Published as: "Learning from Experience, Simply." Lin, Song, Juanjuan Zhang and John R. Hauser. Marketing Science Vol. 34, No. 1 (2015): 1-19.
DOI: 10.1287/mksc.2014.0868

Learning from Experience, Simply

by

Song Lin

Juanjuan Zhang

and

John R. Hauser

May 29, 2014

Song Lin is a PhD Candidate, MIT Sloan School of Management, Massachusetts Institute of Technology, E62-582, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 253-7273, songlin@mit.edu.

Juanjuan Zhang is an Associate Professor of Marketing, MIT Sloan School of Management, Massachusetts Institute of Technology, E62-537, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 452-2790, jjzhang@mit.edu.

John R. Hauser is the Kirin Professor of Marketing, MIT Sloan School of Management, Massachusetts Institute of Technology, E62-538, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 253-2929, hauser@mit.edu.

Acknowledgements: The authors gratefully acknowledge the helpful comments from Andrew Ching, Nathan Fong, Jacob Gramlich, Eric Schwartz, Qiaowei Shen, Duncan Simester, Olivier Toubia, Catherine Tucker; seminar participants at Massachusetts Institute of Technology and University of North Carolina at Chapel Hill; and attendees of the 2012 INFORMS International Conference, 2012 Marketing Science Conference, 2013 Allied Social Science Associations Annual Meeting, and 2013 Marketing Dynamics Conference. The authors thank the Editor, Associate Editor, and reviewers for their constructive comments that improved the paper.

Learning From Experience, Simply

Abstract

There is substantial academic interest in modeling consumer experiential learning. However, (approximately) optimal solutions to forward-looking experiential learning problems are complex, limiting their behavioral plausibility and empirical feasibility. We propose that consumers use cognitively simple heuristic strategies. We explore one viable heuristic – index strategies, and demonstrate that they are intuitive, tractable, and plausible. Index strategies are much simpler for consumers to use but provide close-to-optimal utility. They also avoid exponential growth in computational complexity, enabling researchers to study learning models in more-complex situations.

Well-defined index strategies depend upon a structural property called indexability. We prove the indexability of a canonical forward-looking experiential learning model in which consumers learn brand quality while facing random utility shocks. Following an index strategy, consumers develop an index for each brand separately and choose the brand with the highest index. Using synthetic data, we demonstrate that an index strategy achieves nearly optimal utility at substantially lower computational costs. Using IRI data for diapers, we find that an index strategy performs as well as an approximately optimal solution and better than myopic learning. We extend the analysis to incorporate risk aversion, other cognitively simply heuristics, heterogeneous foresight, and an alternative specification of brands.

Keywords: *forward-looking experiential learning, index strategies, structural models, cognitive simplicity, heuristics, multi-armed bandit problems, restless bandit problems, indexability*.

1. Introduction and Motivation

Considerable effort in marketing is devoted to studying the dynamics by which consumers learn from their consumption experience (e.g., Roberts and Urban 1988; Erdem and Keane 1996; Ching, Erdem and Keane 2013a). As an example, imagine new parents who have to shop for diapers, perhaps with little pre-existing knowledge about this category. As these parents find out more about diaper brands through usage experience, they face a strategic choice. They can *exploit* their knowledge to date and select the most appealing brand. They can also *explore* further, which may entail sampling a currently-less-than-ideal brand, so that they can make a more-informed decision in the future.

Researchers have developed theory-rich models of optimizing forward-looking consumers who balance exploitation with exploration. Pillars of these models include an explicitly specified description of consumer utility and an explicitly specified process by which consumers learn. Most models assume consumers choose brands by solving a dynamic program which maximizes expected total utility taking learning into account. Researchers argue that theory-based models are more likely to uncover insight and be invariant for new-domain policy simulations (Chintagunta et al. 2006, p. 604). However, these advantages often come at the expense of difficult problems and time-consuming solution methods.

The dynamic programs for forward-looking experiential learning models are, themselves, extremely difficult to solve optimally. We cite evidence below that the problems are PSPACE-hard – they are at least as hard to solve as *any* problem that requires PSPACE computational memory.[1] This intractability presents both practical and theoretical challenges. Practically, re-

[1] PSPACE is the set of problems which use polynomial-sized memory—memory proportional to $|\Omega|^n$, where $|\Omega|$ is a measure of the size of the problem and n can be extremely large. PSPACE-hard problems are at least as hard as NP-hard problems, which are themselves suspected of being unsolvable in polynomial time.

searchers have had to rely on approximate solutions. Without explicit comparisons to the optimal solution, we do not know the impact of the approximations on estimation results. Moreover, the well-known "curse of dimensionality" prevents researchers from investigating problems with moderate or large numbers of brands or marketing variables, whereby even approximate solutions may not be feasible. Theoretically, it is reasonable to posit that a consumer cannot solve optimally in his or her head a dynamic problem that requires vast amounts of memory and computation. In fact, well-developed theories in marketing, psychology, and economics suggest that observed consumer decision rules are often cognitively simple (e.g., Payne et al. 1988, 1993; Gigerenzer and Goldstein 1996).

We propose that consumers use cognitively simple heuristics to solve learning problems. As an example of the class of cognitively simple heuristics, we investigate an attractive candidate heuristic, *index strategies*, whereby a consumer develops a numerical score, or an index, for each brand separately and then chooses the brand with the largest index. Index strategies are a solution concept that decomposes an intractable problem into a set of tractable sub-problems. We retain basic pillars of structural modeling such as an explicit description of the decision process and an assumption that consumers seek to optimize. We posit in addition a cost to solving complex problems (e.g., Shugan 1980; Johnson and Payne 1985). We assume the consumer chooses a strategy that optimizes expected discounted utility minus this cognitive cost. While the cost of cognitive complexity might be observable in the laboratory, say through response latency, it is unobservable *in vivo*. Instead, we identify domains where index strategies are nearly optimal in the sense of maximizing expected discounted utility. If, in such domains, index strategies are substantially simpler for the consumer to implement, then it is likely that savings in cognitive costs exceed the slight deviation from optimality and, hence, provide the consumer with greater

utility net of cognitive costs. In the special cases where index strategies provide optimal expected utility, we argue that index strategies are superior as a description of forward-looking learning. Following the same logic, we establish conditions where *myopic learning* strategies (i.e., exploiting posterior beliefs without exploration) suffice to model consumer behavior.

To motivate the viability of index strategies as a descriptive model of consumers we (1) establish whether well-defined index strategies exist, (2) explain why they are intuitive and hence might be used by consumers, (3) investigate when index strategies are (near) optimal solutions to the reduced problem of utility maximization, and whether they are computationally simpler than the approximately optimal solution,[2] and (4) test whether index strategies explain observed consumer behavior at least as well as alternative models.

We address #1 analytically by proving the "indexability" property of canonical forward-looking experiential learning models. (Indexability is hard to establish in general.) We address #2 by examining the form and properties of index strategies and arguing they are behaviorally intuitive relative to the approximately optimal solution assumed in most forward-looking learning models. We address #3 using synthetic data. We address #4 by estimating alternative models using IRI data on the purchase of diapers, a product category where we expect to see forward-looking experiential learning.

Our basic hypothesis is that consumers can use *a* cognitively simple index strategy to solve forward-looking experiential learning problems. Figure 1 is a conceptual summary of our hypothesis. We demonstrate viability by showing that there exists a well-defined index that satisfies the four criteria. We do not argue that consumers actually use this well-defined index. Rather we argue that the well-defined index is a better "as if" description than the (approximately) op-

[2] We use computational simplicity as a surrogate for cognitive simplicity in this paper.

timal solution strategy.[3]

We first describe a canonical learning problem. We next review briefly literatures that address learning dynamics, cognitive simplicity, and related optimization problems. We then examine index strategies from the perspectives of theory, synthetic data, and empirical estimation. We close with extensions.

[Insert Figure 1 about here.]

2. Canonical Forward-Looking Experiential Learning Problem

We consider the following *canonical forward-looking experiential learning problem*. A consumer sequentially chooses from a set A containing J brands. Let j index brands and t index purchase occasions. The consumer's utility, u_{jt}, from choosing j at t has three components. The first component is *quality*, q_{jt}, which can be defined to include enjoyment, fit with needs, weighted sum of brand features, etc. Quality is drawn independently from a distribution $F_j(q_{jt};\theta_j)$ with parameters θ_j. The F_j distributions are independent across j. This independence assumption rules out learning about a brand by choosing another. The consumer, however, does not know the value of the parameters θ_j and observes quality draws to infer the value of θ_j. A quality draw of a brand is only realized *after* the consumer has chosen that brand.

The second component of utility is a set of *observable shocks*, $\vec{x}_{jt}$, such as advertising, price, promotion, and other control variables that are observable to the researcher and consumer.[4] For simplicity, we assume that observable shocks affect utility directly, although the model is extendable to indirect effects through the quality component as in Erdem and Keane (1996), Ack-

[3] An empirical search among heuristics would risk exploiting random variation in the data. Instead we demonstrate that an index strategy, and at least one other cognitively simple heuristic, perform well on the data.

[4] Our consumer learning model treats observable utility shocks as exogenous. However, the same insight applies to endogenous observable utility shocks as long as (1) each "atomic" consumer's learning does not affect these shocks (e.g., a brand's advertising expenditure), and (2) these shocks do not directly convey quality information.

erberg (2003), and Narayanan et al. (2005). The third component of utility is an *unobservable shock*, ϵ_{jt}, which represents random fluctuations in realized utility that are observed by the consumer, but not by the researcher.

Consumer decision-making depends on quality and the weighted sum of observable and unobservable shocks, $\vec{\beta}'\vec{x}_{jt} + \epsilon_{jt}$, where $\vec{\beta}$ is a vector of weight parameters. We refer to this weighted sum as *utility shocks*. We let utility shocks be drawn from a joint distribution, $H_j(\vec{x}_{jt}, \epsilon_{jt}; \phi)$, independently over purchase occasions with parameters, ϕ.[5] The H_j's are independent across j. The consumer knows the distribution H_j and the value of ϕ, observes the current utility shocks *prior to* making a purchase decision, but does not know future realizations of the shocks. Notice that, unlike the quality draws, the utility shocks of a brand are realized regardless of whether the consumer has chosen that brand. We make the conservative assumption that utility shocks are independent of q_{jt} and thus do not help the consumer learn quality directly. However, utility shocks do shape learning indirectly by varying the consumer's utility from exploitation, which in turn affects incentive for exploration.

In summary, we write the consumer's utility from choosing brand j at purchase occasion t as follows:

$$u_{jt} = q_{jt} + \vec{\beta}'\vec{x}_{jt} + \epsilon_{jt} \,. \tag{1}$$

For ease of exposition, in the main analysis, we assume that the consumer is risk neutral. We extend the model to incorporate risk aversion in §8.1.

We model each consumer as if the consumer uses Bayes Theorem to update beliefs about

[5] Observable shocks can be independently distributed over purchase occasions for a number of reasons. For example, firms may intentionally randomize price promotions in response to competition. Such "mixed strategies" can generate observed prices that appear to be independently drawn at each purchase occasion from a known distribution (Narasimhan 1988).

the quality parameter θ_j after each consumption experience (assumed to occur after choice but before the next choice). Let s_{jt} be the information set that summarizes the consumer's beliefs about θ_j at purchase occasion t. At $t = 0$, beliefs about θ_j are summarized by a prior distribution, $B_{j0}(\theta_j; s_{j0})$, where s_{j0} is based on all relevant prior experience. After the t^{th} consumption experience the consumer's posterior beliefs are summarized by $B_{jt}(\theta_j; s_{jt})$. When both F_j and prior beliefs are normal, Bayesian updating is naturally conjugate. We obtain $s_{jt} = (\bar{\mu}_{jt}, \bar{\sigma}_{jt})$ using standard updating formulae. The parameters of posterior beliefs, $s_{jt} \in \Omega$ and the realized utility shocks, $\vec{x}_{jt} \in X$ and $\epsilon_{jt} \in E$, summarize the state of information about brand j. The collection of brand-specific states, $(\vec{s}_t, \vec{x}_t, \vec{\epsilon}_t) = (s_{1t}, s_{2t}, \dots, s_{Jt}, \vec{x}_{1t}, \vec{x}_{2t}, \dots, \vec{x}_{Jt}, \epsilon_{1t}, \epsilon_{2t}, \dots, \epsilon_{Jt})$ represents the set of states relevant to the decision problem at t.

We seek to model a decision strategy, Π: $(\Omega \times X \times E)^J \rightarrow A$, that maps the state space to the choice set. Without further assumptions, the consumer must choose a decision strategy to maximize expected discounted utility:

$$V(\vec{s}_t, \vec{x}_t, \vec{\epsilon}_t) = \max_{\Pi} \mathbb{E}_{\Pi} \left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} \left(q_{j\tau} + \vec{\beta}' \vec{x}_{j\tau} + \epsilon_{j\tau} \right) | (\vec{s}_t, \vec{x}_t, \vec{\epsilon}_t) \right], \tag{2}$$

where δ is the discount factor. The expectation $\mathbb{E}$ is taken over the stochastic process generated by the decision strategy (in particular, the transition between states that may depend on the consumer's brand choice). The infinite horizon can be justified either by consumption over a long horizon or by the consumer's subjective belief that the decision problem will end randomly.

The optimal solution to the consumer's decision problem can be characterized as the solution to the Bellman equation

$$V(\vec{s}_t, \vec{x}_t, \vec{\epsilon}_t) = \max_{j \in A} \left\{ \vec{\beta}' \vec{x}_{jt} + \epsilon_{jt} + \mathbb{E}[q_{jt} + \delta V(\vec{s}_{t+1}, \vec{x}_{t+1}, \vec{\epsilon}_{t+1}) | \vec{s}_t, j] \right\}. \tag{3}$$

While the Bellman equation is conceptually simple, the full solution is computationally difficult because, even after integrating out the utility shocks $\vec{x}_t$ and $\vec{\epsilon}_t$, it evolves on a state space of size $|\Omega|^J$, where $|\Omega|$ is the number of elements in Ω. Not only is $|\Omega|^J$ exponential in the number of brands J, it becomes extremely large if Ω contains many elements, even when the optimal solution is approximated by choosing discrete points to represent Ω, as is common in the literature. We provide an illustrative example in §4.

3. Related Literatures

Before we introduce index strategies, it is helpful to review concepts from literatures on learning dynamics, cognitive simplicity, and related optimization problems.

3.1. Learning Dynamics

Many influential papers study consumer learning dynamics and apply learning models to explain or forecast consumer choices in problems related to the canonical learning problem. For example, using data from automotive consumers, Roberts and Urban (1988) estimate a model in which consumers use Bayesian learning to integrate information from a variety of sources to resolve uncertainty about brand quality. Erdem and Keane (1996) build upon the concept of Bayesian learning and include forward-looking consumers who tradeoff exploitation with exploration. For frequently-purchased goods, their model fits data better than a no-learning model (reduced form of Guadagni and Little 1983) and the myopic learning model of Roberts and Urban.

These papers stimulated a line of research that estimates the dynamics of consumer learning – for a comprehensive review see Ching et al. (2013a). Some models focus on myopic consumers with Bayesian learning (e.g., Narayanan et al. 2005; Mehta et al. 2008; Chintagunta et al. 2009; Narayanan and Manchanda 2009; Ching and Ishihara 2010, 2012), while others explicitly model forward-looking consumers (e.g., Ackerberg 2003; Crawford and Shum 2005; Erdem et

al. 2005, 2008). The computational complexity of forward-looking learning has been one of the reasons that some applications assume myopic learning. However, if a theory is accurately descriptive, more-complex forward-looking models should improve policy simulations.

Because forward-looking choice problems that involve continuous state space generally cannot be solved optimally, significant effort has been spent on developing approximate solutions. For example, Keane and Wolpin (1994) use Monte Carlo integration and interpolation, Rust (1997a) introduces a randomization approach, and Imai et al. (2009) develop an estimator that combines dynamic programming solutions with a Bayesian Markov chain Monte Carlo algorithm.[6] While these solution methods vary in speed, all attempt to approximate the Bellman equation to the overall problems and thus may suffer from the curse of dimensionality ($|\Omega|^J$).

At the same time of technical developments, there is a growing recognition of the need for richer theories of consumer behavior. For example, Chintagunta et al. (2006, p. 614) suggest that "the future development of structural models in marketing will focus on the interface between economics and psychology."

3.2. Cognitive Simplicity

Parallel literatures in marketing, psychology, and economics provide evidence that consumers use decision rules that are cognitively simple. In marketing, Payne et al. (1988, 1993) and Bettman et al. (1998) present evidence that consumers use simple heuristic decision rules to evaluate products. For example, under time pressure, consumers often use conjunctive rules (require a few "must have" features) rather than more complicated compensatory rules. Using simulated thinking costs with "elementary information processes," Johnson and Payne (1985) illus-

[6] There is a related literature on neuro-dynamic programming, which uses neural networks and other approximation architectures to overcome the curse of dimensionality (Bertsekas and Tsitsiklis 1996).

trate how heuristic decision rules can be rational when balancing utility and thinking costs. Methods to estimate the parameters of cognitively simple decision rules vary, but such rules often predict difficult consumer decisions as well as or better than compensatory rules (e.g., Bröder 2000; Gilbride and Allenby 2004; Kohli and Jedidi 2007; Yee et al. 2007; Hauser et al. 2010).

Building on Simon's (1995, 1956) theory of bounded rationality, researchers in psychology argue that human beings use cognitively simple rules that are "fast and frugal" (e.g., Gigerenzer and Goldstein 1996; Martignon and Hoffrage 2002). Fast and frugal rules evolve when consumers learn decision rules from experience. Consumers continue to use the decision rules because they lead to good outcomes in familiar environments (Goldstein and Gigerenzer 2002). For example, when judging the size of cities, "take the best" often leads to sound judgments.[7] In 2010-2011, two issues of *Judgment and Decision Making* were devoted to the recognition heuristic alone (e.g., Marewski et al. 2010).

The costly nature of cognition has also received attention in economics (see Camerer 2003 for a review). A line of research looks to extend or revise standard dynamic decision-making models with the explicit recognition that cognition is costly. For example, Gabaix and Laibson (2000) empirically test a behavioral solution to decision-tree problems, whereby decision-makers actively eliminate low-probability branches to simplify the task. Gabaix et al. (2006) develop a "directed cognition model," in which a decision-maker acts as if there is only one more opportunity to search. In the laboratory, the directed cognition model explains subjects' behavior better than a standard search model with costless cognition. Houser, Keane, and McCabe (2004) provide further evidence that consumers might use heuristic rules to solve dynamic programs.

[7] The take-the-best rule is, simply, if you recognize one city and not the other it is likely larger; if you recognize both use the most diagnostic feature to make the choice.

Cognitive process mechanisms are debated in the marketing, psychology, and economics literatures. Our hypothesis, that consumers use heuristics such as index strategies, need only the observation that consumers favor decision rules that are cognitively simple and that such rules often lead to good outcomes. The simplicity hypothesis assumes that consumers trade off utility gains versus cognitive costs, but does not require explicit measurement of cognitive costs.

3.3. Cognitively Simple Solutions to Complex Optimization Problems

If a ball player wants to catch a ball that is already high in the air and traveling directly toward the player, then all the player need do is gaze upon the ball, start running, and adjust his or her speed to maintain a constant gaze angle with the ball (Hutchinson and Gigerenzer 2005, p. 102).[8] The gaze heuristic is an example where a cognitively simple rule accomplishes a task that might otherwise involve solving difficult differential equations. But the principle is more general: simple solutions often perform well in complex optimization problems.

There are many examples in marketing and economics where descriptive decision rules solve more-complex problems.[9] In domains such as consumer-budget allocation, the choice of which information source to search, and the evaluation of products via agendas, heuristic solutions appear to describe consumer behavior well (Hauser 1986; Hauser and Urban 1986; Hauser, et al. 1993). Rust (1997b) argues that it is likely consumers solve problems requiring an "infeasibly large number of calculations" by using heuristic solutions such as decomposition into sub-problems. He states that "[t]he challenge is to recognize whether or not a problem is nearly decomposable, and if so, to identify its approximately independent sub-problems, [and] determine whether they can be solved separately (p. 18)." This view is closely related to our index approach

[8] Professional athletes use more-complicated heuristics that give them greater range, for example, in baseball, pre-positioning based on prior tendencies and the expected pitch, and the sound as the bat hits the ball.
[9] Of course, the empirical performance of descriptive solutions is not guaranteed. Gilovich et al. (2002) provide a comprehensive survey of human decision heuristics and their possible biases.

to complex forward-looking learning problems.

3.4. Related Optimization Problems: Bandit Problems and Index Solutions

The model we formulate in §2 is closely related to the *multi-armed bandit problem*, a prototypical problem that illustrates the fundamental tradeoff between exploration and exploitation in sequential decision making under uncertainty. In a bandit problem, the decision-maker faces a finite number of choices, each of which yields an uncertain payoff. The decision-maker must make choices, observe outcomes, and update beliefs with a sequential decision rule. The decision-maker seeks to maximize expected discounted values.

The bandit problems was first formulated by the British in World War II, and, for over thirty years, no simple solution was known. Then Gittins and Jones (1974) demonstrated a simple index solution – develop an index for each "arm" (i.e., each choice alternative) by solving a sub-problem that involves only that arm, then choose the arm with the largest index. This index solution reduces an exponentially complex problem to a set of one-dimensional problems. Gittins and Jones (1974) proved the surprising result that the index solution is the optimal solution to the classic bandit problem.[10]

However, the Gittins-Jones' striking result comes at the cost of a strict assumption that the states of the non-chosen choice alternatives do not evolve. When this assumption is violated, say due to random shocks, Gittins' index is no longer guaranteed to be optimal. Such problems are known as *restless bandits* (Whittle 1988) and, in general, are computationally intractable (Papadimitriou and Tsitsiklis 1999). In his seminal paper, Whittle (1988) proposed a tractable

[10] Hauser et al. (2009) apply Gittins' index to derive optimal "website morphing" strategies that match website design with customers' cognitive styles. Urban et al. (2014) field-test morphing for AT&T's banner advertising on CNET and General Motors' banner advertising on a variety of websites. Other well-known applications of index strategies include job-match-learning (Jovanovic 1979; Miller 1984) and pharmaceutical-product learning (Dickstein 2012). See Ching et al. (2013a) for a survey.

heuristic solution. The solution generalizes Gittins' index such that the problems can be solved optimally or near optimally by associating an index, referred to as *Whittle's index*, separately with each alternative and choosing the alternative with the largest index.

The existence of well-defined index solutions relies on a structural property called *index-ability*, which is not guaranteed for all restless bandit problems. Whittle (1988, p. 292) wrote that "One would very much like to have simple sufficient conditions for indexability; at the moment, none are known" (see also Niño-Mora 2001). Gittins et al. (2011, p. 154) also lament that "the question of indexability is subtle, and a complete understanding is yet to be achieved." [11] In an important class of marketing models, choice models, consumer utility tends to be restless over purchase occasions. For example, in most random-utility choice models there is an idiosyncratic "error term" as well as other changes in the choice environment (e.g., McFadden 1986).[12] Without further study, we do not know whether an index strategy is a good solution to such restless problems.

We recognize that the canonical forward-looking experiential learning problem belongs to the general class of restless bandits because of the presence of utility shocks. In §5 we prove that the problem is indexable and, thus, a well-defined index solution exists in the sense of Whittle (1988). Moreover, we explore the key properties of such an index, which shed light on how consumers may behave in solving the learning problem.

4. An Index Strategy in the Absence of Utility Shocks

[11] The indexability of restless bandits is problem-specific. For example, Niño-Mora (2001) takes the achievable region approach (Bertsimas and Niño-Mora 2000) and establishes the indexability of a class of restless bandit problems with linear performance measures (e.g., queue input control). Glazebrook, et al. (2006) show that a special class of restless bandit problems – stochastic scheduling – is indexable. To our knowledge, no general result analogous to Gittins' Index Theorem exists as of today.

[12] The error term has been modeled as an unobserved (to the researcher) state variable in structural applications (Rust 1994, Chapter 51, Sections 3.1 to 3.2). This modeling approach "provides a natural way to 'rationalize' discrepancies between observed behavior and the predictions of the discrete decision process model" (Rust 1994, p. 3101). This is different from the "optimal choice plus noise/measurement error" approach.

The learning problem we examine includes utility shocks, but it is easier to illustrate the intuition of index strategies using a problem without utility shocks. Temporarily assume $\vec{x}_{jt} = \epsilon_{jt} = 0$ for all j and t, although the same result holds when there is no intertemporal variation in $\vec{x}_{jt}$ and ϵ_{jt}. In this special case, the consumer's decision problem is a classic multi-arm bandit.

Gittins' insight is as follows. To evaluate a brand j, the consumer thinks as if he or she is choosing between this brand and a reward λ_j that is fixed for all future purchase occasions. The consumer thus solves a sub-problem at each purchase occasion – the consumer can either sample this brand to gain more information about it, or exploit the fixed reward λ_j. In the latter case, the consumer's belief about brand j ceases to evolve, such that $s_{j,t+1} = s_{jt}$. The optimal solution to this sub-problem is determined by a greatly simplified version of the Bellman equation:

$$V(s_{jt}, \lambda_j) = \max\{\lambda_j + \delta V(s_{jt}, \lambda_j), \mathbb{E}[q_j + \delta V(s_{j,t+1}, \lambda_j) | s_{jt}]\}. \quad (4)$$

Notice that each sub-problem only depends on the state evolution of a single brand, j. The sub-problem is much simpler than the full problem specified in Equation 3.

Gittins' index, $G(s_{jt})$, is defined as the smallest value of λ_j such that the consumer at purchase occasion t is just indifferent between experiencing brand j and receiving the fixed reward. That is, we obtain $G(s_{jt})$ by equating the two terms inside the maximization operator of Equation 4. Gittins proposed that $G(s_{jt})$ could be used as a measuring device for the value of exploring brand j – if there is more uncertainty about a brand left to explore, the consumer will demand a higher fixed reward to be willing to stop exploration. Naturally, Gittins' index is updated when new information arrives.

Gittins' surprising result is the Index Theorem. The optimal solution is to choose the

brand with the highest index at each purchase occasion. A computationally difficult problem has thus been decomposed into J simpler sub-problems.

Index Theorem (Gittins and Jones 1974). *The optimal decision strategy when there are no utility shocks is* $\Pi_G(\vec{s}_t) = argmax_{j\in A} G(s_{jt})$.

Figure 2 illustrates intuitive properties of Gittins' index. We consider one brand. The solid line plots one realization of Gittins' index as it evolves when the brand is chosen repeatedly. The dashed line plots the consumer's posterior mean quality belief. It is updated by brand experience and converges toward the true brand quality. Myopic consumers would exploit experience and choose the brand that yields the highest posterior mean quality. Forward-looking consumers may want to explore further. The dotted curve, which is simply the difference between Gittins' index and the posterior mean quality, measures the value of exploration. This curve declines smoothly with experience because the value of exploration decreases as the consumer learns more about brand quality. When we plot Gittins' index as a function of the consumer's posterior quality uncertainty $\bar{\sigma}_{jt}$ (not shown), it is also intuitive – the index increases with $\bar{\sigma}_{jt}$ because the value of exploration increases with the remaining amount of quality uncertainty. Figure 2 and the simple relationship between Gittins' index and posterior quality beliefs suggest that a consumer might intuit something close to the dotted curve if there were no utility shocks.

[Insert Figure 2 about here.]

5. An Index Strategy in the Presence of Utility Shocks

We now allow utility shocks. Observable shocks $\vec{x}_{jt}$ include effects that researchers observe and model, such as changes in advertising, promotion, or price. Unobservable shocks ϵ_{jt} include effects that researchers do not observe and which do not provide a signal about quality.

The presence of unobservable shocks is central to many empirical consumer choice models. Because shocks enter the utility function regardless of the consumer's decisions, the consumer may, in any purchase occasion, switch among brands.[13]

When the model includes utility shocks, the Gittins-Jones Index Theorem no longer applies because the states of non-chosen brands do not remain constant. With shocks, the consumer's problem belongs to the class of restless-bandit problems as introduced by Whittle (1988). In general, such optimization problems are PSPACE-hard (Papadimitriou and Tsitsiklis 1999, Theorem 4) making the problem extremely difficult, if not infeasible, to solve and making it implausible that the consumer would use a solution strategy based on Equation 3. Among other difficulties, PSPACE-hard problems require extremely large memory – a particularly scarce resource for consumers (e.g., Bettman 1979, p. 140; Lindsay and Norman 1977, p. 306). We develop a theoretical solution to this problem in this section. We will show that the canonical forward-looking experiential learning problem is indexable and index strategies have intuitive properties.

5.1. The Canonical Forward-Looking Experiential Learning Problem is Indexable

Whittle (1988) proposed a solution that generalizes Gittins' index. At each purchase occasion, to evaluate a brand j, the consumer thinks as if he or she must choose between brand j and a reward λ_j that is fixed for all future purchase occasions. The Bellman equation for the j^{th} sub-problem, which now includes utility shocks, becomes

(5) $V\left(s_{jt}, \vec{x}_{jt}, \epsilon_{jt}, \lambda_j\right) =$

[13] Even when there is no learning, a typical empirical model of consumer choices may include a shock, or an idiosyncratic error, ϵ_{jt}, that is treated as unobservable by researchers. Without this shock, the model would predict that the consumer makes the same choice over purchase occasions if all other observable factors remain constant. In the context of learning, incorporating this shock allows for switching among brands even when the consumer has learned much about brand quality.

$$\max\{\lambda_j + \delta\mathbb{E}[V(s_{jt}, \vec{x}_{j,t+1}, \epsilon_{j,t+1}, \lambda_j)],\ \vec{\beta}'\vec{x}_{jt} + \epsilon_{jt} + \mathbb{E}[q_j + \delta V(s_{j,t+1}, \vec{x}_{j,t+1}, \epsilon_{j,t+1}, \lambda_j)|s_{jt}]\}.$$

The index is defined as the smallest value of λ_j such that the consumer at purchase occasion t is just indifferent between choosing brand j and receiving the fixed reward. For such an index to be well-defined and meaningful, the indexability condition need to be satisfied (Whittle 1988). Let $S_t(\lambda_j) \subseteq \Omega \times X \times E$ be the set of states for which choosing λ_j at purchase occasion t is optimal:

$$(6) \qquad S_t(\lambda_j) = \{(s_{jt}, \vec{x}_{jt}, \epsilon_{jt}) \in \Omega \times X \times E : \lambda_j + \delta\mathbb{E}[V(s_{j,t}, \vec{x}_{j,t+1}, \epsilon_{j,t+1}, \lambda_j)] \geq \vec{\beta}'\vec{x}_{jt} + \epsilon_{jt} + \mathbb{E}[q_j + \delta V(s_{j,t+1}, \vec{x}_{j,t+1}, \epsilon_{j,t+1}, \lambda_j)|s_{jt}]\}.$$

Indexability is defined as follows:

Definition: *A brand j is indexable if, for any t, $S_t(\lambda_j) \subseteq S_t(\lambda_j')$ for any $\lambda_j < \lambda_j'$.*

Indexability requires that, as the fixed reward increases, the collection of states for which the fixed reward is optimal does not decrease. In other words, if in some state it is optimal to choose the fixed reward, it must also be optimal to choose a higher fixed reward. Indexability implies a consistent ordering of brands for any state, so an index strategy is meaningful. However, indexability need not always hold in general and can not be taken for granted (Whittle 1988).[14] Thus, before we can posit an index strategy as a consumer heuristic, we must establish indexability for a model that includes utility shocks. In Online Appendix A we prove the following proposition.

Proposition 1. *The canonical forward-looking experiential learning problem defined in §2 is indexable*.

[14] For example, Whittle (1988, p. 297) provides a simple example where indexability fails.

Once the indexability condition is established, then a well-defined strategy is to choose at each purchase occasion the brand with the largest index. The index strategy breaks the curse of dimensionality by decomposing a problem with exponential complexity into J much-simpler sub-problems, each on a state space of $|\Omega|$ after integrating out the utility shocks $\vec{x}_{jt}$ and ϵ_{jt}. With this simplification, it is more plausible that the consumer might use the index strategy. As a bonus, estimation is much faster. The difference in the size of the state space can be dramatic. For example, suppose we were interested in the mean and variance of quality and discretized them with M and N grid points, respectively. With J brands, the state space for index strategies is $M \times N$ for each brand, rather than $(M \times N)^J$ for the original optimization problem given in Equation 3. For $M = N = 10$ and $J = 6$, this is the difference between a state space of 100 (for each of six brands) and $1{,}000{,}000{,}000{,}000$.

5.2. The Index Strategy is Invariant to Scale and Behaves Intuitively

Index strategies dramatically simplify the solution, but can the consumer intuit (perhaps approximately) an index strategy? We expect future laboratory experiments to address this issue empirically. In this paper, we argue that index strategies have intuitive properties and that it is not unreasonable for the consumer to intuit those properties.

An index strategy would be difficult for the consumer to use if the strategy were not invariant to permissible scale transformations. If it is invariant the consumer can intuit (or learn) the basic shape of the index function and use that intuited shape in many situations. Invariance facilitates ecological rationality.[15] The following results hold for fairly general distributions of quality, $F_j(q_{jt}; \theta_j)$, and joint distributions of utility shocks $\vec{x}_{jt}$ and ϵ_{jt}, as long as they have scale and location parameters and the quality belief $B_{jt}(\theta_j; s_{jt})$ is conjugate. To ease interpretation, we

[15] Gittins' index exhibits invariance properties (Gittins 1989).

assume that F_j and B_{jt} are normal distributions with parameters defined earlier: μ_j and σ_j for true quality; $\bar{\mu}_{jt}$ and $\bar{\sigma}_{jt}$ for posterior beliefs about quality; and $\mu_j^{x,\epsilon}$ and $\sigma_j^{x,\epsilon}$ for utility shocks. In Online Appendix B we prove the following proposition.

Proposition 2. *Let $\breve{W}$ be Whittle's index for the canonical forward-looking experiential learning problem computed when the posterior mean quality ($\bar{\mu}_{jt}$) is zero, the mean utility shock ($\mu_j^{x,\epsilon}$) is zero, and the inherent variation of quality (σ_j) is 1. Whittle's index for any values of these parameters is the following simple function of $\breve{W}$.*

$$W_j\left(\bar{\mu}_{jt}, \bar{\sigma}_{jt}, \vec{\beta}'\vec{x}_{jt} + \epsilon_{jt}, \sigma_j, \mu_j^{x,\epsilon}, \sigma_j^{x,\epsilon}, \delta\right) = \bar{\mu}_{jt} + \mu_j^{x,\epsilon} + \sigma_j \breve{W}_j(0, \frac{\bar{\sigma}_{jt}}{\sigma_j}, \frac{\vec{\beta}'\vec{x}_{jt} + \epsilon_{jt} - \mu_j^{x,\epsilon}}{\sigma_j}, 1, 0, \frac{\sigma_j^{x,\epsilon}}{\sigma_j}, \delta).$$

Proposition 2 implies that the consumer can simplify his or her mental evaluations by decomposing the index for each brand into (1) the mean utility gained from myopic learning, $\bar{\mu}_{jt} + \mu_j^{x,\epsilon}$, which reflects the exploitation of posterior beliefs, and (2) the incremental benefit of looking forward, $\sigma_j \breve{W}$, which captures quality information gained through exploration. To assess the value of exploration, the consumer need only intuit the shape of $\breve{W}$ for a limited range of parameter values and scale it by σ_j. Proposition 2 also helps researchers understand which parameters can be identified in the index-strategy model.

To provide further intuition, we prove the following proposition in Online Appendix C. The proposition shows that Whittle's index behaves as expected when the parameters of the problem vary. The consumer likes increases in quality and utility shocks, dislikes inherent uncertainty in quality and utility shocks, but values the ability to learn and, hence, resolve the uncertainty in posterior beliefs about quality.

Proposition 3. *Whittle's Index for the canonical forward-looking experiential learning problem (1) increases with the posterior mean of quality ($\bar{\mu}_{jt}$), the observable utility shocks ($\vec{\beta}'\vec{x}_{jt}$), and the unobservable utility shock (ϵ_{jt}), (2) weakly decreases with the inherent uncertainty in quality (σ_j) and the magnitude of uncertainty in the utility shocks ($\sigma_j^{x,\epsilon}$), and (3) increases with the consumer's posterior uncertainty about quality ($\bar{\sigma}_{jt}$).*

Figure 3 illustrates Whittle's index where we set the posterior mean quality to zero, so that the curve represents the value of exploration. (More generally, Whittle's index fluctuates with the posterior mean quality in a way similar to Figure 2.) As was the case for Gittins' index, the value-of-exploration component of Whittle's index is a smooth decreasing function of experience because experience reduces posterior quality uncertainty. With sufficient experience, the value of exploration converges toward zero implying that, asymptotically, the value of a brand is based on the posterior mean of quality (Proposition 2). Unlike Gittins' index, Whittle's index is a function of the magnitude of utility shocks ($\sigma_j^{x,\epsilon}$). As the magnitude of utility shocks becomes larger, it is less important for the consumer to explore, and the value of exploration deceases as shown in Figure 3. These properties and the shape of the curve itself, are intuitive.

[Insert Figure 3 about here.]

Figure 3 and Proposition 3 suggest that, other things being equal, when the magnitude of the uncertainty in utility shocks is larger, the realized utility shocks are more likely to be the deciding factor in consumers' brand choices. For example, as the depth of price promotions increases, consumers are more likely to base their purchase decisions on price. When $\sigma_j^{x,\epsilon} = 5$ (compared with inherent quality uncertainty normalized as $\sigma_j = 1$), Whittle's index is almost flat implying an almost myopic strategy. To formalize this insight, we state the following Corollary

to Proposition 3:

> **Corollary**. *(1) As the consumer's posterior uncertainty in quality increases relative to the magnitude of the utility shocks, the value to the consumer from looking forward increases. (2) As the magnitude of the utility shocks increases relative to the consumer's posterior uncertainty in quality, the value from looking forward decreases. In this latter case, a myopic leaning strategy (i.e., exploiting posterior beliefs) may suffice, and could be the optimal strategy if it is cognitively simpler than a forward-looking learning strategy.*

These results highlight the intricate relationship between the consumer's uncertainty in quality and uncertainty caused by utility shocks. The two types of uncertainty complement each other in driving the consumer's value of exploitation, but may compete with each other in shaping the consumer's value of exploration. The index solution offers an intuitive description of this relationship. In §6 and §7, we examine the empirical performance of the index strategy.

6. Examination of the Near Optimality of an Index Strategy (Synthetic Data)

We now examine whether an index strategy implies a reasonable tradeoff between optimality and simplicity. Indexability guarantees existence of a well-defined index strategy but does not guarantee its optimality.[16] For the canonical forward-looking experiential learning problem, the performance of the index strategy is an empirical question. Cognitive costs remain unobservable, but §4 and §5 suggest that an index strategy could be substantially simpler than the direct solution of the Bellman equation to the overall problem. To examine whether the loss in utility is small, we switch from analytic derivations to synthetic data because the loss in utility is an issue

[16] Many performance bounds have been developed under different contexts and conditions. See Gittins et al. (2011) for a review of recent developments.

of magnitude rather than direction. Synthetic data establish existence (rather than universality) of situations where index strategies are close to optimal.

For concreteness we examine the special case when F_j and B_{jt} are normal distributions. From the perspective of consumer decision-making, what matters is the joint distribution of observable shocks ($\vec{x}_{jt}$) and unobservable shocks (ϵ_{jt}). Therefore, for the synthetic-data analysis we set observable shocks to zero without loss of generality. Practically, even if there are no observable shocks (e.g., no price promotions), unobservable shocks (e.g., idiosyncratic taste fluctuations) are still likely to prevail in most choice models. We allow for both observable and unobservable shocks in the field-data analysis.

We compare four decision strategies that the consumer might use.

1. **No Learning.** In this strategy the consumer chooses the brand based only on the consumer's *prior* beliefs of quality and the current utility shocks. This strategy provides a baseline to evaluate the incremental value of learning.
2. **Myopic Learning**. In this strategy the consumer chooses the brand based only on the consumer's posterior quality beliefs and the current utility shocks. This strategy exploits the consumer's *posterior* knowledge about brand quality. The Corollary predicts that this strategy will suffice when the magnitude of utility shocks is relatively high compared with posterior quality uncertainty.
3. **Index Strategy**. This strategy assumes the consumer can intuit the shape of Whittle's index. As per Proposition 2, this strategy improves on the myopic learning strategy to take into account the exploration value of learning. Brand choices reflect the consumer's tradeoff between exploitation and exploration.
4. **Approximately Optimal**. The PSPACE-hard forward-looking experiential learning prob-

lem cannot be solved optimally, hence researchers resort to approximate solutions (e.g., Keane and Wolpin 1994; Erdem and Keane 1996; Rust 1997a; Ackerberg 2003; Crawford and Shum 2006; Imai et al. 2009; Ching 2010; Ching et al. 2013b). Although approximation methods vary (see Online Appendix G for a review), discrete optimization is a representative method and should converge to the optimal solution with a larger number of grids (Chow and Tsitsiklis 1991; Rust 1996).

We choose parameters that illustrate the phenomena and are empirically plausible. The simulation requires a finite horizon; we select $T = 50$ purchase occasions. If the discount factor is $\delta = 0.90$, truncation to a finite horizon is negligible. We discretize the state space, $s_{jt} = (\bar{\mu}_{jt}, \bar{\sigma}_{jt})$ into a set of $M \times N$ grid points for each of J brands. We choose $M \times N = 200 \times 50 = 10^4$, which should be close to optimal in the continuous problem.[17] To simplify integration we draw the utility shocks from a Gumbel distribution with parameters $(\mu_j^\epsilon, \sigma_j^\epsilon)$ and normalize the location parameter such that the utility shocks have zero unconditional means (Rust 1987, 1994). Inherent uncertainties in quality for both brands, σ_j, are equal and normalized to 1.

The index strategy evolves on a state-space of size $(M \times N)$ for each of the J brands, whereas the approximately optimal solution evolves on a state-space of size $(M \times N)^J$. We choose $J = 2$ for a conservative test of the relative simplicity of the index strategy.

We vary the parameter values to capture three possibilities: (1) the means and uncertainty both favor one brand, (2) the means are the same but uncertainty favors one brand, and (3) the means and uncertainty favor different brands. Because quality beliefs are relative, we fix the pri-

[17] We choose $M = 200$ grid points for the posterior mean quality. Meanwhile, we fix each brand's prior quality variance. Posterior quality variance evolves deterministically following Bayesian updating formulae. Because there are $T = 50$ purchase occasions, a brand's posterior quality variance has $N = T = 50$ possible values, depending on how many times this brand has been chosen. Therefore, the size of the state space for the index strategy is $M \times N = 200 \times 50 = 10^4$.

or mean quality belief of Brand 1 as $\bar{\mu}_{10} = 0$ and vary the prior mean quality belief for Brand 2 as $\bar{\mu}_{20} \in \{-0.3, 0, 0.3\}$. We normalize the standard deviation of Brand 2's prior quality belief as $\bar{\sigma}_{20} = 1$ and the standard deviation of Brand 1's prior quality belief as $\bar{\sigma}_{10} = 0.5$. Finally, to test the Corollary we allow the uncertainty in shocks to vary from relatively small to relatively large: $\sigma_1^{\epsilon} = \sigma_2^{\epsilon} \in \{0.1, 1\}$.

We compute the indices and the consumer's expected total utilities for 50 purchase occasions under the four decision strategies. Details are provided in Online Appendices D and E. Table 1 summarizes the results.

[Insert Table 1 about here.]

We first examine computation times as surrogates for cognitive complexity. As expected, the no-learning and myopic learning strategies impose negligible computation time, the index strategy requires moderate computation time, and the approximately optimal solution is substantially slower – 600 times as time-consuming as the index strategy even for this basic problem. Faster approximation algorithms would reduce the computational time for the approximately optimal solution (Keane and Wolpin 1994; Rust 1997a; Imai, et al. 2009), but they would also expedite the index strategy because we use the same algorithm for solving the Bellman equations in both models (see Online Appendix G for implementation details). Moreover, faster approximation algorithms do not address the curse of dimensionality. The ratio of computational times in Table 1 could be made arbitrarily large with finer grid points or with a larger number of brands.

We next examine the consumer's expected utilities. In all cases, the no-learning strategy leads to the lowest utility, which suggests that learning is valuable. Furthermore, the index strategy is statistically indistinguishable from the approximately optimal strategy. As long as cognitive simplicity matters even a little, the index strategy will be better on utility minus complexity.

Finally, the results are consistent with the Corollary. When there is relatively low uncertainty in utility shocks (upper panel of Table 1), the index strategy and the approximately optimal strategy generate higher utility than myopic learning, and, in two of the three cases, significantly higher utility. When there is relatively high uncertainty in utility shocks (lower panel of Table 1), the myopic learning model performs virtually the same as either the index strategy or the approximately optimal strategy. The differences are not significant. In this case, the consumer might achieve the best utility minus complexity with a myopic strategy, among the models tested.

Analysis of synthetic data never covers all cases. Table 1 is best interpreted as providing evidence that (1) there exist reasonable situations where an index solution is better than the approximately optimal solution on utility minus complexity and (2) there exist domains where myopic learning is best on utility minus complexity. We now examine field data.

7. Field Estimation of an Index Strategy (IRI Data on Diaper Purchases)

We examine how an index solution fits and predicts behaviors compared with an approximately optimal solution and myopic learning. As a first test, we seek a product category and sample where consumers are likely to be forward-looking. Even if an index solution does no better than an approximately optimal solution, we consider the result promising because an index solution is cognitively simpler. As a test of face validity, we expect learning strategies to outperform no-learning strategies and, because we focus on a situation that favors forward-looking behavior, we expect forward-looking strategies to outperform myopic learning.

7.1. IRI Data on Diaper Purchases

We select the diaper category from the IRI Marketing Dataset that is maintained by the

SymphonyIRI Group and available to academic researchers (Bronnenberg, et al. 2008).[18] Diaper consumers are likely to be learning and forward-looking. Parents typically begin purchasing diapers based on a discrete birth event, and their entry to the category is arguably exogenous (Ching et al. 2010, 2012). Even if the birth is a second or subsequent child, diaper quality may have changed. Informal qualitative interviews suggest that parents learn about whether diaper brands match their needs through experience (with often more than one purchase), that diapers are sufficiently important that parents take learning seriously, and that parents often try multiple brands before settling on a favorite brand. In fact, Ching et al. (2012) find that diaper consumers conduct strategic trials of various brands.[19] There are observable shocks due to price promotions and shocks due to unobservable events. For example, a baby might go through a stage where a different brand is best suited to the parent/child's needs. Finally, diapers have the advantage of being regular purchases, where the no-choice option is less of a concern, and consumers tend to be in the market for many purchase occasions.

To isolate a situation favoring forward-looking learning, we apply the following sample screening criteria. First, to focus on consumers whose purchases are likely triggered by a birth event, we select households whose first purchase occurs 30 weeks after the start of data collection (73% of the entire sample). Second, we focus on frequent buyers. Compared with occasional buyers who might be shopping for a baby shower, frequent buyers are more likely to have both the motivation and the opportunity to explore different diaper brands. Therefore, we select households who have made at least 5 purchases during the observation window (39% of the en-

[18] In comparison, durable goods may induce different learning dynamics. Because of the low purchase frequency, consumers may not have the opportunity to learn by sampling. Also, because the stakes are often high, consumers may have the motivation to acquire other types of information (e.g., Consumer Reports reviews) prior to purchase. The ISMS durables goods dataset (Ni et al. 2012) provides a good resource to study these learning dynamics.

[19] Ching et al. (2012) use a quasi-structural approach, where they model the consumer's expected future payoffs as a function of state variables. Their model detects strategic trial if the coefficients of expected future payoffs are significant and if model fit improves significantly over the myopic model.

tire sample).[20] Third, to focus further, we eliminate any consumers who have purchased private labels and restrict attention to consumers who buy exclusively branded products (64% of the entire sample). To the extent that private label buyers are more price sensitive (Hansen, Singh, and Chintagunta 2006), they may be less interested in learning about product quality. (In §8.4, we re-analyze the data by including private labels.) After applying these screening criteria, the data contain 262 households who made 3,379 purchases (13 purchases per household on average).[21] We randomly select 131 households for estimation and 131 households for validation.

The market is dominated by three major brands, Pampers, Huggies, and Luvs. We aggregate all other branded purchases as "Other Brands" and do not model the no-purchase option. As a first-order view, Table 2a compares market-shared-weighted switching behavior during the first 13 purchases with that after the first 13 purchases.[22] There is a noticeable change in switching patterns. For example, the relative brand loyalty of Huggies increases after 13 purchases. This suggests that consumers may learn about brand quality over purchases. Although the category is chosen as a likely test-bed for consumer learning, high brand loyalty, even during the initial 13 purchases, suggests that there is no guarantee a forward-looking strategy will fit the data.

[Insert Table 2 about here.]

[20] Analyses based on a random selection of buyers rather than frequent buyers are available from the authors. Likely because infrequent buyers have less incentive or opportunity to learn, the myopic learning model does better on this random selection of buyers than on frequent buyers.

[21] The data only record the week, as opposed to the exact time, of purchase. Therefore, if a consumer makes multiple purchases during the same week, we do not observe the sequence of brands purchased. Rather than make potentially erroneous assumptions about the data, we remove consumers who make multiple-brand purchases in any week of the observation window (11% of the entire sample). An alternative analysis strategy might have been to randomize purchase orders. However, there is no reason to expect that removing consumers who make multiple purchases a week will affect the comparison between the index strategy and the approximately optimal solution. We also do not model purchase quantity decisions. Instead we assume that consumers update their quality beliefs after each purchase (and consumption) occasion.

[22] We define market share at the purchase level across the observation window, so that market shares before and after the first 13 purchases add up to 100%. For readers who wish to normalize Table 2 in other ways, the raw counts are obtained by multiplying the percentages in Table 2 by 1,407, the total number of purchases in the estimation sample except the last purchase of each household.

7.2. Empirical Specification

We denote households by i and denote by T_i household i's purchase-occasion horizon. We assume that the quality and quality-belief distributions, F_j and B_{ijt}, are normal and that unobservable shock distributions are Gumbel. For this initial test of an index solution, we limit $\vec{x}_{jt}$ to the weekly average prices. The decision strategies are specified below (β and x_{jt} are now scalars).

No Learning: $\Pi_N = argmax_j\{\bar{\mu}_{ij0} + \beta x_{jt} + \epsilon_{ijt}\}$,

Myopic Learning: $\Pi_M = argmax_j\{\bar{\mu}_{ijt} + \beta x_{jt} + \epsilon_{ijt}\}$,

Index Strategy: $\Pi_W = argmax_j\left\{\bar{\mu}_{ijt} + \mu_j^{x,\epsilon} + \sigma_j \widetilde{W}_j\left(0, \frac{\bar{\sigma}_{ijt}}{\sigma_j}, \frac{\beta x_{jt} + \epsilon_{ijt} - \mu_j^{x,\epsilon}}{\sigma_j}, 1, 0, \frac{\sigma_j^{x,\epsilon}}{\sigma_j}, \delta\right)\right\}$,

Approximately Optimal: $\Pi_A = argmax_j\{\bar{\mu}_{ijt} + \beta x_{jt} + \epsilon_{ijt} + \delta\mathbb{E}[V(\vec{s}_{i,t+1}, \vec{x}_{t+1}, \vec{\epsilon}_{i,t+1})|\vec{s}_{it}, j]\}$.

7.3. Issues of Identification

Although we would like to identify all parameters of the various models, we cannot do so from choice data alone because utility is only specified to an affine transformation, and because the parameters that matter are relative parameters. For the no-learning model we can identify only the relative means of prior beliefs, as well as the price sensitivity parameter β. For the myopic learning model we can identify only the relative means of prior beliefs, the relative uncertainties of prior beliefs, the true means of quality, and price sensitivity. For the no-learning and myopic learning models time discounting does not matter.

For the index strategy and approximately optimal strategy we set the mean prior belief of one brand ($\bar{\mu}_{i10}$) to zero and normalize its variance of quality (σ_1) to one to set the scale of quality. (Only $\bar{\sigma}_{ij0}/\sigma_j$ matters.) We cannot simultaneously identify a brand-specific mean of quality and a brand-specific mean of the unobservable shock, so we set the latter to zero ($\mu_j^{\epsilon} = 0$). The

standard deviation of x_{jt} is observed in the data. We can then compute σ_j^{ϵ} from $\sigma_j^{x,\epsilon}$ because the observable and unobservable shocks are independent. As in most dynamic discrete choice processes (Rust 1994), the discount factor δ is difficult to estimate; we set it to 0.90.[23]

Finally, as in Erdem and Keane (1996), we suppress "parameter heterogeneity" among households. We continue to allow each household's quality beliefs to evolve idiosyncratically, but we do not attempt to estimate heterogeneity in prior beliefs, true mean quality, or the magnitude of utility shocks. We abstract away from parameter heterogeneity for the following reasons. First, there are, on average, only 13 purchases per household. We would overly strain the model by attempting to estimate heterogeneity in all of the parameters.[24] Second, we wish to focus on behavioral heterogeneity that arises endogenously from forward-looking learning. Even if households start with exogenously homogeneous prior beliefs, different quality realizations and utility shocks lead to different posterior beliefs, different exploitation-versus-exploration tradeoffs, and different learning paths (e.g., Ching, et al. 2013a). We seek to evaluate heterogeneous learning dynamics based on the data, rather than using heterogeneous parameters to fit the data. For an initial test of an index strategy, this simplification is conservative because it biases against a good model fit.

We estimate each model's parameters with maximum simulated likelihood estimation. Estimation details are provided in Online Appendices F and G.

7.4. Estimation Results

[23] Sensitivity analyses with other discount rates (e.g., 0.95 and 0.99) yield almost identical log-likelihood statistics and similar parameter estimates for the index strategy model. Anticipating the results of §7.4, we expect a similar lack of sensitivity for the approximately optimal strategy. The ease with which such sensitivity checks can be run is a benefit of the computational tractability of the index strategy model.

[24] Doing so is technically feasible, but would likely over-parameterize the model and exploit noise in the data. More importantly, our goal is to demonstrate that an index solution is a viable representation of cognitive simplicity and that cognitive simplicity is a phenomenon worth studying in structural models. We leave explicit modeling of parameter heterogeneity to future research. §8.3 explores foresight heterogeneity.

Table 3 summarizes the fit statistics for the 1,538 diaper purchases in the in-sample estimation and the 1,841 purchases in the out-of-sample validation. U^2 is an information-theoretic measure that calculates the percent of uncertainty explained by the model (Hauser 1978); AIC and BIC attempt to correct the likelihood function based on the number of parameters in the in-sample estimation, BIC more so than AIC. (There are no free parameters in the out-of-sample validation.)

For comparability, we estimated the index strategy in two ways. The first estimation discretizes the state space in the same manner as the approximately optimal model ($M = N = 5$). This enables an "apples-to-apples" comparison. Then, because the index model does not suffer from the curse of dimensionality, we re-estimate the model with a finer grid ($M = 200, N = 75$). There were only trivial differences. For example, $U^2 = 88.18\%$ for both estimations, and parameter values are not significantly different (nor different from the approximately optimal model). We will report the results associated with the finer grid for the rest of the paper.

[Insert Table 3 about here.]

First, on all measures there are sizable gains to learning – all learning models explain and predict brand choices substantially better than the no-learning strategy. Second, the index strategy improves in-sample fit and out-of-sample predictions relative to myopic learning. The likelihood is significantly better (Vuong test significance is $p = 0.0002$ in-sample and 0.0429 out-of-sample).[25] This result is consistent with our expectation that frequent buyers of branded diapers are forward-looking. Third, the index strategy performs as well as the approximately optimal solution in terms of both in-sample fit and out-of-sample predictions. This result is consistent with the synthetic-data analysis – when two strategies yield almost the same expected utilities and

[25] We use the Vuong test to compare non-nested models (Vuong 1989).

hence predict almost the same brand choices, they are observationally equivalent and statistically indistinguishable.

As a further visualization of model fit, Table 2b reports the predicted market-share-weighted switching patterns. The predicted switching patterns are qualitatively similar to actual switching patterns in Table 2a. For example, the index strategy model picks up the fact that consumers are more loyal to Huggies than to the other brands because, as we will discuss below, the true mean quality is higher for Huggies and it is likely more rewarding to learn about Huggies ($\sigma_j^{x,\epsilon}$ being relatively small). Although predictions are not perfect and could be improved if other x-variables were observed, the overall MAE is within 0.8% of actual switching. Moreover, the predicted switching patterns from the index strategy model are virtually identical to those from the approximately optimal solution model (reported in Online Appendix H). The market-share-weighted mean absolute error (MAE) is approximately $3/100^{\text{ths}}$ of 1%.

Table 4 summarizes the estimated parameter values. As expected, the price sensitivity coefficient is negative in every model. Across all learning models, all four brands increase in mean quality relative to prior beliefs, which implies that diaper buyers learn to purchase these brands more through experience. These results are consistent with the switching patterns in Table 2a.

[Insert Table 4 about here.]

Forward-looking models identify the magnitude of utility shocks relative to inherent quality uncertainty (last panel of Table 4). Because the relative shock uncertainty varies across brands, the index curve implies different behavior than myopic learning for those brands. This explains why forward-looking models fit and predict better than myopic learning. For example, Huggies has lower relative shock uncertainty than other brands, which may provide greater in-

centives for consumers to explore Huggies. Because the myopic learning model ignores this difference, it compensate by overestimating the mean prior belief of Huggies. Managerially, Huggies has a higher true mean quality than Pampers and Luvs, but also higher inherent relative uncertainty in quality across consumption. (The table reports the ratio of shock uncertainty to quality uncertainty – a smaller number means higher relative quality uncertainty.)

Both the index strategy and the approximately optimal strategy lead to similar parameter estimates. Parameter estimates of either model are usually within confidence regions of the alternative model. This result is consistent with the synthetic-data analysis, which suggests that both strategies lead to near optimal utility. The index strategy will be a more plausible description of consumer behavior if it is cognitively simpler. We explore this last point below.

Computation time in the embedded optimization problem is one surrogate for cognitive complexity. The last row of Table 3 reports the time necessary to compute one likelihood function in each model. For the index strategy model we report the computation time for both the original grid ($M = N = 5$) and the finer grid ($M = 200, N = 75$) – the latter is in parentheses. Consistent with the synthetic-data analysis, the index strategy is substantially faster than the approximately optimal strategy (74-to-1 ratio based on the same grid density of $M = N = 5$).

The size of the state space is another surrogate for cognitive complexity (e.g., a consumer's memory). The state space for the approximately optimal strategy is 15,625 times as large as the state space for the index strategy given the same grid density of $M = N = 5$. Computational-time ratios are not equal to state-space ratios due to computational overhead. Nonetheless, if we were to attempt to use the finer grid of $M = 200, N = 75$ for the approximately optimal strategy, we would increase the state space of the approximately optimal solution by a factor of 130 billion. It is unlikely that approximately optimal computations would be feasible for the finer grid.

Detailed calculations are presented in Online Appendix G.

In summary, using IRI data on diaper purchases we find that (1) learning models fit and predict substantially better than the no-learning model; (2) forward-looking learning models fit and predict significantly better than the myopic learning model; (3) the index strategy and the approximately optimal solution achieve similar in-sample fit and out-of-sample forecasts, as well as reasonably close parameter estimates; and (4) computational (and cognitive) simplicity favors the index strategy model relative to the approximately optimal model.

8. Further Explorations

We have shown that the canonical forward-looking experiential learning model is index-able and that an index strategy performs well. We now extend the analysis to explore consumer risk aversion, other cognitively simple heuristics, heterogeneous consumer foresight, and private-labels.

8.1. Risk Aversion

For ease of exposition, in previous sections we assumed that consumers are risk neutral. However, risk aversion can be an important issue for decision-making under uncertainty (see Ching et al. 2013a for a review). We generalize our model to incorporate risk aversion following the standard discounted-utility approach (e.g., Samuelson 1937; Erdem and Keane 1996). At each purchase occasion t, the consumer maximizes $\sum_{\tau=t}^{\infty} \delta^{\tau-t} u(w_\tau)$, where w_τ is the net payoff the consumer receives at purchase occasion τ, and $u(\cdot)$ is the consumer's utility function. Utility increases with net payoff (i.e., $u' > 0$). In addition, the curvature of the utility function captures general risk preferences: the consumer is risk neutral if $u'' = 0$, risk averse if $u'' < 0$, and risk seeking if $u'' > 0$. The Bellman equation for the sub-problem of the j^{th} brand (Equation 5) is generalized as:

$$(7)\quad V(s_{jt},\vec{x}_{jt},\epsilon_{jt},\lambda_j) = \max\begin{Bmatrix} u(\lambda_j)+\delta\mathbb{E}\left[V(s_{jt},\vec{x}_{j,t+1},\epsilon_{j,t+1},\lambda_j)\right], \\ \mathbb{E}\left[u(\vec{\beta}'\vec{x}_{jt}+\epsilon_{jt}+q_j)|s_{jt}\right]+\delta\mathbb{E}\left[V(s_{j,t+1},\vec{x}_{j,t+1},\epsilon_{j,t+1},\lambda_j)|s_{jt}\right] \end{Bmatrix}$$

We prove in Online Appendix A.2 that the generalized canonical forward-looking experiential learning model is indexable. This is true for all consumer utility functions satisfying $u' > 0$.

The indexability result allows us to test for risk aversion at low computational costs. We do so using the diaper data. To parameterize the test, we assume that consumers exhibit constant absolute risk aversion: $u(w) = 1 - e^{-rw}$, where $r > 0$ measures the degree of risk aversion (e.g., Roberts and Urban 1988).[26] Based on this utility function, we re-estimate the index strategy model. In addition, we re-estimate the myopic learning model to see whether general risk preferences as opposed to the exploration incentive suffice to explain consumer choices.[27]

Table 3 reports the fit statistics. Allowing for risk aversion brings little improvement to the likelihood and U^2, and worsens the AIC and BIC because of the extra risk-aversion parameter. Table 4 reports the parameter estimates. The risk-aversion parameter is insignificant for the myopic learning model; it is marginally significant for the index strategy model but the magnitude is small. Diaper buyers in our sample do not seem to be strongly risk averse. Because the approximately optimal model provides parameter estimates that are close to the index model for the risk neutral case, we expect similar results if we were to estimate the approximately optimal model for the risk averse case.

8.2. Other Cognitively Simple Heuristics

The canonical forward-looking experiential learning model assumes that consumers have perfect foresight. But the degree of foresight is an empirical question. Ho and Chong (2003) find

[26] For constant risk aversion, $u(w) \to w$ as $r \to 0$. Erdem and Keane (1996) express risk aversion with a quadratic utility function.
[27] The risk aversion parameter cannot be separately identified from the mean prior beliefs in the no-learning model.

that a parsimonious myopic model accurately describes and predicts SKU demand.[28] Models in which the decision-maker looks one period ahead sometimes explain choices well (Hauser et al. 1993; Gabaix et al. 2006; Che et al. 2007). In the bandit literature, Ny and Feron (2006) explore one-period look-ahead heuristics as approximate solutions to restless bandits with switching costs. A one-period look-ahead model is arguably simpler than the full dynamic optimization problem, and is a heuristic consumers might use. For a one-period look-ahead model, the Bellman equation (Equation 3) is modified as:

$$V(\vec{s}_t, \vec{x}_t, \vec{\epsilon}_t) = \max_{j \in A} \left\{ \vec{\beta}' \vec{x}_{jt} + \epsilon_{jt} + \mathbb{E}\left[q_{jt} + \delta \max_{k \in A} \{ q_{k,t+1} + \vec{\beta}' \vec{x}_{k,t+1} + \epsilon_{k,t+1} \} \middle| \vec{s}_t, j \right] \right\}. \quad (8)$$

Tables 3 and 4 report the empirical results. The one-period look-ahead model has worse in-sample fit than the index strategy model (Vuong test $p = 0.0244$) and approximately the same out-of-sample prediction (Vuong test $p = 0.2529$). The one-period look-ahead model fits better than the myopic learning model both in-sample (Vuong test $p = 0.0033$) and out-of-sample (Vuong test $p = 0.0009$). These results suggest that diaper consumers are not myopic, although they may not be perfectly forward-looking.

We could easily estimate a variety of cognitively simple heuristics including T_l-period look-ahead models for $T_l < T$, Gittins'-index models modified to allow for utility shocks,[29] and various heuristics such as those proposed by Bertsimas and Niño-Mora (2000). For example, a modified-Gittins'-index model ($U^2 = 88.09\%$ in-sample; $U^2 = 88.84\%$ out-of-sample) does better than myopic learning, but not as well as the index strategy model (which is based on Whittle's index). We strongly caution against choosing a best-predicting model based on a single da-

[28] The model was used by Procter & Gamble to predict SKU purchases.

[29] Specifically, the modified Gittins' index assumes that the consumer i at purchase occasion t chooses the brand with the highest value of $G(s_{jt}) + \vec{\beta}' \vec{x}_{jt} + \epsilon_{jt}$, where $G(s_{jt})$ is Gittins' index derived from the optimization problem in the absence of utility shocks (see §4). The modified Gittins' index is an ad hoc solution compared with the Whittle's-index model.

taset. Unrestricted search among models would likely exploit random variation. However, from the good fit and predictive ability of the three tested heuristics (Whittle's index, one-period look-ahead, and modified Gittins' index), we are comfortable in our hypotheses that (1) cognitively simple heuristics are plausible alternatives to modeling forward-looking behavior and (2) an index strategy is one viable model.

8.3. Heterogeneous Foresight

In an alternative approach we allow for heterogeneous consumer foresight. We assume there are two latent consumer segments that represent the two "extremes" of the foresight spectrum. One segment engages in myopic learning and the other segment has perfect foresight. Because the index strategy and the approximately optimal solution are observationally indistinguishable, we assume that the perfect-foresight segment follows the computationally favorable index strategy. We use the latent class method (Kamakura and Russell 1989) to estimate the fraction of consumers belonging to each segment, as well as the set of parameters associated with each segment.

Not surprisingly, as Table 3 shows, the latent class model generates higher likelihood and U^2 than both the myopic learning model ($p < 0.0001$ in-sample; $p < 0.0001$ out-of-sample) and the index strategy model ($p = 0.0003$ in-sample; $p = 0.018$ out-of-sample). The flexibility of the latent class model comes at the cost of extra parameters. It produces a slightly better AIC but a worse BIC than the index strategy model.

The last two columns of Table 4 report the parameter estimates of the latent class model. The parameter estimates associated with the respective segments are comparable to values in the homogeneous models. Meanwhile, 77% of diaper buyers are forward-looking. This finding echoes the result from the one-period look-ahead model that the average diaper buyers is neither myopic nor perfectly forward-looking.

8.4. Private Labels

Our primary analyses eliminated any consumer who purchased a private label during the observation window. This restriction allowed us to focus on a case where we expected forward-looking learning. The decision was also driven by the curse of dimensionality inherent in the approximately optimal solution – adding another brand increases the size of the state space by $M \times N = 25$ times. As a robustness check, we repeat our estimations replacing "other brands" with private labels. Table 5 presents the fit statistics. The relative fit and predictive accuracies are the same as in Table 3. Furthermore, the (unreported) parameter estimates for Pampers, Huggies, and Luvs are not significantly different when comparing the index strategy and approximately optimal models in Tables 3 and 5.

9. Summary, Conclusions, and Future Research

Models of forward-looking experiential learning are important to marketing. These theory-driven models examine how consumers make tradeoffs between exploiting and exploring brand information. Managerially, these models enable researchers to investigate effects due to quality uncertainty, learning, and the variation in utility shocks. However, the consumer problem in these models is computationally intractable (PSPACE-hard). Existing solutions via the Bellman equation require vast computational resources (time and memory) that may contradict cognitive simplicity theories of consumers.

In this paper we propose that consumers use cognitively simple heuristics to solve forward-looking experiential learning problems. We explore one viable heuristic – index strategies. Index strategies represent a solution concept that decomposes a complex problem into a set of much simpler sub-problems. We prove analytically that an index strategy exists for canonical forward-looking experiential learning models and that the index function has simple properties

that consumers might intuit. Using synthetic data, we demonstrate that a well-defined index solution achieves near optimal expected utility and is fast to compute. Using IRI data on diaper purchases, we show that at least one index solution fits the data and predicts out-of-sample significantly better than either a no-learning model or a myopic learning model. Compared with an approximately optimal solution, the index strategy fits equally well, produces similar estimation results (and hence managerial implications), requires significantly lower computational costs and, we believe, is more likely to describe consumer behavior.

We address many issues, but many issues remain. We do not model advertising as a quality signal (the IRI dataset for the diaper category does not track advertising). The consequence of incorporating advertising signals depends on how consumers learn. We abstract away from inventory problems. Inventory effects are found to be insignificant in previous research (Ching et al. 2012), but nevertheless add a dimension to consumers' dynamic planning. We study standard settings where consumers do not learn from non-chosen alternatives. It would be interesting to model correlated learning or extend index strategies to incorporate hypothetical reinforcement of non-chosen options (Camerer and Ho 1999). Technically, it would also be interesting to examine the indexability of learning models when there are switching costs.[30]

Diaper buyers are likely forward-looking, but consumers in other product categories may not be. Our theory suggests that consumers are most likely to be forward-looking when shock uncertainty is small compared to quality uncertainty; we expect myopic learning models to do well when shock uncertainty is large. This prediction is testable using cross-category analysis. For instance, shock uncertainty may be large in hedonic goods categories where consumption

[30] Bank and Sundaram (1994) prove that there is no consistent way to define an optimal index in the presence of switching costs among choice alternatives. However, a bandit problem with switching cost can be re-formulated as a restless bandit problem, which could be indexable (Glazebrook et al., 2006, Niño-Mora 2008).

value swings with idiosyncratic mood. Shock uncertainty may also be dominant in markets characterized by volatile marketing mix variables. The recent rise of flash sales introduced remarkable price volatility to categories such as food, gadgets, and apparel. It will be interesting to study whether this change serves to promote myopic purchase behaviors.

Finally, an index solution appears to be a reasonable tradeoff for diaper consumers, but our basic hypothesis is that consumers use cognitively simple heuristic strategies. Other cognitively simple heuristics might explain consumer behavior even better than index strategies. §8.2 suggests testable alternatives. Future research can explore these and other heuristics using either field data or laboratory experiments.

References

Ackerberg, D. A. 2003. Advertising, learning, and consumer choice in experience good markets: an empirical examination. *International Economic Review,* 44 (3), 1007–1040.

Banks, J. S. and R. K. Sundaram. 1994. Switching costs and the Gittins index. *Econometrica*, 62 (3), 687–694.

Bertsimas, D. and J. Niño-Mora. 2000. Restless bandits, linear programming relaxations, and a primal-dual index heuristic. *Operations Research*, 48 (1), 80–90.

Bertsekas, D. and J. Tsitsiklis. 1996. *Neuro-Dynamic Programming*. Athena Scientific Press, Cambridge, MA.

Bettman, J. R. 1979. *An information processing theory of consumer choice*, Addison-Wesley Publishing Company, Reading MA.

Bettman, J. R., M. F. Luce, and J. W. Payne. 1998. Constructive consumer choice processes. *Journal of Consumer Research,* 25 (3), 187–217.

Bröder, A. 2000. Assessing the empirical validity of the "take the best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26 (5), 1332-1346.

Bronnenberg, B. J., M. W. Kruger, and C. F. Mela. 2008. The IRI marketing data set. *Marketing Science,* 27 (4), 745 – 748.

Camerer, C. F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction (Roundtable Series in Behavioral Economics)*. Princeton University Press.

Camerer, C. F. and T. H. Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica*, 67 (4), 827 – 874.

Che, H., K. Sudhir, and P. B. Seetharaman. 2007. Bounded rationality in pricing under state-dependent demand: do firms look ahead, and if so, how far? *Journal of Marketing Research*, 44 (3), 434-449.

Ching, A. 2010. A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review*, 51 (4), 1175-1207.

Ching, A., T. Erdem, and M. Keane. 2010. How much do consumers know about the quality of products? Evidence from the diaper market. Working paper, University of Toronto.

Ching, A., T. Erdem, and M. P. Keane. 2012. A simple approach to estimate the roles of learning, inventory and experimentation in consumer choice. Working paper, University of Toronto.

Ching, A., T. Erdem, and M. P. Keane. 2013a. Learning models: an assessment of progress, challenges and new developments. *Marketing Science*, 32 (6), 913-938.

Ching, A., T. Erdem, and M. Keane. 2013b. Online appendix of "Learning models: as assessment of progress, challenges and new developments". *Marketing Science*. Available at: http://pubsonline.informs.org/doi/suppl/10.1287/mksc.2013.0805

Ching, A., and M. Ishihara. 2010. The effects of detailing on prescribing decisions under quality uncertainty. *Quantitative Marketing and Economics,* 8 (2), 123–165.

Ching, A., and M. Ishihara. 2012. Measuring the informative and persuasive roles of detailing on prescribing decisions. *Management Science,* 58 (7), 1374–1387.

Chintagunta, P., T. Erdem, P. E. Rossi and M. Wedel. 2006. Structural modeling in marketing: review and assessment, *Marketing Science*, 25 (6), 604–616.

Chintagunta, P., R. Jiang and G. Z. Jin. 2009. Information, learning, and drug diffusion: the case of Cox-2 inhibitors. *Quantitative Marketing and Economics,* 7 (4), 399–443.

Chow, C.-S. and J. N. Tsitsiklis. 1991. An optimal one-way multigrid algorithm for discrete-time stochastic control. *IEEE Transactions on Automatic Control*, AC-36 (8), 898-914.

Crawford, G. S. and M. Shum. 2005. Uncertainty and learning in pharmaceutical demand. *Econometrica,* 73 (4), 1137–1173.

Dickstein, M. 2012. Efficient provision of experience goods: Evidence from antidepressant choice. Working paper, Stanford University.

Erdem, T. and M. P. Keane. 1996. Decision-making under uncertainty: capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science,* 15 (1), 1–20.

Erdem, T., M. P. Keane, T. Öncü, and J. Strebel. 2005. Learning about computers: an analysis of information search and technology choice. *Quantitative Marketing and Economics,* 3 (3), 207-247.

Erdem, T., M. P. Keane, and B. Sun. 2008. A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science,* 27 (6), 1111 – 1125.

Gabaix, X., and D. Laibson. 2000. A boundedly rational decision algorithm. *American Economic Review*, 90 (2), Papers and Proceedings of the 125th Annual Meeting of AEA, 433-438.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg. 2006. Costly information acquisition: experimental analysis of a boundedly rational model. *American Economic Review*, 96 (4), 1043–1068.

Gigerenzer, G. and D. G. Goldstein. 1996. Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review,* 103 (4), 650–669.

Gilbride, T. J. and G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, 23 (3), 391-406.

Gittins, J. 1989. *Multi-armed bandit allocation indices*. John Wiley & Sons, NY.

Gittins, J., K. Glazebrook, and R. Weber. 2011. *Multi-armed bandit allocation indices*, 2nd edition, John Wiley & Sons.

Gittins, J. and D. Jones. 1974. A dynamic allocation index for the sequential design of experiments. In J. Gani, K. Sarkadi, and I. Vince (Eds.), *Progress in Statistics*, North-Holland. Amsterdam, NL, 241–266.

Glazebrook, K., D. Ruiz-Hernandez, and C. Kirkbride. 2006. Some indexable families of restless bandit problems. *Advances in Applied Probability*, 38 (3), 643–672.

Gilovich, T., D. Griffin, and D. Kahneman. 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press.

Goldstein, D. G. and G. Gigerenzer. 2002. Models of ecological rationality: the recognition heuristic. *Psychological Review,* 109 (1), 75–90.

Guadagni, P. M. and J. D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Science,* 2 (3), 203–238.

Hansen, K., V. Singh, and P. Chintagunta. 2006. Understanding the store-brand purchase behavior across categories. *Marketing Science*, 25 (1), 75-90.

Hauser, J. R. 1978. Testing the accuracy, usefulness and significance of probabilistic models: an information theoretic approach. *Operations Research*, 26 (3), 406-421.

Hauser, John R. 1986. Agendas and consumer choice. *Journal of Marketing Research*, 23 (3), 199-212.

Hauser, J. R., O. Toubia, T. Evgeniou, R. Befurt and D. Dzyabura. 2010. Disjunctions of conjunctions, cognitive simplicity, and consideration sets. *Journal of Marketing Research*, 47 (3), 485-496.

Hauser, J. R. and G. L. Urban. 1986. The value priority hypotheses for consumer budget plans. *Journal of Consumer Research,* 12 (4), 446–462.

Hauser, J. R., G. L. Urban, G. Liberali, and M. Braun. 2009. Website morphing. *Marketing Science*, 28 (2), 202–223.

Hauser, J. R., G. L. Urban, and B. D. Weinberg. 1993. How consumers allocate their time when searching for information. *Journal of Marketing Research*, 30 (4), 452-466.

Ho, T-H. and J-K. Chong. 2003. A parsimonious model of stockkeeping-unit choice. *Journal of Marketing Research*, 40 (3), 351-365.

Houser, D., M.P. Keane, and K. McCabe. 2004, Behavior in a dynamic decision problem: An analysis of experimental evidence using Bayesian type classification algorithm, *Econometrica*, 72 (3), 781-822.

Hutchinson, J. M. C., and G. Gigerenzer. 2005. Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. *Behavioural Processes*, 69 (2), 97-124.

Imai, S., N. Jain, and A. Ching. 2009. Bayesian estimation of dynamic discrete choice models. *Econometrica*, 77 (6), 1865-1899.

Jovanovic, B. 1979. Job matching and the theory of turnover. *Journal of Political Economy*, 87 (5), 972–990.

Johnson, E. J. and J. W. Payne. 1985. Effort and accuracy in choice. *Management Science*, 31 (4), 395-414.

Kamkura, W. and G. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure, *Journal of Marketing Research*, 26 (4), 379-390.

Keane, M. and K. Wolpin. 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *The Review of Economics and Statistics*, 76 (4), 648-672.

Kohli, R. and K. Jedidi. 2007. Representation and inference of lexicographic preference models and their variants. *Marketing Science*, 26 (3), 380-399.

Lindsay, P. H. and D. A. Norman. 1977. *Human information processing: an introduction to psychology*. Academic Press. New York, NY.

Marewski, J. N., R. F. Pohl, and O. Vitouch. 2010. Recognition-based judgments and decisions: introduction to the special issue (Vol. 1). *Judgment and Decision Making*, 5 (4), 207–215.

Martignon, L. and U. Hoffrage. 2002. Fast, frugal, and fit: simple heuristics for paired comparisons. *Theory and Decision*, 52 (1), 29-71.

McFadden, D. 1986. The choice theory approach to market research. *Marketing Science*, 5 (4), 275-297.

Mehta, N., X. J. Chen, and O. Narasimhan. 2008. Informing, transforming, and persuading: disentangling the multiple effects of advertising on brand choice decisions. *Marketing Science,* 27 (3), 334 – 355.

Miller, R. 1984. Job matching and occupational choice. *Journal of Political Economy*, 92 (6), 1086–1120.

Narasimhan, C. 1988. Competitive promotional strategies. *Journal of Business*, 61 (4), 427-449.

Narayanan, S. and P. Manchanda. 2009. Heterogeneous learning and the targeting of marketing communication for new products. *Marketing Science,* 28 (3), 424 – 441.

Narayanan, S., P. Manchanda, and P. K. Chintagunta. 2005. Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research,* 42 (3), 278–290.

Ni, J., S. A. Neslin, and B. Sun. 2012. Database submission -- the ISMS durable goods data sets. *Marketing Science*, 31 (6), 1008-1013.

Niño-Mora, J. 2001. Restless bandit, partial conservation laws and indexability. *Advances in Applied Probability*, 33 (1), 76–98.

Niño-Mora, J. 2008. A faster index algorithm and a computational study for bandits with switching costs. *INFORMS Journal on Computing*, 20 (2), 255-269.

Ny, J. L. and E. Feron. 2006. Restless bandits with switching costs: linear programming relaxations, performance bounds and limited lookahead policies. *Proceedings of the 2006 American Control Conference*, Minneapolis MN.

Papadimitriou, C. H. and J. N. Tsitsiklis. 1999. The complexity of optimal queuing network control. *Mathematics of Operations Research,* 24 (2), 293–305.

Payne J. W., J. R. Bettman, and E. J. Johnson. 1988. Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14 (3), 534-552.

Payne J.W., J. R. Bettman, and E. J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press: Cambridge UK.

Roberts, J. H. and G. L. Urban. 1988. Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Science,* 34 (2), 167–185.

Rust, J. 1987. Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. *Econometrica*, 55 (5), 999-1033.

Rust, J. 1994. Structural estimation of Markov decision processes. *Handbook of Econometrics*, Chapter 51. Elsevier: Maryland Heights, MO.

Rust, J. 1996. Numerical dynamic programming in economics, *Handbook of Computational Economics*, Chapter 14. Elsevier: Maryland Heights, MO.

Rust, J. 1997a. Using randomization to break the curse of dimensionality. *Econometrica*, 65 (3), 487- 516.

Rust, J. 1997b. Dealing with the complexity of economic calculations, working paper, Yale University.

Samuelson, P. A. 1937. A note on measurement of utility. *The Review of Economic Studies*, 4 (2), 155-161.

Shugan, S. M. 1980. The cost of thinking. *Journal of Consumer Research,* 7 (2), 99–111.

Simon, H. A. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics,* 69 (1), 99–118.

Simon, H. A. 1956. Rational choice and the structure of the environment. *Psychological Review,* 63 (2), 129–38.

Urban, G. L., G. Liberali, E. MacDonald, R. Bordley, and J. R. Hauser 2014. Morphing banner advertising. *Marketing Science*, 33 (1), 27-46.

Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57 (2), 307–333.

Whittle, P. 1988. Restless bandits: activity allocation in a changing world. *Journal of Applied Probability,* 25, 287–298.

Yee, M., E. Dahan, J. R. Hauser, and J. Orlin. 2007. Greedoid-based noncompensatory inference. *Marketing Science*, 26 (4), 532-549.