

**Empirical Models of Learning Dynamics:
A Survey of Recent Developments**

by

Andrew T. Ching
University of Toronto

Tülin Erdem
New York University

Michael P. Keane
University of Oxford

October 21, 2015
Revised October 31, 2015
Revised June 6, 2016
Revised July 16, 2016

*Acknowledgements: Keane's work on this project has been supported by Australian Research Council grants FF0561843 and FL110100247. Ching's work has been supported by SSHRC. But the views expressed are entirely our own. We thank Shervin Shahrokhi Tehrani for providing excellent research assistance. The article was previously circulated under the title "Models for Marketing Dynamics and Learning: A Survey of Recent Developments."

1. Introduction

There is now a very large literature on dynamic models in marketing. In a narrow sense, dynamics can be understood as a mechanism whereby past product purchases affect a person's current evaluation of the utility he/she will obtain from buying a product. Most of the prior literature has focussed on three mechanisms that may generate such a causal link from past to current purchases: learning, habit persistence, and inventory dynamics.¹ This work has been reviewed extensively in papers by Ching, Erdem and Keane (2013) and Keane (2015).

However, dynamics can be more broadly defined as encompassing any process whereby the prior history of a consumer or market affects current utility evaluations. For example, it is clear that, aside from past purchase, a consumer's perception of a product may be influenced by experiences of friends or other social network members ("social learning"), experience with related products ("correlated learning" or "information spillovers"), examination of publicly available information or expert opinion ("search"), inferences about product attributes that may be drawn from the purchase decisions of others, etc.. In the present chapter we focus on the rapidly growing literature that deals with this broader view of dynamics and learning.

We begin in Section 2 with a description of the basic structural learning model developed by Erdem and Keane (1996). In this model, consumer trial and advertising are the only sources of information about a product. It is important to understand this model before we proceed, as most of the subsequent work in this area can be understood as extending Erdem-Keane to allow for additional information sources. Then, in Sections 3 through 6, we focus on (i) learning from others (social learning), (ii) learning and strategic interactions, (iii) information spillovers and

¹ Given habit persistence or learning, past purchase creates exposure to a product, which directly affects a consumer's perceived utility of a product. In an inventory model, past purchase matters because it affects current inventory, but also, more subtly, because the prices at which past purchases are made affect the reference price of the product.

correlated learning, and (iv) learning and search. In section 7 we consider the use of heuristics to capture aspects of learning. In section 8 we consider recent work that uses exogenous events and policy changes to study learning behavior. Section 9 discusses directions for future research, in particular attempts to relax some of the strong assumptions of the Bayesian learning model. Section 10 concludes.

2. The Basic Bayesian Learning Model

In this section, in order to provide a background for the subsequent discussion, we describe a basic consumer learning model, similar to Erdem and Keane (1996). Of course, there were important antecedents to their paper. In particular, the seminal papers by Roberts and Urban (1988) and Eckstein et al. (1988) also developed structural learning models. Roberts-Urban modelled learning by risk-averse but myopic consumers, while Eckstein et al. modelled learning by forward-looking but risk-neutral consumers.² The key innovation of Erdem and Keane (1996) was to develop a framework that could accommodate both risk aversion with respect to product attributes and forward-looking behavior.

The key feature of the Erdem-Keane model is that consumers do not know the attributes of brands with certainty. Before receiving any information via use experience, consumers have a normal prior on brand quality:

$$(1) \quad Q_j \sim N(Q_{j1}, \sigma_{j1}^2), \quad j = 1, \dots, J.$$

This says that, prior to any use experience, consumers perceive that the true quality of brand j

² In order to implement their model, Erdem and Keane (1996) used the approximate solution methods for dynamic programming models developed in Keane and Wolpin (1994). See Ching, Erdem and Keane (2013) for a complete explanation of the procedure. The simplifying assumptions in Eckstein et al (1988) allowed them to use the Gittin's index to find the solution of their model (see Appendix A of Ching, Erdem and Keane, 2013).

(Q_j) is distributed normally with a mean of Q_{j1} and variance σ_{j1}^2 . The values of Q_{j1} and σ_{j1}^2 may be influenced by many factors, such as reputation of the manufacturer, advice from friends, etc.

Use experience does not fully reveal quality because of “inherent product variability.” This has two main sources: First, quality of different units of a product may vary. Second, and more importantly, a consumer’s experience of a product will vary across use occasions.³

Given inherent product variability, there is a difference between “experienced quality” by consumer i for brand j on purchase occasion t , which we denote Q_{ijt}^E , and true quality Q_j .

Assume the experienced quality delivered by use experience is a noisy signal of true quality, as in:

$$(2) \quad Q_{ijt}^E = Q_j + \varepsilon_{ijt} \quad \text{where} \quad \varepsilon_{ijt} \sim N(0, \sigma_\varepsilon^2)$$

Here σ_ε^2 is the variance of inherent product variability, which we call “experience variability.” It should be noted that all brands have experience variability, so (2) holds for all j .

Note that we have conjugate priors and signals, as both the prior on quality (1) and the noise in the quality signals (2) are assumed to be normal. The posterior for perceived quality at $t=2$, after a single use experience signal is received at $t=1$, is given by the updating formulas:

$$(3) \quad Q_{ij2} = \frac{\sigma_{j1}^2}{\sigma_{j1}^2 + \sigma_\varepsilon^2} Q_{ij1}^E + \frac{\sigma_\varepsilon^2}{\sigma_{j1}^2 + \sigma_\varepsilon^2} Q_{j1}$$

$$(4) \quad \sigma_{ij2}^2 = \frac{1}{(1/\sigma_{j1}^2) + (1/\sigma_\varepsilon^2)}$$

Equation (3) describes how a consumer’s prior on quality of brand j is updated as a result of the

³ For instance, a diaper may hold all of a baby’s urine on some occasions but not on others (depending on how much milk the baby drank), so one use may not fully reveal its quality.

experience signal Q_{ij1}^E . Note that the extent of updating is greater the more accurate is the signal (i.e., the smaller is σ_ε^2). Equation (4) describes how a consumer's uncertainty declines as he/she receives more signals. The variable σ_{ijt}^2 is often referred to as the "perception error variance."

Equations (3) and (4) generalize to any number of signals received. Let $N_{ij}(t)$ denote the total number of use experience signals received by person i before he/she makes a purchase decision at time t . Then we have that:

$$(5) \quad Q_{ijt} = \frac{\sigma_{j1}^2}{N_{ij}(t)\sigma_{j1}^2 + \sigma_\varepsilon^2} \sum_{s=1}^t Q_{ijs}^E d_{ijs} + \frac{\sigma_\varepsilon^2}{N_{ij}(t)\sigma_{j1}^2 + \sigma_\varepsilon^2} Q_{j1}$$

$$(6) \quad \sigma_{ijt}^2 = \frac{1}{(1/\sigma_{j1}^2) + N_{ij}(t)(1/\sigma_\varepsilon^2)}$$

where d_{ijt} is an indicator for whether brand j is bought/consumed at time t by person i .

In (5), the perceived quality of brand j at time t , Q_{ijt} , is a weighted average of the prior and all quality signals received up through time t , $\sum_{s=1}^t Q_{ijs}^E d_{ijs}$. Perceived quality is random across consumers, as some receive, by chance, better quality signals than others. So the learning model endogenously generates heterogeneity across consumers in their perceptions of products.

Let $I_{it} = \{Q_{it}, \sigma_{it}^2\}$ denote consumer i 's information set (i.e., all the past signals he/she has received), where $Q_{it} = \{Q_{i1t}, \dots, Q_{ijt}\}$ and $\sigma_{it}^2 = \{\sigma_{i1t}^2, \dots, \sigma_{ijt}^2\}$.⁴ As Eq (6) indicates, the variance of perceived quality around true quality declines as more signals are received, and in the limit perceived quality converges to true quality.⁵

The Erdem and Keane (1996) model generalizes the model in (1)-(6) by including

⁴ Note that this is equivalent to say I_{it} consists of all past signals consumer i has received before he/she makes the purchase at time t , given the Bayesian learning framework.

⁵ Still, heterogeneity in S_{it} may persist over time, because: (i) both brands and consumers are finitely lived, (ii) as people gather more information the value of trial purchases diminishes, and so eventually learning about unfamiliar products will become slow; (iii) there is a flow of new brands and new consumers entering a market.

advertising as a second signal of quality. It is fairly easy to modify equations (1)-(6) to accommodate two (or more) signals. We do not need to go into such complications here, as our goal is just to explain the basics of the framework. But it is important to note that many extensions of the Erdem-Keane model that we discuss below rely on the fact that it is fairly straightforward to extend the Bayesian learning framework to accommodate multiple sources of information.⁶ We give some concrete examples in subsequent sections.

Finally, to complete the model we need to assume a particular functional form for utility. For instance, we could assume consumer i 's (conditional indirect) utility of consuming brand j is:

$$(7) \quad U_i(Q_{jt}^E, P_{jt}) = f(Q_{ijt}^E) - w_P P_{ijt} + e_{ijt},$$

where P_{ijt} is price, w_P is marginal utility of income and e_{ijt} is an idiosyncratic brand, time and person specific error, distributed *iid* extreme value.⁷ If we then assume that $f(Q_{ijt}^E)$ takes the constant absolute risk aversion form (CARA), then the expected utility is given by:

$$(8) \quad E[U(Q_{ijt}^E, P_{ijt})|I_{it}] = -\exp\left(-r\left(Q_{ijt} - \frac{r}{2}(\sigma_{ijt}^2 + \sigma_\varepsilon^2)\right)\right) - w_P P_{ijt} + e_{ijt},$$

where $r > 0$ captures risk aversion with respect to variation in product quality.

One can see from (8) that a higher perception error variance σ_{ijt}^2 reduces the expected utility of purchase of brand j , *ceteris paribus*. Thus, purchase of an unfamiliar brand is risky, which lowers its expected utility for a risk averse consumer. But on the other hand, trial of an unfamiliar brand has the benefit that it generates new information that lowers σ_{ijt}^2 .

⁶ Ching, Erdem and Keane (2013) explain how to extend the above basic framework to allow consumers to learn from multiple information sources (such as advertising, word-of-mouth and so on).

⁷ In almost all cases no purchase is also an option. Erdem and Keane (1996) denote the no purchase option as $j=0$, and simply set the expected utility of no purchase to $E[U_{i0}|S_{it}] = e_{i0t}$.

Finally we note that Bayesian learning models can be classified as “forward-looking” or “myopic” depending on whether consumers take the benefit of future information into account when they make current purchase decisions. Erdem and Keane (1996) found that accounting for forward-looking behavior led to only modest improvement in fit in the detergent category, but Ching, Erdem and Keane (2014a) find a more important role for forward-looking behavior in the diaper category. The literature we review below contains many models of both types.

3. Learning from Others (Social Learning)

Learning from others encompasses a wide range of activities. Some examples are word-of-mouth” learning, the source of which is typically friends and relatives, consulting expert opinions (as in movie reviews, magazine articles, and news coverage), or reading online reviews by other consumers. Erdem, Keane and Sun (2008) allow for four sources of information when studying scanner panel data for consumer package goods, and Erdem, Keane, Oncu and Strebel (2005) allow for five sources (including word-of-mouth, salespeople, and articles in computer magazines and other magazines) when studying the PC purchase decisions.

Alternatively, one may consider an environment where there exists a special agent (or information source) who serves as an information aggregator who pools consumers’ experiences together. For instance, in the context of choosing between brand-name drugs or generic counterparts, Ching (2010a, 2010b) assumes that a random sample of individual’s experiences can be observed by everyone via physician networks or consumer watch groups. The Ching (2010a, 2010b) model considers a representative physician who learns about the average quality of a generic drug (relative to the quality of its brand-name originator).

Specifically, let S_t be the random sample of individual experience signals that are revealed to the representative physician at time t . Let q_t be the quantity of the generic drug sold

at time t and let κ be the proportion of individual experience signals revealed in each period. Then $\text{card}(S_t) = \kappa q_t$.⁸ Assume the experience signals are distributed *iid* across patients with mean Q_j and variance σ_ε^2 . Finally, let \bar{Q}_{ijt}^E denote the mean of the κq_t individual experience signals observed by the representative physician regarding drug j at time t . The representative physician's updating process can be described by extending equations (3) and (4) as follows:

$$(9) \quad Q_{jt+1} = \frac{\sigma_{jt}^2}{\sigma_{jt}^2 + \sigma_\varepsilon^2 / (\kappa q_t)} \bar{Q}_{ijt}^E + \frac{\sigma_\varepsilon^2 / (\kappa q_t)}{\sigma_{jt}^2 + \sigma_\varepsilon^2 / (\kappa q_t)} Q_{jt},$$

$$(10) \quad \sigma_{jt+1}^2 = \frac{1}{(1/\sigma_{jt}^2) + ((\kappa q_t)/\sigma_\varepsilon^2)}$$

where $\bar{Q}_{ijt}^E | \kappa q_t \sim N\left(Q_j, \frac{\sigma_\varepsilon^2}{\kappa q_t}\right)$ by the central limit theorem.

Another very interesting problem that has received attention lately is learning about product quality in environments where consumers only observe the choices of others (without specifically observing their experience signals). This is known as “observational learning.” To our knowledge, the first structural empirical model of observational learning is Zhang (2010), who extends observational learning to a dynamic setting in order to explain consumers' decisions to accept a donated organ for transplant – specifically a kidney.

Zhang (2010) considers an environment where patients wait in line to receive a kidney for transplant. However, it is not uncommon for a patient to choose not to accept a kidney and wait for a better match. Hence, in the model, when a patient receives a kidney offer, he/she needs to choose whether to accept it or decline it and continue to wait.

Prior to making this accept vs. decline-and-wait decision, the first patient in the line

⁸ $\text{card}(\cdot)$ is the cardinality of the set in question. It measures the number of elements in the set.

“examines” his match with the kidney and obtains a noisy signal distributed around the true quality of the kidney. When making his/her decision, all the first patient can rely on is his/her signal and initial prior belief. Hence, the posterior expected value of the kidney is simply given by Equation (3). If the expected utility of receiving the kidney is higher than the expected future value of waiting, the first patient will accept. Otherwise, he/she will decline-and-wait.

However, the decision facing the second patient is more complex. When he/she decides whether to accept the kidney or not, the second patient must take into account not only his/her own signal, but also the fact that the first patient declined. As the first patient’s decision was a function of the signal that he/she observed, the first patient’s choice reveals that the first signal must have been below a certain cutoff (or reservation) value. When the second patient updates his/her belief, he/she should take this fact into consideration, rather than relying purely on his/her own personal signal. Similarly, when it is the third patient’s turn, he/she takes into account that the first two patients have declined the kidney.

A simplified version of the Zhang (2010) model can be described as follows. Let Q_i^P be the private signal received by i -th consumer, where i indexes the position of the consumer on the waiting list. Then,

$$(11) \quad Q_i^P = Q + \varepsilon_i \quad \text{where} \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2)$$

For $i=1$ (the first patient in line), the decision problem is the same as in the Erdem-Keane set up. Specifically, patient 1 uses his/her own noisy signal Q_1^P to update his/her belief about Q . If the expected utility of accepting the kidney is higher than that of declining (and waiting for the next offer), then patient 1 accepts. It is easy to show that $E[Q|Q_1^P]$ is monotonically increasing in Q_1^P . Hence, the first patient’s decision rule can be characterized by a cutoff rule – i.e., there exists a

B_1 such that if $Q_1^P \geq B_1$, patient 1 accepts (i.e., $d_1 = 1$); otherwise he/she declines (i.e., $d_1 = 0$).

But for the second patient in line ($i=2$) the situation is more complex. Specifically, for the second patient the information set is given by $I_2 = \{d_1 = 0; Q_2^P\} = \{Q_1^P < B_1; Q_2^P\}$. The key is to find the conditional distribution, $p(Q|I_2) = p(Q|Q_1^P < B_1; Q_2^P)$. Zhang (2010) assumes that each patient draws an independent signal. Hence, it follows from the Bayes' rule that,

$$(12) \quad p(Q|Q_1^P < B_1; Q_2^P) \propto p(Q_1^P < B_1; Q_2^P|Q) \cdot p(Q),$$

where $p(Q)$ is the initial prior belief about Q . Moreover,

$$(13) \quad p(Q_1^P < B_1; Q_2^P|Q) = \Phi\left(\frac{B_1 - Q}{\sigma_\varepsilon}\right) \cdot \phi\left(\frac{Q_2^P - Q}{\sigma_\varepsilon}\right),$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function and probability density function of standard normal, respectively. One can then use (12) and (13) to obtain

$$(14a) \quad E(Q|Q_1^P < B_1; Q_2^P) = \frac{\int p(Q|Q_1^P < B_1; Q_2^P) \cdot Q \cdot dQ}{\int p(Q|Q_1^P < B_1; Q_2^P) \cdot dQ},$$

where the denominator is a normalizing factor to ensure that the posterior density of Q is proper.

Note that it is straightforward to extend the logic above to the decision problem of the i -th

consumer, for $i > 2$.⁹ This leads to higher order conditioning, of the form:

$$(14b) \quad E(Q|Q_1^P < B_1, \dots, Q_{i-1}^P < B_{i-1}; Q_i^P) = \frac{\int p(Q|Q_1^P < B_1, \dots, Q_{i-1}^P < B_{i-1}; Q_i^P) \cdot Q \cdot dQ}{\int p(Q|Q_1^P < B_1, \dots, Q_{i-1}^P < B_{i-1}; Q_i^P) \cdot dQ}.$$

Then, solving for this expectation typically requires using Monte Carlo simulation methods, such

⁹ Hendrick et al. (2012) propose a similar framework to study how consumers choose a product among $J > 2$ alternatives. Newberry (2016) extends this framework to study the role of pricing in observational learning using data from an online market for music.

as the recursive conditioning simulator developed in Keane (1994).

Many other problems are similar in structure to the observational learning problem discussed here. Examples are decisions of whether or not to accept a job offer when the decisions of prior individuals who were offered the same position can be observed. This tends to be the case for high-profile positions such as deanships, coaching positions, key executive positions and so on. Alternatively, one might consider the decision of whether or not to extend a job offer to an applicant, given knowledge of the set of offers and rejections that he/she has so far received (as is often the case in organized job markets like that for assistant professors).

Returning to the medical example, Ching and Ishihara (2010) consider an alternative situation where patients can only obtain qualitative information about a signal. Often times, a product review may simply reveal whether product A is better product B, without revealing the exact realization of the quality signal. In this case, consumers can only infer that the signal lies within a certain range.

Of course, social learning has become an area of great applied interest in recent years. Thus, many other new papers on social learning are notable for their substantive (as opposed to methodological) contributions.¹⁰ An important substantive topic that has recently received attention is how consumers may learn from each other through online reviews. Zhao et al. (2013) estimate a Bayesian learning model with myopic consumers which allows for consumers to learn about product quality both through their own experiences with the same type of product (e.g., a book genre), as well as through product reviews posted by other consumers. Furthermore, in addition to learning from others about the same product, they also allow for what is known as “correlated learning.” That is, other consumers’ experiences with books of the same genre can

¹⁰ The models in these papers largely adopt the framework discussed so far, and hence we will not devote space to explicitly discussing their structure.

allow one to update his/her belief about other books in that genre. The Zhao et al. (2013) model also incorporates learning about the credibility of product reviews posted by others (captured as the precision of the information provided). The model is estimated on book purchases of a panel of consumers. The results indicate that consumers learn more from online reviews of book titles than their own experiences with other books of the same genre.

Similarly, Wu et al. (2015) study the economic value of online reviews to consumers, as well as to restaurants, using a dataset from Dianping.com, a leading Chinese website providing user-generated reviews. The proposed Bayesian learning model with myopic consumers allows for different reviews to be of different informational value to different consumers. It also allows consumers to learn about their own preferences for multiple product attributes, as well as learn the mean and variance of consumption experiences in the population. The findings indicate that the majority of the created value comes from reviews on the quality of the restaurants, and that contextual comments are more valuable than numerical ratings in reviews.

Using field experiments, Godlonton and Thornton (2013) study the impact of others' testing on individual perceptions of AIDS risk and subsequent decisions to practice safe sex in rural Malawi. In this context, it appears that individuals tend to overestimate the underlying prevalence of HIV incidence. Godlonton-Thornton measure the response to others' HIV testing, which alters individuals' beliefs about the underlying prevalence. They measure the causal effect of others' testing by utilizing an experiment that randomly offered incentives to individuals to learn about their HIV test results at randomly located results centers. They use the village-level average of these incentives and distance from results centers to instrument for the proportion of community testing. They find robust evidence of downward revision of beliefs about HIV prevalence, and subsequent changes in sexual behavior (e.g., reduced condom use).

Knicht and Schiff (2010) develop and estimate a social learning model to study voters' decisions in US presidential primary elections – a system under which States vote sequentially, so voters in later State elections can learn about candidates from earlier State results. The advantage of a sequential system is that it provides late voters with valuable information, but its drawback is that it exaggerates the influence of early States.

Hummel and Knight (2015) use this model to conduct a counterfactual experiment to compare outcomes under simultaneous vs. sequential elections. The advantage of a simultaneous election is that it weighs States equally. However, it also places great weight on voter priors, creating a large advantage for front-runners. Thus, simultaneous (sequential) elections are preferred if the front-runner advantage is small (large). The quantitative welfare analysis of presidential primaries reported in Hummel and Knight (2015) suggests that simultaneous systems would slightly outperform sequential systems.

Lee and Bell (2013) model social learning through neighbors' past purchases. They estimate a Bayesian learning model with myopic consumers on combined data from consumer purchases on Bonobos, a leading online fashion retailer in the US, and the Social Capital Community Benchmark Survey (SCCBS). SCCBS has data relating to two dimensions of social capital, that is, trust among local neighbors and the frequency of interaction. Utilizing these data, they estimate a model where consumers update their beliefs about experience attributes through a social learning mechanism where local neighbors' purchases serve as information signals. The results indicate that social capital improves the learning process and therefore indirectly drives sales when the information that is communicated is favorable.

Huang et al. (2015) model social learning in the context of crowdsourcing new product ideas. They propose a Bayesian learning model that accounts for consumers' learning about the

potential of their ideas, as well as the cost structure of the firm. The model is estimated on data from IdeaStorm.com (a crowdsourced ideation initiative affiliated with Dell). The findings suggest that individuals overestimate the potential of their own ideas and underestimate the firm's costs. They learn about both, but they learn a lot faster about the potential of their own ideas than about the firm's cost structure.

In a reduced form study, Ching, Clark, Horstmann and Lim (2016) find evidence that in the prescription drug market, where physicians can learn from different sources, the interaction between detailing (i.e., pharmaceutical firms send sales representatives to visit doctors) and publicity (news coverage) could have non-trivial outcomes, depending on the complexity of the information. In the setting of anti-cholesterol drugs, some information is simple to describe (e.g., side-effects or the extent to which a drug can reduce cholesterol), and some can be much more complicated (e.g., a drug's ability to reduce heart disease risks). They argue that news coverage is subject to tight space constraints and hence, even though the source is credible, it cannot report all the relevant information related a clinical trial that documents reducing heart disease risks.

But physicians may treat news coverage as corroborative evidence that supports what sales representatives claim. In particular, when physicians see that sales representatives' claims are consistent with news coverage, they may give more time to the sales rep and let them explain the details of new clinical trials. This may cause detailing and publicity about complex information to be complements. On the other hand, simple information is much easier for physicians to verify, and hence credibility is less of an issue. As a result, different sources of simple information are substitutes. This research suggests the importance of distinguishing the complexity of information, and modeling the idea of corroborative evidence in the context of learning – areas that the structural learning literature has not tackled yet.

Several recent papers also focus on learning about entertainment products. Lovett and Staelin (2016) model social learning in the context of TV shows. They focus on how paid (e.g., advertising) and owned media (e.g., a TV network's own website) differ from social media in influencing consumers' utility of watching a TV show. Lovett and Staelin (2016) decompose these influences into three channels: learning, reminding and social engagement. The unique feature of their data is that it contains consumer's stated expectations about TV shows. They use these to calibrate the informative effect of these media, and allow the rest of their impacts to be picked up by their reminding and social engagement functions.

Liu and Ishihara (2015) study consumer learning about new video games, using product level sales data and critic and user reviews from the US video game market. Their model allows heterogeneity in consumer tastes (leading to horizontal differentiation of games), and controls for the spurious correlation that is likely to exist between review ratings and demand. Specifically, they use the market shares of pre-order data (period 0) to measure consumers' initial priors (about games). After critic reviews become available in period 1, consumers' update their priors about games. Then, the market share in period 1 shows the impact of critic ratings.

The Liu-Ishihara model departs from standard learning models by also incorporating the psychological theory of reference points. Consumers use critic and user reviews over time to update their prior and form a reference quality for games. Consumer reviews then depend on how experienced quality compares to the reference point. For instance, if experience is worse than expected it may lead to a very poor review, while if experience is good but similar to what was expected it may lead to only a mildly positive review.

Liu and Ishihara (2015) find that a 20% discount for pre-orders will increase immediate profits, but reduce future profits via lower consumer review ratings. The latter happens because

the pre-order discount attracts consumers with lower product valuations, and these consumers' reviews pull down the average consumer review rating (compared to the situation where only consumers with high product valuation purchase and write consumer reviews). Ignoring such consumer review endogeneity when conducting profit simulations might cause one to overstate the effectiveness of the pre-order discount in improving profitability.

Wei (2015) models how movie studios decide which potential projects to invest in when facing uncertainty about the potential of novel types of movies. He hypothesizes that in addition to observable characteristics (e.g., budget, actors, director, genre), studios can rely on the past performance of a new line of movies to update their beliefs about its latent quality distribution. This in turn will help the studios to compute the expected profits of adding such a project to their existing production portfolio.

In a different type of application, Liu et al. (2012) model learning from others using panel data obtained from a series of laboratory experiments. Specifically, they study agent's strategic behavior in an entry limit pricing game where firms use price to signal costs (Milgrom and Roberts, 1982). In experimental economics, researchers often provide subjects with choices made by others as feedback, in order to speed up the convergence of one's strategy. Thus, Liu et al. (2012) argue that peer group effects should be very pronounced in experiments run in this style (which is typical). To capture learning from peers, they extend dynamic discrete choice panel data models (Heckman, 1981) by introducing a time-lagged social interactions variable.¹¹ Their results indicate that learning from peers is important in this experiment.

Like Liu et al. (2012), Chan, Li and Pierce (2014) also model learning from peers, but they use field data. They make use of a unique data set which consists of all sales persons'

¹¹ The likelihood of their model involves multiple integrals because the explanatory variables include lagged latent dependent variables and serially correlated errors, but they show that the GHK simulator remains tractable for this generalized framework (see Keane 1994).

performance (i.e., their actual weekly sales) from all cosmetic product counters in a department store for a period of four years. The authors use this data set to measure the impacts of learning from others within the same counter, and from adjacent counters. They are able to identify these effects because workers are assigned to different shifts (there are three shifts) “randomly” based on fairness instead of their productivity. This provides an excellent opportunity to measure the effects of learning from others based on different sources, and disentangle them from self-learning. They argue that learning from peers can be fundamental to knowledge spillovers and explaining organizational learning curves.

4. Learning and strategic interaction

Ching (2010b) is the first paper that incorporates consumer and firm learning in a dynamic oligopoly structural model. In the model, firms are forward-looking and set price in each period. The demand side of the model is taken from Ching (2010a). The dynamic oligopoly model is developed to study the competition between a brand-name drug and its generic counterparts. The key innovation here is that firms are also uncertain about the true quality of generic drugs, and they can use price to control the rate of learning. In particular, assuming consumers are risk-averse, generic firms may have an incentive to price low to encourage more consumers to try their products and resolve the uncertainty (measured by the variance of their posterior belief).

To explain the main features of the Ching (2010b) model, we make the simplifying assumption that there is one brand-name firm and one generic firm.¹² Let p_{bt} and p_{gt} be the brand-name price and generic price at time t , respectively. The per period profit for firm $j \in \{b, g\}$ is $\pi_j = (p_j - mc) * q_j(p_b, p_g)$, where $q_j(p_b, p_g)$ is determined by the discrete choice

¹² Ching (2010b) allows for multiple generic firms. In addition, generic firms’ entry decisions are also endogenous. But since the focus of this chapter is learning, we abstract away the entry decisions when describing the model.

model described in equations (1)-(8).¹³ The model assumes that in each period the brand-name firm acts as the Stackelberg leader, and the generic firm is the follower. Let $S_t = (Q_{gt}, \sigma_{gt}^2)$ be the state variables at time t , which evolve according to equations (3) and (4), respectively. The generic firm's dynamic problem can be characterized using dynamic programming as follows:

$$(15) \quad V_g(S_t) = \max_{p_{gt}} [\pi_g(S_t, p_{bt}, p_{gt}) + \beta E[V_g(S_{t+1}) | S_t, q_{gt}(p_{bt}, p_{gt})]], \quad \text{for } t < T;$$

$$(16) \quad V_g(S_T) = \max_{p_{gT}} \pi_g(S_T, p_{bT}, p_{gT}).$$

Similarly, the brand-name firm's dynamic problem can be characterized as follows.

$$(17) \quad V_b(S_t) = \max_{p_{bt}} [\pi_b(S_t, p_{bt}, p_{gt}^*(p_{bt})) + \beta E[V_b(S_{t+1}) | S_t, q_{gt}(p_{bt}, p_{gt})]], \quad \text{for } t < T;$$

$$(18) \quad V_b(S_T) = \max_{p_{bT}} \pi_b(S_T, p_{bT}, p_{gT}^*(p_{bT})).$$

Note that the main difference between the problem faced by the brand-name and generic firms is that the brand-name firm takes into account that its price will influence the generic price; on the contrary, the generic firm simply takes the brand-name price as given. The equilibrium concept is Markov Perfect Nash equilibrium (i.e., the pricing function only depends on payoff relevant state variables contained in S_t). Because this model has a final period, T , it can be solved using backward induction. Although this model is conceptually tractable, it is computationally very challenging to solve. This is because the state space is continuous, and unlike other discrete choice problems, the firms are choosing a continuous variable to maximize their total discounted profits.

¹³ It is the choice probability of choosing product j multiplied by the total number of potential patients in this market.

Zou (2014) develops and estimates an equilibrium model of intertemporal pricing of new products. The model extends Ching (2010b) by allowing consumers to be heterogeneous in their information sets. The learning process in the model naturally generates such an outcome, similar to Erdem and Keane (1996). In contrast, Ching (2010b) assumes there is a representative consumer and hence there is only one information set. Similar to Ching (2010b), this model allows for forward-looking firms but myopic consumers. In addition, Zou (2014) also allows for additional state dependence beyond what consumer learning implies. He applies his model to data from the Yogurt category, for a time period covering the entry of Chobani. Using counterfactual analysis, Zou finds evidence that the firm's introductory pricing strategy is mainly driven by positive state dependence rather than learning.

Huang, Luo and Xia (2015) use a structural learning model to study how dealers set prices for used cars over time. They argue that used cars are hard to price because, unlike new cars, they differ in multiple dimensions (depending on mileage, year, and maintenance), and it is not clear how consumers trade-off these dimensions a priori. Facing this uncertainty about the unobserved demand factor, in every period a dealer sets the price for a used car and then consumers decide whether to buy it. If consumers choose not to buy the used car, this gives the dealer a signal about its unobserved demand component, and he then updates his belief accordingly. If a dealer is forward-looking, he has an incentive to price the used car high early on, because not selling the car gives him an opportunity to obtain better information about the demand for the car.

Using a panel dataset of used-car sales from CarMax, Huang et al. (2015) find that their structural model can explain demand and pricing patterns well. As in Ching (2010b), learning is the main source of dynamics that determines how a firm sets their prices dynamically over time

in this paper. But Huang et al. (2015) focus on a dynamic monopoly problem, while Ching (2010b) studies a dynamic oligopoly problem. Moreover, Ching (2010b) has a symmetric two-sided learning environment (both firms and consumers are equally uncertain about the quality of generic drugs), while Huang et al. (2015) have an asymmetric one-sided learning environment (consumers know their demand, and only firms are uncertain about demand conditions).

A recent paper by Chen, Sun and Singh (2009) investigates the interaction between learning and addiction in the tobacco market by estimating a forward-looking structural learning model. They model addiction using the brand loyalty variable of Guadagni and Little (1983) (GL), i.e., exponential smoothing of the past choices. They use their model to study the effects of Marlboro's permanent price cut that happened on April 2, 1993 (Marlboro Friday) as a reaction to the continuous loss of market share to generic brands.¹⁴ They find that by permanently lowering the price, consumers who previously bought only generics became willing to experiment. Their estimation results suggest that there is positive interaction between expected quality and the GL variable, and that Marlboro has a higher quality than generic brands. This implies that the brand loyalty effect due to GL is stronger for Marlboro. As a result, when the permanent price cut induces consumers to try Marlboro, most of these new consumers stay with it because of Marlboro's stronger GL effect.

It is interesting to compare the Marlboro permanent price cut strategy with how brand-name drugs increase prices in response to generic competition (Ching 2010a, 2010b). At first, these two situations seem similar (both of them face generic competition). But a closer examination reveals some key differences of the environments considered in Chen et al. (2009) and Ching (2010b). The paper by Ching (2010b) assumes that by the time the patent expires,

¹⁴ Although this paper does not explicitly model a dynamic game, the dynamic demand model is very useful in evaluating the consequences of Marlboro's strategic response to the competition of generic brands.

most consumers know the quality of the branded product. But they need to learn about the quality of generics. The story is that consumers who are price-sensitive would slowly switch to generics as they learn and become increasingly more confident that they are safe over time. But this implies that the demand faced by the brand-name firm becomes more price inelastic over time. Of course, this effect cannot last forever, and the theory suggests that at some point, the price for the brand-name drug should come down. But, as long as there is a mass of consumers loyal to the brand-name drug who do not update their belief at all, that may be sufficient to keep the brand-name drug price high permanently.

On the contrary, in the tobacco case, Chen et al. (2009) hypothesize that some consumers have not tried Marlboro (or other premium brands) before. So a permanent price cut allows Marlboro to regain some market share by gaining new customers. Quality here refers to taste (i.e., Marlboro could taste better than other generic brands). So in the tobacco case, some consumers may discover that they actually like Marlboro more in a complete information situation.

Structural learning has also been introduced in other problems of strategic interactions. Yang (2016) introduces learning from others in an incomplete information discrete dynamic game with entry and exit (similar to Aguirregabiria and Mira 2007). The model captures the idea that firms have uncertainty about the market potential of their products. There are two ways a firm can resolve this uncertainty: (i) learn directly by entering the market; or (ii) learn from other firms which are operating in the market. The learning mechanism is similar to Ching (2010a). Yang (2016) finds that learning from others can partially offset the negative business-stealing effects of rivals (because learning from others helps a firm find out sooner if it should exit the market).

Ho, Park and Su (2015) expand on standard models of iterative thinking by introducing a Bayesian level-k model,¹⁵ in which players perform Bayesian updating of their beliefs about opponents' rule levels, and best-respond with different rule levels over time. The authors apply this sophisticated learning model to experimental data on p -beauty contest and price matching games and find evidence for this type of sophisticated learning.

5. Information Spillovers and Correlated Learning

As we noted in the introduction, information spillovers and correlated learning refer to situations where one can learn about a given product via experience with related products. The standard learning model can be easily extended to study such environment. We first specify a more general prior belief to capture the idea that consumers may believe the qualities of products are correlated. Using vector notation, one can modify equation (1) as:

$$(19) \quad Q \sim N(Q_{t=1}, \Sigma_{t=1}),$$

where $Q_{t=1}$ is the $J \times 1$ initial prior mean vector and $\Sigma_{t=1}$ is the $J \times J$ initial prior variance-covariance matrix. Consider a two-product case. With off-diagonal elements greater than zero, an information signal for product 1 will be used to update ones belief about product 2, and vice versa. The updating formula can be generalized as follows.

$$(20) \quad \Sigma_t = \begin{bmatrix} \sigma_{1,t}^2 & \pi_t \\ \pi_t & \sigma_{2,t}^2 \end{bmatrix}.$$

¹⁵ The basic idea of this model is that players in a game vary in their depth of strategic thinking. A completely naive player will choose actions by completely ignoring the presence of other players (level zero). A level one player believes that other players will not react to his choice, and his action is the best response with respect to this belief. A level two player believes that all other players are level one, and so on and so forth. This model captures bounded rationality, and can explain players' behavior in games that cannot be rationalized by standard game theory.

When receiving an information signal for product 1 at time t , the updating for $Q_{1,t+1}$ and $\sigma_{1,t+1}^2$ will be the same as the standard learning model explained in equations (3) and (4). But the information signal for product 1 will also be used to update the consumer's belief about product 2 as follows:

$$(21) \quad Q_{2,t+1} = Q_{2t} + \frac{\pi_t}{\sigma_{2t}^2 + \sigma_\varepsilon^2} (Q_{1t}^E - Q_{1t}).$$

The variance-covariance matrix for posterior beliefs becomes:

$$(22) \quad \sigma_{2,t+1}^2 = \sigma_{2,t}^2 - \frac{\pi_t^2}{\sigma_{2t}^2 + \sigma_\varepsilon^2},$$

$$(23) \quad \pi_{t+1} = \frac{\pi_t \sigma_\varepsilon^2}{\sigma_{2t}^2 + \sigma_\varepsilon^2}.$$

Erdem (1998) was the first paper to apply this framework to show that, in the case of umbrella brands, consumers learn about brand quality in one category through their experiences with the same brand in another category. Recent years have witnessed a marked increase in papers that focus on this topic. For instance, expanding on Erdem (1998)'s idea about cross-brand learning, Szymanowski and Gijsbrechts (2012) study whether experiences with a private label affects consumer quality perceptions about other private labels. They estimate a Bayesian learning model with myopic consumers on household scanner panel data on dish soap and breakfast cereals. Their results indicate there is cross-retailer learning among standard private labels regardless of their name and quality differences.

Ching and Lim (2016) significantly extend Erdem's framework to an environment where firms are selling similar but differentiated competing products. To explain why sometimes late entrants can easily surpass incumbents, they propose a new consumer theory of correlated

learning and indirect inference. They apply their model to study the market for anti-cholesterol drugs, where the late entrant, Lipitor, overtook long-time incumbents within six quarters of its entry. They argue that physicians use each drug's ability to lower cholesterol, and their prior belief about its efficiency ratio,¹⁶ to infer each drug's ability to reduce heart disease risks. They argue that correlated learning happens when new clinical trials provide evidence about another drug's efficiency ratio. Therefore, even without any direct clinical evidence to show it can reduce heart disease risks, Lipitor was perceived to be the best drug for this purpose because of its superior ability to lower cholesterol. The physicians act as if they believe that evidence on an individual drug's ability to reduce heart disease risks can be generalized to the whole class of statins. This is the first paper that provides a structural explanation for a *late mover advantage*.

Finally, Che et al. (2015) use a forward-looking dynamic demand model to examine how brand preferences evolve when consumers are new to a market and their needs change over time. They allow for strategic sampling behavior of consumers under quality uncertainty, and they also allow for strategic sampling to increase periodically when consumers' needs change. The proposed model differs from previous work on forward-looking consumer Bayesian learning by allowing for (1) spill-over learning effects across different versions of products (or products in different product categories that share a brand name); and (2) duration-dependence in utility for a specific version of a product or product class to capture systematic periodic changes in consumer utility. Che et al. estimate their model using scanner data for the disposable diaper category. Here, it is likely that use experience with a particular size of a brand provides noisy information about another size of the same brand. And consumers' size needs change exogenously over time as the baby grows older and needs to change diaper sizes. The proposed model is useful in

¹⁶ The efficiency ratio measures how well a drug converts reduced *cholesterol* levels to reduced *heart disease* risks.

assessing the extent of use experience spillover effects and the degree to which information from past use experience is retained when consumers migrate across classes, versions, and the like.

6. Models Incorporating Both Learning and Search

Both the learning literature and the search literature focus on consumer choice under uncertainty. Search models are usually applied to explain dispersion in prices (or wages). In its simplest form, this class of models usually assumes there are a large number of retailers that sell the same product. Conditioning on a consumer who has already made up his mind to buy this product, his/her objective is to buy it from a retailer that offers the lowest price. But before visiting a retailer, he/she does not know what price it offers. However, it is typically assumed that consumers know the distribution of prices in the market. There is a cost to visiting a retailer (search cost). Finally, assume that consumers are forward-looking and conduct a sequential search. Then the dynamic programming problem solution implies that consumers' decision whether to continue to search (or to make a purchase now) is an optimal stopping problem. The solution is characterized by a reservation price: a consumer's decision rule is to reject any price above the reservation price, and accept any price that is below the reservation price.

Several papers have tried to relax the assumption that the price distribution is known. Rothschild (1974) proposes the first theoretical model to characterize the decision rule under such an environment. Recently, Koulayev (2013) and De los Santos et al. (2013) have extended Rothschild's model to an empirical setting. Koulayev (2013) uses Dirchlet distribution priors to model the uncertainty about the price distribution, while De los Santos et al. (2013) use Dirchlet process priors.

With Dirchlet distribution priors, Koulayev (2013) is able to derive closed form choice probabilities, and the characterization only relies on: (i) the identity of the second-best product

among the discovered set; (ii) the number of searches to date. This allows Koulayev (2013) to estimate the model using only market share data. The Dirchlet process considered in De los Santos et al. (2013) is more general (i.e., it is the infinite dimensional version of the Dirchlet prior). Thus, to estimate their model, they need to observe price histories of each consumer.¹⁷ They employ a moment inequality approach to estimate bounds on the parameters.

Interestingly, while the standard search model implies that consumers always buy on the last search, these learning and search models are able to explain why some consumers return to a previous search (as is supported by empirical patterns). This is because, in a learning and search model, reservation price is decreasing with the number of searches (conditional on continuing search). The reason is that, if a consumer continues to search, the price he/she just sampled must be higher than his reservation price. Thus, when the consumer uses this last observed price to update his prior, it must drive up his perceived “average” price (based on the prior belief he/she held right before seeing the last observation). The price distribution thus shifts up slightly, resulting in a higher reservation price. But with a higher reservation price, it is possible that some previously seen prices actually fall below it. That’s why it may make sense for consumers to return to a previous search.

Roos et al. (2015) develop a structural model of hyper-media search and consumption while accounting for features unique to this consumption context: the rapid refresh of information (and consumers' concomitant uncertainty about its relevance and availability), the role played by linked excerpts in signaling the relevance and availability of new information on

¹⁷ Note that the basic idea of search models is that consumers need to compare the expected gain from searching vs. the cost of search. In contrast, standard choice models with learning assume that consumers learn about an attribute by buying the product multiple times because information signals are noisy. These models also assume there are a fixed number of alternatives to choose from. In search models with an unknown price (or attribute) distribution, consumers learn about the parameters that characterize the distribution. For a normal distribution, that would be simply learning its mean and standard deviation. But, for a Dirichlet distribution, a consumer needs to use the whole history of price realizations and the initial parameters that characterize the prior to construct his posterior.

other sites, and the potential novelty or redundancy of information across sites. They estimate this new search model using panel data on consumer celebrity blog browsing and information scraped from sites regarding the links between them. Their results indicate that celebrity blogs are differentiated horizontally by their degree of sexually-oriented content and that links are a useful signal of the linked sites' content. Moreover, in many cases, links decrease (increase) visits to the linked (linking) sites.

7. Heuristic and Approximation Approaches to study Consumer Incentives to Explore

In most structural models of learning, the literature uses a dynamic programming approach to capture the idea that consumers make their choices while taking into account the benefit of exploration (or experimentation). In order to illustrate the set up, note that one can express the value of choosing alternative j as follows:

$$(24) \quad V(j, t|I_t) = E[U(Q_{jt}^E, P_{jt})|I_t] + \beta EV(I_{t+1}|I_t, j) \quad \text{for} \quad j=0, \dots, J$$

Here we suppress the person subscript i for notational convenience. As one will recall from the setup in equations (1)-(8), experienced utility depends on experienced quality, Q_{jt}^E . This may depart from true quality due to experience variability. Furthermore, the consumer must also account for the fact that he/she does not know true quality of brand j with certainty. Rather, the consumer uses his/her information set I_t to infer the subjective distribution of brand j quality. In forming expected utility, $E[U(Q_{jt}^E, P_{jt})|I_t]$, the consumer must account for both of these sources of uncertainty.

Finally, $EV(I_{t+1}|I_t, j)$ is the expected future value of choosing product j , which takes into account how the choice of j changes the information set at $t+1$. The parameter β is the discount factor. The “alternative specific value function” $V(j, t|I_t)$ simply adds together the current and

discounted future payoffs from choosing brand j . A complete solution of the consumers dynamic optimization problem would give the values of the $EV(I_{t+1}|I_t, j)$ at every possible state point. This would enable a researcher to construct the alternative specific value functions $V(j, t|I_t)$, from which one could construct the choice probabilities and the likelihood function.¹⁸

To be more concrete, if we substitute using equation (7) we obtain:

$$(25) \quad V(j, t|I_t) = Ef(Q_{jt}^E|I_t) - w_P P_{jt} + e_{jt} + \beta EV(I_{t+1}|I_t, j)$$

This expression makes clear that the current payoff depends on (i) the subjective distribution of quality, which depends on the information set I_t , (ii) price (as well as other possible covariates like promotion that we might choose to add), and (iii) transitory taste shocks.

Now, suppose we compare the value of choosing two brands j and k . We obtain:

$$(26) \quad V(j, t|I_t) - V(k, t|I_t) = [Ef(Q_{jt}^E|I_t) - w_P P_{jt} + e_{jt}] - [Ef(Q_{kt}^E|I_t) - w_P P_{kt} + e_{kt}] + G(j, k, I_t)$$

where:

$$(27) \quad G(j, k, I_t) \equiv \beta [EV(I_{t+1}|I_t, j) - EV(I_{t+1}|I_t, k)]$$

Equations (26)-(27) make clear that the value of choosing j over k can be decomposed into (i) the difference in expected current payoffs (which is all that matters in a static model) and (ii) the information advantage of choosing j over k , which we denote by $G(j, k, I_t)$. Intuitively, if k is a very familiar brand while j is new, we would expect $G(j, k, I_t) > 0$ as there is more information to be gained by trying j , which might turn out to be better than k . The existence of the G function is what generates the incentive for strategic trial in dynamic learning models.

Erdem and Keane (1996), along with most subsequent dynamic structural learning

¹⁸ This is in contrast to static models, where the current expected utilities, $E[U(Q_{jt}^E, P_{jt})|I_t]$, alone determines choice probabilities.

models, obtain the expected future value functions $EV(I_{t+1}|I_t, j)$ by solving a dynamic programming problem. However, it is not feasible to solve for $EV(I_{t+1}|I_t, j)$ at every possible state point (I_t, j) , so the usual approach is to approximate the solution. For instance, the Keane and Wolpin (1994) approximation technique involves solving for the $EV(I_{t+1}|I_t, j)$ at a subset of state points, and interpolating to the other points (see Ching, Erdem and Keane (2013) for further details). But we now discuss a number of alternative heuristic and approximation approaches that have been proposed more recently:

A recent trend in the economics/marketing literatures has been to model consumer learning without using the full dynamic optimization and Bayesian updating framework. This may be done either by assuming (or allowing) that consumers use heuristics, or that consumers solve their “true” underlying dynamic optimization problem by approximation using heuristic methods. Gabaix and Laibson (2000) argue that since cognition is costly, sophisticated decision-makers should adopt heuristics or short cuts to reduce cognitive burden. They use a decision tree model where agents systematically prune away low probability paths. They estimate the model on lab data, using students as subjects. The model successfully captures the decisions students made in the experiment. The authors note that further research should attempt to identify a parsimonious set of parameterized algorithms, and provide a theory that describes how the parameters adjust across problems. Natural adjustment candidates include reinforcement learning and expected-payoff maximization subject to constraints on calculation and memory.

Geweke and Keane (2000) (hereafter, GK) develop a method to approximate the solution to DP problems by replacing the “future component” of the value function with a flexible function of the state variables. Specifically, they rewrite (24) as:

$$(28) \quad V(j, t|I_t) = E[U(Q_{jt}^E, P_{jt})|I_t] + F[I_{t+1}(I_t, j)|\pi_t] \quad \text{for} \quad j=0, \dots, J$$

Here $F[I_{t+1}(I_t, j)|\pi_t] \approx \beta EV(I_{t+1}|I_t, j)$ is a flexible polynomial in the state variables that approximates the “future component” of the value function. And π_t is a vector of reduced form parameters that characterize the future component. The structural parameters of the current payoff function are then estimated jointly with the reduced form parameters of this polynomial approximation. GK showed that this method, which involves no greater computational burden than estimating a static discrete choice model, uncovers estimates of the structural parameters that exhibit negligible bias.

It is interesting to note that the GK approach is equivalent to directly assuming a simple parameterization of the G function in (27). For instance, one might assume that the information gain from choosing j over k is an increasing function of the perception error variance for j relative to that of k (see equation (6)).

Houser et al. (2004) extend GK to introduce a new Bayesian procedure for drawing inferences about both the nature and number of decision rules that are present in a population of subjects, where each subject is confronted with a dynamic decision problem. More specifically, a game is designed for an experiment in which participants play for money. Participants are allowed to practice to learn the game before playing. Data from the experiment shows that some participants make close to optimal decisions, while others appear to use simple heuristic rules (some of which are less accurate than others). The main take-away of this paper is that there is significant heterogeneity in how consumers solve dynamic problems.

Ching, Erdem and Keane (2014a) use the GK approach to study consumer’s incentive to experiment with unfamiliar brands in the diaper category. They also provided some new insights on identification in the GK framework. Notice that, as F in (28) is just a flexible function of the

state variables, all that is assumed in the GK approach is that consumers understand the laws of motion of the state variables (i.e., how $(I_{t+1} | I_t, j)$ is formed). They need not form expectations based on the true model. The approach is also agnostic about whether consumers use Bayesian updating or some other method. In general, identification of π_t requires either: (i) observing current payoffs,¹⁹ or (ii) exclusion restrictions, such that some variables enter the future component F but not current utility. Ching et al (2014a) point out that such exclusion restrictions arise naturally in dynamic learning models, because the updated perception error variances $\sigma_{ij,t+1}^2$ only affect future payoffs, not current utility.

As one can see from equation (25), when the full structure is not imposed, we will not be able to identify the discount factor. The β is subsumed as a scaling factor for the parameters π_t of the F function. However, one can test whether $\pi_t = 0$, which is a test for forward-looking behavior (or “strategic trial”). Although this test makes weak assumptions about F , it is *not* non-parametric, as a functional form must be chosen for the current payoff function. As Ching et al (2014) show, given the current payoff function, the π_t are identified in the learning model because different current choices lead to different values of next period’s state variables (e.g., the posterior variances in equation (6)). Ching et al (2014a) find evidence of forward looking behavior in the diaper category.

Another way to reduce computational burden in dynamic models is to rely on “index strategies.” These include the Gittin’s index developed by Gittins and Jones (1974) and Whittle’s index developed by Whittle (1988). To understand how these approaches work, return to equations (26)-(27), but now assume the choice is between brand j and a hypothetical certain alternative (denoted by 0) that delivers a fixed payoff λ_j and that, when chosen, leads to no gain

¹⁹ In labor economics, researchers may argue that wages capture much of the current payoff. Or, researchers can control current payoffs in a lab experiment (e.g., Houser, Keane and McCabe, 2004).

of information. Then we can write that:

$$(29) \quad V(j, t|I_t) - V(0, t|I_t) = [Ef(Q_{jt}^E|I_t) - w_P P_{jt} + e_{jt}] - \lambda_j + G(j, 0, I_t)$$

where:

$$(30) \quad G(j, 0, I_t) \equiv \beta[EV(I_{t+1}|I_t, j) - EV(I_t)]$$

Here $EV(I_t)$ is simply the expected value of arriving in the next period with no more information than one has today. One can see that the consumer will be indifferent between alternative j and the hypothetical certain alternative if:

$$(31) \quad \lambda_j \equiv [Ef(Q_{jt}^E|I_t) - w_P P_{jt} + e_{jt}] + G(j, 0, I_t)$$

That is, the sure payoff λ_j from choosing the hypothetical alternative is equal to the expected payoff from alternative j plus the value of information gained by choosing j . The value of λ_j in (31) is known as Whittle's index.

Whittle's (1988) result is that, under certain conditions, it is optimal in each period to choose the brand j that has the highest λ_j . This greatly simplifies the dynamic optimization problem because, instead of a dynamic problem with J choices, one only has to solve a set of J simple and independent optimal stopping problems. In each of those simple problems, agents chose in each period between a single brand j (for $j = 1, \dots, J$) and the certain option 0.²⁰

However, for this simplification to work, the key condition that must be satisfied is that, if brand j is chosen at time t , the information sets for all brands $k \neq j$ must remain unchanged from period

²⁰ Each sub-problem can be characterized as follows. A consumer either chooses a fixed reward in each period forever, or chooses brand j this period. If he/she chooses brand j this period, a noisy quality signal about brand j will be revealed, and then the consumer faces these two choices again next period. The reasons the index method provides significant computational gains are: (a) it reduces the size of the state space from N^J to $J \times N$, where N is the number of state points associated with each alternative, and (b) solving for the index strategy for J optimal stopping problems is much less costly compared with solving one J dimensional dynamic programming problem.

t to $t+1$. This rules out exogenous sources of information (such as the advertising signals in Erdem and Keane (1996), word of mouth, etc.), as well as correlated learning across brands.

The Gittin's index is basically a simplified version of the Whittle index that rules out exogenous shocks to current utilities of the various brands. The first paper in marketing to use an index strategy to solve a dynamic model was Eckstein, Horsky and Raban (1988), who used the Gittin's index. When consumers are risk-neutral (i.e., Q_{jt}^E enters the utility function linearly) and there are no random shocks to the values of the alternatives, a solution based on choosing the alternative with the highest Gittin's index is exactly the same as solving a full-fledged dynamic programming problem. However, in the typical random utility framework used in marketing and economics, a solution based on Gittin's index may not be optimal. Moreover, if consumers are risk-averse, the Gittin's index may not exist. Even if it exists, it does not necessarily lead to the optimal choice.

Lin, Zhang and Hauser (2015) show how to use the Whittle index to deal with dynamic random utility models that include learning through use experience as well as risk aversion and both observed and unobserved shocks to the utilities of alternatives (the literature classifies these as "restless-bandit" problems). They allow for direct persuasive effects of advertising, but do not allow advertising to convey information about brand quality (which, as noted above, would violate the Whittle index assumptions). Lin et al. (2015) apply their model to IRI diapers data and find evidence for forward-looking behavior (i.e., strategic trial). That is, a forward-looking learning model fits the data better than a myopic learning model.

Lin et al. (2015) also show that the Whittle index provides a solution to the DP program that is very close to the Keane and Wolpin (1994) approximate solution, but at lower computational cost. They also argue that an index strategy would be intuitive to consumers, so

that it is plausible that consumers follow heuristics that are close to an index strategy. This argument can be understood by looking at equations (26)-(27). Clearly, the optimal decision rule is equivalent to the static decision rule except that a value of the gain from acquiring information is added on to the value of each alternative. It seems intuitive that consumers understand there is some value to the information gained by trying out unfamiliar brands, and that they would try to take this into account when making purchase decisions.

Notably, the GK approach is equivalent to a simplified index strategy where the analyst directly chooses a functional form for the gain from gathering information.²¹ The GK method then involves directly inferring from the data the (possibly suboptimal) index rule or heuristic that rationalizes consumer choice behavior.

Sauer (2015) develops what he calls a “hybrid” approach that combines GK with Keane and Wolpin (1994). He assumes that consumers can look one period ahead – that is, they can backsolve a dynamic programming problem optimally from one period ahead – but at $t+2$ they use the GK approach to approximate the future expected value functions as simple functions of the state variables. An advantage of this approach is that it allows one to estimate the discount factor.

Tehrani and Ching (2016) propose another heuristic concept called the Value of Perfect Information (VPI), which dates back Howard (1966). The basic idea of this concept is to capture the expected gain of finding out the true value of choosing alternative j . To illustrate how to obtain the VPI for alternative j , let’s first reorder the alternatives such that the best myopic choice is 1 based on $I(t)$, that is,

²¹ To see this, compare (28) with (29)-(30). Clearly, the GK approach amounts to choosing a parameterization for the $G(j, 0, I_t)$ function.

$$(32) \quad E[U_1|I(t)] > E[U_2|I(t)] > \dots > E[U_j|I(t)].$$

Let's consider alternative $j=1$ first. Suppose the true quality is Q_j^* . This knowledge is valuable if it reveals that the original best myopic choice is no longer the best, i.e., $U_1(Q_1^*) < E[U_2|I(t)]$ (because that will lead the consumer to choose alternative 2 instead of 1); otherwise, the new knowledge does not change choice and hence renders no gain. Similarly, for alternatives $j>1$, the knowledge of true Q_j^* is valuable only if $U_j(Q_j^*) > E[U_1|I(t)]$ as this will trigger the consumer to choose alternative j over the original best myopic choice. To illustrate how to obtain VPI, let's define a gain function as follows:

$$(33) \quad Gain_{jt}(Q_j^*) = \begin{cases} E[U_2|I(t)] - U_1(Q_1^*) & \text{if } j = 1 \quad \text{and} \quad U_1(Q_1^*) < E[U_2|I(t)]; \\ U_j(Q_j^*) - E[U_1|I(t)] & \text{if } j > 1 \quad \text{and} \quad U_j(Q_j^*) > E[U_1|I(t)]; \\ 0 & \text{otherwise.} \end{cases}$$

However, the consumer is uncertain about Q_j^* . Hence, he can only compute the expected $Gain_{jt}(Q_j^*)$ based on his/her prior belief at t , $f_{jt}(\cdot)$, and this gives us the VPI associated with alternative j :

$$(34) \quad VPI_{jt} = \int_{-\infty}^{\infty} Gain_{jt}(x) f_{jt}(x) dx.$$

Tehrani and Ching (2016) propose including VPI_{jt} as an additively separable variable to the current expected utility associated with alternative j . In a sense, VPI_{jt} is a replacement for the expected future value, which we normally obtain by solving a dynamic programming problem.

One way to interpret VPI_{jt} is that consumers look one period ahead and assume that all uncertainty will be resolved by one trial. It does not take into account that learning could be slow (as when signals are noisy, so that one-trial-learn-everything cannot be achieved). This

shortcoming can be addressed by modifying the definition of VPI_{jt} above by using the Bayesian updating formula to take the variance of the noisy signals into account. It is worth highlighting that the computational burden of solving for VPI_{jt} is relatively light, because it only involves solving a one-dimensional integration instead of a dynamic programming problem. In fact, it is much easier to implement the VPI approach compared with the index strategy which still requires solving for J dynamic programming problems (optimal stopping problems). Therefore, it is conceivable that consumers may adopt a heuristic like VPI. Tehrani and Ching (2016) provide evidence that the VPI approach can explain the brand choice dynamics well in the diaper category.

The concept of learning has also been treated from the viewpoint of a company, which learns how to match the “look and feel” of a web site to the cognitive styles of consumers (a process known as “website morphing”). Hauser et al. (2009) use clickstream data to infer cognitive styles. Their proposed model balances exploration (learning how morphing affects consumer choice probabilities) with exploitation (maximizing short term sales) by solving a dynamic program using Gittin’s index. The authors apply their Bayesian updating and dynamic programming model to an experimental British Telecom web site. Findings reveal that adaptation of such approaches can lead to substantial additional revenues.

Last but not least, Dzyabura and Hauser (2011) develop and test an active machine-learning model to identify heuristic decision-making. They illustrate their algorithm using data from a web-based survey conducted by an American automotive manufacturer to study vehicle consideration. The conjoint experiment included 872 respondents and 53 feature-levels. The authors conclude that active machine learning is an effective methodology to select questions

adaptively in a conjoint context in order to identify consideration heuristics. But many challenges remain open in this area.

One point that should be stressed about all heuristic approaches is they are subject to the Lucas-Marschak critique of using reduced form models to predict the effect of policy changes. This is because a heuristic that works well in one environment may perform poorly in another. Thus, if the policy environment changes, people may change the heuristics that they use.

For example, consider demand for diapers. In an environment where stores keep prices fairly stable, except they put diapers on sale on most Fridays, consumers may well hit on a (close to cost minimizing) heuristic that simply says "Buy diapers on Friday." However, if stores start to randomize the day of sales, consumers will presumably change their heuristic. One would need a structural model to forecast the new heuristic that consumers adopt in the new context.

Similarly, several papers we have discussed are motivated by the idea that consumers use heuristics to reduce information processing costs. But such costs are a function of the market environment. For example, say a consumer faces a choice between only two product varieties. In this case, he/she may compare all their attributes carefully. But if the variety in the market expands to 50 items this becomes infeasible. Then, the consumer may adopt a lexicographic rule (e.g., first screen by price range, then by color, etc.) as a simplifying heuristic. Again, one would need a structural model to predict the point at which the choice environment gets sufficiently complex that people switch from a compensatory to a lexicographic decision rule.²²

Similar arguments apply to the use of approximate/heuristic approaches to solving dynamic optimization problems. An approximation that is accurate in one context may be

²² Another example is that the advent of internet retailing has made it possible to do comparison shopping from home, thus arguably reducing the costs of gathering information. An interesting hypothesis is that this change in the environment may have caused consumers to engage in more comparison shopping.

inaccurate in another. For example, as noted by Keane and Wolpin (1994), the expected maximum of several alternative specific value functions (the “E max”) is well approximated by maximum of the expected values of those same value functions (the “max E”) *provided* that the choice specific error variances are small. But this approximation breaks down in an environment with more uncertainty. So, while the use of heuristic-based models may ease the burden of econometric estimation, they are unlikely to substitute for more structural approaches when it comes to the problem of predicting behavior under very different policy regimes.

These caveats should not be taken to diminish the potential importance of heuristic-based models in providing valuable information about how consumers actually behave in particular real world choice environments. Our point is simply that heuristic-based models (like reduced form models) should only be used to predict behavior in response to policy changes if it is plausible to maintain that the choice heuristic is invariant to the policy change.²³

8. Using Exogenous Events and Policy Changes to Study Learning

In recent years, there has been an increased interest in studying the nature of consumer learning by exploring the impact of exogenous events or policy changes. For example, one event that has generated a great deal of attention is the implementation of the Medicare prescription

²³ One way to interpret our argument is that only structural models *attempt* to predict what decision rules consumers will adopt in a new environment (indeed, this is precisely what structural models are designed to do). But that doesn’t mean their predictions will necessarily be correct. It is important to keep in mind the point that a structural model is only invariant to all conceivable environmental changes if it is perfectly correctly specified – that is, if it is in fact the “true model.” As all models are ultimately false (as they are simplifications), a completely policy invariant model is an aspirational goal, not a reality. The best we can do in practice is to incrementally validate a structural model by showing that it predicts well across a range of policy environments. This may give us confidence in using the model to predict in a new environment. But we can never be certain that the new environment won’t be the one that reveals the flaw in the model! As a practical matter, the best we can hope for is to build structural models we are confident in using for certain types of policy predictions, but perhaps not for others (i.e., it is perfectly possible that a structural model can reliably predict responses to some types of policy changes but not others – just as we see with commonly used models in the physical sciences and engineering). See Keane (2010) for further discussion of these issues.

drug plan in the US (known as Medicare “Part D”). This program went into effect in 2006, at which point the federal government created (via heavy subsidies) a new private market in prescription drug plans for people 65 and over. This event created a unique opportunity to study consumer learning behavior, because, essentially over-night, a large number of alternative prescription drug plans were suddenly made available to senior citizens.

On average, each senior citizen faced a choice among roughly 50 drug plans, offered by 20 different insurers. These plans vary in terms of premium, out-of-pocket costs, formulary, the ease of seeking reimbursement, and customer service. With such a degree of complexity, it is likely that a significant portion of consumers make uninformed decisions. Other than uncertainty about plan’s attributes, consumers may also be uncertain about their state of the world in the coming year (i.e., how sick he/she could become and hence the type of drugs needed). Because the timing of decisions is clear (open enrolment happens once a year), this market provides an excellent opportunity to study how consumers decide to consider switching (Ching, Erdem and Keane 2009; Ketcham, Lucarelli and Powers 2015; Ching and Lim 2016). Moreover, the data on their initial choice, subsequent choices, and actual spending patterns potentially provide researchers with information about how consumers learn about which plan fits them best over time.

Ketcham et al. (2012) find evidence that consumers are in fact “learning” about these plans over time. They study whether Medicare Part D enrollees improved over time in terms of reducing overspending. They find that the mean cost difference between individuals’ actual choices and their cheapest option fell by about \$330 from 2006-2007 and this average reduction in overspending is in part due to individuals who chose to switch plans. The likelihood of

switching plans for 2007 increased substantially with the amount of overspending in 2006. They attribute these results to participants learning about the costs and benefits of different plans.

However, given the complexity of the market, it is possible that a simple Bayesian learning model may still miss many important features in the data. As pointed out in Ching, Erdem and Keane (2013), having stated preference data could potentially enhance researchers' ability to build a better structural model as a closer proxy to actual behavior.

To address this research agenda, Ketcham, Kuminoff and Powers (2015) have linked the claim data made available by The Center of Medicare and Medicaid Services with the Medicare Current Beneficiary Survey conducted three times per year. Using this data set, Ketcham et al. (2015) are able to separate informed and uninformed consumers. Their main research question is to estimate how consumers may change their choices and the welfare consequences under several counterfactual policies. But this data set can also potentially tell us how confused or uninformed consumers are. However, so far researchers in this area have mainly relied on a simple static multinomial logit model to draw inferences. It would be very interesting to develop a model with limited foresight and formation of consideration sets to understand how consumers choose in such a complicated environment.

Turning to a different type of example, Sudhir and Yang (2015) explore free upgrade events in car rentals, and argue that they provide exogenous "random" assignment of car types to consumers, independent of their preferences. As a result, they argue that stickiness of preferences after such consumption experiences can be used to draw causal inferences about state dependence.

Similarly, Larcom et al. (2015) use another exogenous event to study the stickiness of choice that leads to suboptimal experimentation. The February 2014 London underground train

strike temporarily shut down some stations, and that forced some commuters to look for alternative routes to get to their destination. Larcom et al. (2015) argue that this provides them with an opportunity to investigate whether commuters were using the best route prior to the strike, or whether they had not explored all the options yet and were using a suboptimal route. The strike forced some commuters to experiment, but because not all commuters were affected (some stations operated during the strike), Larcom et al. (2015) are able to use the unaffected commuters as a control group, and apply the difference-in-difference approach to test their hypothesis. Interestingly, they find that a majority of commuters return to the original routes after the strike; but a small percentage switched to the new routes, suggesting they were using a suboptimal route before.

Gallagher (2014) uses an event study framework to estimate the effect of large regional floods on the take-up of flood insurance. He finds that insurance take-up spikes the year after the flood and then declines steadily to baseline. Residents in non-flooded communities in the same TV media market increase take-up at one third rate of the flooded communities. Gallagher's findings are consistent with a Bayesian learning model with forgetting and/or incomplete information about past floods. Thus, the form of belief updating is an area where it may be important to relax or generalize the behavioral assumptions of standard Bayesian learning models (a point we return to in Section 9).

Finally, Davis (2004) measures the impact of an outbreak of pediatric leukemia on local housing values. A model of location choice is used to describe conditions under which the gradient of the hedonic price function with respect to pediatric leukemia risk is equal to household marginal willingness to pay to avoid risk. The equalizing differential is estimated using property-level sales records from a county of Nevada where residents experienced a severe

increase in pediatric leukemia. Housing prices are compared before and after the increase with a nearby county acting as a control group. The results indicated that housing values decreased 15.6 percent during the period of maximum risk. Using lifetime estimates of risk derived from a Bayesian learning process, the results imply the statistical value of pediatric leukemia is \$5.6 million. The approach adopted in this study suggests avenues for future research on quantifying the trade-offs between money and various risks, such as health risks.

9. Future Research Directions

In this section we discuss some largely unexplored territories for future research and/or areas that are under-researched. In Section 7, we discussed papers that approximate agents' optimization decisions with heuristic methods, as well as papers that attempt to capture agents' use of heuristics. Such use of heuristics and/or behavioral phenomena that deviate from typical rational decision-making may be important for understanding choice in many contexts.

One important way that people may deviate from the behavioral assumptions of standard Bayesian learning models is in how they update beliefs in response to new information. There is, for example, evidence that consumers either over or under-react to information. There is also evidence suggesting that consumers may over-react to experience signals when learning opportunities happen infrequently, and then "forget" that experience rather quickly over time. Agarwal et al. (2013) measure learning and forgetting in the credit card market, using a panel with four million monthly credit card statements. They find that paying a fee last month, which is a negative feedback, reduces fee payment in the current month 40%. However, the authors also find evidence for recency effects and a 10% or more depreciation of knowledge per month.²⁴ The

²⁴ Interestingly, higher-income borrowers learn twice as fast, and forget twice as slowly, as lower-income borrowers.

paper by Gallagher (2014) on flood insurance that we discussed in Section 8 also presents results that appear consistent with forgetting.

Another area of particular interest is choice in very complex situations. What we mean by a “complex” choice situation includes cases where:

- a) The object under consideration is complex, in that it has many attributes, or some attributes that are difficult to understand or evaluate;
- b) The choice set is complex because there are a very large number of alternatives;
- c) Choice requires evaluating probabilities and/or making intertemporal allocations.

Good examples of what we mean are choices in areas such as health/life insurance, retirement plans or investments. These types of choices all arise in the area of optimal life-cycle planning, and they exhibit all three factors that contribute to complexity of the decision task as described above. However, optimal life-cycle planning is an area largely ignored in the standard learning literature, and the marketing literature in general.²⁵

Optimal life-cycle planning requires the solution of a complex dynamic programming (DP) problem. But actual decision making in the domain of such planning (e.g., retirement planning, saving for college education of children, etc.) often departs in obvious ways from this normative principle, and people often seem to react to the difficulty of the problem with the use of simple heuristics, or even with delays and procrastination (see Keane and Thorp (2015) for a review).

Interestingly, methods that appear to be relevant for such planning problems have already been developed for the closely related problem of inventory planning. Specifically, Ching, Erdem and Keane (2009, 2014a) – henceforth CEK – develop a model of consumer demand for a

²⁵ An exception is Yang and Ching (2014) who develop and estimate a consumer life-cycle model to explain the adoption decision of a new technology.

storable (or quasi-durable) branded commodity. In this context, optimal behavior involves: (i) checking the prices of all brands of a product in every period, and (ii) solving a DP problem to determine both (ii-a) the reservation price for purchase of each brand, and (ii-b) the optimal quantity to buy in the event that the price of a brand is below its reservation price. Of course, the reservation prices and optimal purchase quantities both evolve in a complex way with inventories.

CEK argue that a normative model is unrealistic for two reasons: (i) For most products, consumers presumably do not have the time, interest or mental capacity to check all prices in every period, and (ii) in those periods when consumers do pay close attention to a category, they presumably make decisions using more or less sophisticated rules of thumb, not by literally solving a DP problem. Thus, CEK develop a two-stage model of demand for a storable branded product. In the first stage, consumers decide whether or not to pay attention to the product category.²⁶ If they do decide to pay attention then, in stage two, they use a rule of thumb that may or may not provide a good approximation to the DP problem (depending on parameter estimates), as in Geweke and Keane (2000, 2001). In CEK's empirical applications, the decision whether to consider a category is modelled as a simple probit or logit discrete choice model, where the factors that drive consideration are cues like advertising, displays and low inventory.²⁷

It seems fairly clear how one might apply the CEK framework to financial products like annuities, life insurance or choice of retirement plans. As discussed in Keane and Thorp (2015), there is clear evidence that most consumers are averse to thinking about these products on a

²⁶ CEK's work was originally motivated by the observation that brand choice *conditional* on category purchase is very sensitive to price, while the decision to make a purchase in a category is quite insensitive to price. CEK showed that these seemingly contradictory facts could be explained if consumers only occasionally look at (i.e., consider) a category.

²⁷ In the optimal solution consumers should consider a category in every period regardless of their inventory. Even if inventory is high, a low enough price would make it optimal to stock up even more.

regular basis. For example, as is well-known, the typical consumer does not engage in a frequent re-balancing of his/her stock portfolio as the state of the world changes. It is natural to think of a framework where, in a first-stage, consumers decide on, say, a quarterly or annual basis whether to consider financial products in a certain category. The decision to consider could be driven by advertising cues, as well as by major life events such as retirement, children leaving home, a spouse passing away, selling a house and/or moving house, or reaching a milestone birthday. In the second stage, it would again be optimal to estimate a behavioral rule of thumb from the data, rather than imposing an optimal DP solution.

Given such an estimated model, one could simulate behavior under the model vs. under a normative solution to the planning problem. One could then evaluate whether or not the wealth losses from following the simplified decision process rather than the optimal DP solution are substantial (as in Houser et al. 2004).

Another under-researched area is the endogenous formation of preferences. Malmendier and Nagel (2011) present evidence that risky asset returns experienced over the course of an individual's life have significant effects on the willingness to take financial risks. People who have experienced high stock market returns report lower aversion to financial risks, are more likely to participate in the stock market, and allocate a higher proportion of their liquid asset portfolio to risky assets. While individuals put more weight on recent returns than on more distant realizations, the impact fades only slowly with time. Malmendier and Nagel remain agnostic whether the experience effects on risk taking arise from experience-dependent beliefs or from endogenous risk preferences. Their results show, however, that there is dependence on "experienced data" – as opposed to "available data" – in standard rational and boundedly rational learning models. The implications of this in a variety of contexts (from exploring micro issues

such as how heterogeneity arises among agents to macro issues, such as the dynamics of asset prices in this specific case) are another interesting avenue for future research.

Finally, another avenue for future research is to bring new data to bear on the problem of identifying learning processes – that is, information that goes beyond just the history of signals received and choice made. For example, Erdem, Keane, Oncu and Strebler (2005), who combined psychometric data on product ratings with revealed preference to help identify parameters of the learning process. In another example, Ching, Erdem and Keane (2014b) develop a novel way to measure the extent to which consumers are uncertain about the true quality of quasi-durable goods. Using data on diapers, they argue that data on consumer buying decisions reveals their subjective perceived qualities, while data on inter-purchase spells reveals the objective quality (at least in terms of durability). The difference of these two measures tells us to what extent consumers are uncertain about the quality of the products.

10. Conclusion

In this review paper, we discussed recent developments in the literature on consumer choice dynamics and learning. Erdem and Keane (1996) showed that a simple Bayesian learning model was quite successful in explaining the observed dynamics in consumer choice behavior. Their application was to learning via use experience and advertising about frequently purchased consumer goods. Since that time, a large literature has developed that extends learning models both in terms of (i) the sources of information and types of learning that take place and (ii) the nature of the objects that agents are learning about.

In particular, recent years have seen learning models extended to include learning from experiences of friends or other social network members (“social learning”), experience with related products (“correlated learning” or “information spillover”), examination of publicly

available information or expert opinion (“search”), and inferences about product attributes from purchase decisions of others (“observational learning”). Similarly, learning models have been successfully applied to many different types of products and activities including high-tech durables, drugs, medical procedures, movies, books, video games, restaurants, sexual behaviors, new product ideas, health insurance plans, and so on.

In this review we have summarized papers that find evidence of learning as one of the main mechanisms that explains dynamics in consumer choice behavior for a wide range of products or activities in a wide variety of settings. We have also summarized papers that find evidence of learning from many different sources. Learning models have also been applied to model firm learning about consumer willingness to pay for goods and market entry potential.

One important recent development has been the adoption of heuristic approaches to model consumer and firm learning and decision-making in complex environments. The basic idea is to relax some of the normative assumptions of Bayesian learning models to allow for cognitive limitations and/or behavioral biases on the part of consumers. While we have discussed some work in this area, it remains an under-researched topic that should be a fruitful avenue for future research.

Another key area for future research is to develop models that integrate learning with other potentially important sources of choice dynamics, such as inventories, switching costs, habit persistence, as well as behavioral factors like inattention/procrastination. For example, in Ching, Erdem and Keane (2014a) we develop a choice framework that integrates learning, inventories and rational inattention, while relaxing some assumptions of the Bayesian learning model. But much more work remains to be done in this area.

In conclusion, learning models have proven to be a fruitful area of research activity for the past 20 years, and the level of activity in this area has been growing substantially. This is reflected in the fact that the bulk of our references are from just the past few years. We expect this literature to continue to grow rapidly as learning models are applied to more and more domains, and as researchers continue to generalize the structure of the original Bayesian learning models in Roberts and Urban (1988), Eckstein et al. (1988) and Erdem and Keane (1996) in many interesting new directions.

References

- Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson. 2013. Learning in the Credit Market. Working Paper.
- Aguirregabiria, V. and P. Mira. 2007. Sequential Estimation of Dynamic Discrete Games. *Econometrica*, 75(1): 1-53.
- Chan, T., J. Li and L. Pierce. 2015. Learning from Peers: Knowledge Transfer and Sales Force Productivity Growth. *Marketing Science*, 33(4): 463-484.
- Che, Hai, Tülin Erdem and T. Sabri Öncü. 2015. Consumer Learning and Evolution of Consumer Brand Preferences. *Quantitative Marketing and Economics*, 13(3), 173-202.
- Chen, T., B. Sun and V. Singh. 2009. An Empirical Investigation of the Dynamic Effect of Marlboro's Permanent Pricing Shift. *Marketing Science*, 28(4), 740-758.
- Ching, A.T. 2010a. Consumer learning and heterogeneity: dynamics of demand for prescription drugs after patent expiration. *International Journal of Industrial Organization*, 28(6): 619-638.
- Ching, A.T. 2010b. 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. 2009. The Price Consideration Model of Brand Choice, *Journal of Applied Econometrics*, 24(3): 393-420.
- Ching, A.T., T. Erdem and M. Keane. 2013. Learning Models: An Assessment of Progress, Challenges and New Developments, *Marketing Science*, 32(6): 913-938.
- Ching, A.T., T. Erdem and M. Keane. 2014a. A Simple method to Estimate the Roles of Learning, Inventories and Category Consideration in Consumer Choice, *Journal of Choice Modelling*, 13: 60-72.
- Ching, A.T., T. Erdem and M. Keane. 2014b. How Much Do Consumers Know About the Quality of Products? Evidence from the Diaper Market. Working paper, Rotman School of Management, University of Toronto.
- Ching, A.T., R. Clark, I. Horstmann and H. Lim. 2016. The Effects of Publicity on Demand: The Case of Anti-cholesterol Drugs, *Marketing Science*, 35(1): 158-181.
- Ching, A.T. and M. Ishihara 2010. The Effects of Detailing on Prescribing Decisions under Quality Uncertainty, *Quantitative Marketing and Economics*, 8(2): 123-165.

- Ching, A.T. and H. Lim 2016. A Structural Analysis of Correlated Learning and Late-Mover Advantages: The Case of Statins. Working Paper, Rotman School of Management, U of Toronto.
- Davis, L. 2004. The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster. *American Economic Review*, 94(5): 1693-1704.
- De los Santos, B., A. Hortacsu, M. R. Wildenbeest. 2013. Search with Learning. Working paper, Kelley School of Business, Indiana University.
- Dzyabura, D. and J. R. Hauser. 2011. Active Machine Learning for Consideration Heuristics, *Marketing Science*, 30(5): 801–19.
- Eckstein, Z., D. Horsky, and Y. Raban. 1988. An Empirical Dynamic Model of Brand Choice. Working paper 88, University of Rochester.
- Erdem, T. 1998. An Empirical Analysis of Umbrella Branding. *Journal of Marketing Research*, 35(3): 339-351.
- Erdem, T. and M. 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. Keane, S. Oncu and J. Strebel. 2005. Learning About Computers: An Analysis of Information Search and Technology Choice, *Quantitative Marketing and Economics*, 3(3): 207-246.
- Erdem, T., M. Keane and B. Sun. 2008. A Dynamic Model of Brand Choice when Price and Advertising Signal Product Quality, *Marketing Science*, 27(6): 1111-25.
- Gabaix, X., and D. Laibson. 2000. A Boundedly Rational Decision Algorithm. *American Economic Review*, 90(2): 433–38
- Gallagher, J. 2014. Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States. *American Economic Journal: Applied Economics* 6(3); 206-233.
- Geweke, J. and M. Keane. 2000. Bayesian Inference for Dynamic Discrete Choice Models without the Need for Dynamic Programming. *Simulation Based Inference and Econometrics: Methods and Applications*, Mariano, Schuermann and Weeks (eds.), Cambridge University Press, 100-131.
- Geweke, J. and M. Keane. 2001. Computationally Intensive Methods for Integration in Econometrics. *Handbook of Econometrics: Vol. 5*, J.J. Heckman and E.E. Leamer (eds.), Elsevier Science B.V., 3463-3568.

- Gittins J. and D. Jones. 1974. A Dynamic Allocation Index for the Sequential Design of Experiments. In Gani, J., K. Sarkadi and I. Vince (eds). *Progress in Statistics* (North-Holland, Amsterdam), 241-266.
- Godlonton, S. and R. L. Thornton. 2013. Learning from Others' HIV Testing: Updating Beliefs and Responding to Risk. *American Review Papers & Proceedings*, 103(3): 439-444.
- Guadagni, P.M. and John D.C. Little. 1983. A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science*, 2(3), 203-238.
- Hauser, J. R., G. L. Urban, G. Liberali and M. Braun. 2009. Website Morphing. *Marketing Science*. 28(2), 202-223.
- Heckman, J.J. 1981. Statistical Models for Discrete Choice Panel Data. In *Structural Analysis of Discrete Data with Econometric Applications*, C.F. Manski, D. McFadden (eds). MIT Press: Cambridge, MA: 114-178.
- Hendrick, K., A. Sorensen, and T. Wiseman. 2012. Observational Learning and Demand for Search Goods. *American Economic Journal: Microeconomics*, 4(1): 1-31.
- Ho, T.-H., S.-E. Park and X. Su. 2015. A Bayesian Level- k Model in n -Person Games. Working Paper.
- Houser, D., M. Keane, and K. McCabe. 2004. Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm. *Econometrica*, 72(3): 781-822
- Howard, R.A. 1966. Information Value Theory. *IEEE Transactions on System Science and Cybernetics*, vol.2(1): 22-26.
- Huang, G., H. Luo and J. Xia. 2015. Invest in Information or Wing it? A Model of Dynamic Pricing with Seller Learning. Working paper, Carnegie Mello University.
- Huang, Yan, Param Vir Singh and Kannan Srinivasan. 2014. Crowdsourcing New Product Ideas Under Consumer Learning. *Marketing Science*, 60(9): 2138-2159.
- Hummel, P. and B. Knight. 2015. Sequential or Simultaneous Elections? A Welfare Analysis. *International Economic Review*, 56(3): 851-887.
- Keane, M. 1994. A Computationally Practical Simulation Estimator for Panel Data. *Econometrica*, 62(1): 95-116.
- Keane, M. 2010. Structural vs. Atheoretic Approaches to Econometrics. *Journal of Econometrics*, 156(1): 3-20.
- Keane, M. 2015. Panel Data Discrete Choice Models of Consumer Demand. Oxford Handbook of Panel Data, B.Baltagi (ed.), Oxford University Press, Chapter 18, p. 548-582.

- Keane, M. and S. Thorp. 2015. Complex Decision Making: The Roles of Cognitive Limitations, Cognitive Decline and Ageing, *The Handbook of Population Ageing*, Elsevier, J. Piggott and A. Woodland (eds), forthcoming.
- Keane, M. and K. Wolpin. 1994. The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation: Monte Carlo Evidence. *Review of Economics and Statistics*, 76(4): 648-672.
- Ketcham, J. D., C. Lucabell, E. J. Miravete, and M. C. Roebuck. 2012. Sinking, Swimming, or Learning to Swim in Medicare Part D. *American Economic Review*, 102(6): 2639-2673.
- Ketcham, J. D., C. Lucabell and C.A. Powers. 2015. Paying Attention or Paying Too Much in Medicare Part D. *American Economic Review*, 105(1): 204-233.
- Ketcham, J. D., N. V. Kuminoff and C.A. Powers. 2015. Estimating the Heterogeneous Welfare Effects of Choice Architecture: An Application to the Medicare Prescription Drug Insurance Market. Working paper, Arizona State University.
- Knight, B. and N. Schiff. 2010. Momentum and Social Learning in Presidential Primaries, *Journal of Political Economy*, 118(6): 1110-1150.
- Koulayev, S. 2013. Search with Dirichlet Priors: Estimation and Implications for Consumer Demand. *Journal of Business and Economic Statistics*, 31(2): 226-239
- Larcom, S., F. Rauch and T. Willems. 2015. The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network. Working paper, University of Oxford.
- Lee, J. Y. and D. R. Bell. 2013. Neighborhood Social Capital and Social Learning for Experience Attributes of Products. *Marketing Science*, 32(6): 960-976.
- Lin, S., J. Zhang, and J. R. Hauser. 2015. Learning from Experience, Simply. *Marketing Science*, 34(1): 1-19.
- Liu, Y. and M. Ishihara. 2015. Causal Impact of Online Reviews on Durable Goods Adoption: The Case of the U.S. Video-Game Market. Working paper, Stern School of Business, NYU.
- Liu, X., J.H. Kagel and L.F. Lee. 2012. Learning from Peers in Signaling Game Experiments. *Journal of Applied Econometrics*, 27, 1037-1058.
- Lovett, M. J. and R. Staelin. 2016. The Role of Paid, Earned, and Owned Media in Building Entertainment Brands: Reminding, Informing, and Enhancing Enjoyment. *Marketing Science*, 35(1): 142-157.
- Malmendier, U. and S. Nagel. 2011. Depression Babies: Do Microeconomic Experiences Affect Risk-Taking? *The Quarterly Journal of Economics*, 126(1): 373-416.

- Milgrom, P. and J. Roberts. 1982. Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis. *Econometrica*, 50(2): 443-459.
- Newberry, P.W. 2016. An Empirical Study of Observational Learning. *RAND Journal of Economics*, 47(2): 394-432.
- Roberts, J. and G. Urban. 1988. Modelling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice. *Management Science*, 34(2): 167-185.
- Roos, J.M.T., C.F. Mela and R. Shachar. 2015. The Effect of Links and Excerpts on Internet News Consumption. Working paper, Rotterdam School of Management, Erasmus University.
- Rothschild, M. 1974. Searching for Lowest Price when the Distribution of Prices is Unknown. *Journal of Political Economy*, 82(4): 689-711.
- Sauer, Robert M. 2015. Does it Pay for Women to Volunteer? *International Economic Review*, 56(2): 537-564.
- Sudhir K. and N. Yang. 2015. Exploiting the choice-Consumption Mismatch: A New Approach to Disentangle State Dependence and Heterogeneity. Working Paper.
- Szymanowski, M. and E. Gijsbrechts. 2012. Consumption-Based Cross-Brand Learning: Are Private Labels Really Private? *Journal of Marketing Research*. XLIX (April), 231-246.
- Tehrani, Shervin S. and A.T. Ching. 2016. A Heuristic Approach to Explore: Value of Perfect Information. Working paper, Rotman School of Management, University of Toronto.
- Wei, Y. 2015. Imitation vs. Innovation: Product Similarity Network in the Motion Picture Industry. Working paper, University of Pennsylvania.
- Whittle, P. 1988. Restless Bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25(2): 287-298.
- Wu, C., H. Che, T.Y. Chen and X. Lu. 2015. The Economic Value of Online Reviews. *Marketing Science*, 34(5): 739-754.
- Yang, B. and A.T. Ching. 2014. Dynamics of Consumer Adoption of Financial Innovation: The Case of ATM Cards. *Management Science*, 60(4): 903-922.
- Yang, N. 2016. Learning in Retail Entry. Working paper, McGill University.
- Zhang, J. 2010. The sound of silence: observational learning in the U.S. kidney market. *Marketing Science*, 29(2): 315-335.

Zhao, Y., S. Yang, V. Narayan and Y. Zhao. 2013. Modeling Consumer Learning from Online Product Reviews. *Marketing Science*, 32(1): 153-169.

Zou, D. 2014. Intertemporal Pricing of New Products: Incentivizing Consumer Learning and Inertia. Working paper, Booth School of Business, University of Chicago.