we could assume that, for $m = 1, \ldots, M_i$:

$$\boldsymbol{\beta_{im}^t} = \boldsymbol{\beta_{im}^0} + e^{-\delta \cdot (t-s+1)} \cdot \left(\boldsymbol{\beta_{im}} - \boldsymbol{\beta_{im}^0} \right) \ .$$

This function satisfies the differential equation: $\boldsymbol{\beta_{im}'} = -\delta \cdot (\boldsymbol{\beta_{im}} - \boldsymbol{\beta_{im}^0})$. We call δ the *rate of decay* of the agent's preference memory.

If this is the case, then the result that the learned set L^t is absorbing no longer holds. Parameters for each good will vacillate between their updated values and their prior values. At any time t, Alice likely has incorrect beliefs about the parameters for all of the goods. The only exception is that goods that start in W^0 will retain stable parameters. Alice will be no more likely to mis-categorize a good as a bad than she would with persistent memory—again, this mis-categorization will only happen if goods not in W are in W^0. But at any finite time, she is more likely to lose welfare by choosing a bundle that is not optimal.

Since the decay process occurs over time, we can make the following inferences about welfare loss:

Proposition 4.

If the agent's memory of her preference parameters decay between experiences with a good, then the expected welfare loss from a consumption choice in time t is larger if:

(i) She is more forgetful, i.e., the decay rate δ is larger.

(ii) The good is rarer, i.e., it appears in choice baskets less often (q_i is smaller).

(iii) The true preference parameters are more unexpected, i.e., the true parameters $\boldsymbol{\beta_i}$ are farther from the prior parameters $\boldsymbol{\beta_i^0}$. □

Because memory decay implies an increased likelihood that parameters are wrong and therefore an increasingly large chance that bundles of goods are suboptimal as time progresses since last experience, forgetting increases the welfare loss Alice experiences, and that welfare loss is exacerbated by these factors. If forgetting is made more likely or more intense by factors in Alice's life like cognitive stress (including that caused by poverty, as discussed in Mani et al, 2013), this welfare loss will again be larger.

3.2 Relaxing Separable Learning

In Section 2, we also assumed that when Alice samples one good, she only updates her preference parameters for that good (Axiom 5). In reality, Alice forms her beliefs about her preference parameters for a good based on both experience with that good and thought experiments informed by experience with other goods. Thus, let us relax the assumption of separable learning to allow Alice to predict her tastes for an as-yet unlearned good based on her experiences with other goods.[9]

Instead of separable learning, we will allow Alice to believe that goods are similar to some other goods, and to update her believed parameters for one good based on experience with those similar goods. We are agnostic about how she judges the similarity between goods, but this judgment is based only on characteristics and information that are observable to her.

For simplicity, let us assume that her respective value functions $v_i(.)$ and $v_j(.)$ have the same functional form (with $M = M_i = M_j$ parameters), whereby each parameter β^t_{im} (for $m = 1, \ldots, M$) of good i is a constant multiple of the corresponding parameter for good j, that is:

$$\beta^t_{im} = \alpha^t_{ijm} \cdot \beta^t_{jm} \tag{2}$$

The new parameter α^t_{ijm} is the level of "correlation" that Alice believes at time t to exist between good i and good j for parameters $m = 1, \ldots, M$. The parameter vectors α_{ij} are updated by experience in line with Axiom 6 and Axiom 7. This does not require that all goods are connected with each other, but it does allow for a case where the agent believes that more than two goods are correlated.

[9] These thought experiments might also be the process by which she generates prior beliefs before any goods are tried.

For example, Alice may have never tasted an Asian pear, but may conjecture that the Asian pear is very like the Bartlett pear and a little like an apple, and thus may have an α between the Bartlett and Asian pears that is close to one and an α between the apple and the Asian pear that is far from one, so that she believes she ranks the Asian pear close to the Bartlett but not to the apple.

Following Axiom 7, Alice updates her good i preference parameters $\boldsymbol{\beta_i^t}$ at time t to $\boldsymbol{\beta_i}$ when she has had sufficient experiences with this good. If good $j \notin L^t$ is connected to good i through parameter vector α_{ij}, then $\boldsymbol{\beta_j^t}$ will be updated at time t according to Equation (2). Conversely, once Alice learns her true preference parameters for good j as well, she will update $\boldsymbol{\beta_j^t}$ to $\boldsymbol{\beta_j}$, and as a result find her true correlation parameters α_{ijm} as $\alpha_{ijm} = \beta_{im}/\beta_{jm}$ for all $m = 1, \ldots, M$.

Thus, before Alice ever tastes an Asian pear, she updates the parameters in her Asian pear value function as she learns her tastes for Bartlett pears and apples. If she finally does taste an Asian pear, she will then know exactly how like it is to a Bartlett pear and to an apple and adjust those "correlation" parameters accordingly (although they are useless to her now that she knows her values for each relevant good).

Learning spillovers of this type can make it easier to learn preferences in the sense that they may make the naïve guess about the value of an as-yet-unexperienced good more accurate (i.e. $\boldsymbol{\beta_i^t}$ is closer to $\boldsymbol{\beta_i}$ than is $\boldsymbol{\beta_i^0}$). Since we noted above that welfare loss is greater the farther the parameters are from their true values, we can say that learning spillovers may reduce welfare loss from misunderstood preferences. However, it need not make the discovery process occur any faster: we have assumed that that takes the number of tries required to mathematically identify the parameters of the value function regardless of how much the true parameters differ from their believed values. Further, this type of learning spillover reduces welfare loss only if the believed similarities between goods are relatively accurate.

Learning spillovers also mean that W^t is no longer an absorbing set, although L^t still is. If some good i (like bananas, in our example in the last section) is mistakenly characterized as a bad, e.g. because of incorrect priors, then if Alice believes that good's parameters are correlated with those of another good that she does consume (perhaps a plantain), she may be able to update good i's parameters in a way that moves it out of W^t. Once it moves out of W^t, she may try it and learn its parameters precisely.

However, by the same token, learning spillovers may cause Alice to relegate to W^t an untried good i that she would actually like. This would happen if she tried a good x_j that she believed was correlated with good i and found that good j belonged in W. This is possible if Alice is wrong about the correlations between the goods. For example, Alice believe plantains to be strongly correlated with bananas, and may categorize both outside of W^s for times $s < t$. She may then try plantains in time t and find that they belong in W and demote them there. She may not realize that bananas are much sweeter, and thus she may demote bananas as well. In this way, the result that all goods that start outside of W^0 are discovered (Lemma 3) no longer holds; even though her initial impression of bananas was not bad, she may end up mischaracterizing them for her whole life without ever trying them.

3.3 Uncertain Preference Beliefs: the Agent Knows She Doesn't Know

In our main model, we assumed that Alice had point estimate beliefs of her preference parameters. As a result, if she had not yet learned her preference parameters for a good, she did not know that; in other words, she had no idea whether a good was in L^t or not.

Instead, we generally know whether we've tried something. Let us therefore assume that for any good $i \notin L^t$, Alice believes that her parameter vector $\boldsymbol{\beta_i^t}$ is an M_i-dimensional random variable with joint probability density function $f_i^t\left(\boldsymbol{\beta_{i1}^t}, \boldsymbol{\beta_{i2}^t}, \ldots, \boldsymbol{\beta_{iM_i}^t}\right)$.

The learning process in this case lets her find the precise point value of each of her parameters. Distributional beliefs like this yield two differences from our main results.

First, Alice will sample some goods with negative expected marginal utility implied by her point estimate of their preference parameters. This is because some such goods will have a distribution of possible parameters with a resulting distribution of possible levels of marginal utility that extends into the positive region. If Alice has an intertemporal utility function that allows her to consider possible future utility gains, then she can benefit from "experimental consumption" of the good (Kihlstrom et al, 1984). If she does not like the good, she can avoid it for the rest of her life, but there is a possibility that it will afford her a stream of positive utility into the future. Depending on her belief about the distribution of possible values, her discount rate, how often the good appears, and how long she expects to live, she may choose to consume the good.

Second, as a result of this, fewer goods will be forever-incorrectly-categorized in W^t. However, it is still possible for a good to be incorrectly assigned to W^t forever. This is more likely to occur the lower her prior belief of the utility the goods will afford, the narrower the distribution of possible utilities she thinks it can afford, the greater her discount rate, the rarer the good, and the shorter her expected remaining life (she might think "life's too short" to risk trying the good).

Thus, a sophisticated agent that knows what she doesn't know may be less likely to lose welfare from suboptimal choices caused by incompletely learned preferences, but some welfare loss is still likely. As she gets older, she refines her perception of more and more goods for which she is willing to undertake experimental consumption. By the time she is quite experienced, any goods that she has not yet tested are goods for which her believed distribution looks quite unfavorable; by the same token, if she is forced to consume one of these goods, she is more likely to receive a positive than a negative surprise.

3.4 Preferences for Lotteries

So far we have assumed that all goods are deterministic. How will preference discovery change when stochasticity is a feature of the goods?

Let's say that Alice is in a position to choose between lotteries over goods, where lottery i offers quantity x_{i_1} of a good with probability ρ_{i_1} and quantity x_{i_2} with probability ρ_{i_2}. (We start with a story that revolves around a lottery of different possible quantities of the good, and turn in a moment to a parallel story of a lottery over different qualities of a good.)

If Alice gets utility from a lottery according to expected utility theory, then if Alice knows the possible lottery outcomes and probabilities,[10] her preferences at time t over lotteries over quantities x_i can be inferred immediately from her value function for $v_i(x_i|\boldsymbol{\beta_i^t})$ for that good. Thus, if Alice has already experienced different quantities of x_i deterministically and has discovered $\boldsymbol{\beta_i}$, she also fully understands her own risk preferences over that good. For example, if she has already eaten one orange and two oranges, she

[10] Again, here we distinguish between "institutional learning" and "value learning" (Braga and Starmer, 2005), where learning the probabilities is a part of institutional learning that has been addressed in the literature (see, e.g., the two-armed bandit problem in Rothschild, 1974).

knows exactly how happy a lottery with 50-50 odds of getting one or two oranges will make her.

Learning one's tastes for risk might be more difficult than that, though. In each case, the fact that tastes for risk are harder to learn than tastes for deterministic outcomes means that it is more likely that Alice has incorrect beliefs about her tastes for risk, and as a result she's more likely to lose utility by making suboptimal choices. Some of these biases are pessimistic (believing she dislikes risk more than she actually does) like the biases that result from our basic model, but not all of them are.

First, if Alice has a different utility function for risk than for certainty, as Andreoni and Sprenger (2012) suggest, then she must commit the whole learning process over again for risk even if her entire utility function is parameterized for certainty.

Second, if Alice has the same utility function for risk as for certainty but derives utility according to cumulative prospect theory, then there is another parameter for her to learn that is separate from her value function parameters: the probability weighting parameter. As a result, no amount of experience with certain outcomes will let Alice learn her risk preferences. In other words, choice under risk is harder in this case because there is an additional parameter to learn.

In either of the two preceding cases, since there is effectively more to learn about tastes for risk, the implication is that Alice will have unlearned preferences for a longer short-term period during the learning process, and since there are more parameters to mis-estimate, these parameters may also lead her to mischaracterize some lotteries as belonging in W^t and thus more long-term welfare loss may ensue as well.

A third bias may arise if some goods vary by quality. We represent that as a case in which a good appears only as an outcome of a lottery over quality levels of the good. In this case, Alice may never learn her taste for the high quality versions of the good if she *a priori* ranks the quality lottery relatively low and the lower quality outcomes turn out to be unpleasant. Imagine she knows that some bananas are sweet and tasty and others are mushy and terrible, and she knows she has 50-50 odds of getting each kind of experience when she picks up a banana. She might not initially rank the lottery banana as a member of W^0 and thus she might be willing to eat an uncertain banana. However, if the mushy banana tastes worse than she thought, an experience with a mushy banana might cause her to demote the lottery banana into W^t. If she also believes the good banana tastes less good than it actually does, this is another route to welfare loss: she might unknowingly really love good bananas, and thus her true value for the lottery might be high, but her bad banana experience may cause her to refuse to eat any bananas again.[11]

Further, one might conjecture that stochasticity interacts with other features of interest. In particular, if memory persistence fails, then it might be hard for Alice to mentally construct the utility she derives from a lottery.

4 Conclusion

The idea that an agent like Alice might not know all of her tastes perfectly at birth but might have to learn them through experience is an intuitive one and related ideas have been explored in psychology. Plott (1996) introduced the idea of preference discovery into economics, but it has never been formalized. We have traced out a general theory of preference discovery that justifies the issue as worthy of economists' attention by demonstrating the potential for a persistent, non-self-correcting loss of welfare. We have further shown that some cases should yield particularly large welfare loss: people early in their consumption lives, goods rarely experienced, goods that are quite unfamiliar or unexpected, and people who forget their tastes.

Experimental studies can confirm features of the learning and perhaps the forgetting process with a combination of no-experience trials of goods, mandatory experience, and later choice opportunities. The implications of discovered preferences for applied microeconomics remain to be explored, but our model's predictions indicate that cross-sectional data may include observations in which people are making choices over goods for which their preferences have not been discovered. Panel data exhibiting choice reversals can be interpreted in a new light: in addition to error and experimental consumption, the process of preference discovery is an additional explanation for erratic choice behavior. Moreover, an appearance of consistent choice need not imply utility maximization. Since nearly all of public economics depends on an assumption that a person's welfare is best inferred from what she chooses for herself, the implications are not inconsiderable. The extent to which preference discovery affects economic predictions and conclusions should vary across choice domains in predictable ways.

[11] Our discussion assumes that she can tell whether she the banana she tasted was a good or a bad banana; if she can't tell, then the potential for error is even larger.

If confirmed, preference discovery would lend credence to common parenting techniques: requiring a child to try a new food at least three times before declaring it unpalatable, for example. It might also advise gently paternalistic policies that nudge people to try goods they might believe to be undesirable, particularly if there are reasons to believe people's priors are mistaken. Indeed, this may be why many public health, awareness, and advertising campaigns urge consumers: "Try it: you'll like it!"

We have used discovery of preferences for foods as a discursive example throughout, but we should be careful not to trivialize the implications of our model. The model we present describes preference discovery for all consumption experiences, including the most important choices in people's lives. In many of these weighty matters, it is very unlikely that people have enough experience to know what they like, including choice of a college major, profession, home, or even a spouse or partner. Because uncertainty can exacerbate the learning problems, particularly for inexperienced people, we might see people taking on too much undesirable risk (e.g. credit card debt) and too little beneficial risk (e.g. entrepreneurship).

Even so, caution must be taken in translating results from this theory and work that follows it into policy advice. While our results show that people may not always know what's best for them, we take this to mean that gentle and "libertarian" paternalism (Thaler and Sunstein, 2008) might provide some benefits and that policy could find ways to encourage people to learn their tastes and fight their own biases, rather than that the state should take a more strongly paternalistic role in individuals' lives.

Acknowledgements

For helpful comments, we thank Yongsheng Xu, Annemie Maertens, and participants at FUR 2012, SABE/IAREP/ICABEEP 2013, and seminars at Williams College and George Mason University, and we particularly thank CeMENT 2014 participants Brit Grosskopf, Muriel Niederle, J. Aislinn Bohren, Angela de Oliveira, Jessica Hoel, and Jian Li for detailed feedback. We gratefully acknowledge funding from the Georgia State University Center for Economic Analysis of Risk.

5 References

References

Andreoni J, Sprenger C (2012) Risk preferences are not time preferences. American Economic Review 102(7):3357–76, DOI 10.1257/aer.102.7.3357, URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.102.7.3357

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

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

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

Kihlstrom RE, Mirman LJ, Postlewaite A (1984) Experimental Consumption and the 'Rothschild Effect.'., Studies in Bayesian Econometrics, vol. 5. New York; Amsterdam and Oxford: North-Holland; distributed in U.S. and Canada by Elsevier Science, New York, pp 279 − 302

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

Levitt SD (2016) Heads or tails: The impact of a coin toss on major life decisions and subsequent happiness. Working Paper 22487, National Bureau of Economic Research, DOI 10.3386/w22487, URL http://www.nber.org/papers/w22487

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

Mani A, Mullainathan S, Shafir E, Zhao J (2013) Poverty impedes cognitive function. Science 341(6149):976–980, DOI 10.1126/science.1238041

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

Rothschild M (1974) A two-armed bandit theory of market pricing. Journal of Economic Theory 9(2):185–202

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, http://cdp.sagepub.com/content/14/3/131.full.pdf+html

A Proofs

A.1 Proof of Lemma 1

Let $i \in W^t$. Then the solution to the agent's optimization problem specifies that $x_i = 0$ at time t. Hence, $i \notin Z^t$. □

A.2 Proof of Lemma 2

(a) Let $t \geq 0$ and let $i \in \{1, \ldots, N\} \setminus Z^t$. Axiom 6 states that $\boldsymbol{\beta}_{\boldsymbol{i}}^{t+1} = \boldsymbol{\beta}_{\boldsymbol{i}}^{t}$ unless the agent has a relevant experience with any good at time t that would lead to an updating of her preferences for good i. Axiom 5 implies that any relevant experience the agent has of any *other* good will not impact $\boldsymbol{\beta}_{\boldsymbol{i}}^{t+1}$. And since $i \notin Z^t$, the agent cannot have any relevant experience of good i at time t that would cause her to update her preferences, either. Hence, $\boldsymbol{\beta}_{\boldsymbol{i}}^{t+1} = \boldsymbol{\beta}_{\boldsymbol{i}}^{t}$.

(b) Let $i \in W^s$ for some time $s \geq 0$. Then, by Lemma 1, $i \notin Z^s$. It follows by Lemma 2(a) that $\boldsymbol{\beta}_{\boldsymbol{i}}^{s+1} = \boldsymbol{\beta}_{\boldsymbol{i}}^{s}$. Therefore, if $i \in W^s$, we also have $i \in W^{s+1}$. And thus, by induction, $i \in W^u$ for any $u \geq s$.

(c) Let $s \geq 0$ and let $i \in L^s$. Axiom 6 states that the agent's perceived preferences of good i can only change through relevant experiences with this or any other good. However, if $i \in L^s$, then the agent already knows her true preferences for good i and no experience at time s will make her change her mind. Therefore, $\boldsymbol{\beta}_{\boldsymbol{i}}^{s+1} = \boldsymbol{\beta}_{\boldsymbol{i}}^{s} = \boldsymbol{\beta}_{\boldsymbol{i}}$, and thus $i \in L^{s+1}$. By induction, therefore $i \in L^u$ for all $u \geq s$. □

A.3 Proof of Lemma 3

We first show that if $i \notin W^0$ and $i \in W^t$ for some time $t > 0$, then $i \in L^t$. To prove this, let us assume that $i \in \{1, \ldots, N\} \setminus W^0$, and that there is a finite time $t > 0$ such that $i \in W^t$. Then there has to be a time $s \in \{0, \ldots, t-1\}$ such that $i \notin W^s$ but $i \in W^{s+1}$. Therefore, $\boldsymbol{\beta}_{\boldsymbol{i}}^{s+1} \neq \boldsymbol{\beta}_{\boldsymbol{i}}^{s}$. Axiom 6 states that this change is the result of a relevant experience the agent had regarding a good at time s, and according to Axiom 5 would have to be of good i. In fact, we know from Lemma 2(a) that $i \in Z^s$. Axiom 7 states that experiencing good i leads to a change from $\boldsymbol{\beta}_{\boldsymbol{i}}^{s}$ to $\boldsymbol{\beta}_{\boldsymbol{i}}^{s+1}$ only if her time-s experience of good i is sufficient for her to discover her true preferences, in which case $\boldsymbol{\beta}_{\boldsymbol{i}}^{s+1} = \boldsymbol{\beta}_{\boldsymbol{i}}$, and therefore $i \in L^{s+1}$. Finally, Lemma 2(c) ensures that $i \in L^t$ since $t > s$.

Since items appear in choice sets independently of time and of each other, and with a probability (q_i) that is strictly between 0 and 1, and since there is a finite number of goods (N), good i has a positive probability of being the only item in the agent's time-t choice set for any $t \geq 0$, namely

$$r_i := \mathbb{P}\left(G^t = \{i\}\right) = q_i \cdot \prod_{j \neq i} (1 - q_j) > 0\,.$$

If N_i^t denotes the number of times that good i has appeared alone in the agent's choice set prior to time t, then N_i is a random variable that follows a binomial distribution with parameters t and r_i. In particular, $\mathbb{P}\left(N_i^t \geq M_i\right) \to 1$ as $t \to \infty$.

The result from the first part of the proof ensures that for any time $t \geq 0$, if $i \notin W^0$ and if $i \notin L^t$ yet, then $i \notin W^t$. That is, the agent will choose to consume (a positive amount of) good i each time good i appears alone in the agent's choice set. The second part of the proof ensures that with probability 1 there exists a time t such that this event will have occurred M_i times prior to time t. Combining the two statements, we see that with probability 1 the agent will have had M_i experiences of good i prior to time t, and therefore—by Axiom 7 and our earlier discussion—$i \in L^t$. Thus, as $t \to \infty$, $\mathbb{P}\left(i \in L^t\right) \to 1$. □

A.4 Proof of Proposition 1

Recall that

$$\Delta u^t = \max_{x_i (i \in G_t)} u\left(x_1, \ldots, x_N \mid \boldsymbol{B}\right) - u\left(x_1^*, \ldots, x_N^* \mid \boldsymbol{B}\right)\,,$$

subject to Equation (1), and with $x_1^*, \dots, x_N^*$ as the solutions to the agent's time-t optimization problem, that is based on parameter specifications $\boldsymbol{B}^t$.

Clearly, $\Delta u^t \geq 0$. Moreover, Axiom 1 and Axiom 3 together ensure that the solution to the constrained optimization problem

$$\max_{x_i (i \in G_t)} u\left(x_1, \dots, x_N \mid \boldsymbol{B}\right) \quad \text{s.t. Equation (1)}$$

exists and is unique. Therefore, $\Delta u^t = 0$ if and only if the agent chooses the same quantity x_i for each $i \in G^t$ under parameters $\boldsymbol{B}^t$ as under parameters $\boldsymbol{B}$. This requires that $\boldsymbol{\beta_i^t} = \boldsymbol{\beta_i}$ for any good i of which the agent would choose a positive amount x_i under parameters $\boldsymbol{B}$; and that the agent chooses $x_i = 0$ under both $\boldsymbol{B}$ and $\boldsymbol{B}^t$ for all other goods i.

Since G^t is random and could equal any subset of $\{1, \dots, N\}$ with a positive probability, the agent would only choose $x_i = 0$ for certain under $\boldsymbol{B}$ if $i \in W$. Therefore, to ensure that $\Delta u^t = 0$ (irrespective of G^t and m^t), we require that $W^t = W$, and that for all goods $i \in \{1, \dots, N\} \setminus W$, $\boldsymbol{\beta_i^t} = \boldsymbol{\beta_i}$. That is, the agent has achieved full relevant discovery by time t.

Conversely, if the agent has achieved full relevant discovery at time t, then $\boldsymbol{\beta_i^t} = \boldsymbol{\beta_i}$ for all $i \in \{1, \dots, N\} \setminus W$ and $W^t = W$. Therefore, the agent will make the same optimal consumption choices under $\boldsymbol{B}^t$ as under $\boldsymbol{B}$. Hence, $\Delta u^t = 0$. □

A.5 Proof of Proposition 2

(i) Lemma 3 states that the agent will eventually learn her preferences for any good $i \in \{1, \dots, N\} \setminus W^0$. That is, with probability 1, there exists $t \geq 0$ such that $\{1, \dots, N\} \setminus W^0 \subseteq L_t$. Moreover, by Lemma 2(b), $W^0 \subseteq W^t$. Therefore: $\{1, \dots, N\} \setminus W^0 \supseteq \{1, \dots, N\} \setminus W^t$, which implies $L_t \supseteq \{1, \dots, N\} \setminus W^t$. That is, as $t \to \infty$, the agent will achieve full voluntary discovery.

(ii) We prove the equivalence of the two expressions separately in each direction.

First, suppose that $W^0 \subseteq W$. Let $i \in \{1, \dots, N\} \setminus W$. Thus, $i \notin W_0$, which implies (by Lemma 3) that $\lim_{t \to \infty} \mathbb{P}\left(i \in L^t\right) = 1$. That is, there exists (with probability 1) a time $t \geq 0$ such that $L^t \supseteq \{1, \dots, N\} \setminus W$. Moreover, we have to show that there exists (with probability 1) a time $t \geq 0$ such that $W^t = W$.

We already know that exists $t \geq 0$ (with probability 1) such that for any $i \notin W$, $i \in L^t$. That is, $\boldsymbol{\beta}_{\boldsymbol{i}}^t = \boldsymbol{\beta}_{\boldsymbol{i}}$, and thus if $i \notin W$, then also $i \notin W^t$. This implies that $W^t \subseteq W$. Similarly, if we choose any $i \in W$, we either have $i \in W^0$, in which case $i \in W^t$ for all $t \geq 0$, or $i \in W \setminus W^0$. In the latter case, since $i \notin W^0$, Lemma 3 assures that there exists (with probability 1) a time $t \geq 0$ such that $i \in L^t$. Thus, $\boldsymbol{\beta}_{\boldsymbol{i}}^t = \boldsymbol{\beta}_{\boldsymbol{i}}$, and since $i \in W$, we must also have $i \in W^t$. This proves that $W \subseteq W^t$, which—together with our earlier result—ensures that $W = W^t$, and therefore there exists a time $t \geq 0$ with probability 1 so that the agent achieves full relevant discovery by time t. This completes the first direction of the proof.

Conversely, suppose that $W^0 \not\subseteq W$. Then there exists a good $i \in W^0 \setminus W$. Since $i \in W^0$, Lemma 2(b) implies that $i \in W^t$ for all $t \geq 0$. Thus, for all $t \geq 0$, $W^t \neq W$.

(iii) If $W^0 = \varnothing$, then $\lim_{t\to\infty} \mathbb{P}(i \in L^t) = 1$ for all $i \in \{1, \dots, N\}$, according to Lemma 3. This implies that the agent will (eventually) achieve full discovery of her preferences for all goods. □

A.6 Proof of Proposition 3

(i) The more time passes, the more consumption opportunities will include good i in the available basked. As a result, at later times the agent is more likely to have experienced a good sufficiently to learn her true preferences.

(ii) As discussed above, the solution to the agent's optimization problem requires that she equates the (decreasing) marginal utility per dollar for each good to the (decreasing) marginal utility of money, when possible. On the other hand, she does not consume any good that has a lower marginal utility even for its first unit. A larger value of m^t reduces the marginal utility of money, which may make it optimal for the agent to consume a good that she had not consumed if she had less money at her disposal.

(iii) The more frequently the goods are available for the agent's consumption, the sooner she will learn her preferences for the goods. □

A.7 Proof of Proposition 4

(i) The larger the δ, the larger the decay from the true value in each period since experiencing the good.

(ii) The rarer the good, the less frequently it is experienced, and therefore at any given time t the expected amount of time since the good was last experienced $E(t - s)$ is larger.

(iii) The larger the difference between $\boldsymbol{\beta_i}$ and $\boldsymbol{\beta_i^0}$, the larger the size of the decay in any discrete time period, and therefore a larger initial gap will yield the same result as a larger δ. □