Discovered Preferences for Risky and Non-Risky Goods

Jason J. Delaney[a], Sarah A. Jacobson[b,1], Thorsten P. Moenig[c]

[a] *School of Business, Georgia Gwinnett College, 1000 University Center Lane, Lawrenceville, GA 30043, USA, jdelaney@ggc.edu*
[b] *Department of Economics, Williams College, 24 Hopkins Hall Dr., Williamstown, MA 01267, USA, sarah.a.jacobson@williams.edu*
[c] *Department of Mathematics, OSS 201, University of St. Thomas, Saint Paul, MN 55105, USA, thorsten@stthomas.edu*

Abstract

We develop an axiomatic theory that integrates the discovered preference hypothesis into neoclassical microeconomic choice theory. A theory in which preferences must be discovered through experience can explain patterns observed in choice data, including preference reversals, evolution of or instability in risky choice, and errors that decline with repetition as seen in contingent valuation data. With reasonable assumptions, we show that preferences for common, high-ranked, and non-stochastic choice items are learned quickly and thus should appear stable. However, initially low-ranked choice items may remain persistently mis-ranked. Preferences for choice items with stochastic outcomes are difficult to learn, so choice under uncertainty is subject to error. At finite time, a choice item is more likely to be mis-ranked if it has stochastic outcomes, if it is initially low-ranked, or if it appears rarely in choice sets. The existence of a default option may or may not render correct ranking more difficult. Undiscovered preferences can lead to real welfare loss

[1] Corresponding author

as agents make choices not congruent with their true preferences. This theory is amenable to tests using laboratory experiments. Preference discovery has implications for policy, and the process of discovery may contaminate choice data in a variety of contexts.
Keywords: discovered preferences, preference stability, learning, risk preferences
JEL: 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

How do we know what we like? We may be born with complete and correct knowledge of our preferences; on the other hand, perhaps we have to learn them, as suggested by the discovered preference hypothesis (Plott, 1996). Microeconomic theory is built upon an assumption of stable and well-behaved preferences, but if a person must discover those preferences through experience, choice may appear error-prone or unstable over time. This could particularly affect risk preferences: learning one's taste for risk is harder than learning one's taste for other things because an *ex ante* risk is fundamentally different from the *ex post* experience a person has with that risk. The process of preference discovery may explain preference reversals as well as other signs of instability of and evolution in choice that have been demonstrated in studies of risk preferences (e.g., Cox and Grether, 1996; Hey, 2001; Loomes et al, 2002; Loomes and Sugden, 1998; Thaler et al, 1997). It may also explain puzzles in repeated choice in other settings, such as errors that decline with repeated contingent valuation choices (e.g., Kingsley and Brown, 2010; Ladenburg and Olsen, 2008; Bateman et al, 2008) and repeated play of a strategic game even without feedback (Weber, 2003). If preference learning is an important phenomenon, we need to understand how it works because learning may confound the analysis of choice data from the lab and field and because imperfectly-learned preferences may lead people to lose welfare through erroneous choices—and this welfare loss may be avoidable if people can be helped to learn their preferences.

We develop an axiomatic model of preference learning to explore implications of the discovered preference hypothesis. We show that under reasonable assumptions, most choice items will eventually be experienced by the agent and become properly ranked. The exception is that the agent may retain an un-self-correcting negative bias for a small number of items that seem *a priori* unattractive. Items with stochastic outcomes, like lotteries, are harder to learn. At any time, a choice item is more likely to be ranked incorrectly if it is stochastic, appears rarely, or initially seems undesirable. Preferences for rare and undesirable goods are hard to learn because they are tried less often. However, preferences for risky goods are hard to learn because a risky good is an *ex ante* distribution but when the agent tries it she only experiences one *ex post* outcome, so multiple experiences are needed to fully understand her taste for it. These learning difficulties may be exacerbated or mitigated if a default option is available in all choice sets. The effect of a default depends on the how how mis-rankings occur relative to the position of the default.

If preferences must be learned, then choices made at finite time may not represent a person's true preferences, especially in situations that arise infrequently or that involve uncertainty. This means that choice data from an individual in the lab or the field may not reflect a single coherent set of preferences. Most modern economic analysis assumes stable preferences, but learning could confound or mask phenomena of interest in nearly every setting. Preference learning would also mean that people frequently fail to make the choice that would make them happiest, and particularly so in high-stakes situations like housing, labor, and insurance markets where decisions are infrequent and involve substantial risk.

This paper proceeds as follows. In the next section, we lay the foundation for the model. In the section that follows, we explore preference learning when there is no uncertainty. In the following section, we extend the model to include choice items that yield stochastic outcomes. Next, we discuss the relationship between the model and existing literature and we explore some of the model's assumptions, highlighting useful variants for future study. Finally, we conclude with implications of the model and discuss how the model can be tested.

2. Modeling Approach

We model an agent's decision-making. There is a set X of $n \in \mathbb{N}$ distinct choice objects $\{x^1, \ldots, x^n\}$, which we call strategies or choice items. They may be goods, actions, or stochastic experiences. At discrete points in time t, the agent faces a choice set $Y_t \subseteq X$ which contains k_t mutually exclusive strategies. The choice set Y_t will always contain at least $k_{min} \geq 1$ strategies; that is, $k_t \geq k_{min}$ for all t. We require that each possible choice set of size k_{min} has a strictly positive probability of occurring at each time t.[2]

At time t, the agent chooses a strategy $x_t \in Y_t$ according to binary preference relation $\succsim_t$. We assume that she knows what items are in the

[2] This ensures that no strategy x^i's likelihood of appearing is perfectly correlated with another strategy x^j's likelihood of appearing. Such a correlation would not change our fundamental results, but would expand the set of strategies that can remain forever untried.

available choice set when she makes this decision. We denote indifference as $\sim_t$. The relation $\succsim_t$ is an ordering over the choice universe X. The preference relation is subscripted because it may change with experience, but the universe remains constant. After time t (when the agent experiences the strategy x_t that she chose), her preference relation changes to $\succsim_{t+1}$. We assume that at each time t, the preference relation $\succsim_t$ satisfies some of the standard choice axioms, as stated below, though we do not require continuity, monotonicity, or convexity.

Axiom 0. *(Well Behaved Preferences) The preference relation $\succsim_t$ at all times t satisfies the standard axioms of completeness, reflexivity, and transitivity.*

These properties must hold for every time, but since the preference rankings $\succsim_t$ and $\succsim_s$ need not be the same at two times t and $s \neq t$, they need not hold across times. For example, it could be that $x^i \succsim_1 x^j \succsim_1 x^k$ and therefore $x^i \succsim_1 x^k$ but $x^k \succsim_2 x^j \succsim_2 x^i$ so that $x^i \not\succsim_2 x^k$ if $\succsim_1 \neq \succsim_2$.

We also assume separability such that the agent's experience with one strategy does not change her relative ranking between two other strategies. We consider implications of this in the Discussion section.

Axiom 1. *(Separability) Trying a strategy has no causal effect on the relative ranking of two other, different strategies. That is: for any $x^i, x^j \in X$, either a) $x^i \succsim_{t+1} x^j$ for all $x_t \in X \setminus \{x^i, x^j\}$, or b) $x^j \succsim_{t+1} x^i$ for all $x_t \in X \setminus \{x^i, x^j\}$, or c) both (a and b) for all $x_t \in X \setminus \{x^i, x^j\}$.*

This allows rankings to change at time t even for goods that are not experienced at t. If such changes occur, they must happen regardless of which alternative good x_t was chosen, so that there is no causal link between the choice of some good and subsequent rankings of unrelated goods. They might change, for example, if learned preferences are forgotten, as addressed in the Discussion.

These axioms govern preferences $\succsim_t$. The discovered preferences hypothesis requires that there be an ordering $\succsim_\infty$ that is the agent's true, underlying preference ranking.

Axiom 2. *(True Preference Stability) There exists a true preference ranking $\succsim_\infty$, and this true ranking does not change over time.*

A true preference ranking is the ranking that the agent would choose according to if she fully knew her tastes for all choice objects. If the agent's ordering represents true preferences, then she can gain no information about her tastes that would cause her to change any pairwise rankings within that ordering. That is, if $x^i \succsim_\infty x^j$ and $x^i \succsim_t x^j$ and either x^i or x^j is chosen as x_t, that does not cause the ranking to change: in the simplest case, it should be true that $x^i \succsim_{t+1} x^j$.[3] The existence of true preferences also implies that the learning of tastes through experience should generally occur in the direction of correctness: the axioms we define later require that experience with x_t moves ranking $\succsim_t$ toward true preference $\succsim_\infty$.

At time $t = 0$, the subject is fully inexperienced: she has never tried any strategies in the choice universe. By completeness, however, she can

[3]This would not be true in two cases. First, if the agent does not yet know the goods and ranking $\succsim_t$ is only "accidentally" correct, experiencing one of the items could induce learning that causes the ranking to swap, although experiencing both of them would cause the ranking to remain the same. Second, if learning can be forgotten, the agent could forget her taste for the good not experienced in a way that causes the ranking to swap.

rank all choice items. This ranking is an ordering of choice items based on the relative satisfaction she believes each will provide. Does the agent necessarily know how much satisfaction to associate with each item? If she has undertaken a strategy before, she may remember how well she liked it. For example, for food, there is evidence that people learn and retain their tastes for significant consumption items; see Rozin (1982) and Rozin and Vollmecke (1986). However, if she has not previously consumed an item, she has a belief that may be an imperfect prediction of how it will please her. This belief may be informed by an assessment of the strategy's characteristics or by comparisons to similar strategies that she has tried. We make no assumptions about the source of initial rankings and the nature of beliefs the agent has about strategies with which she is not wholly familiar.

After choosing some strategy x^i at time $t = 0$ (so $x_0 = x^i$), the agent experiences x^i. That experience will yield information about how much she likes x^i. She can now better guess how it will please her, and this lets her update her rankings from $\succsim_0$ to $\succsim_1$. In $\succsim_1$, strategy x^i may be ranked higher or lower relative to other strategies as compared to the initial rankings. The assumption of separability ensures that through experience with strategy x^i she learns nothing about strategies x^j, where $j \neq i$, but such "learning spillovers" are a possible extension of this model.

If she fully understands of how much she likes a strategy x^i, then she can compare it correctly with other strategies that she also knows this well. If she fully understands her preference for two strategies x^i and x^j then, faced with a choice between them, she chooses according to true preferences $\succsim_\infty$.

We define C_t as the set of all strategies that the agent has already experienced enough at time t to have fully learned her relative rankings of them (so that choices made according to $\succsim_t$ between elements of C_t are correct, i.e. are indistinguishable from choices governed by $\succsim_\infty$).

Definition 1. *(Set of Correctly Ranked Strategies) C_t is the set of all strategies with which the agent has had sufficient experience at time t so that she ranks them correctly. That is:*

(a) C_0 is the empty set ($C_0 = \varnothing$).

(b) If $x^i \in C_t$, then $x_t = x^i \Longrightarrow x^i \in C_{t+1}$.

(c) $x^i, x^j \in C_t \Longrightarrow x^i \succsim_t x^j \Leftrightarrow x^i \succsim_\infty x^j$.

The first part of the definition of C_t eliminates rankings that start out "accidentally" correct.[4] The second part excludes items whose rankings become "accidentally" correct. "Accidental" correctness occurs when $x^i \succsim_\infty x^j$ and $x^i \succsim_t x^j$ but at time t items x^i and x^j have not been fully learned, so that more experience with x^i might cause the ranking to swap until more experience with x^j is gained, at which point the ranking would be re-corrected. The third part of the definition states that the set contains strategies ranked as they would be ranked by the agent's true preferences.

If the agent were fully aware of her preferences at time t, the preference relation according to which she chose would represent her true preference rankings, and all strategies in choice universe X would be in C_t. However,

[4]Without this requirement, the composition of C_t would sometimes be ambiguous. However, this requirement means that if the agent's initial guess about her rankings is completely correct she still has to try and "learn" each item to fill C_t over time.

in this model we allow her preference relation at time t to differ from her true preferences. In other words, the preference relation $\succsim_t$ that governs her choice at any time t may be incorrect because she may not have learned her true preferences $\succsim_\infty$. The purpose of this model is to demonstrate the functioning of undiscovered preferences and of the learning process by which the operative ordering is updated with experience to reflect true preferences.

Finally, we assume that once preferences are learned, this learning persists. That is, once experience has fully updated the agent's preference ranking of a subset of strategies, she will retain that true ranking and not forget it between experiences with those strategies. In the Discussion section, we consider the possibility of forgetting.

Axiom 3. *(Persistent Memory) The relative ranking of any two strategies does not change between experiences with either of them. That is: if $x^i \succsim_t x^j$, and if $x_t \neq x^i, x^j$, then $x^i \succsim_{t+1} x^j$.*

For generality, we use preference relations rather than a valuation function. However, a valuation function could be used instead. In such an approach, there would be a static utility function at each time t. For stochastic goods, the utility function could be based on expected utility, prospect theory, or any other theory. Experience with a strategy would make the agent update its subjective value, and this updated value could change the ranking of that strategy with regard to other strategies in the choice universe.

3. Preference Learning under Certainty

Let the choice alternatives facing the deciding agent be non-stochastic consumption items, and let her satisfaction from each also be deterministic.

That is, each strategy will provide her a particular level of satisfaction with probability 1. For example, she may be facing a basket of apples, oranges, pears, and bananas. Each apple is of a consistent quality and yields the same consumption experience (and this is also true of the other fruit).[5]

If at time t the agent experiences deterministic-outcome strategy x^i, the satisfaction it yields will be the same satisfaction she can look forward to every time she experiences it. In other words, a single experience suffices for complete learning of her feelings about any deterministic-outcome strategy.

Axiom 4D. *(Full Updating for Deterministic Strategies) For any deterministic-outcome strategy, a single try is sufficient experience for that strategy to become correctly ranked. That is: for any strategy x^i with a deterministic outcome: $x_t = x^i \Longrightarrow x^i \in C_{t+1}$.*

Thus, if the agent experiences two deterministic-outcome strategies, she knows her true ranking of those two strategies. Given Persistent Memory, this also implies that once a strategy is experienced it is at all future times correctly ranked against other strategies (according to her current taste for the other strategies, which may not be her true taste). It also implies that once two deterministic-outcome strategies are experienced, they are at all future times correctly ranked against each other. That is, if deterministic-outcome strategies x^i and x^j have both been experienced by time s, the agent will for all $t > s$ rank them against each other as she would under $\succsim_\infty$.

[5]Obviously, in reality, all apples are stochastic experiences.

3.1. Full and Partial Learning of the Choice Set

At some time T, an agent has had *full experience* when she has experienced all strategies in X. The agent may instead have *full relevant discovery*, in which she may not have experienced all strategies but her expressed preferences have nonetheless converged to her stable underlying preferences.

Definition 2. *(a) (Full Experience) The agent has had full experience if she has experienced all strategies. That is, at time T, she has full experience if for all $x^i \in X$, it is true that $x_t = x^i$ for some $t < T$.*

(b) (Full Relevant Discovery) At time T, the agent has had full relevant discovery if for any choice set $Y_T \subseteq X$ that may appear at time T, it will be the case that the x_T she chooses is truly preferred to all other elements of the choice set Y_T: $x_T \succsim_\infty x^i$ for all $x^i \in Y_T$.

With Persistent Memory and Full Updating for Deterministic Strategies, full experience implies that $\succsim_t = \succsim_\infty$ (and that $C_t = X$) for all $t \geq T$, and that the agent knows this. In other words, once full experience is achieved, all choices will be made according to true preferences $\succsim_\infty$.

Full relevant discovery is a weaker condition. Once full relevant discovery is achieved, as with full experience, all choices will be made according to true preferences $\succsim_\infty$. However, the agent may not have tried all strategies. This can occur at any time (even $t = 0$) without full experience if the agent has a correct guess about rankings of all untried items, in which case there are no mis-rankings. Full relevant discovery also allows for some mis-rankings, but only if all mis-ranked strategies are truly ranked so low that she would never choose them. We define W_t as the set of lowest-ranked strategies at time t.

Definition 3. *(Set of Lowest Ranked Strategies)*

(a) *Let* $W_t \subset X$ *be the set of strategies ranked low enough that they will not be tried at time* t*, even if they appear in the time-*t *choice set. That is: For any* $w^i \in W_t$*, and in every possible choice set, there exists some* $x^j \in X \setminus W_t$ *such that* $x^j \succsim_t w^i$*.*

(b) W_∞ *denotes the set of all truly least preferred strategies, such that for any* $w^i \in W_\infty$*, and in every possible choice set, there exists some* $x^j \in X \setminus W_\infty$ *such that* $x^j \succsim_\infty w^i$*.*

(c) $\bar{w}_t$ *denotes the highest-ranked member of* W_t*. That is,* $\bar{w}_t \succsim_t w^j$ *for all* $w^j \in W_t$*.*

Strategies in W_t may not be tried in future periods, as discussed below, but strategies in $X \setminus W_t$ will generally be tried eventually. In full relevant discovery, the agent may incorrectly rank elements of W_t relative to each other, but by definition this would not affect her choices.

If, some strategy $w^j \in W_t$ would be preferred according to $\succsim_\infty$ to some strategy $x^i \in X \setminus \ W_t$, unless that w^j is removed from W_s at future time s, this incorrect ranking could persist forever. We call the case in which the incorrect ranking persists *false full relevant discovery.*

Definition 4. *(False Full Relevant Discovery) At time* T*, the agent has had false full relevant discovery if for all* $x^i, x^j \in X \setminus \ W_T$*,* $x^i \succsim_T x^j \iff x^i \succsim_\infty x^j$ *but for some* $x^i \in X \setminus \ W_T$ *and some* $w^j \in W_T$*,* $w^j \succsim_\infty x^i$*.*

In false full relevant discovery, the agent is stuck in a condition where she will never learn that she has misjudged some choice items. The distinction between full relevant discovery and *false* full relevant discovery is crucial. Both situations lead to consistency in observed choice, and correspond to what Plott refers to as "stage two" of preference discovery (Plott, 1996) in which choice is stabilized. However, the former yields optimal choices, while the latter yields a stable mis-ordering of strategies that may cause significant welfare losses. How likely is false full relevant discovery? Since all choice items appear in choice sets with positive probability, all choice items $x^i \in X \setminus W_t$ will eventually be tried. Therefore, if W_t is small, then nearly all strategies are tried, learned, and correctly ranked eventually. W_t will be small when k_{min} is small and, if there is a default strategy, if that default strategy is unattractive. The following subsections formalize these points.

3.2. No Default Strategy

Assume that there is no default strategy—no strategy that appears in every choice set. In this case, W_t is of size $k_{min}-1$. If $k_{min} > 1$, W_t is nonempty. As time progresses, the agent will choose and experience strategies, but at any given time t she will never choose a member of W_t.

Members of W_t can leave the set over time, as shown in Figure 1. This happens if experience with some other strategy $x^i \in X \setminus W_t$ teaches the agent that she likes x^i less than $\bar{w}_t$. When this happens, x^i enters W_{t+1}, and $\bar{w}_t$ gets bumped out of W_{t+1}. If some w^j is ranked m places below $\bar{w}_t$ (in Figure 1, $m = 2$), $m+1$ strategies like $x^i, x^k \in X \setminus W_t$ must be demoted to ranks below w^j before w^j is bumped out of W_t. If no $x^i \in X \setminus W_t$ is ever demoted into W_{t+1}, the agent never experiences any members of W_0 (the initially lowest-ranked strategies).

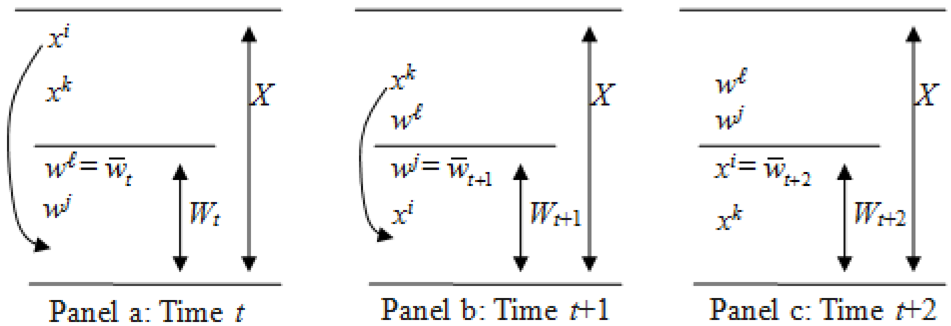

Figure 1. Bumping strategies out of W_t

The following proposition states that eventually, nearly all strategies will be tried and nearly all rankings will become correct.

Proposition 1. *Assume a choice universe that contains only deterministic-outcome strategies and that has no default option.*

(a) All strategies initially ranked outside of the bottom $k_{min}-1$ strategies will be eventually tried. That is: $x^i \notin W_0 \Longrightarrow x^i = x_t$ for some $t \geq 0$.

(b) As $t \to \infty$, at most $k_{min}-1$ items remain incorrectly ranked (and remain outside C_t).

(c) If at some time s, $W_s = W_\infty$, full relevant discovery will be achieved as $t \to \infty$ with probability 1.

(d) If at time r, a strategy x^i is placed in W_{r+1} but is not in W_∞, there will be some time $s > r$ such that x^i will be promoted out of W_{s+1}. It will not return to W_u for any $u > s$.

Proof.

(a) First, imagine that $W_t = W_0$ for all t. Since each choice set contains at least k_{min} strategies, the agent will (with probability approaching 1 as

$t \to \infty$) eventually encounter a choice set containing only W_0 and each strategy x^i outside that set of lowest-ranked strategies for all $x^i \in X \setminus W_0$. Since by definition the members of W_0 are dominated by nonmembers of W_0, she will be forced to choose every $x^i \in X \setminus W_0$. Now note that while W_t does change over time, a strategy cannot enter W_t except through experience with that strategy. Therefore no strategy outside W_0 can remain untried (and thus incorrectly-ranked, by Full Updating for Deterministic Strategies) as $t \to \infty$. $\square$

(b) Since W_0 is of size $k_{min} - 1$, the largest number of strategies that can be incorrectly ranked forever is $k_{min} - 1$. $\square$

(c) The only strategies that can remain untried forever are the strategies in W_0. Therefore, if the strategies in W_0 include only truly-least-preferred strategies, then all other strategies must be tried and their rankings learned. Rankings among members of W_0 may be incorrect, but since the agent is never forced to choose between those strategies, then all of the agent's decisions will be in accordance with $\succsim_\infty$. The same logic follows if $W_s \subseteq W_\infty$ for any $s \geq 0$. $\square$

(d) Because x^i is in W_{r+1} is not in W_∞, the fixed size of W_t for all t implies that another strategy $x^k \in W_\infty$ is not in W_{r+1}. Since $x^k \notin W_{r+1}$, x^k must be tried at some future time $v \geq r + 1$. Since x^i has already been tried, at all $t \geq v + 1$, $x^i, x^k \in C_t$ by Persistent Memory. If $x^i \succsim_\infty x^k$, x^i will be bumped upward when x^k is demoted, although x^i need not be immediately bumped out if $x^i \neq \bar{w}_v$. If instead $x^k \succsim_\infty x^i$, another strategy $x^m = \bar{w}_v$ is bumped out of W_{v+1}. In this way, any strategy that is in W_t during this interval but is incorrectly ranked above x^i will eventually be bumped out of W'_r at some future $t = r'$, which implies that eventually at some time v' it will be tried and found to be less desirable than x^i. Therefore at some time s, x^i must be bumped out of W_{s+1}. This is inevitable because as long as $x^i \in W_t$, there will always be some strategy $x^{k'} \notin W_t$ such that $x^{k'} \in W_\infty$ and therefore trial of $x^{k'}$ will reveal that $x^i \succsim_\infty x^{k'}$, so x^i cannot stop being bumped upward until it leaves W_t. Once x^i is bumped out of W_{s+1}, it will never return for any $t > s$ because, by Persistent Memory, x^i and all of the items the agent incorrectly believed to be better than it will remain in C_t forever, and at no time t will a strategy x^l be moved to a ranking above x^i unless x^l is tried and therefore $x^l \in C_t$ and therefore $x^l \succsim_\infty x^i$, and correct ranking changes like this cannot incorrectly demote x^i into W_t. $\square$

This proposition implies several things. First, if minimal choice set size $k_{min} = 1$, all strategies are eventually tried and preferences eventually converge to true preferences $\succsim_\infty$. Second, the larger k_{min}, the more strategies may remain untried, and the more likely it is that the agent only ever achieves false full relevant discovery. Relatedly, smaller choice sets force fuller and perhaps faster discovery of preferences across the choice universe. Third, there is no reason that the truly most preferred strategy should ever be tried: it could be a element of W_0 that is never bumped out, so that the agent forever misjudges her truly favorite item. Fourth, not all strategies are equally likely to be incorrectly ranked. At any finite time t, the strategies most likely to be tried and therefore correctly ranked appear in many choice sets and are initially high-ranked relative to other strategies. Fifth, when a strategy is persistently mis-ranked (except for mis-rankings among pairs within W_t), the agent always believes it to be less preferred than it actually is—positive biases are corrected with experience, so persistent biases are always negative. Finally, the agent may converge to full relevant discovery even if some elements of W_0 do not belong in that set; but there is no assurance that this will happen because there those strategies need not be "bumped" out of W_t. When at time t the agent tries some $x^i \in X \setminus W_t$ and finds it undesirable, even if it belongs in W_∞, she might demote it to a position above some elements w^j of W_t that are truly preferred to it such that $x^i \succsim_{t+1} w^j$ but $w^j \succsim_\infty x^i$.

To illustrate, suppose the universe consists of different fruits and each choice set is a fruit basket. If the agent must choose a fruit on every choice occasion and choice sets may be as small as one fruit, then she will eventually try all fruit and learn her true ranking of all fruit. Imagine instead that every fruit basket contains at least three fruits, and that she believes durian fruit to be the worst fruit and noni fruit to be the second-worst fruit. She will try and learn to correctly rank all other fruits, but she will never try durian or noni. If this is to her true preference ranking, or even if her true ranking of noni and durian is swapped, then she has achieved full relevant discovery. If instead her true ranking of noni is higher and persimmon (currently at the third-worst spot) belongs in the second-worst spot, then unless a taste of persimmon convinces her that (as-yet-untried) noni must be better, she will maintain an incorrect ranking forever. This is false full relevant discovery: she will never realize her ranking is wrong but her choices will appear internally consistent. She could be losing real welfare, because it is possible that she would get great pleasure from noni if she tried it.

More broadly, we note that choice occasions of this type in life are frequent, so we expect people with some life experience to be able to properly judge their preferences for most choice items.

3.3. Default strategy

Let us define a default strategy.

Definition 5. *(Default strategy) A default strategy x^d is a strategy present in all choice sets, such that $x^d \in Y_t$ for all times t.*

A common default is inaction: instead of choosing a fruit, eating nothing; instead of choosing an insurance policy, driving uninsured; and so on. A default option adds realism to the model. For a default to be non-trivial, it must be voluntarily chosen sometimes; therefore, we only consider default strategies that are ranked outside of the bottom $k_{min}-1$ lowest-ranked strategies according to $\succsim_0$ and $\succsim_\infty$. We now require $k_{min} > 1$ to ensure that the default strategy is never the only option, we require all possible sets of size $k_{min} - 1$ containing members of $X \setminus \ x^d$ to have a positive probability of appearing in the choice set alongside x^d at all times t.

Let t^d be the first time t at which x^d is tried ($x_t = x^d$). Some odd situations can arise in the time before x^d is tried,[6] but we focus on times $t > t^d$, since agents generally have early experience with natural defaults.

Since x^d appears in all choice sets, any strategy that the agent believes to be ranked below x^d at $t > t^d$ will never be tried. Therefore, the set W_t

[6]For example, the order in which strategies appear in choice sets in the periods before t^d can affect which strategies are eventually learned.

no longer has a fixed size: it must have at least $k_{min} - 1$ members, but it now contains all strategies that the agent ranks below x^d at time t. As we show in the proposition below, for $t > t^d$ no strategy can leave W_t since the "bumping" previously discussed depended on the fixed size of W_t when no default option existed. Strategies can still be added to W_t. If a default option is present, full relevant discovery requires the correct ranking of all strategies truly preferred to the default option but does not require correct ranking of any strategies to which the default is truly preferred, since those strategies need never be chosen.

The proposition below outlines properties of preference discovery when a default option is present.

Proposition 2. *Assume a choice universe that contains only deterministic-outcome strategies and that has a default option.*

(a) Strategies can never leave W_t for any $t > t^d$. That is: $W_s \supseteq W_t$ for all $s > t > t^d$.

(b) All strategies ranked above to the default strategy at any time $t > t^d$ will be eventually tried. That is: $x^i \succsim_t x^d \Longrightarrow x_s = x^i$ for some $s \geq t > t^d$ with probability 1.

(c) If at some time $s > t^d$, $W_s \subseteq W_\infty$, full relevant discovery will be achieved as $t \to \infty$ with probability 1.

(d) Full relevant discovery will eventually occur with probability 1 if and only if $\left(x^i \succsim_\infty x^d \Longrightarrow x^i \succsim_{t^d+1} x^d\right)$.

Proof.

(a) A strategy x^i is a member of W_t if $x^d \succsim_t x^i$. In any choice set in which x^i appears, x^d also appears (because it is in all sets). Since $x^d \succsim_t x^i$, x^i would not be chosen at time t. Since $t > t^d$, when the agent tries x^d the experience cannot change her ranking of x^d and x^i. Thus x^i will still be in W_{t+1}. Since this is true for all $t > t^d$, all $x^i \in W_t$ must remain in all W_s for $s > t > t^d$. □

(b) All strategies appear in all choice sets with positive probability. All strategies x^i initially preferred to the default strategy will eventually appear in a choice set with only x^i and x^d (and some member(s) of W_t if $k_{min} > 2$). If $x^i \succsim_t x^d$ for some $t > t^d$ then that ranking will persist at least until x^i is tried. Therefore x^i will be chosen when it appears in such a set. □

(c) The proof of Proposition 2.b shows that all strategies preferred at any time $s > t^d$ to x^d will eventually be tried. Therefore, if at time $s > t^d$ it is true that $W_s \subseteq W_\infty$, all members of X outside W_∞ will eventually be tried (and some members of W_∞ may also be tried). By Persistent Memory and Full Updating for Deterministic Strategies, this achieves full relevant discovery. □

(d) If $x^i \succsim_\infty x^d \implies x^i \succsim_{t^d+1} x^d$ then $W_{t^d+1} \subseteq W_\infty$ since some strategies may be preferred to x^d at $t^d + 1$ but not according to $\succsim_\infty$. Since all strategies outside W_{t^d+1} are eventually tried, then strategies outside W_∞ (and possibly some in W_∞) must eventually be tried. By Persistent Memory and Full Updating for Deterministic Strategies, trying these strategies at some time means that they become and remain fully learned (they are in C_t for all future t). This is full relevant discovery. If, on the other hand, $x^i \succsim_\infty x^d \not\Longrightarrow x^i \succsim_{t^d+1} x^d$, there are some $x^i \succsim_\infty x^d$ where $x^d \succsim_{t^d+1} x^i$—that is, some $x^i \in W_{t^d+1}$ for which $x^i \succsim_\infty x^d$. As shown above, no member of W_t for any $t > t^d$ can ever leave. Therefore, such an x^i will never be tried and will remain forever mis-ranked, but its correct ranking is necessary for full relevant discovery. $\square$

These results show that a default can have a large impact on discovery outcomes, and this impact is greater the more attractive that default option is. Note the key role that the Axiom of Separability plays in these results, particularly the finding that no strategy that starts in set W_t to escape that set.

A default option changes the nature of the set W_t of worst options: it is now absorbing and can achieve virtually unlimited size. Other behavioral research has highlighted the importance of a carefully-chosen default option, for example in investment plans (Madrian and Shea, 2001) and in organ donation (Johnson and Goldstein, 2003). One way to interpret those results in light of preference learning is that agents may undervalue choice items that are not presented as the default.

A key implication of our results is that a default option may render full relevant discovery less likely and false full relevant discovery more likely (and renders full experience impossible unless no choice objects start out less preferred than the default option). A default option may therefore reduce the agent's welfare by letting her avoid sampling items she would actually enjoy. On the other hand, the presence of a default option reduces the number of strategies that must be tried for full relevant discovery to be achieved in the same way that increasing k_{min} does when no default exists. This makes false full relevant discovery less likely and full relevant discovery more likely. The discovery outcome is worse with a default than without if some truly-more-preferred items are initially ranked below the default, and is better with a default than without if $k_m in$ is small and the truly-lowest-ranked strategies are mis-ordered.

To return to our fruit basket example, imagine that the agent may always choose the default of "no fruit." For any choice set size, she will never try any fruit she believes to be worse than eating no fruit. In this case, even if fruit baskets can be as small as a single fruit, she may avoid a whole host of fruits: say, durian, noni, lychee, and kumquat. This ranking could be wrong (reflecting false full relevant discovery), and she could be losing welfare without knowing it. To both the agent and an observer, her choices would appear consistent and stable.

On the other hand, the "no fruit" default partitions the fruit universe into fruits that must and that need not be tried for full relevant discovery. If durian, noni, lychee, and kumquat are truly less preferred than "no fruit," the agent can achieve full relevant discovery without correctly ranking any of them against each other. Without a default, if k_{min} was less than four, at least the truly-most-preferred of these fruits would have to be correctly ranked for full relevant discovery.

It seems likely, however, that the welfare gain from the positive effect of a default is smaller than the welfare loss from the negative effect of a default, since the former involves mis-ranking among low-ranked strategies while the latter could include the mis-ranking strategies that are truly high-ranked. In other words, while we have shown that preference discovery functions mostly just as well if a default option exists in all choice sets, the default may on net cause more welfare loss from missed opportunities.

3.4. Other Facets of Simple Preference Learning

We have shown so far that if preferences need to be learned through experience, preferences for most deterministic-outcome items will be eventually discovered. This means that preferences should look consistent and stable for people with some life experience when they face most choice problems in their lives. However, some rankings take longer to discover—namely items that appear more rarely or are low-ranked—and for these, choice may appear unstable as the agent learns her preferences, revealing symptoms like preference reversals. And some items may remain mis-ranked forever if the agent can avoid experience with them because they appear so unattractive to her. This problem may be exacerbated or mitigated by the presence of a default option. We make several additional observations.

First, the quality of the agent's initial guessed ranking is crucial to her ability to achieve correct rankings. The better she is at guessing her ranking, the closer $\succsim_0$ is to $\succsim_\infty$; the worse she is at guessing, the more likely she will persistently mis-rank items, and the more welfare she will lose from suboptimal choice.

Second, consider a scenario in which an agent can remove items from her choice universe so that she can never choose them in the future, and in which there is some cost to retaining an item in the choice universe—either an explicit or an opportunity cost. For example, to be able to date Chris, the agent may need to sacrifice the opportunity to date Pat in the future.

She will only foreclose on an item she believes to be low-ranked. Once an item has been removed from the choice universe, rankings with regard to that item can never be updated. Since these low rankings may be mistaken, such an excision may eliminate an option that is truly high-ranked.

Third, one might read the model to imply that the agent is myopic in that she will only pick the strategy that she believes gives most satisfaction right now. In our model, in contrast with a model of experimental consumption (Kihlstrom et al, 1984), the agent's belief about the ranking of an item is certain so it may appear that she will not sample untried items to learn about them. However, experimental consumption could be built into the prior rankings in our model. We have been agnostic about the determinants of the agent's rankings at any time; they need not depend solely on instantaneous gratification. They could include a premium that boosts an as-yet-untried item up in the ranking, where the premium represents the value of experimental consumption. Such a premium would make it less likely that any good remains untried forever, but it is no guarantee; there still may be strategies that the agent will still never try because her expectation of the satisfaction they could yield is too low, and this expectation may be incorrect.

The opposite could also be true. Because of risk aversion, ambiguity aversion, status quo bias, or inertia, the agent may avoid strategies that have not yet been tried precisely because of the uncertainty in their proper ranking positions. This would give a discount, rather than a premium, to novel items. In this case, untried items would be less likely to be tried in the future. This would increase the chance that strategies remain improperly ranked.

4. Preference Learning for Stochastic Items

Assume now that the choice universe X contains both simple strategies x^i that yield deterministic outcomes and more complex strategies y^i that yield stochastic outcomes. If the agent chooses a stochastic-outcome strategy y^i, she faces a probability distribution over m possible outcomes z^j: $y^i = (z^1, z^2, \ldots, z^m; p^1, p^2, \ldots, p^m)$. Let these outcomes z^j all be deterministic and let them be common across all strategies (although some may occur with probability zero for some strategies). They may be amounts of money, characteristics of fruit, meal qualities, or other goods. For each stochastic-outcome strategy y^i, the agent knows the probability distribution over the set of potential outcomes. For example, y^i may be a lottery over money outcomes, in which case we assume that the agent knows all possible prizes z^j and the probabilities p^j with which she will win each. Alternatively, y^i may be a restaurant. Each restaurant provides a random quality because the staffs and menus vary in ways unpredictable to the agent. The agent knows the possible meal outcomes z^j she may get at each restaurant and she knows how likely (p^j) each meal outcome is at each restaurant.

We assume that the previously-defined axioms of Well Behaved Preferences, Separability, True Preference Stability, and Persistent Memory still apply when stochastic items exist. We also assume that, even though X contains some stochastic-outcome strategies, Full Updating for Deterministic Strategies holds for the deterministic-outcome strategies in the universe.

We assume that all outcomes z^i are relatively high-ranked members of universe X so that they are never in W_t. By propositions 1 and 2, all z^i are eventually experienced and correctly ranked relative to other experienced strategies. Define time $t = t^z$ to be the time at which all z^i's have been experienced.[7] For example, if stochastic-outcome strategy y^i is a money lottery, at all $t > t^z$ the agent knows how she feels about receiving each money prize for sure. If it is a restaurant, she can rank meal outcomes. We will focus on times $t > t^z$ to demonstrate difficulties in learning preferences for stochastic-outcome strategies even under the most congenial circumstances.

When the agent experiences a new stochastic-outcome strategy y^i, how is her preference learning process different from the process of learning preferences for a new deterministic-outcome strategy x^i? It differs because when she experiences y^i, she does not experience the full *ex ante* distribution that defines y^i. She experiences some anticipation informed by her knowledge of the strategy's objective characteristics (the probability distribution over outcomes), but a single experience of y^i only yields one outcome z^j. In other words, a stochastic-outcome strategy is fundamentally an *ex ante* item, whereas experience with it is largely *ex post*.

For example, if the strategy x_t she chooses at time t is a stochastic restaurant y^i, she will experience anticipation and then the meal will be realized as a meal of particular type and quality z^j. An unsophisticated agent may then update her ranking for y^i so that $y^i \sim z^j$. However, if the agent understands the stochastic nature of y^i, she should consider the anticipation as well as the possible outcomes that did not occur. It need not necessarily even be the

[7] If a default option exists, let $t = t^z$ when all z^i's and x^d have been tried.

case that the realization of a highly ranked outcome (a good meal) causes her to move the item up in her ranking or that a poorly ranked outcome (a bad meal) causes her to move it down. She may realize that a good meal was not worth the uncertainty, or that a bad meal is less unpleasant because it could have turned out to be good.

A sophisticated agent must consider the whole distribution of possible outcomes of a strategy, but on any single experience only one realization of that distribution occurs. Therefore, while we assumed Full Updating for Deterministic Strategies, we cannot reasonably make such an assumption for stochastic-outcome strategies. We instead put structure on the preference-updating process for stochastic strategies in the following axiom.

Axiom 4S. *(Partial Updating for Stochastic Strategies) Let y^i be a strategy with a stochastic outcome. Let q^i be the number of times stochastic-outcome strategy y^i has been experienced. As $q^i \to \infty$, strategy y^i enters set C_t with probability approaching 1.*

This axiom ensures that the concept of true underlying preferences is meaningful in the case of stochastic experiences. If an agent has true preferences but can never learn them, then the idea of preferences has little meaning. This axiom does not preclude full perfect updating, but it embraces the possibility that updating is not full or immediate for stochastic strategies. We put no stronger structure on preference learning for stochastic-outcome strategies, but we note that a sophisticated agent who experiences a risky strategy should update her ranking of that strategy based on the most recent outcome and on her remembered history of outcomes, as well as her knowledge of the objective properties of the uncertainty.

When stochastic items are present, the thoroughness of the eventual learning process is similar to the case with only deterministic outcome strategies.

Proposition 3. *Assume a choice universe that contains deterministic- and stochastic-outcome strategies.*

(a) Assume that the universe does not contain a default option. Propositions 1.a and 1.b hold: all strategies initially ranked outside of the bottom $k_{min} - 1$ strategies will be eventually tried; and no more than $k_{min} - 1$ items will remain incorrectly ranked as $t \to \infty$.

(b) Assume that a deterministic-outcome default option x^d exists. Propositions 2.a, 2.b, and 2.d hold: strategies can never leave W_t; all strategies preferred at any time $t > t^d$ to the default strategy will be eventually tried; and full relevant discovery will eventually occur with probability 1 if and only if ($x^i \succsim_\infty x^d \Longrightarrow x^i \succsim_{t^z+1} x^d$).

Proof.

(a) The proofs are the same as the proofs of Propositions 1.a and 1b. □

(b) The proofs are the same as the proofs of Propositions 2.a, 2.b, and 2.d. □

What is different now? Updates may be incomplete, and rankings may move in the wrong direction after an experience (although not after significant repeated experience) with a stochastic-outcome strategy. Assume that at time $t > t^z$ some deterministic-outcome strategy x^i has been experienced but stochastic-outcome strategy y^j is not fully learned. Imagine further that

$y^j \succsim_t x^i$ but $x^i \succsim_\infty y^j$. If strategy y^i is experienced at time t, it need not move to its proper ranking relative to x^i, and indeed it need not even necessarily move down in the rankings. The agent could have a false impression of her taste for this strategy; for example, if an unusually favorable outcome occurs, she might be fooled into thinking that she likes the risk inherent in the strategy more than she actually does.

Because of this possibility of updating in the wrong direction, experience at time $t > t^z$ can cause a stochastic strategy that belongs outside of the set W_∞ of lowest-ranked strategies to be demoted into W_{t+1} and, unlike the result in Proposition 1.d, that strategy need not ever promoted out of W_s at some future $s > t$. While a try of a deterministic-outcome strategy causes that strategy to be fully learned (to enter C_{t+1}, because of Full Updating for Deterministic Strategies), that is not the case for a stochastic-outcome strategy. In other words, trying all other strategies in the universe would certainly correct the ranking of a deterministic-outcome strategy that had been tried and incorrectly demoted, but need not correct the ranking of a stochastic-outcome strategy similarly maligned. Relatedly, if a default strategy exists, a deterministic-outcome strategy can never for $t > t^z$ be incorrectly demoted into W_{t+1} but a stochastic-outcome strategy can.

Thus, the possibility of mistaken demotion renders it even more difficult to properly rank stochastic-outcome strategies, and renders mis-ranking and the associated welfare-reducing errored choice more likely, even at infinite time. This problem is exacerbated if a default option exists, since in that case strategies can never be bumped out of W_t.

Because of the possibility of incorrect demotion, Propositions 1.c and 2.c do not have parallels for a universe with stochastic items: if $W_s \subseteq W_\infty$ for some time $s > t^z$, we are no longer assured that full relevant discovery will occur. This also means that stochastic-outcome strategies have a greater chance than deterministic-outcome strategies of being forever mis-ranked.

Lemma 1. *Consider a deterministic-outcome strategy x^i and a stochastic-outcome strategy y^i such that $x^i \sim_0 y^i$ and $x^i \sim_\infty y^i$ and that both x^i and y^i are equally likely to appear in any choice set. If there exists $t > t^z$ such that $y^i, x^i \notin W_t$, then for all $u > t$, $Pr(x^i \in C_u) \geq Pr(y^i \in C_u)$.*

Proof. Let S_{x^i} denote the random time when the strategy x^i is experienced for the first time after t (that is, $S_{x^i} = \min\{s > t : x_s = x^i\}$), and let $f_{S_{x^i}}(.) \geq 0$ denote the probability mass function of S_{x^i}. Similarly define S_{y^i}. Propositions 1 and 2 ensure that $Pr(S_{x^i} < \infty) = 1$, and Proposition 3 ensures that $Pr(S_{y^i} < \infty) = 1$. Once S_{x^i} is realized, by Full Updating for Deterministic Strategies, the deterministic-outcome strategy is learned instantaneously: $Pr(x^i \in C_s | S_{x^i} = s) = 1$, and thus (by Persistent Memory) $Pr(x^i \in C_u | S_{x^i} = s) = 1$ for all $u \geq s$. This is not true for the stochastic-outcome strategy. Hence, $Pr(y^i \in C_u | S_{y^i} = s) \leq 1 = Pr(x^i \in C_u | S_{x^i} = s)$ for all $u \geq s$. And therefore, for any $u > t$,

$$\begin{aligned} Pr(y^i \in C_u) &= \textstyle\sum_{s=t+1}^{u} Pr(y^i \in C_u | S_{y^i} = s) \cdot f_{S_{y^i}}(s) \\ &\leq \textstyle\sum_{s=t+1}^{u} Pr(x^i \in C_u | S_{x^i} = s) \cdot f_{S_{x^i}}(s) \\ &= Pr(x^i \in C_u). \end{aligned}$$

□

We can now show that there are differences in the completeness of the agent's expected learning between a universe with only deterministic-outcome strategies and one with stochastic-outcome strategies as well.

Proposition 4. *Assume a choice universe that contains deterministic- and stochastic-outcome strategies. The universe may or may not contain a default option.*

(a) At any time $t > t^z$, a strategy is more likely to be incorrectly ranked if it has stochastic outcomes rather than a deterministic outcome.

(b) Assume that a deterministic-outcome default option x^d exists. Consider two universes: X consists of deterministic-outcome strategies and $\tilde{X}$ is identical except that some subset of the deterministic-outcome strategies have been replaced with stochastic-outcome strategies between which our agent is indifferent under both $\succsim_0$ and $\succsim_\infty$. The set $\tilde{W}_t$ of all strategies believed to be less preferred than x^d will tend to be larger as $t \to \infty$ in the universe which includes stochastic-outcome strategies than will the set W_t in the universe with only deterministic-outcomes strategies.

Proof.

(a) Denote the deterministic-outcome strategy by x^i and its stochastic counterpart by y^i. First, if $x^i, y^i \notin W_{t^z+1}$, then by Lemma 1, $Pr(x^i \in C_s) \geq Pr(y^i \in C_s)$ for all $s > t$. This implies that $Pr(x^i \in C_s) \geq Pr(y^i \in C_s)$. Conversely, suppose that $x^i, y^i \in W_{t^z+1}$. Since either strategy can only be bumped out by experience with other strategies outside W_t at $t > t^z$, both types of strategies are equally likely to be bumped out of W_t. If they always remain in W_t for all $t > t^z$, then neither of them will be experienced, and $Pr(x^i \in C_s) = Pr(y^i \in C_s) = 0$. On the other hand, if both x^i and y^i leave W_t at some time $t > t^z$, then Lemma 1 ensures again that $Pr(x^i \in C_s) \geq Pr(y^i \in C_s)$. If only one of them leaves W_t at a future t, that strategy will become learned while the other one will not, but each is equally likely to be the one that leaves and is learned. $\square$

(b) W_0 for X is the same size as $\tilde{W}_0$ for $\tilde{X}$. The W_t set grows as strategies are demoted into it. By Propositions 2.a and 3.b, strategies that enter W_t when $t > t^z$ will never leave. Both W_t and $\tilde{W}_t$ will expand based on negative but correct impressions of strategies that are tried, but their growth will differ in two ways. First, stochastic-outcome strategies can be incorrectly demoted into $\tilde{W}_t$ for $t > t^z$, while this is not possible for deterministic-outcome strategies. This will tend to make $\tilde{W}_t$ larger than W_t. Second, a stochastic-outcome strategy that belongs in $\tilde{W}_\infty$ may be tried at time t and incorrectly *not* demoted into $\tilde{W}_{t+1}$, and this is also not possible for deterministic-outcome strategies. This will tend to make $\tilde{W}_t$ smaller than W_t. However, by Proposition 2.a (and 3.b), strategies that enter $\tilde{W}_t$ and W_t cannot leave those sets, but by Propositions 2.b and 3.b strategies outside those sets will be tried. Therefore, the second factor (incorrect failure to demote) will be corrected over time while the first factor (incorrect demotion) will not. Therefore, while $\tilde{W}_t$ may be smaller or larger than W_t, as $t \to \infty$ $\tilde{W}_t$ must be larger than W_t.

For the deterministic-outcome universe X, the maximum size of W_t as $t \to \infty$ is the size of W_{t^z+1} plus the number of strategies that are ranked higher than x^d according to $\succsim_{t^z+1}$ but are inferior to x^d according to $\succsim_\infty$. For the universe $\tilde{X}$ with stochastic-outcome strategies, $\tilde{W}_t$ may be arbitrarily larger than that. $\square$

As in the case of deterministic-outcome strategies, a strategy is less likely to be tried and therefore to be properly ranked if it appears infrequently in choice sets or is initially low-ranked. We have shown that a strategy is less likely to be properly ranked if it has stochastic outcomes. Further, this is true at all finite times t and is also true as $t \to \infty$. Thus we can see that false full relevant discovery is more likely when stochastic-outcome strategies are present.

We make a few additional observations about preference learning as regards stochastic-outcome choice items.

First, the presence of a default option interacts with the difficulties associated with learning preferences for stochastic-outcome strategies. For similar reasons, if it is costly to keep items in the choice set, the agent is more likely to "incorrectly" remove a truly-high-ranked stochastic-outcome strategy from the choice set than a similar deterministic-outcome strategy.

Second, regardless of taste for risk, stochastic-outcome strategies are subject to persistent under-ranking in a way that is not possible for deterministic-outcome strategies. This would make an agent appear more risk-averse after extensive experience than her true preferences would dictate. However, at finite time, she may appear either more or less risk-averse; it would depend on her preferences and whether her prior ranking and her updating process show any kind of a bias.

Third, stronger assumptions about the partial learning process would not change our results, but, if justified, could add nuance. For example, we could assume that full learning occurs when the observed frequency distribution of the strategy's realized outcomes approaches the true probability distribution.

Similarly, we could make assumptions about how the agent forms a ranking position for stochastic-outcome strategy y^i based on the set of experiences she has had with it. She could assess this set of experiences as a temporally extended outcome (with elements that occur at different times), forming a ranking position as an average of peak and end sensations as shown in Kahneman et al (1997). Another possibility is some sort of fractional updating, as in Sarin and Vahid (1999). However, such operations cannot be directly applied to a model of preference rankings—only to a model of value functions.

Fourth, some strategies with stochastic outcomes may take more experience than others to fully learn, and these differentials would be informed by assumptions on the learning process. If full learning happens as the observed frequency of outcomes approaches the strategy's true probability distribution, then strategies with longer odds would be harder to properly rank. For example, a coin-flip bet to win \$1 or lose \$1 would be easier to rank than the Powerball lottery, with odds of 1 in 175,223,510. Similarly, if home insurance claims occur at a rate of about 4% each year, then the gamble of living without home insurance may be difficult to correctly rank.

Finally, if the agent has not fully learned her preferences for a strategy, then experience with that strategy will cause some learning, possibly changing her operative ranking. Behaviorally, learning will make the agent look like she has unstable preferences. This could generate preference reversals, for example. Since a stochastic outcome strategy is less likely to be properly ranked than a comparable deterministic outcome strategy at finite time t, preferences for risky items will look more unstable than preferences for deterministic items.

5. Discussion

In preceding sections, we developed a model to formalize the discovered preference hypothesis, which was first laid out in Plott (1996). Our model is congruent with experimental and non-experimental data that generally support assumptions of well-behaved preferences but show some evidence of instability, including preference reversal-type phenomena, especially in unfamiliar domains like environmental valuations and in particular when choosing over stochastic (risky) items. In this section, we connect our model and results to existing literature, and we then explore the implications of some of our assumptions with an eye to possible extensions of the model.

5.1. Literature on Stability of Preferences

Economists tend to reject the idea that preferences are constructive (context-dependent) (as advocated in, e.g., Lichtenstein and Slovic, 2006), favoring models of stable preferences in which constraints may change. This position is exemplified in Stigler and Becker (1977). Stable preferences are analytically tractable and have considerable predictive power.[8]

Stability of preference is not, however, sufficient to ensure stability of choices. People may have stable preferences but may need to learn these preferences. The idea of discovered preferences was articulated in Plott (1996), who proposed that if preferences must be discovered, then people always try

[8]Stability of expressed preference after experience is also possible in the "coherent arbitrariness" framework of Ariely et al (2003). In coherent arbitrariness, however, the value "imprinted" on an agent making a first decision regarding a good is arbitrary and does not represent a fundamental value the agent has for the good.

to optimize but at first know too little to be successful. Repeated choice with feedback would then allow people to arrive at consistent and stable *choices* that correspond to their consistent and stable preferences. Some theories have considered preferences that are context- or state-dependent or that are influenced by institutions but otherwise well-behaved, such as Karni (1985), Bowles (1998), and Pattanaik and Xu (2013). As argued in Andersen et al (2008), the distinction between stationary but state-contingent preferences and preferences that can change is a fine one. Ours is a model of stationary preferences in which choices may appear unstable at times, and in which stationary choices may or may not perfectly correspond to the stable underlying preferences.[9] A model of preference discovery is behaviorally distinct from these models because, among these models, only preference discovery should yield convergence to stable expressed preferences.

One must distinguish preference discovery from learning of objective facts (facts outside the decision-making agent). Braga and Starmer (2005) describe the former as "value learning" and the latter as "institutional learning." Preference discovery is a process of learning individual-specific subjective values. Cubitt et al (2001) suggest that discovered preferences be interpreted in the context of risky choice by positing that there is a true relationship between an experience and a person's affect relating to that experience, but that people may need to learn about that relationship through repeated

[9] Another model that allows expressed preferences (particularly those for risk) to change with experience is proposed in Cohen et al (2008). However, that model suggests that the agent's value function changes: the agent becomes more or less pessimistic (by changing her probability weighting) in reaction to past experience.

consumption.[10]

If preferences are undiscovered in some cases, then in those cases we do not know how much we like the choice items we encounter. As discussed by Kahneman et al (1997), Scitovsky (1976) argued that people are bad at predicting utility from a prospective choice, while Becker (1996) argued the opposite. Kahneman and Snell (1990) note that when experiences are familiar and immediate, people seem fairly good at predicting utility. However, many results from psychology and economics support Scitovsky's claim. Loewenstein and Adler (1995) find people fail to predict changes in tastes, and Wilson and Gilbert (2005) review extensive evidence showing systematic systematic errors in forecasting happiness. Well-known biases aside, we may be particularly bad at predicting utility when experienced utility is largely divorced from the good over which the choice is made (or is not immediate) and when the choice item is unfamiliar (Kahneman and Snell, 1990). We suggest that this may be in part because preference learning is difficult in such cases. Choices over risk and environmental valuation have these characteristics, and preference discovery may be important in these domains.

Experiments using repeated choice provide suggestive evidence of the importance of preference discovery in some situations. With repeated choice, people can converge to a true value that has been induced for them (Noussair et al, 2004). In more complicated cases, errors and biases often decline with

[10]The role of affect in the learning of preferences is distinct from the role of affect in forming preferences. The latter is discussed in affect literature like Isen (2005), who shows that moods affect choice behavior. Such a perspective allows exogenous moods to change an agent's ranking, rather than allowing an agent to learn her own true ranking.

repeated choice. This has been observed with the gap between willingness to pay and willingness to accept (e.g., Coursey et al, 1987; List, 2003; Shogren et al, 2001, 1994)[11], with non-dominant bidding behavior (List and Shogren, 1999), and in strategic games (Weber, 2003). Several contingent valuation (environmental valuation) studies find that errors decline with repeated trials: Kingsley and Brown (2010) find that intransitive choice, preference reversals, and estimated error decline (although apparent preference remains stable), Ladenburg and Olsen (2008) find that starting point bias declines, and Bateman et al (2008) find that the gap between values elicited through single-bounded and double-bounded procedures declines. In the realm of risk preferences, over repeated trials without feedback, Cox and Grether (1996) show that preference reversals decline, although Braga et al (2009) show that further repetition may cause other anomalies. Several studies of repeated lottery experience without feedback (Birnbaum and Schmidt, 2009; Hey, 2001; Loomes et al, 2002) find that estimated error rates or choice inconsistencies decline with repetition. Keren and Wagenaar (1987) and Loomes and Sugden (1998) find a decrease in expected utility violations with repetition; however, Bone et al (1999) find an increase therein. The overall picture from the literature is that repeated choice may, but need not, improve consistency of choice.

Plott (1996) highlights the importance of feedback in repeated choice to allow preference learning. If risk preferences are hard to learn, feedback between choices may be particularly essential for risky choice. However, the

[11]However, Knetsch et al (2001) show that experience in an auction may not decrease but actually increase the WTA-WTP gap.

few risk preference studies that use feedback generally require the agent to learn probabilities (which are institutional features). As we discuss later, it is not clear how learning of the institution should interact with learning of preferences. Thaler et al (1997) find that more frequent feedback causes repeated risky choice to become more risk averse. Barron and Erev (2003) and Hertwig et al (2004) find that subjects who are told lottery odds perform differently than those who learn about the lotteries through experience with feedback. This sort of learning of the value of uncertain prospects is modeled and simulated in March (1996). Jessup et al (2008) describe probabilities to subjects and find that subjects who get feedback move toward choosing according to objective probabilities while those without feedback overweight small probabilities. Van de Kuilen and Wakker (2006) find that repeated trials without feedback don't reduce Allais paradox violations but with feedback the violations do decrease. Weber (2003) find that repeated plays of a strategic game without feedback do exhibit the appearance of learning, although less than that observed with feedback. The contingent valuation studies cited above showed reduced error over repeated choice when feedback is clearly impossible in that context. Overall, feedback seems to increase the consistency of choices, as the discovered preference hypothesis suggests, although improvements happen without it. How can any learning happen at all when there is no feedback, as in the case of contingent valuation studies? We conjecture that introspection, perhaps in the form of mental simulation, is a partial substitute for feedback.

Another hallmark of preference learning could (but need not) be an appearance of preference instability over time. McGlothlin (1956), in what might be the first study of preference stability, finds some stability in aggregate betting on horse races. With regard to risk preferences, Horowitz (1992) finds individual temporal instability but aggregate stability over six weeks, but Harrison et al (2005) find a fair amount of individual stability over four weeks. Zeisberger et al (2012) find that 45% of subjects show some instability in risk preference parameters over a month, although again aggregate statistics appear stable. Over a three month period, Baucells and Villasís (2010) find that 63% of subjects change their answer to at least one of three risk preference questions. Andersen et al (2008) find strong correlation across a person's risk preference parameters across 17 months but significant within-person variance, including some sensitivity to the current financial state of the respondent. Sahm (2012) finds a raw correlation of only 0.18 between gamble decisions within a respondent over a period of years, although some of this variation is certainly noise. Thus, existing evidence shows some preference stability but some drift or change over time, and the changes are sometimes large.

We add to the literature by creating a model of preference discovery that is consistent with the methods and approaches of neoclassical economics but that can present the appearance of unstable preferences because of incomplete learning, particularly in the domain of risk. This model allows careful tests of preference discovery. If preference learning is important, this has implications for researchers, particularly in the realm of risk and uncertainty, and suggests the existence of real welfare losses that could be avoided.

5.2. Literature on Learning and Search

We draw upon literature from economics and psychology about learning processes. Thorndike (1898) established the "law of effect" that underlies theories of learning. Important models include reinforcement learning (Bush and Mosteller, 1955) and fictitious play (Brown, 1951). However, in this context they have mostly been used to model learning of probabilities or dynamics of strategic interactions. Sarin and Vahid (1999) lay out a model of learning in which valuations are updated with experienced outcomes; our model is similar but when applied to risky assets considers the *ex ante* nature of risk. Sampling plays an important role in the learning that occurs in case based decision theory (Gilboa and Schmeidler, 1995), in which an agent confronted with a structurally challenging problem of uncertainty integrates over remembered experiences to determine the best act. This model was more descriptive than reinforcement learning in Ossadnik et al (2012). Sampling is also an essential element of decision field theory (Busemeyer and Townsend, 1993), in which an agent facing risk chooses by "sampling" her impressions of how she will feel if she picks particular actions and particular states of the world result. Preference discovery fits well with these psychology-informed ideas: it includes a sampling of past experience, and can account for a sophisticated agent who understands the objective properties of the choice item. Further, it brings these ideas into a framework that allows for optimization, rather than satisficing (as in case based decision theory) or threshold-breaching (as in decision field theory).

Our model of learning of preferences bears a relationship to models of experimental consumption (Kihlstrom et al, 1984) and the two-armed bandit problem (Rothschild, 1974). In these models, an agent facing unknown objective circumstances samples items to gain information, and may or may not eventually gain full information. Our model differs in that our agent is learning subjective, rather than objective, information, and our agent may not know whether she knows her true preferences.

Our model of learning of subjective "information" (the agent's own preferences) can easily be transformed into a model of learning of objective information about the characteristics of choice items, such as quality or price. As Stigler (1961) discusses, imperfect information and nonzero search costs can have important market implications, and Stiglitz (1979) shows that these include implications for market structure. Our model in such a setting would share the implication of such models that some information is simply never discovered; our model would further imply that producers who are able to reduce the variance of the product characteristic being sampled by the buyer may have a market advantage even if on average they offer a higher price or lower quality than competitors with higher variance.

5.3. Implications and Extensions of the Discovered Preferences Model

First, our model focused on a situation in which the agent acts as if she does not know that she is mis-ranking some strategies: at any time t her ranking is unambiguously $\succsim_t$. The agent could instead be uncertain about how to rank strategies with which she has little experience. Her uncertainty might manifest as a subjective probability distribution of the likelihood a particular ranking is true. If she has not tried x^i or x^j at time t, she may believe that $x^i \succsim_t x^j$ is true with probability p and that the reverse is true with probability $1 - p$. If the agents' preferences were represented by a value function, then an incompletely-learned strategy's believed value would be diffuse rather than degenerate. This is adds a layer of uncertainty to judgments about all unlearned strategies, whether they be deterministic-outcome or stochastic-outcome. As long as the agent behaves as if she has a well-defined subjective probability for any given ranking, a model that accounts for this compounded uncertainty should offer predictions similar to those we have derived, although risk- or ambiguity-averse (or compounding-averse) agents may rank unlearned items lower because of their novelty (as discussed in Section 3.4).

On a related note, our model deals with strategies whose outcomes are stochastic but are specified in terms of a known probability profile—that is, strategies whose outcomes are risky but not ambiguous. However, if the agent does not know the probability distribution across outcomes, experience will yield both value learning (learning of her own rankings) and institutional learning (learning of the probabilities). This is like the case of the two-armed bandit (Rothschild, 1974), and Thaler et al (1997), Barron and Erev (2003), and Hertwig et al (2004) also examine situations of this type, although none of these studies account for a need to learn preferences. It seems likely that for an ambiguous prospect, the agent can never reach fully correct rankings until she knows the strategy's true probability distribution over outcomes. It is unclear how these two learning processes might interact.

We have also assumed throughout that learning persists, i.e. that the agent never forgets preferences she has learned. This model could allow for a decay of what has been learned between experiences with the choice item; this could be a probabilistic decay from the learned ranking toward the original ranking. This will cause rankings to appear unstable over time. Strategies that are more frequently updated to their correct ranking positions will have a greater tendency to be correctly ranked. These strategies are strategies that are more common (appear more frequently in choice sets), are higher-ranked, and are deterministic-outcome rather than stochastic-outcome. Learning of preferences becomes particularly hard for stochastic-outcome strategies when forgetting is possible because incomplete learning and the process of forgetting together make it difficult to reach or maintain correct rankings. On the other hand, if experience at time t moves a stochastic-outcome strategy into the set W_{t+1} where it does not belong, the process of forgetting makes it possible to resurrect that strategy by forgetting this mis-ranking.

The model we develop in this paper intentionally avoids consideration of learning spillovers and failures of separability. Learning spillovers may be important if experience with x^i helps the agent learn about her preference for similar strategies. For example, the agent's first bite of a mango might cause her to update her ranking of (as-yet-untasted) papaya relative to apples and bananas if she thinks of mango and papaya as similar. Separability could fail in other ways, including complementarity or substitutability between goods experienced at the same time or at different times. Consumption of a cookie might change the ranking of milk relative to other goods such as chocolate cake. Also, the agent's starting level of happiness could change her rankings, as suggested by Maslow's hierarchy of needs (Maslow, 1943). In the case of learning spillovers, the true underlying ranking between x^j and x^k does not change (the current preference ranking is simply updated toward that true ranking); with non-separability, the true ranking does change.

While general separability failures would introduce technical issues but provide no additional interesting results, learning spillovers deserve study. If preferences for item x^i must be learned, they can most likely be learned not just through experience with x^i but also through experience with other items and introspection. The importance of these ways of learning is supported by evidence cited above that some learning can happen without feedback, although (as discussed above) feedback clearly is the most effective tutor.

Finally, consider preference discovery in a framework of utility maximization, in which the agent would be learning her (ordinal) value for each item. The agent's utility function gives some coherent structure across values, at least for different quantities of the same good. If a sophisticated agent is not aware of this coherence, perhaps she just thinks of her satisfaction in terms of local approximations. On the other hand, if she is aware of her own coherence, she should essentially only need limited experience with a small number of quantities of each good to learn the parameters of her utility function, and then she should be able to extrapolate to other untried strategies. If goods are simply bundles of characteristics or services and it's only these characteristics or services that provide utility, then much of what the agent learns experiencing one good can also be applied to other goods as long as characteristics are well-known. Learning of preferences could be nearly trivial, although such a model may merely shift the role of experience from value learning to institutional learning—the process of correctly associating characteristics with goods.

6. Conclusion

Preference discovery is a middle ground that balances the analytical power of neoclassical microeconomics with the nuanced view of critiques of assumptions of stable preferences. Our model of preference discovery shows that while normal adults facing common decisions will usually exhibit stable preferences, at other times people might make mistakes because they don't know or are in the process of learning what they like. This model can explain some situations in which individual choice appears unstable, appears to evolve with experience, or seems to indicate choice error, such as preference reversals.

Errors should occur most often in cases in which preferences are harder to learn. We argue in this paper that preferences are hard to learn for stochastic items like lotteries. Preferences should also be hard to learn for environmental assets or characteristics. The choices that respondents are asked to make in environmental valuation studies involve no possibility of feedback and often present states of the world that the agent has never experienced. Therefore, our theory of preference discovery could explain some of the choice anomalies observed in environmental valuation studies.

Unlearned preferences could cause make mistakes that cost people real welfare: people fail to optimize because they don't know what they would most like. We would expect such errors in rarely-made decisions and when facing randomness. Housing markets, insurance choice, and career path choice all have these characteristics—but all three are situations in which a single choice can have very large consequences. If people have unlearned preferences, policy could help reduce welfare loss, including through careful engineering of default options and through training programs that help people introspect through difficult choices.

Preference learning could also interfere with analysis of human behavior. When analysts study observational or experimental data, some choices could be influenced by errors, and some trends could be contaminated by a process of learning. This is particularly a problem if subjects undertake repeated tasks. Because of this, Cubitt et al (2001) use one-task tests to demonstrate violations of expected utility theory that are robust to any discovery process. Without such measures, preference discovery could cause error and even bias in the analysis of wholly unrelated phenomena.

These consequences suggest that preference discovery merits empirical study. Plott (1996) noted that we can observe the discovery process as it happens but that once preferences are learned (in "phase two"), the agent should appear perfectly consistent and rational, so that evidence of preference discovery can no longer be observed. He argues that social institutions and incentives should play important roles in the discovery process, and that the process likely involves repeated experience with feedback. This has two implications.

First, a test of the theory of discovered preferences, as proposed in Plott (1996) and formalized in this paper, should stimulate a learning process using repeated experience with feedback. It should focus on choice items for which an agent might not yet have learned her preferences, such as lotteries. It first should seek evidence of change in choice with experience, and then convergence to a stable choice pattern after more experience. We plan a series of experiments along these lines.

Second, if preferences need to be learned, it may be possible to help agents make better decisions in their everyday lives. We must understand more about the learning process to understand precisely how to help people. Training programs with simulations of the situations of interest could help a homeowner decide what kind of home insurance to buy; interactive tools could help patients assess their willingness to bear risk when choosing among medical procedures.

Acknowledgements

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

7. References

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

Barron G, Erev I (2003) Small feedback-based decisions and their limited correspondence to description-based decisions. Journal of Behavioral Decision Making 16(3):215–233

Bateman IJ, Burgess D, Hutchinson WG, Matthews DI (2008) Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. Journal of Environmental Economics and Management 55(2):127–141

Baucells M, Villasís A (2010) Stability of risk preferences and the reflection effect of prospect theory. Theory and Decision 68(1-2):193–211

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

Birnbaum MH, Schmidt U (2009) The impact of experience on violations of independence and coalescing

Bone J, Hey J, Suckling J (1999) Are groups more (or less) consistent than individuals? Journal of Risk and Uncertainty 18(1):63–81

Bowles S (1998) Endogenous preferences: The cultural consequences of markets and other economic institutions. Journal of Economic Literature 36(1):pp. 75–111, URL http://www.jstor.org/stable/2564952

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

Braga J, Humphrey SJ, Starmer C (2009) Market experience eliminates some anomalies–and creates new ones. European Economic Review 53(4):401–416

Brown GW (1951) Iterative solution of games by fictitious play. In: Koopmans TC (ed) Activity analysis of production and allocation, vol 13, Wiley, New York, pp 374–376

Busemeyer JR, Townsend JT (1993) Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. Psychological Review 100(3):432–459

Bush RR, Mosteller F (1955) Stochastic models for learning. Wiley, New York

Cohen M, Etner J, Jeleva M (2008) Dynamic decision making when risk perception depends on past experience. Theory and Decision 64(2-3):173–192

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

Cubitt RP, Starmer C, Sugden R (2001) Discovered preferences and the experimental evidence of violations of expected utility theory. Journal of Economic Methodology 8(3):385–414

Gilboa I, Schmeidler D (1995) Case-based decision theory. The Quarterly Journal of Economics 110(3):605–639

Harrison GW, Johnson E, McInnes MM, Rutstrm EE (2005) Temporal stability of estimates of risk aversion. Applied Financial Economics Letters 1(1):31–35

Hertwig R, Barron G, Weber EU, Erev I (2004) Decisions from experience and the effect of rare events in risky choice. Psychological Science 15(8):534–539

Hey JD (2001) Does repetition improve consistency? Experimental Economics 4(1):5–54

Horowitz JK (1992) A test of intertemporal consistency. Journal of Economic Behavior and Organization 17(1):171–182

Isen AM (2005) Positive Affect, John Wiley & Sons, Ltd, pp 521–539

Jessup RK, Bishara AJ, Busemeyer JR (2008) Feedback produces divergence from prospect theory in descriptive choice. Psychological Science 19(10):1015–1022

Johnson EJ, Goldstein D (2003) Do defaults save lives? Science 302(5649):1338–1339

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

Karni E (1985) Decision making under uncertainty: the case of state-dependent preferences. Harvard University Press, Cambridge, MA, USA

Keren G, Wagenaar WA (1987) Violation of utility theory in unique and repeated gambles. Journal of Experimental Psychology: Learning, Memory, and Cognition 13(3):387–391

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

Kingsley DC, Brown TC (2010) Preference uncertainty, preference learning, and paired comparison experiments. Land Economics 86(3):530–544

Knetsch JL, Tang FF, Thaler RH (2001) The endowment effect and repeated market trials: Is the Vickrey auction demand revealing? Experimental Economics 4(3):257–269

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

Ladenburg J, Olsen SB (2008) Gender-specific starting point bias in choice experiments: Evidence from an empirical study. Journal of Environmental Economics and Management 56(3):275–285

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

List JA, Shogren JF (1999) Price information and bidding behavior in repeated second-price auctions. American Journal of Agricultural Economics 81(4):942–949

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

Loomes G, Sugden R (1998) Testing different stochastic specifications of risky choice. Economica 65(260):581–598

Loomes G, Moffatt PG, Sugden R (2002) A microeconometric test of alternative stochastic theories of risky choice. Journal of Risk and Uncertainty 24(2):103–130

Madrian BC, Shea DF (2001) The power of suggestion: Inertia in 401(k) participation and savings behavior. The Quarterly Journal of Economics 116(4):1149–1187, DOI 10.1162/003355301753265543, URL http://qje.oxfordjournals.org/content/116/4/1149.abstract, http://qje.oxfordjournals.org/content/116/4/1149.full.pdf+html

March JG (1996) Learning to be risk averse. Psychological Review 103(2):309–319

Maslow AH (1943) A theory of human motivation. Psychological Review 50(4):370

McGlothlin WH (1956) Stability of choices among uncertain alternatives. The American Journal of Psychology 69(4):604–615

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

Ossadnik W, Wilmsmann D, Niemann B (2012) Experimental evidence on case-based decision theory. Theory and Decision pp 1–22

Pattanaik PK, Xu Y (2013) Rationality and context-dependent preferences. Andrew Young School of Policy Studies Research Paper Series (13-13)

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

Rozin P (1982) Human food selection: the interaction of biology, culture and individual experience. In: Barker L (ed) The psychobiology of human food selection, AVI, Westport, Conn, pp 225–254

Rozin P, Vollmecke TA (1986) Food likes and dislikes. Annual Review of Nutrition 6(1):433–456

Sahm CR (2012) How much does risk tolerance change? The Quarterly Journal of Finance 2(04)

Sarin R, Vahid F (1999) Payoff assessments without probabilities: A simple dynamic model of choice. Games and Economic Behavior 28(2):294–309

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

Stigler GJ (1961) The economics of information. Journal of Political Economy 69(3):pp. 213–225, URL http://www.jstor.org/stable/1829263

Stigler GJ, Becker GS (1977) De gustibus non est disputandum. American Economic Review 67(2):76–90

Stiglitz JE (1979) Equilibrium in product markets with imperfect information. American Economic Review 69(2):pp. 339–345, URL http://www.jstor.org/stable/1801670

Thaler RH, Tversky A, Kahneman D, Schwartz A (1997) The effect of myopia and loss aversion on risk taking: An experimental test. The Quarterly Journal of Economics 112(2):647–661

Thorndike EL (1898) Animal intelligence: An experimental study of the associative processes in animals. Psychological Monographs: General and Applied 2(4)

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

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

Zeisberger S, Vrecko D, Langer T (2012) Measuring the time stability of prospect theory preferences. Theory and Decision 72(3):359–386, DOI 10.1007/s11238-010-9234-3, URL http://dx.doi.org/10.1007/s11238-010-9234-3