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*

December 2018

Abstract

Is the assumption that people automatically know their own preferences innocuous? We present a theory and an experiment that study the limits of preference discovery. Our theory shows 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, expensive, must be bought in large quantities, or are initially judged negatively relative to other items. Preferences will eventually stabilize, but they need not stabilize at true preferences. A pessimistic bias about untried goods 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.

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

Results from an online experiment show that learning occurs as predicted and with the expected biases, but with even more error than our theory suggests. 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
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

Do you know what you like? Neoclassical microeconomic choice theory is grounded in the assumption that people choose according to a stable ranking that represents their true preferences. 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 not fully discovered, people may make choices that don't maximize utility. If this is the case, some standard results of neoclassical microeconomic theory come into question. In this paper, we start from an assumption that preferences must be learned from experience, and, using a model of preference discovery and an experiment simulating preference learning, we explore implications for choice patterns and welfare, focusing on the extensive margin: what is and is not learned.

Consider an encounter with a new food. For example, one of this paper's authors had not eaten celeriac until a few years ago. She *a priori* believed it untasty. When she tried it, she discovered that she likes it. Experience yielded a more accurate assessment of her preferences, and she now enjoys a more efficient level of celeriac consumption. Still, because of her initial misperception, she might have missed out on a lifetime of celeriac appreciation had she not been induced to try it—indeed, she is likely missing out on other delicious vegetables due to mistaken beliefs and a lack of experience. Our model and experiment show that the need to learn preferences through experience can generally cause persistent welfare loss.

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, so our work is informed by past studies from both fields. Preference discovery has been little studied in either field because psychological models often do not feature stable underlying preferences (Ariely et al, 2003; Lichtenstein and Slovic, 2006), while models in economics typically implicitly assume stable preferences that are known to the decision-maker.[1]

Preferences might not need to be discovered through experience if people can simply predict what they will like. As discussed by Kahneman et al (1997), Scitovsky (1976) argued that people are bad at predicting their 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 their own tastes, and Wilson and Gilbert (2005) review extensive evidence showing systematic errors in forecasting happiness.

A few papers have studied preference discovery from an economic perspective, but they all focus on the intensive margin (the updating process) while we focus on the extensive margin (what is and is not learned). Several theoretical studies explore the process by which people will sample consumption items if they must learn them from experience, including Easley and Kiefer (1988), Aghion et al (1991), Keller and Rady (1999), Piermont et al (2016), and Cooke (2017). However, these all focus on the experimentation and updating process, and each either includes assumptions that ensure full learning by making learning effectively costless, or does not focus on the completeness of learning.[2] Armantier et al (2016) use theory and a lab experiment to study preference discovery as well, and in that sense their study is closest to ours. However, they, too, focus on the experimentation and updating process, testing different theories of learning, and do not consider the potential incompleteness of learning.

Plott (1996) noted that feedback should help the learning process, so we can find suggestive evidence about preference discovery in lab experiments that demonstrate unstable choices that are ameliorated over time by feedback. For example, van de Kuilen and Wakker (2006) find that repeated

[1] We distinguish between learning about objective circumstances and learning 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).

[2] Brezzi and Lai (2000) show, in another theoretical study, that learning when facing a multiple-armed bandit will be incomplete, but it is for a different reason (discounting) than what we study.

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.

Other economic experiments on repeated choice with feedback provide further suggestive evidence of preference discovery. Preference instability may be a marker of preference discovery, and there is a large literature debating the importance and interpretation of "preference reversals," e.g., Cox and Grether (1996). Noussair et al (2004) find that with repeated choice, people can converge to a true induced value. Similarly, errors and biases often decline with repeated choice, as observed in the gap between willingness-to-pay and willingness-to-accept, non-dominant bidding behavior, and strategic games (Coursey et al, 1987; Shogren et al, 1994, 2001; List, 2003).

By the same token, preference stability over longer periods is sometimes taken as evidence that discovery is not happening, though we argue it should not be. While studies like Eckel et al (2009) show that preferences are affected by outside conditions (mediated by psychological affect), other studies (including Andersen et al, 2008 and Dasgupta et al, 2017) look over longer time periods and find some evidence of stability and some evidence that preferences depend on conditions in predictable ways.[3] However, our theory shows that eventual stability in choices is expected even with preference discovery, and need not indicate fully discovered preferences: if you stop trying new things, you stop learning and your choice behavior appears stable.

Our contribution is to develop and test 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 as a result she will intentionally sample some goods with the goal of learning. However, learning has an opportunity cost, and since the benefits of learning are finite, learning will not be complete. We show that in finite time, the agent will exhibit choice reversals as she learns her preferences. We also demonstrate that she will

[3]Chuang and Schechter (2015) find, in developing country contexts, very little stability in preferences within a person over years, except in survey measures of self-reported social preferences. However, their interpretation is that the experimental measures they study are not good measures of preferences in these contexts.

learn her preferences over time for many goods, thus reducing her welfare loss from bad choices. However, she may not fully learn her preferences. Learning failure is more likely for people who are impatient, risk averse, have low income, or have a short lifespan, and for preference objects that are initially undervalued relative to other goods, expensive, rare, or that require a large minimum purchase. 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. Less formally, we discuss that as the agent lives and learns, she should become more pessimistic, since optimistic errors in prior beliefs are more likely to be corrected than pessimistic errors.

In an online experiment, we directly test most of our theoretical predictions, finding support for all that we test. There is ample evidence that subjects are intentionally learning, but the learning process is fraught with error. This suggests that practical considerations beyond our model would make preference learning even more difficult, thus amplifying the concerns we raise about unstable preferences and welfare loss. Nevertheless, we find support for our main predictions about tastes that remain unlearned and welfare loss that declines but not to zero. We also observe the biases we expected: subjects retain pessimistic errors as life progresses, and fail to try those goods exhibiting characteristics predicted by the model. Our results also show that unlearned goods and welfare loss appears to be persistent, as we predict.

This paper proceeds as follows. First, we informally outline our model's setup and results. (The full workings of the theory are contained in appendices.) We then describe the experiment design. Our experiment results follow, and we conclude.

2 A Model of Preference Discovery

We propose a model in which a sophisticated agent, whom we call Alice,[4] is born not knowing her preferences but must learn them through experience. Our goal is to study not the updating process but the extensive margin: when goods are tried and, by extension, in what cases preferences are never learned. We set up a model that makes it as easy as possible for Alice to learn her

[4] By assonantal coincidence, Cooke (2017) also calls his preference-discovering agent Alice.

preferences. We do this to show that, if we also make reasonable assumptions that create an opportunity cost for trying new things, incomplete learning is the result. This section outlines the assumptions and results of the model. The formal model is set up in Appendix A, with results in Appendix B and proofs in Appendix C.

Alice is a rational expected utility maximizer. She consumes goods that give her utility according to a utility function. Her inborn preferences are stable, but until she has learned her preferences for each good, she doesn't know precisely how consumption of that good translates into her utility. We operationalize this by assuming she has some probability distribution for the parameter value for each good. However, we assume that when she consumes a bundle, she instantly learns how much utility she got from each good she consumes in a meaningful quantity (a concept we will explain more fully below).

For convenience, Alice has an additively separable well-behaved utility function with one parameter per good, and we assume that the goods are non-stochastic in the sense that each gives her deterministic utility. We also assume that once she learns preferences, she never forgets them. More complicated goods, utility functions, or temporal patterns should simply make it harder to learn preferences. Since our goal is to give learning its best shot so we can highlight its failures, this simplest model makes our point most strongly.

The preference learning process proceeds as Alice tries goods throughout her life. There is a finite universe of goods; in each time period, each good is available with some fixed probability and there is an outside option that is always available. We assume Alice has some fixed income each period and prices are constant over time.

Thus, in each period, Alice chooses the bundle that maximizes the present value of her stream of expected utility. If she is myopic, she will choose a bundle that maximizes current-period expected utility; if she is forward-looking, she is willing to sacrifice a finite amount of current period expected utility to try an as-yet-untried good that would not be in the myopic bundle if the potential gain from learning preferences is high enough. The cost of doing so is the best utility she could have otherwise gotten from spending the money she's using to purchase the as-yet-untried good. The benefit is the potential stream of utility from consuming the new good in the future if it turns out to be good; if it does not turn out to be good, she need never consume it again, so this stream is positive-biased. As a result, Alice might choose to consume enough of a good to learn her taste for it even if her prior belief is that, in expectation, it is not the best use of her resources. This is experimental consumption (Kihlstrom et al, 1984).

One more assumption is crucial to the results that follow. To learn her taste for a good, Alice must consume a non-vanishingly small quantity of that good. If she could learn from consuming any arbitrarily small quantity of a good, the opportunity cost to trying a new good would be approximately zero, because she could sample such a small quantity that the cost would be vanishing so that *any* positive benefit would justify it. But in reality, having an atom of an apple on your tongue does not teach you your taste for apples, so we assume that the minimum sampling for a meaningful learning experience is finite.

Built from these assumptions, our theory yields the results in Table 1.

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 mean prior that is 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 noted in 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 negative.

As we note in Table 1, our experiment tests most of the predictions of the theory. We did not intend to test the obvious Proposition 1 (unaffordable goods are not tried), and we did not include elicitations of risk and intertemporal preferences so we did not test Propositions 3a and 3d about the influence of those kinds of preferences. We did not vary prices, so did not test Proposition 3h (that more expensive goods are less likely to be tried), as this should follow from the same process as Proposition 3c (that agents with lower income are less likely to try a good). However, we tested and found support for all of the other results of our model.

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,

Table 1: Theory Results

Theory Result	Description	Experimental Result
Proposition 1	Unaffordable goods are never learned	*not tested*
Proposition 2a	Goods with priors better than outside option are eventually learned	(partially tested in tests of 3a-k)
Proposition 2b	A good will never be tried if the opportunity cost of a meaningful taste outweighs the expected gain from future optimized consumption	(partially tested in tests of 3a-k)
Proposition 3	A good is less likely to be learned if:	
Proposition 3a	The agent is more impatient	*not tested*
Proposition 3b	The agent has a shorter lifetime	Some support Sec. 4.3
Proposition 3c	The agent has a smaller income (for normal goods)	Supported Sec. 4.3
Proposition 3d	The agent is more risk averse	*not tested*
Proposition 3e	The good's mean prior is low	Supported Sec. 4.3
Proposition 3f	Given a low prior mean, the agent's belief is less diffuse	Supported Sec. 4.3
Proposition 3g	Given a high prior mean, the agent's belief is more diffuse	Supported Sec. 4.3
Proposition 3h	The good is more expensive	*not tested*
Proposition 3i	The minimum "nibble" size to learn the good is larger	*not tested*
Proposition 3j	The good appears less frequently	Supported Sec. 4.3
Proposition 3k	Other goods seem more attractive	Supported Sec. 4.3
Proposition 4	Preferences eventually become stable	Supported Sec. 4.4
Informal result	The average parameter belief error becomes negative (pessimistic) over time	Supported Sec. 4.4
Proposition 5a	Choice reversals occur	Supported Sec. 4.2
Proposition 5b	The rate of choice reversals declines to zero over time	Supported Sec. 4.2
Proposition 6a	Unlearned preferences may cause welfare loss	Supported Sec. 4.5
Proposition 6b	Welfare loss weakly declines over time	Supported Sec. 4.5
Proposition 6c	Welfare loss need not approach zero as time passes	Supported Sec. 4.5

so Proposition 6c is in that way our most important result. Also, Propositions 4 (that preferences 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. 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.

3 Experiment Design

We present an experiment in which individuals face a decision environment based on our model above. The experiment tests the extensive margin of preference learning (what is and is not learned) using induced values for fictitious goods instead of actual consumption of goods. We do this to avoid satiation, to ensure subjects are at the same level of preference learning at the start of the experiment, and because homegrown preferences for actual goods will vary significantly across people, be complicated by preferences for moderation and the potential for variation in access to complementary and substitute goods, and be difficult to observe, thus limiting our ability to test the model's precise predictions.

In the experiment, the subject plays through a series of T rounds. In each round t, she has a budget y to spend and is confronted with a basket of available goods, which are randomly chosen from the universe of N goods: each good i appears in each round with probability q_i. She has an induced utility function that is converted to dollars to determine her experiment earnings. The utility function has fixed parameters. The subject starts out not knowing these parameters but is given noisy guesses about them. Each guess is updated to the true value when she has sufficient experience with that good.

Specifically, her utility is linear in the goods:
$u(x_1, x_2, ..., x_N) = zx_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_N x_N$.

The values $\hat{\beta}_i$ for the goods are randomly chosen for each subject, and they remain fixed for that subject for all rounds. There is a numeraire good x_1 that is available in all rounds. It gives a known return z and costs 1 per unit. Half of the non-numeraire goods appear with low probability and the rest with high probability. The goods have fixed prices $p_i = 1$ and she has a fixed income y. She cannot save or borrow across rounds. A nibble (minimum meaningful consumption experience) is $m_i = 1$ for all goods.

While she makes her decision in each round t, she sees her true or guessed value β_i^t for each available good. When the experiment starts, these are the priors (guessed values) we assign to her, and as she learns values over the course of the experiment, priors are replaced with true values. We generate each prior by adding an independent random disturbance to the true value. For each subject and each good, the random disturbance is drawn from a discrete uniform distribution over $[-\sigma, +\sigma]$. We call these "starting guesses" and tell the subject that each starting guess value is her true value plus a positive or negative random number, so that it is related to, but generally not the same as, the true value.

In each round, from the set of available goods, the subject must choose a bundle that costs y or less. This decision is time-limited by our software: if she does not choose an affordable bundle within a minute then she consumes zero of all goods, earning zero for the round. After the round, the software tells her what her total utility is in that round and reminds her what bundle she chose. For each good, it also tells her what its value or her guess of its value is. The software automatically updates with the correct value each good of which she consumed at least m_i in that round. Since we do not seek to study the subjects' ability to infer parameters of multivariate functions but, rather, whether and when some goods will be tried, our design reduces the "learning" problem to a "tasting" problem.

The subject's earnings in a round come from her utility in that round. After all rounds of the experiment are complete, the subject sees a summary of her earnings in each round and the sum of those rounds' earnings in points and in dollars. She then is presented with a short questionnaire about herself and about the experiment. Her total earnings for the experiment are the sum of her earnings in all rounds, converted to dollars with a conversion rate c, plus an additional \$0.50 for completing the questionnaire.

As shown in Table 2, we experimentally vary income y, lifetime T, and noisiness in priors σ, so that our experiment has eight cells. Across all cells, all subjects have the same number of goods, likelihood of each good appearing, numeraire value, maximum disturbance size, conversion rate, and distribution from which values are drawn.

We gave each good the name of a fake fruit and we called the numeraire good "bread" to make the experiment more engaging and game-like while still limiting their importation of beliefs and tastes from outside the experiment. See Appendix D for full instructions.

We programmed the experiment in oTree (Chen et al, 2016) and deployed

Table 2: Experiment Parameters

Variable	Description	Fixed or varied?
N	Number of goods in the universe	Fixed: 10
q_i	Probability good i appears in a round	Fixed: 25% or 50%
p_i	Price of good i	Fixed: 1
y	Income	Experimentally: 3 or 6
T	Lifetime (number of rounds)	Experimentally: 10 or 20
z	Value of numeraire good	Fixed: 65
β_i	True value of good i	Random integer in [50, 80]
σ	Max disturbance in "starting guesses"	Experimentally: 25 or 49
m_i	Meaningful consumption experience	Fixed: 1
c	Conversion rate, points to dollars	Fixed: 1000

it on Amazon's Mechanical Turk (mTurk). Subjects were screened on being US-based and having successfully completed a large number of past mTurk tasks.

4 Experiment Results

We ran the experiment in February 2018. In all, 1,252 potential subjects signed up to participate, of which 646 completed the experiment.[5] Table 3 shows the number of subjects in each treatment condition. Among these 646 subjects, subjects earned an average of 4,797 experimental points, or $4.80 plus a $0.50 participation payment. The first quartile of earnings was $3.50 and the third quartile was $7.34.[6]

Our analysis proceeds as follows. First, in Section 4.1, we validate the experiment and model by showing that subjects choose according to their

[5]Problems with the server caused fatal timeouts for some potential subjects. Of the 606 who did not complete the experiment, 547 (90.3%) had made no choices by the time they stopped. Most of these likely had server timeouts.

[6]The post-experiment questionnaire asked a comprehension question that posed a simplified version of the experiment's choice problem. 82.43% of subjects answered correctly. Including only those who answered correctly produces qualitatively identical results except that the Mann-Whitney test for the effect of noise on efficiency becomes insignificant and the effect of noise on efficiency becomes significant at the 10% level in the Tobit regression for $T = 10$. This paper reports results from the full sample of subjects.

Table 3: Number of Subjects in Each Treatment Cell

Lifetime $T = 10$	Income $y = 3$	Income $y = 6$
Noise $\sigma = 25$	76	91
Noise $\sigma = 49$	85	95
Lifetime $T = 20$	**Income $y = 3$**	**Income $y = 6$**
Noise $\sigma = 25$	74	71
Noise $\sigma = 49$	76	78

beliefs often, but not always, with some deviations consistent with learning and others consistent with error. We next, in Section 4.2, show that choice reversals exist and decline with time, as predicted in Proposition 5. We then show in Section 4.3 that most, but not all, goods are tried, with evidence supportive of Proposition 2 and all of the elements of Proposition 3 that our experiment can test. Then in Section 4.4 we show how this engenders different degrees of eventual preference learning. Next, we show that believed preferences become increasingly stable, as shown in Proposition 4. We demonstrate the pessimism bias predicted in our informal hypothesis. Finally, in Section 4.5 we show that welfare loss occurs and declines over time, but importantly, as predicted in Proposition 6, it does not decline to zero.

4.1 Consistency with Believed Preferences

To show the extent to which subjects choose according to their believed preferences, we construct a dummy variable for each subject for each round, and we give it a value of 1 if the subject chose the bundle that maximizes believed utility. The subject's believed utility is based on the parameters of goods they have learned up to that round, and the point estimates ("starting guesses") they have for goods they have not yet tried. For goods that are not yet learned, the assumption that a rational person will maximize utility by maximizing current-period expected value based on these parameter point estimates is only strictly true for risk neutral people, as risk averse or risk loving people will have a bias against (or, respectively, in favor of) untried goods; and experimental consumption by definition will cause people to diverge as well, but this is a starting point for our analysis; we will discuss divergences from this simple myopic optimization below.

Pooled across all treatments and rounds, people maximize believed expected earnings in this way 61.0 % of the time. In round 1, subjects choose in accordance with their believed preferences at a rate of 43.2% for $T = 10$ and 41.8% for $T = 20$. At the end of experimental lifetimes, that value is 65.4% in round 10 for $T = 10$ and 69.2% in round 20 for $T = 20$, a significant increase (within-subject sign-rank test: $p < 0.001$ in both cases, $n_{T=10} = 347$, $n_{T=20} = 299$).[7]

Recall that experimental consumption explains some choices that don't maximize believed preferences. Since experimental consumption has no further value in the final round, why would a subject make a non-myopic-maximizing choice in her last round? One potential reason is that subjects who are risk averse may choose a learned good with a lower parameter value over an as-yet-unlearned good with a higher "starting guess" to avoid uncertainty. Subjects who are risk loving may do the opposite, choosing an as-yet-unlearned good with a lower "starting guess" over a learned good with a higher parameter value.

Neither risk preferences nor experimental consumption can explain non-maximizing choices among goods that have already been learned. In 21% of all choices we observe, subjects choose a good they have tried before with a lower known value than another available good with a higher value (in some cases also forgoing untried available goods with higher "starting guesses"). These choices likely indicate error. As noted above, these errors decline significantly over time. Further, the magnitude of most of these errors is small: of these choices, 63.7% choose a good that's only dominated by a small amount (between 1 and 9 units of absolute value). This means that in 92.3% of all choices, an error cannot be identified (although they may choose something with a lower prior) or we identify only a "small" error.

Further, subjects who make non-myopically-maximizing choices in the final period do not seem to suffer in our study: on average, those who behave inconsistently achieve 95.13% efficiency while those who behave perfectly consistently achieve 94.96% efficiency (not statistically different, $p = 0.908$).[8]

[7]Subjects made other non-maximizing choices as well. Of 103,950 good-round pairs, subjects chose a value between 0 and 1 (less than a meaningful consumption experience) 322 times (or 0.31% of the time), and a value less than 0 a total of 14 times (0.01% of the time). 99.63% of the time, subjects chose an integer between 0 and 6.

[8]Choices inconsistent with believed preferences have little impact on our experiment's results. Excluding subjects who make these non-myopic-maximizing choices yields qualitatively similar results with only a few changes: In Table 4, the difference in efficiency

These results show that subjects are engaging in some optimizing choice as proposed in our model, but that they are quite a bit less sophisticated than our hypothetical Alice. To the extent to which error enters into our subjects' decision process, that introduces noise that makes it harder for us to detect the empirical results we find in the remainder of this paper.

We made no theoretical predictions about the effects of treatment variables (T, y, and σ) on subjects' tendency to choose in a way consistent with current beliefs, except that we point out that experimental consumption depends on current period sacrifice and discounted expected potential gain therefrom (Lemma 3(iv)). In Table 4, for tests pooled across rounds, we see that subjects with longer lifetimes and lower incomes made fewer choices that are myopically inconsistent with their beliefs. If experimental consumption happens more in early than in later rounds, then a longer lifetime should give more rounds (as compared to a shorter lifetime) in which little experimenting is happening, thus explaining why longer lifetimes are associated with choices more consistent with beliefs. Higher income yielding more choices inconsistent with beliefs could happen because higher incomes should yield more experimentation, as argued in the proof of Proposition 3. We return to the rest of the results in Table 4 later in this section.

We can seek evidence that some choices that diverge from maximizing believed utility are experimental consumption by regressing the dummy for deviation from believed preference maximization on factors that should affect the value of experimental consumption. Table 5 shows OLS results. (Logit and probit results are similar.) Belief-inconsistent choices increase with remaining lifetime and endowment and decrease with overall lifetime. These results are consistent with subjects making inconsistent choices early as they discover their preferences, and then increasing consistency as their understanding of their preferences improves. We return to the rest of Table 5 in the next subsection.

---

across income levels becomes marginally significant ($p = 0.090$), while the difference in efficiency across noise levels ceases to be statistically significant ($p = 0.118$). In Table 5, the number of remaining rounds becomes significant at $p < 0.001$ in the second model, while in the third model, noise becomes marginally significant ($p = 0.064$) as does income ($p = 0.063$).

Table 4: Nonparametric Tests of Treatment Effects on Learning Outcomes

Lifetime T	Choices inconsistent with beliefs	Full discovery	Efficiency
10	0.444	0.159	0.846
20	0.358	0.502	0.896
p-value	0.000	0.000	0.000
Income y	**Choices inconsistent with beliefs**	**Full discovery**	**Efficiency**
3	0.375	0.270	0.878
6	0.432	0.361	0.861
p-value	0.001	0.013	0.355
Noise σ	**Choices inconsistent with beliefs**	**Full discovery**	**Efficiency**
25	0.405	0.340	0.876
49	0.404	0.296	0.863
p-value	0.941	0.237	0.035

All variables are aggregated to the subject level. N's can be inferred from Table 3. "Full discovery" captures whether a subject has tried every good by the end of the experiment. "Choices inconsistent with beliefs" is the proportion of rounds in which a subject's choices do not maximize expected utility given beliefs. "Efficiency" is the utility achieved as a proportion of the maximum achievable. p-values are from Mann-Whitney tests.

Table 5: Drivers of Utility Maximization Deviations and Choice Reversals

	Non-maximizing choice	Choice reversal (all rounds)	Choice reversal (rounds > 5)
Remaining rounds	0.017***	0.000	0.006***
	(0.001)	(0.001)	(0.002)
Lifetime $T = 20$	-0.170***	-0.037**	-0.081***
	(0.015)	(0.016)	(0.000)
Noise $\sigma = 49$	-0.003	0.009	0.027
	(0.014)	(0.014)	(0.017)
Income $y = 6$	0.053***	0.050***	0.050***
	(0.014)	(0.014)	(0.017)
Constant	0.340***	0.296***	0.295***
	(0.016)	(0.017)	(0.019)
R^2 (overall)	0.0399	0.0044	0.0087
n subjects	646	646	646
n subject-rounds	9,450	8,804	6,220

Robust standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < 0.1$. Random effects OLS panel regressions at the subject-round level with errors clustered at the subject level. For treatment variables, we use dummies that are equal to 1 for the higher value.

4.2 Choice Reversals

Now we turn to choice reversals. For each subject in each round, we infer whether her choice contradicted the ranking implied by a past choice, and we call this contradiction a choice reversal. In other words, if goods A and B were available in round 1 and the subject chose more of A than of B, but in round 2 when both were available she chose more of B than of A, that is a choice reversal. In the theory, we used a more limiting definition of choice reversals; we use a slightly different definition here because in our subjects' finite experimental lifetimes, the odds of the exact basket of available goods reappearing are quite small, so we would have little power to observe the kind of reversals we describe in our theory in a reasonably rich experiment.

Proposition 5 held that choice reversals would be observed and would decline over time. We test this prediction in the latter two columns of Table 5, using OLS panel regressions at the subject-round level. In the experiment, each subject needs some time to build up a choice profile that can be contradicted. In the first round, it is impossible to observe a choice reversal because there is nothing to contradict. If the sets of goods available in rounds 1 and 2 are disjoint, then it is also impossible to witness a choice reversal in round 2. For this reason, the second column presents a regression model that includes all rounds except the first, and the third column includes only rounds 6 and up (when half or a quarter of the subjects' lifetimes have passed) to allow a choice profile to be established. This third column is our preferred specification.

While the specification in Table 5 that includes all rounds does not show an effect of remaining rounds on choice reversal rate, our preferred specification (which excludes the first five rounds) does. The latter indicates that choice reversals decline over time, as predicted, and the former indicates that this is confounded by the mechanical difficulty in observing reversals in early rounds.

We made no theoretical predictions about the relationship between our treatment variables and choice reversals. However, we show in Table 5 that a longer lifetime reduces the rate of choice reversals, while a higher income increases the rate of choice reversals. Thus, the same experimental factors that drive inconsistent-with-belief choices also drive choice reversals, giving further evidence of experimentation.

4.3 Trying Goods

Next, we examine subjects' tendency to try goods. Most subjects try most, but not all, goods that they have the opportunity to try. Figure 1 plots over time the proportion of all goods that are chosen in at least a nibble (the minimum quantity to learn) as well as the proportion tried of all goods that have appeared (and thus could be chosen). The raw proportion of total goods tried increases at a decreasing rate until it levels off at around 87% approaching Round 20. The proportion of possible goods tried shows a similar trend. By the end of 20 rounds, subjects had been presented with 99.85% of all possible goods on average, and 89.52% of subjects were presented with all 11 goods at some point.

Tendency to try more goods as time progresses (but at a decreasing rate) is not proof of our model, as even random choice would yield this outcome, so we move on to more interesting results.

If subjects were myopic, i.e., if there was no experimental consumption, life length would not affect tendency to try goods, in which case the $T = 10$ and $T = 20$ lines would coincide, but they do not. Those with shorter lifetimes try more of the available goods in the first round ($p < 0.001$, rank-sum test at subject level), but have tried a smaller proportion of the available goods in the tenth round, which is the last round for the subjects with shorter lifetimes but only half-way through for those with longer lifetimes (again $p < 0.001$, rank-sum test at subject level). This crossing of the lines is not consistent with a perfectly forward-looking model of learning, which predicts that people with longer lifetimes get a higher benefit for trying new goods early in their lifetime: specifically, while our theory does predict the fact that the $T = 20$ line is steeper than the $T = 10$ line, it does not predict that the $T = 10$ line starts at a higher intercept. We confirm this pattern using regression methods below.

For both lifetime lengths, both curves end far short of 100%. For subjects with a lifetime of 10 rounds, 15.85% try every good by the last round. This increases to 50.17% for subjects with a lifetime of 20 rounds. These values differ, $p < 0.001$, based on a rank-sum test at the subject level. While subjects in our experiment have finite experimental lifetimes, the fact that learning seems to flatten out is supportive of Proposition 2's implication that some goods will never be tried even in infinite time.

Our Proposition 3 predicted that agents with longer lifetimes and larger incomes would be more likely to learn their preferences, and the nonpara-

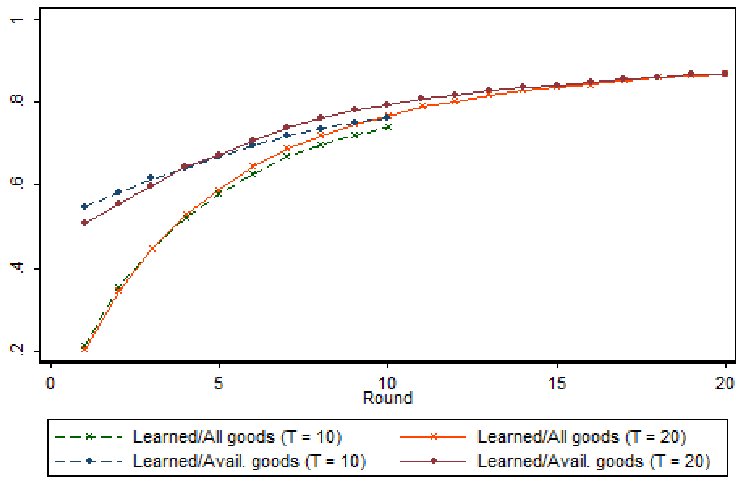

Figure 1: Percent of Goods Learned by Round
The vertical axis the proportion of all goods (or of goods that have appeared) that subjects have by the given round chosen at least $m_i = 1$ unit of, averaged across subjects.

metric tests in Table 4 confirm these predictions.

Proposition 3 also gave characteristics that should predict whether a good is tried. In Table 6, we report a panel regression with one observation per subject per good per round, where the outcome variable is a dummy indicating whether this subject has learned their preference for this good as of this round, i.e., whether by this round she has tried it in at least the minimum size needed to learn. We show results from an OLS regression; results are similar for logit and probit. These regressions include the numeraire as a good.[9] Our preferred specification is III, which includes the numeraire and round-Lifetime interactions (which we discuss below), but results are consistent across specifications.

We find again that subjects with longer lives and larger incomes try more goods. In Model II, we see that subjects have tried more goods as time passes, which appears to account for the effect of the longer lifetime; this is not one of our key results but is a reasonable sanity check. In Model III we examine time trends in the learning process by separately estimating the

[9]Recall that the numeraire value is known with certainty and thus is always "learned." Results are similar excluding the numeraire. We also considered a version of specification III that used individual period dummies and found qualitatively the same result.

Table 6: Factors driving whether a good is learned

	I	II	III
Lifetime $T = 20$	0.145***	-0.0142	
	(0.014)	(0.014)	
Income $y = 6$	0.0653***	0.0654***	0.0654***
	(0.014)	(0.014)	(0.014)
Noise $\sigma = 49$	0.155***	0.155***	0.155***
	(0.045)	(0.045)	(0.045)
Prior	0.00678***	0.00678***	0.00678***
	(0.001)	(0.001)	(0.001)
Prior x σ	-0.00238***	-0.00238***	-0.00238***
	(0.001)	(0.001)	(0.001)
True value	-0.00141***	-0.00141***	-0.00141***
	(0.001)	(0.001)	(0.001)
Average of	-0.00554**	-0.00554**	-0.00554**
other values	(0.003)	(0.003)	(0.003)
Probability of	0.422***	0.423***	0.423***
appearance	(0.019)	(0.019)	(0.019)
Round		0.0319***	0.0549***
		(0.001)	(0.001)
L: $T = 20$			-0.0134
First 10 rounds			(0.014)
L: $T = 20$			0.448***
Last 10 rounds			(0.019)
L: $T = 20$			0.00449**
Round x First 10 rounds			(0.002)
L: $T = 20$			-0.0463***
Round x Last 10 rounds			(0.001)
Constant	0.351*	0.176	0.0490
	(0.195)	(0.195)	(0.195)
R^2 (overall)	0.1250	0.2013	0.2190
n subjects	646	646	646
n subject-goods	7,106	7,106	7,106
n subject-good-rounds	103,950	103,950	103,950

Robust standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < 0.1$
Random effects OLS panel regressions at the subject-good-round level with errors clustered at the subject level. For treatment variables, we use dummies that are equal to 1 for the higher value. Model III includes two dummies for the Lifetime $T = 20$ treatment group, one for the first ten periods and one for the last ten. It also includes interactions of these dummies with the round.

intercept and the time trend during the first ten rounds for the two Lifetime treatments using additional dummy variables and their interactions with the round. The coefficient on the "Round x First 10 rounds" term indicates that those with a longer life try goods more quickly than do those with shorter lifetimes ($p = 0.014$ on the slope interaction coefficient). They do not appear to try a larger number of goods at the outset ($p = 0.353$ on the the "First 10 rounds" dummy-intercept) despite, in principle, a higher payoff to learning. In other words, we have no theoretical explanation for why people with a longer lifetime don't try more goods in the first round as compared to people with a shorter lifetime; however, after that, people in the longer-lifetime treatment do have a faster rate of learning as predicted by Proposition 3(b).

We also find evidence consistent with the theoretical predictions that goods are more likely to be learned if they have a higher prior belief (as predicted by Proposition 3(e)) or probability of appearance (Proposition 3(j)). The value of other goods decreases the likelihood that a given good has been tried (Proposition 3(k)).

Consider now the interaction between noise and priors. The prior can range from 1 to 129. Based on Specification III, the effect of noise ranges between $0.155 + (-0.00238) * 1 = 0.153$ for the lowest possible prior and $0.155 + (-0.00238) * 129 = -0.152$ for the highest possible prior. This confirms our prediction (from Proposition 3(f) and (g)) that goods with low priors would be more likely to be tried with more noise, and goods with high priors would be more likely to be tried with less noise (which prediction was conditional on the agent being risk averse). This result is very particularly tied to our model of learning.

4.4 Learning Benchmarks, Stability, and Pessimism

This tendency to try some but not all goods has predictable ramifications for the levels of discovery that our subjects achieve. We define *full discovery* as the state in which the agent has learned her preferences for all goods. We define *full relevant discovery* as the state of having learned all goods that are better than the numeraire; if one achieves this benchmark then one will not lose welfare because of misunderstood preferences. By the end of the experiment, only 25.94% achieve full relevant discovery for $T = 10$ while 57.86% achieve full relevant discovery for $T = 20$. A weaker benchmark is *full voluntary discovery*, a state in which the agent has learned all goods with priors better than the numeraire; from this state, the agent can still

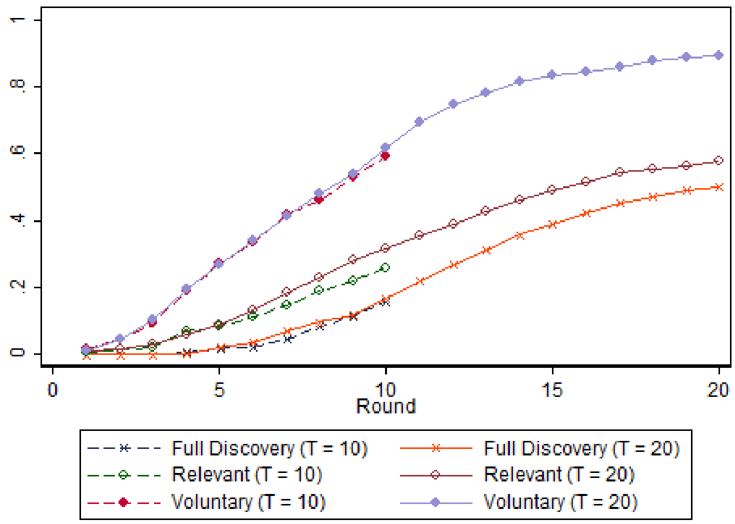

Figure 2: Achievement of Learning Benchmarks

lose welfare from misunderstood preferences, but when she is in this state, her believed values, including her knowledge gained from consumption and her "starting guesses" for untried goods, provide no basis to try new goods and exit the state. Only 59.37% achieve full voluntary discovery for $T = 10$, while 89.30% achieve full voluntary discovery for $T = 20$. Figure 2 shows the proportion of subjects who reach full discovery, full relevant discovery, and full voluntary discovery over time.

Recall that an agent who has achieved full voluntary discovery may stop trying goods she has not already learned. We declare a subject a candidate for persistent welfare losses if she has reached full voluntary discovery but not full relevant discovery. This is a relatively conservative definition, since given the flattening out of the learning curve, we infer that some subjects who have not achieved full voluntary discovery by our definition may be unwilling to sample new goods. This may be in part due to risk aversion. At the end of their experimental lives, 35.45% of subjects with $T = 10$ and 32.78% of subjects with $T = 20$ are candidates for persistent welfare loss. These are not significantly different ($p = 0.476$ from a subject-level rank-sum test). This implies that a sizable proportion of subjects, regardless of their experimental lifespan, may have reached a point at which they are done experimenting in spite of incomplete learning.

The leveling off of the learning curves in Figures 1 and 2 supports the idea that believed preferences eventually become stable even in our subjects' finite experimental lifetimes, as argued in Proposition 4 for infinite time, but we can test that hypothesis explicitly. We construct a variable for each subject for each round (starting at round 2) that indicates how many parameters changed between this round and the preceding round. While the average number of changes in rounds 2-6 is 0.911 for $T = 10$ and 0.972 for $T = 20$, the average number of changes in round 10 for $T = 10$ is 0.219 and in round 20 for $T = 20$ is 0.033. The difference is significant in both cases (sign-rank test at the subject level: $p < 0.001$ in both cases). Of the 299 subjects with $T = 20$, 289 (96.67%) chose no new goods in the final round, and 274 (94.81%) chose no new goods in the final two rounds.

We have shown, then, that subjects in our experiment are learning their preferences but are not learning them completely over the course of finite but long lifetimes. We hypothesized that this would lead to a pessimistic bias over time because positive misperceptions would be more likely to be corrected by experience. We test this by constructing a variable for each subject for each round that averages the subject's parameter belief errors, where each error is her current believed value minus her true value. In round 1, this error averages -0.156 across all subjects. This, as expected, is not significantly different from zero (t-test $p = 0.551$). At the end of subjects' experimental lifetimes, this value is significantly negative: -2.893 in round 10 for $T = 10$ and -2.038 in round 20 for $T = 20$. These values are significantly different from zero (t-test $p < 0.001$ in both cases). This is another result that is particular to our model.

Figure 3 shows how this pessimism evolves. The average error declines quickly as positive errors correct themselves faster than negative ones. The average error then levels off and starts to climb as subjects choose goods with small negative errors, correcting these errors. If we measure the bias among only undiscovered goods (thus eliminating the zero errors from the average), the average error is an even larger negative number. As a result, subjects' average beliefs about goods they have never tried steadily diverge from the true value, and display a persistent pessimistic bias.

4.5 Efficiency

Finally, we turn to the welfare implications of the learning process and its failures. We calculate an efficiency measure for each subject for each round

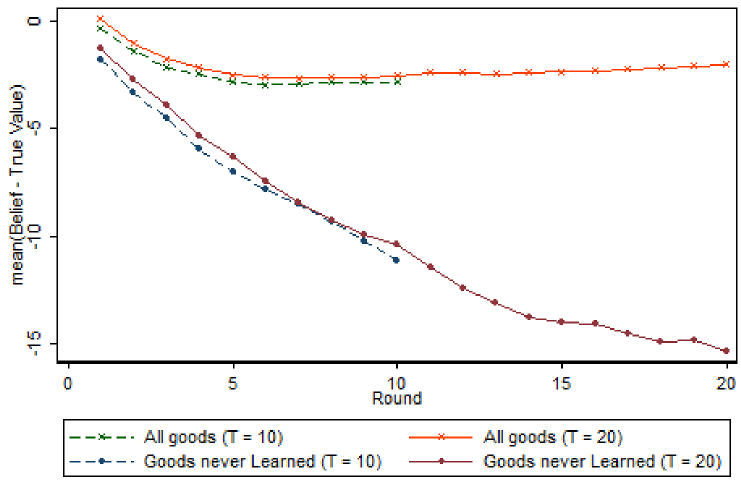

Figure 3: Average Error in Beliefs
The vertical axis is the mean error by round: the difference between subjects' prior value and true value, averaged across goods.

as the utility achieved in that round divided by the maximum she could have achieved if she had chosen according to her true preferences. Averages of this measure by treatment pooled across rounds are shown in Table 4. Longer lifetimes and lower noise in priors both yield higher efficiency, which accords with theory and results we have already shown about learning in those cases. Income does not affect efficiency (and we did not predict that it would).

To look at how welfare evolves across rounds in the different treatments, we run a panel Tobit regression at the subject-round level, which we report in Table 7. As time passes, efficiency loss declines, as predicted in Proposition 6. The effect of time is nonlinear, however: efficiency improves at a decreasing rate over time. Once we control for round number, life length ceases to have an effect, and in our regression, we see that the effect of income is only significant for $T = 20$.

A declining welfare loss is not particularly surprising and could have obtained as a result of other processes, as long as those processes involve optimization. The most important point from our theory is that this welfare loss need not decline to zero even as time approaches infinity. While our experiment subjects are not infinitely-lived, as we show in Section 4.4, choices become quite stable by the end of our subjects' experimental lifetimes, par-

Table 7: Determinants of Efficiency

	Pooled	$T = 10$	$T = 20$
Lifetime $T = 20$	0.003		
	(0.020)		
Income $y = 6$	-0.028	-0.011	-0.048*
	(0.019)	(0.027)	(0.027)
Noise $\sigma = 49$	-0.025	-0.034	-0.016
	(0.019)	(0.027)	(0.027)
Round	0.057***	0.072***	0.058***
	(0.002)	(0.008)	(0.003)
Round^2	-0.001***	-0.003***	-0.001***
	(0.0001)	(0.0007)	(0.0001)
Constant	0.698***	0.666***	0.702***
	(0.021)	(0.031)	(0.027)
Number censored at 0	554	242	312
Number censored at 1	3,979	1,063	2,916
n subjects	646	347	299
n subject-rounds	9,450	3,470	5,980
χ^2	1,607.14	464.17	1,093.86

Robust standard errors in parentheses. *** $p < .01$, ** $p < .05$, * $p < 0.1$. Tobit panel regressions at subject-round level with bootstrapped standard errors.

ticularly for those in the $T = 20$ treatment, and yet subjects do not discover their preferences for all goods that are better than the numeraire; therefore, if welfare is still being lost in the last period, this would confirm our most important theoretical result. Indeed, we find that in period 10 of the $T = 10$ treatment, efficiency is 90.7% and in period 20 of the $T = 20$ treatment, efficiency is 95.4%. Thus, we have shown that our warning that unlearned preferences could cause loss forever is borne out by our experiment.

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 of 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 formalized theory to identify 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. The results of an online experiment support the predictions of our model, and show that even in our simplistic setting, rationality errors make learning outcomes even worse than our theory predicts.

Our model shows that people may not fully learn their preferences even under the most congenial circumstances. 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 may forget her preferences for a good after learning them. If 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. If learning is not separable, this might make learning easier by letting each consumption experience have spillovers but also should create more parameters (such as coefficients that govern relationships between goods), thus increasing the dimensionality of the learning problem and making it harder, so the net effect is ambiguous.

Preference discovery processes 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 by the agent. Moreover, stable choice behavior does not indicate that agents are choosing according to their true underlying preferences: they may simply have stopped experimenting. While goods in our study could 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, 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.

The preference discovery process must be studied in more detail and in more settings to understand how factors internal and external to the agent affect learning and thus welfare loss. It is possible that an agent's mental simulation of consumption can allow some learning without consumption, and 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 contexts in which learning one's preferences through direct experience is very difficult, our model and experimental results indicate that losses could persist; if the choices are important, like choices regarding a house or a job, the losses could be large, and, as Thaler and Sunstein (2008) note, policy-relevant. If agents are aware of the problems we identify, for important decisions, they may 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 identify whether people do this and whether it seems to be welfare-enhancing, and could study whether specific nudges can help the preference learning process or can effectively replace it.

If we must learn through experience to know our own preferences, the implications are large. On the one hand, this model can provide new insights on how to get people to try new things, whether in the case of a company marketing a product or a government or non-profit promulgating a green 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 helpful comments from the editor and two anonymous referees. For advice early on, 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 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

Armantier O, Lévy-Garboua L, Owen C, Placido L (2016) Discovering preferences: A theoretical framework and an experiment

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

Brezzi M, Lai TL (2000) Incomplete learning from endogenous data in dynamic allocation. Econometrica 68(6):1511–1516, DOI 10.1111/1468-0262.00170, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00170

Chen DL, Schonger M, Wickens C (2016) otree: An open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9:88–97

Chuang Y, Schechter L (2015) Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. Journal of Development Economics 117:151 – 170, DOI https://doi.org/10.1016/j.jdeveco.2015.07.008, URL http://www.sciencedirect.com/science/article/pii/S0304387815000875

Cooke K (2017) Preference discovery and experimentation. Theoretical Economics 12(3):1307–1348

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

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

Eckel CC, El-Gamal MA, Wilson RK (2009) Risk loving after the storm: A bayesian-network study of hurricane katrina evacuees. Journal of Economic Behavior and Organization 69(2):110 – 124, DOI https://doi.org/10.1016/j.jebo.2007.08.012, URL http://www.sciencedirect.com/science/article/pii/S0167268108001741, individual Decision-Making, Bayesian Estimation and Market Design: A Festschrift in honor of David Grether

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

Keller G, Rady S (1999) Optimal experimentation in a changing environment. The Review of Economic Studies 66(3):475–507, URL http://www.jstor.org/stable/2567011

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

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

Piermont E, Takeoka N, Teper R (2016) Learning the krepsian state: Exploration through consumption. Games and Economic Behavior 100:69 – 94, DOI https://doi.org/10.1016/j.geb.2016.09.002, URL http://www.sciencedirect.com/science/article/pii/S0899825616300896

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

A Axiomatic 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; any failures we highlight in our simple model will 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 C.

A.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 goods in 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 some random basket of fruits is available to choose from.

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.

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

□

We can therefore represent Alice's tastes with a utility function $u(.)$.[10] Alice knows the form of $u(.)$, but may not know its precise shape. In partic-

[10] 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."

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 true values. These beliefs are not point estimates because she is sophisticated enough to know she has not yet learned her tastes: her beliefs are probability distribution functions over possible values. Therefore, we represent Alice's time-t 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, more formally,

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

The random variable β is characterized by a N_1-dimensional probability density function (p.d.f.) $f^{(\beta)}(b) : \mathbb{R}^{N_1} \to \mathbb{R}_0^+$. 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 beliefs $\hat{\beta}$ 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 $\mathbb{P}(b = \hat{\beta}) = 1$—but $\hat{f}(b) = 0$ for all other b.

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 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:

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

Thereby, for $i = 1, \ldots, N$, the real-valued vector b_i is 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 $\mathbb{B}_i$ is the set of all random variables of the dimension of $\hat{\beta}_i$ that assign a positive probability to (the neighborhood of) $\hat{\beta}_i$, akin to Equation (1). Now 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. 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 the positive or negative domain: 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, 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 $x_i \geq 0$.

□

We further simplify our analysis by restricting each β_i to be one-dimensional (which implies that $N_1 = N$):

Axiom 5. Single Parameter Sub-Utility Functions.
For each good $i \in \{1, \ldots, N\}$, $u_i(.)$ is characterized by a single parameter.

□

Lastly, we make two additional assumptions for ease of exposition: 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 $b_i \in \mathbb{R}$. Second, we specify that larger parameter values always imply (weakly) larger utility; that is, for each good i, $\frac{\partial u_i(x_i;b_i)}{\partial b_i} \geq 0$.

A.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 standby 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 $p_i > 0$.

A.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 and instead at time t she chooses according to a utility function parameterized by random beliefs β^t (with density function $f^{(\beta^t)}(.)$), starting from 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.[11] 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 6. Experience of Utility.
If she consumes a bundle with 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 her 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 the utility she got from the overall bundle.

□

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. If no good was (like the numeraire) available with probability 1 in each time, the opportunity cost of consuming a good would sometimes be zero. If we did not require at least a nibble to learn, then Alice could learn her tastes by purchasing an infinitesimally small

[11]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 very small, so that the utility gained is also very 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, given more-than-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.

Axiom 5 implies a unique mapping between utility received from a good and the parameter value for that good. Because of that implication and Axiom 6, Alice should update her beliefs about her preferences based on the utility she experienced 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.

Axiom 7. Separability of Learning.
Experiencing a good has no effect on the agent's perceived parameters of any other good.

□

Axiom 7 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 $\beta \in \mathbb{B}$:

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

Moreover, once learned, parameters are not forgotten.

Axiom 8. 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 $s \geq t$.

□

Together, Axiom 7 and Axiom 8 ensure that believed parameters for some good i only change with experience with good i. This implies that:

Lemma 1. Updating of Preferences.
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*,* $f_i^{(\beta^t)} \in \left\{ f_i^{(\beta_i^0)}, \hat{f}_i \right\}$*.*

□

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.

A.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 by $x^t = (x_1^t, \ldots, x_N^t)$. 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

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

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, 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, $x_i = 0$) 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 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.

Axiom 9. 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 9, 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 by 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 the random variable X based on the information available at time t, that 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 $x_i = m_i$ when $x_i < m_i$ maximizes myopic utility. 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).[12] We define for $i \in G^t$ with $f_i^{(\beta_i^t)} \equiv f_i^{(\beta_i^0)}$:

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

subject to

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

$U_i(.)$ is Alice's time-t expected utility from consuming a nibble of good i and allocating the rest of her money optimally among the remaining 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 learn her taste if it does not, we focus on the case in which it is a loss.

[12] 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.

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, so $f_i^{(\beta_i^t)} \equiv f_i^{(\beta_i^0)}$. 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 ϕ_i^{t+1}.
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 a non-increasing function in t.

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

□

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 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 preferences for all other goods remain the same.

□

B Theory Results

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

B.1 Preference Learning

Let us first explore what goods Alice will and will not learn her tastes for in 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 t. That is, $i \in L^t$ if and only if $f_i^{(\beta_i^t)} \equiv f_i^{(\beta_i^0)}$. 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 weakly 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. Lastly, *full voluntary discovery* is the state in which she has learned all the goods that she would ever voluntarily consume at least a nibble of; which goods fall in this category will depend on Alice's preferences and the factors that influence ϕ. We do not define full voluntary discovery here in a formal, general sense since we will only refer to it in our experiment results section, where the definition is straightforward.

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.

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 be eventually tried. Other goods, like perhaps (for Alice) the mangosteen, are more prospective: goods with lower expected marginal utility can only be discovered through experimental consumption, and that can 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) For $T = \infty$, 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) $\max_{G \in \mathbb{G},\, \beta \in \mathbb{B}_i'} \delta \cdot \phi_i^1(f^{(\beta)}) - U(f^{(\beta)}, G) + U_i(f^{(\beta)}, G) < 0,$

where

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

□

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.

A good can meet condition (b)(i) but not (b)(ii). These 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 generally 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 either of the following conditions:*

(a) *She discounts future consumption more heavily (smaller* δ*).*

(b) *She has a shorter lifespan (smaller* T*).*

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

(d) *She is more risk averse.*

(e) *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).*

(f) *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.*

(g) *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* $\left.\frac{d\,u_i\left(x_i;E[\beta_i^0]\right)}{dx_i}\right|_{x_i=m_i} > z \cdot \frac{p_i}{p_1}$.

(h) *The good is more expensive (larger* p_i*).*

(i) *A larger nibble is required to constitute a meaningful learning experience (larger* m_i*).*

(j) *The good appears less frequently (smaller* q_i*).*

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

□

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 people will make errors when they are inexperienced in the sense of having unlearned preferences if their 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(j)), or goods she'd buy rarely because they are expensive (Proposition 3(h)), or goods that require more consumption to learn (Proposition 3(i)). In our model, being poorly informed is eventually self-correcting unless Alice is poorly informed in the negative direction (Proposition 3(e)).

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 Alice's preference parameters will stabilize, albeit not necessarily at her true preferences.

Proposition 4. Eventual Preference 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 theoretical prediction shows those results are not evidence against preference discovery.

We also suggest the hypothesis 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: goods i for which the true $\hat{\beta}_i$ lies to the left within the distribution $f_i^{(\beta_i^0)}$; let us call this 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 $f_j^{(\beta_j^0)}$; let us call this 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 never learns 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, in each case, she'll taste the good eventually and will 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(e). By the same token, goods with a pessimistic bias will be less likely to be ever 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 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 B.2, we study how observers may see evidence of the learning process in action. In Section B.3, we study how Alice loses welfare because of undiscovered preferences.

B.2 Choice Reversals

Consider now the phenomenon of choice reversals, as discussed in work such as Cox and Grether (1996).[13] 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 in finite time, but not as $t \to \infty$.

Proposition 5. Choice Reversals.

(a) If there exists a good $i \in \{2, \ldots, N\}$ for which $p_i m_i \leq y$, then for any $\hat{\beta} \in \mathbb{B}$ there exists a prior $\beta^0 \in \mathbb{B}$ such that for any time t, $\mathbb{P}(x^{t+1} \neq x^t | G^{t+1} = G^t) > 0$.

(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, as found in Cox and Grether (1996).

B.3 Welfare Implications

Recall that $U(\hat{f}, G)$ denotes the maximum myopic utility Alice can attain with the goods available in set G. As a result, 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 reduction in utility she experiences from not choosing according to her true preferences at time t:[14]

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

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

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

[14] Since utility is not cardinal, it is usually preferable to define welfare losses in terms of compensating or equivalent variation. However, since we restrict our attention to a single agent, utility loss is equally appropriate here.

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 it need not decline to zero. We can thus draw the following conclusions about the agent's welfare loss:

Proposition 6.

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 $t \to \infty$.

□

Believed and true 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 will not eliminate the failure to try some goods with true values that would render them part of myopically optimal bundles nor the resulting welfare loss, as long as the correlation between beliefs and true values is not perfect.

C Appendix: Technical Proofs

Proof of Lemma 1

(a) Axiom 6 implies that Alice updates her preferences to the true $\hat{\beta}_i$ upon her meaningful consumption experience at time t. That is, $f_i^{(\beta_i^{t+1})} \equiv \hat{f}_i$. Then, by Axiom 8, she will maintain these true preferences into perpetuity.

(b) According to Axiom 6, if $x_i^t < m_i$, Alice has no reason to update her preferences for good i at that time. Axiom 7 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, $f_i^{(\beta_i^{t+1})} \equiv f_i^{(\beta_i^t)}$.

(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

Let us 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 maximizing 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) $\phi_i^t \geq 0$ because it is the sum of non-negative numbers.

(b) For $T < \infty$,
$$\phi_i^t(f) - \phi_i^{t+1}(f) = \delta^{T-t-1}\alpha_i^{T-t}(f) \geq 0 .$$

(c) Similarly, for $T = \infty$,
$$\phi_i^t(f) - \phi_i^{t+1}(f) = 0 .$$

□

Proof of Lemma 4

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 $f_i^{(\beta_i^s)} \equiv f_i^{(\beta_i^t)}$, 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.

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.

Alternatively, Alice might choose to *experimentally* consume m_i units. For this to happen, according to Lemma 3, it must be true that for $\tau \in \{t, s\}$:

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

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 $f_i^{(\beta_i^s)} \equiv f_i^{(\beta_i^t)}$ 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, we show that $x_i^s \leq x_i^*$ for $s = t$ 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 expense of a myopically-optimal mix of good i and the numeraire. If strict convexity holds, then this requires Alice to buy less 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 i. Thus, $x_i^s \leq x_i^*$.

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.

□

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 probably converges to 1, which would obviously be equally true if other cases gave 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 these conditions:

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

(ii) 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

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

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

(b) 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 (iii), making experimental consumption less likely.

(c) Because the budget constraint is tighter in future periods, future consumption is lower, which reduces ϕ_i^{t+1} and thus the right side of the inequality in (iii), making experimental consumption less likely. If i is normal, then a smaller y means that the optimal myopic choice in the current period is smaller and could fall below the nibble size, so myopic consumption (ii) might cease to select learning this good; 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. These points are unambiguous 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. A smaller y can make the inequality in (i) fail to hold, so that a nibble of i becomes unaffordable. 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 (iii), making experimental consumption less likely.

(d) If Alice is more risk averse, then for all $i \in \{2, \ldots, N\}$ and for any $x_i > 0$, $Eu_i(x_i; \beta_i^0)$ is smaller due to her increased disutility from the uncertainty about β_i^0. This makes her consume less of good i in her myopic choice—relative to the numeraire good as well as other available goods that she has already learned and that she therefore has no uncertainty over (so that increased risk aversion does not devalue the utility from these goods). This makes myopic consumption per (ii) less likely. By the same token, increased risk aversion increases the current period sacrifice required for experimental consumption, increasing the left side of the inequality in (iii) and making experimental consumption less likely.

(e) A lower prior, that is a left-shifted $f_i^{(\beta_i^0)}$, by our informal 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 (ii), and increases the current-period sacrifice for experimental consumption, which increases the left side of the inequality in (iii) and makes experimental consumption less likely.

(f) 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, 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 (iii) and making experimental consumption less likely.

(g) 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. Then $\left.\frac{d\,u_i\left(x_i;E[\beta_i^0]\right)}{dx_i}\right|_{x_i=m_i}$ 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) can fall below the right side, so that good i will not be learned through myopic consumption, i.e., so condition (ii) does not hold. The level of dispersion of β_i^0 and her believed preferences for the other goods could make this expected utility so low that the current period utility sacrifice, the left side of the inequality in (iii), 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.

(h) 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 (iii), making experimental consumption less likely. It also renders a nibble of the good less likely to be affordable, so it could make (i) cease to be met.

(i) 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 (h).

(j) 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 (ii)) good i, she may not encounter that basket during her life. In addition, since there are fewer future consumption opportunities with this good in which decisions can be optimized, ϕ_i^{t+1} is reduced, which reduces the right side of the inequality in (iii) and makes experimental consumption less likely.

(k) 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. 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.

□

Proof of Proposition 4

Let $T = \infty$. Suppose that at some time t, Alice's preferences 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 preferences change. Therefore, there exists some set G and some good $i \notin L^t$, such that under her current preferences 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. We conclude 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 changed her preferences between time t and τ due to some other consumption experience, so the learned set must expand based on that preference change; or (ii) preferences 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 an unstable preference will result in a preference change as a good is added to the learned set by 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 are 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 paragraph of this proof, but since that scenario involved either the good i under consideration or some other unknown good, we can't identify which good and thus which probability to use, so we are using 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 preferences are currently unstable, there is a positive, time-invariant probability (greater or equal to ρ) in each future period that she will change her preferences 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 preferences 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 preferences, 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.

Since there are only $N-1 < \infty$ goods to be discovered, and since preferences for each good remain stable once discovered (Lemma 1), there can be at most $N-1$ preference 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.) The time of her final change of preferences is thus no greater than $T_1 + \ldots + T_{N-1}$. 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) $N-1$ events to occur.

Therefore, if T^* denotes a random variable that follows a negative binomial distribution with probability parameter ρ and frequency parameter $N-1$, we can conclude that:

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

The first inequality follows from our earlier discussion that preferences will not change after time $T_1 + \ldots + T_{N-1}$ (and possibly sooner). 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^* itself. This is reflected in the second inequality. Lastly, the convergence is a property of the negative binomial cumulative distribution function.

Note that this essentially proves that any good that will be eventually learned will be learned eventually. This is not true for all goods, 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 for which $p_i m_i \leq y$, as assumed in the proposition. Let x_i' denote the quantity of good i that Alice chooses to consume as the solution to her myopic choice problem if the set of available goods in that period is $G_i = \{1, i\}$.

Consider first the case in which under her true preferences $\hat{\beta}_i$, she would choose $x_i' \neq m_i$. If we choose her prior for the good, $\beta_i^0 \in \mathbb{B}_i$, such that $\left.\frac{d\, Eu_i\left(x_i; \beta_i^0\right)}{dx_i}\right|_{x_i = m_i} = z \cdot \frac{p_i}{p_1}$, then the following sequence of events will ensure a choice reversal between some time t and $t+1$:

(i) $i \notin G^s$ for any $s < t$;

(ii) $G^t = \{1, i\}$; and

(iii) $G^{t+1} = \{1, i\}$.

This situation is possible: since there is a finite number of goods, the probability for (i), (ii), and (iii) to occur jointly is positive, since each occurs with some positive probability independently of the others.

Part (i) ensures that $f_i^{(\beta_i^t)} \equiv f_i^{(\beta_i^0)}$. Thus, based on (ii) and given our chosen prior, Alice will choose to consume $x_i^t = m_i$ at time t.[15] By Lemma 1, Alice will then learn her true preferences for good i, that is $f_i^{(\beta_i^{t+1})} \equiv \hat{f}_i$. Then, at time $t+1$, Alice knows her preferences for all available goods, so that 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. Thus in this case, facing the same basket of available goods in two times, she chooses different bundles.

Secondly, for the alternative case where under $\hat{\beta}_i$ she would choose $x_i' = m_i$, we choose a prior $\beta_i^0 \in \mathbb{B}_i$ such that $x_i' = y/p_i > m_i$ and follow the same logic as in the first case.

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 preferences become stable—which Proposition 4 guarantees to happen eventually—Alice will always choose her consumption in order to maximize her myopic expected utility. Since preferences 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:

[15]If convexity is only weak, then Alice will consume some amount $x_i^t \geq m_i$, since myopically she will be indifferent between different combinations of i and the numeraire but consuming at least m_i gives a dynamic benefit from learning. Since i will still be learned, the same conclusions will hold.

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

Thus, by Proposition 2(b), good i will never be learned. As a result—and due to condition (i)—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 every time the choice set $G_i = \{1, i\}$ appears. 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 preferences 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 becomes less prevalent, because fewer parameters will be unknown and because the benefits of learning will be diminished (in expectation, 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 is a (weakly) decreasing function of t.

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

$\square$

D Appendix: Experiment Screens

Appendix B: Experiment Screen Shots

Welcome to the experiment! You will first read some instructions, then make a series of decisions, then answer a quick survey. This will take up to 10-20 minutes. Your earnings depend on your decisions and on chance, but will be between $0.50 and $10.

Next

Instructions: Decisions

You will make decisions over a series of 10 rounds. In each round, you will start with some money: 3 francs. You will spend your money buying fruits, which each has a price of 1 franc, and/or bread, which also costs 1 franc. You will have up to a minute to make your decision in each round! If you do not submit a decision in a round, you will buy no goods and thus earn nothing for that round.

You will get value (in points) from the foods you buy. Bread always earns you 50 points points per franc you spend on it. There are several fruits; these are not normal fruits you see every day, but fruits with names we made up. Each fruit gives you a particular value, and each fruit's value stays the same for the whole experiment. Some fruits will appear more often than others, but the chance that a given fruit will appear stays the same across all rounds. Not all the fruits will be available in all rounds.

Your food earnings in a round is the sum of the values you get from all the fruits and/or bread you buy. Your total earnings in the experiment is the sum of your food earnings in each round plus $0.50 for filling out a short survey at the end of the experiment.

For example, imagine you have 4 francs. Imagine that an apple gives you 200 points of value, and an orange gives you 100 points of value. Bread, as stated above, gives you 50 points points of value. If you buy one apple, two oranges, and one bread, how much value do you earn in points?

Next

Instructions: Value

In the example above, we told you what your values were for each fruit. But in the experiment, you will not be told those values.

At the beginning of the experiment, you will be given a "starting guess" for the value for each fruit. That will be related to the fruit's real value for you: it will be the actual value plus or minus some random number.

At the end of each round, you will learn how much value you got from bread and from each of the fruits you chose. You will only learn your value from any fruit you chose at least 1 unit of. Your guesses for each fruit will be updated with these values.

In future rounds, to help you make your decisions, you will see all of the values for goods whose values you have learned, and you will see your starting guess for the goods whose values you haven't learned yet.

Next

Instructions: Summary and Earnings

In summary, in each of 10 rounds, you will choose how to spend your 3 francs buying fruits and/or bread. Bread is always available, but whether each fruit appears depends on chance. You will earn money based on the values you get for each fruit and/or bread you buy. You will start out not knowing for sure the values you get for each fruit, but you will be given starting guesses that are related to your actual values (they are the true values plus or minus a random number). Bread always gives you 50 points per franc.

After you choose how much you want to spend on each fruit and/or on bread, you will learn the value you got from any fruit you bought at least 1 unit of, and that will update your guesses with these actual values.

In each round, you will have one minute to make your food choice, and 30 seconds to review the information on values, so make sure you're paying attention! If you do not choose some fruit and/or bread by the time a round ends, you will get none of the fruit and no bread and thus earn no value that round.

Your earnings in points for each round is the value you get from your fruit and/or bread plus the bonuses you earn. Your earnings for your decisions are calculated as: the sum of your earnings in each round times the conversion factor of 0.001 dollars per point. After your decisions, you will answer a survey that will take a few minutes, and you'll receive another $0.50 for your completion of the survey.

For example, if you earned 4,000 points across all of the rounds, that would give you 4,000*0.001 = $4 for your decisions, plus $0.50 for the survey, for a total payment of $4.50. (Your payment will be rounded to the nearest cent if necessary.)

Decision

Time left to complete this page: **0:44**

This is round 1. You will play this game for 10 rounds in total.

Instructions reminder: spend all your money buying fruits and/or bread for 1 franc each. Bread always gives you 50 points points per franc; you start out with guesses about how many points per franc you get for each fruit. Your starting guesses before you try the good are your actual values from the fruits plus or minus random numbers. At the end of the round you'll learn your earnings from each fruit you buy at least 1 of. Your values for those fruits will be updated for you to see in future rounds. Your payment for this experiment depends on the values you earn in each round!

You have 3 francs to spend.

Choose how many of each of the foods you would like to buy:

Food name	**Value or guess**	**Guess?**	**How much would you like to buy?**
Merooki	48 points	guess	Merooki: 0.0
Bread	50 points		Bread: 0.0

Next

Decision Round Report

Time left to complete this page: **0:16**

This was round 1.

You earned 0 points.

Here are your updated values for all of the fruits. If the word “guess” appears, the value is your starting guess. If it does not appear, this is a value you’ve learned in this or a past round.

Food name	Guess?	Value
Frutana	guess	60 points
Jojofruit	guess	90 points
Banello	guess	69 points
Niblunda	guess	80 points

Decision Round Report

Time left to complete this page: **0:17**

This was round 2. You bought:

3.0 Niblunda

You earned 165 points.

Here are your updated values for all of the fruits. If the word “guess” appears, the value is your starting guess. If it does not appear, this is a value you’ve learned in this or a past round.

Food name	Guess?	Value
Frutana	guess	60 points
Jojofruit	guess	90 points
Banello	guess	69 points

This was round 8. You bought:

3.0 Banello

You earned 225 points.

Here are your updated values for all of the fruits. If the word "guess" appears, the value is your starting guess. If it does not appear, this is a value you've learned in this or a past round.

Food name	Guess?	Value
Frutana	guess	60 points
Jojofruit		69 points
Banello		75 points
Niblunda		55 points
Danutia	guess	40 points
Yegrevy	guess	45 points
Merooki		67 points
Oggerydot	guess	65 points
Zellitan	guess	44 points
Valavoo	guess	54 points
Bread		50 points

Next

End of Decisions Report:

Here are your earnings from the decision rounds:

Round	Food Earnings
1	0 points
2	165 points
3	201 points
4	234 points
5	225 points
6	156 points
7	207 points
8	225 points
9	234 points
10	201 points

Your total food earnings are the sum of your earnings in those rounds, which is 1848 points. Since you earn 0.001 dollars per point, this means your food earnings are worth $1.85. Your total earnings are that plus $0.50 for filling out the survey that you are about to start, or a total of $2.35.

Click Next to start the survey!

Next

Questionnaire page 1

Please answer all of the questions below. Your answers will not affect your payment but will help us understand our results.

1. What do you think the experimenters will learn from this experiment?

2. Imagine you have 100 francs. If apples and bread each cost 1 franc, you know that apples earn you 5 points per franc and bread earns you 4 points per franc, how much would you buy of each if you want to earn as many points as possible?

Apples:

Bread:

3. Imagine we are throwing a five-sided die (with sides numbered 1, 2, 3, 4, and 5) 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)?

4. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir?

   (Please indicate the probability in percent):

5. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each

6. In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with probability of 5%. What is the probability that a poisonous mushroom in the forest is red?

   (Please indicate the probability in percent):

Next

Questionnaire page 2

Please answer all of the questions below. Your answers will not affect your payment but will help us understand our results.

7. What is your gender?

   --------- ▾

8. What country were you born in?

   --------- ▾

9. What is your age?

Next