

PRICE SEARCH ACROSS STORES AND ACROSS TIME

By

Navid Mojr and K. Sudhir

March 2014

Updated March 2016

COWLES FOUNDATION DISCUSSION PAPER NO. 1942



**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY**

Box 208281

New Haven, Connecticut 06520-8281

<http://cowles.econ.yale.edu/>

Price Search across Stores and across Time

Navid Mojir and K. Sudhir*

Yale School of Management

March 2016

* Navid Mojir (navid.mojir@yale.edu) is a PhD candidate in Marketing and K. Sudhir (k.sudhir@yale.edu) is James L. Frank Professor of Private Enterprise and Management at the Yale School of Management, 165 Whitney Ave, New Haven, CT 06520. The paper is the first essay of the first author's doctoral dissertation at Yale University. We thank participants at the 2013 UTD FORMS, Marketing Dynamics and Marketing Science Conferences; marketing camps at Minnesota and Northwestern; and seminars at LBS, Georgia Tech, MIT, Tilburg, Toronto, Utah, Washington University at St. Louis, Wisconsin Madison, Yale Faculty Lunch, Yale Marketing and Yale IO prospectus seminars. We thank Ahmed Khwaja for many constructive discussions and feedback. We thank Steven Berry, Stephan Seiler and Che-Lin Su for comments.

Price Search across Stores and across Time

Abstract

In response to price dispersion across stores and price promotions over time, consumers search across both stores (spatial) and time (temporal), in many retail settings. Yet there is no search model in extant research that jointly endogenizes search in both dimensions. We develop a model of spatiotemporal search that nests a finite horizon model of spatial search across stores within an infinite horizon model of inter-temporal search. The model is estimated using an iterative procedure that formulates it as a mathematical program with equilibrium constraints (MPEC) embedded within an E-M algorithm to allow estimation of latent class heterogeneity. The empirical analysis uses data on household store visits and purchases in the milk category. In contrast to extant research, we find that omitting the temporal dimension underestimates price elasticity. We attribute this difference to the relative frequency of household stock outs and purchase frequency in the milk category. Further, contrary to the conventional wisdom that promotions increase store switching and reduces store loyalty, we find that in the presence of search frictions, price promotions can be a store loyalty-enhancing tool.

1 Introduction

Price dispersion across stores and price promotions across time is widespread in retail settings. In response, consumers can search across stores (spatial) and across time (temporal) to avail the best possible prices. Depending on their cost of search, ability to time (delay or accelerate) purchases, relative preferences for stores, and household locations with respect to stores, there is empirical evidence that consumers choose different search strategies along the space and time dimensions (Gauri, Sudhir, & Talukdar, 2008). While the Gauri et al. analysis was for grocery products involving repeat purchases, spatiotemporal search is widespread even for one-time purchases. For instance, a potential car or household appliance buyer may search for a sufficiently low price across several stores and repeat the search at these stores over many months before making a purchase and exiting the market. Though spatiotemporal search occurs often in the real world, there are no theoretical or empirical models that endogenize search on *both* the spatial and temporal dimensions. In this paper, we therefore develop and estimate a structural empirical model that endogenizes search across stores and across time.

There is a vast literature in economics and marketing on price search, both theoretical and empirical. Much of this research is focused on search around a one-time purchase in the presence of price dispersion across stores, but with no price promotions. Two types of search models dominate the search (across stores) literature. The first is the fixed sample size search model proposed by Stigler (1961), where faced with price uncertainty, consumers search at a fixed sample of stores and choose the lowest priced alternative. The second and more widely used type of model is the sequential search model proposed by McCall (1970) and Mortensen (1970), which argues that a consumer will not find it optimal to search a pre-determined fixed set of stores, when the marginal cost of the additional search may not exceed the benefit. Other notable contributions to the theoretical sequential search literature include Weitzman (1979), who introduces a dynamic programming approach to model search across stores. Consumers buy after sampling prices in the fixed sample size price search, or when they decide not to search any further in sequential search. As these models abstract away from price promotions, search along the temporal dimension (as in waiting and searching again at the stores for a low price) is never optimal in these models.

In marketing, the literature on consideration sets is based on the fixed sample size model (Roberts & Lattin, 1991; Mehta, Rajiv, & Srinivasan, 2003). Honka (2013) assumes a fixed sample size model; a reasonable assumption in the context of her study of insurance purchases. In contrast, Kim et al. (2010) assume a sequential search model to rationalize price dispersion in a differentiated

product market as does Koulayev (2009). There has been some recent work testing which of the two search models fit the data better. Using online data on price dispersion, Hong and Shum (2006) are not able to empirically assess the superiority of the two types of search models using their data. Using more detailed data on the sequence of searches across online book stores, De los Santos et al. (2012) finds that in the context of the online book retailing, there is greater support for the fixed sample size model because unlike the prediction of the sequential search model, consumers do not always purchase at the last store. To address situations, where the sequence of search is not known, but price and consideration sets only are available, Honka and Chintagunta (2013) develop an identification strategy to distinguish between sequential and simultaneous search. Bell et al. (1999) model choice between EDLP and High-low Price formats, based on the fixed cost of shopping (that does not depend on basket size) and variable costs of shopping (that does depend on basket size). But, their paper does not account for forward looking behavior.

There is also a literature on price search over time in the presence of periodic price promotions. Theoretical models include Salop and Stiglitz (1982), Conlisk, Gerstner and Sobel (1984) and Besanko and Whinston (1990). In recent years, there have been many empirical models of intertemporal price search, building off the descriptive evidence on purchase acceleration in response to price promotions using scanner data (e.g., Neslin, Henderson and Quelch 1985). For example, Erdem, Imai and Keane (2003), and Hendel and Nevo (2006) structurally model price search behavior over time allowing consumers to have the flexibility to time their purchases by either accelerating or decelerating purchases by holding inventory, or by postponing consumption itself. Some papers recognize the fact that consumers do visit and make purchase at multiple stores, but make the simplifying assumption that store visits occur due to an exogenous process (e.g., Erdem, Imai and Keane 2003; Hartmann and Nair 2010; Seiler 2013). Hartmann and Nair (2010) study the problem of inter-temporal demand estimation of tied goods (razors and razor blades) across multiple store formats, treating store visits as exogenous. Seiler (2013) studies the problem of inter-temporal price search for detergents treating store choice as exogenous, but endogenously models whether consumers will search for the price of detergents (prices of all brands are revealed if the consumer incurs the search cost) when at the store, allowing him to estimate search costs for price information in the category, conditional on visiting the store. The model provides a structural “search cost” based framework for the “price consideration” model in Ching, Erdem and Keane (2009). Our paper extends the literature by endogenizing search along both the spatial and temporal dimensions.

There are a number of modeling issues and challenges that we need to address in developing a model of search across stores and across time and applying it to frequently purchased consumer goods. First, this is a unique setting, in which we nest a dynamic model of sequential search and purchase across stores in a time period within another model of repeated purchases across time. Since the number of grocery stores that consumers search is finite, we nest a finite horizon store search problem within a larger infinite horizon problem of search across time. Second we need to allow for stockpiling and stockouts in the category, where consumer purchases are stored and consumed over multiple periods, and they may suffer from stockouts when a trip is not feasible, or the prices are high when the household runs out of inventory. As we note in the introduction, even with one-time purchases, if there is price dispersion and price promotions, our modeling framework of nesting a finite horizon model of store search embedded in an infinite horizon model of temporal search will be applicable. But without repeat purchases, stockpiling and stockout issues, it reduces to an optimal stopping problem. Finally, we need to account for the fact that store visits are driven by factors unrelated to the focal category. Extant temporal search models abstract away from this issue by assuming that store visits are exogenous.

We estimate the dynamic structural model allowing for discrete heterogeneity. Given that we model visits and (not just purchases as in extant models), the number of events included in each household's visit and purchase sequence is large enough that the likelihood of each household's observed sequence falls below machine precision. Without heterogeneity, this is not an issue as one can work by summing over the log-likelihoods of each observation. But when accounting for unobserved heterogeneity, one needs to use the weighted sum of the household's observed sequence likelihood, based on the probability of belonging to different segments. This becomes computationally infeasible when the sequence is made of a large number of visit and purchase events as in our setting. An EM algorithm similar to Arcidiacono and Jones (2003) allows us to address this issue. We solve the dynamic program using the MPEC approach. Overall, we therefore embed an MPEC based estimation within an EM Algorithm in estimating the dynamic programming model with unobserved heterogeneity.

We estimate the structural model using household visit and purchase choices in the milk category. With the highest level of penetration and the second highest (after soda) level of spend among groceries and high frequency of purchases, milk is an ideal category for studying prices search across stores and time. Our key findings are as follows: First, we find three segments of consumers that vary in their level of search costs and price sensitivity and therefore exhibit different patterns

of search across stores and time. The largest segment (49%) has high cost of spatial search, and low price sensitivity. Therefore, they search little across stores and visit stores less frequently. Yet, they still can get low prices by searching temporally within their preferred store. A second segment (22%) has relatively low search cost for its primary store, hence visits the preferred store often, and prefers to shop during weekdays. The third segment (29%) has the lowest search cost; this segment searches both spatially and temporally and obtains the lowest prices. They also prefer to shop over weekends. The implicit search costs for a visit to a store varies from \$4.36 for the low search cost segment to \$26.30 for the high search cost segment. Second, not accounting for the time dimension of search leads to underestimated search costs and price elasticity; the direction of the bias in estimate is opposite to what has been reported in the literature (e.g., Erdem, Imai and Keane 2003; Hendel and Nevo 2006). We explain that this difference is because previous literature focused on categories with potentially high levels of consumer stockpiling, while there is more concern about stockouts and not having milk readily available in the milk category. Finally, we find that increasing the frequency of price promotions reduces store switching, and increases loyalty to their preferred store even for low search cost consumers. This result questions the conventional wisdom that price promotions induce greater cherry picking behavior among consumers.

The rest of the paper is organized as follows: Section 2 describes the model and Section 3 describes the estimation. Section 4 describes the data, while Section 5 describes the results of the structural model and biases induced by omitting time dimension of search. Section 6 describes the counterfactual on how price promotions can induce greater store loyalty. Section 7 concludes.

2 The Model

We model household buying behavior in a frequently purchased non-durable category for which consumers can hold inventory.¹ A household can purchase the good from a finite set of stores that are differentiated both spatially and in terms of retail characteristics. By holding inventory, households can decouple purchase timing from consumption timing; allowing the consumer to either advance purchase when there is a price promotion or delaying purchase till there is a price promotion. A household can also choose to forego consumption in the category, if the utility from consuming

¹ We abstract away from brand choice. Given our empirical application is for the milk category, where most purchases are on private labels, this allows us to focus on the cross-store, cross time dimension of price search without making the estimation computationally intractable due to another dimension of choice beyond “when” and “where” in a dynamic model. The modeling framework of course can be easily conceptually extended to accommodate brand choice as computational speed increases.

an outside good is higher than the expected benefit of purchasing at a higher price within the category. We recognize that store choice for frequently purchased consumer goods is not driven exclusively by the “focal” category of interest and allow for the possibility that other factors affect a household’s decision to visit stores. As mentioned earlier, we develop a finite horizon, dynamic programming model for the sequential search across stores and embed this finite horizon model in an infinite horizon dynamic programming model of search over time to model the timing of repeated purchases. This allows us to include both the spatial and temporal dimensions of search in our model.

2.1 The Basic Set Up

A household h can search across a finite consideration set of stores denoted by Ω_h at time t . Let N_h^{max} be the number of stores in Ω_h ; then, there are potentially a maximum of N_h^{max} stages of store search in any given period t until all stores in the consideration set Ω_h are exhausted. Let the tuple (t, n) represent the time and store dimensions of the search process; n representing the store search stage at time period t . Let Ω_{htn} denote the set of unvisited stores for household h at time period t at spatial search stage n .

Figure 1a: Schematic of model at period t and non-final store search stage $n < N_h^{max}$

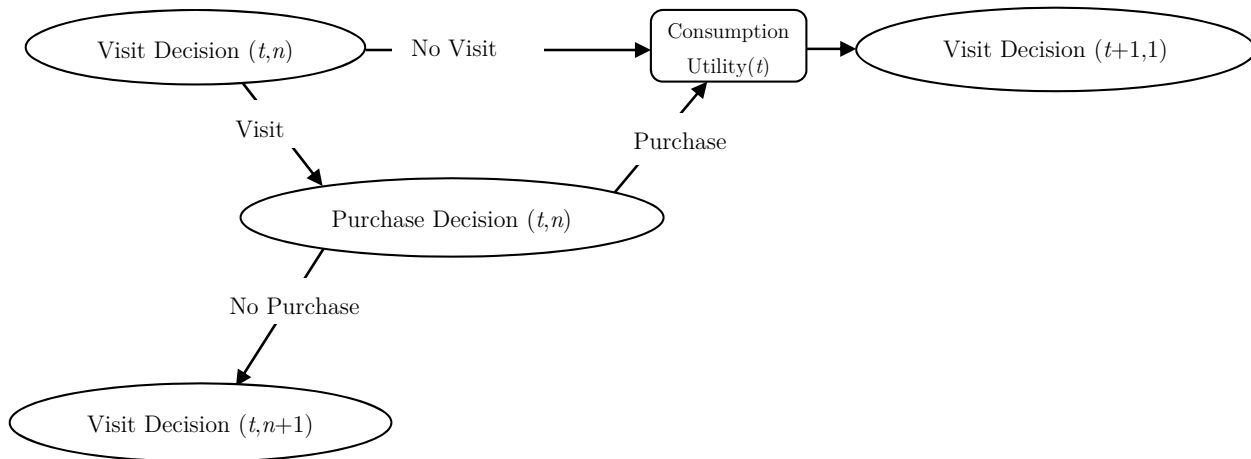


Figure 1a represents one stage of store search (store search stage n at time period t), for a non-final store search stage $n < N_h^{max}$. Each store search stage involves two decisions by the household: a store visit decision and a category purchase decision.

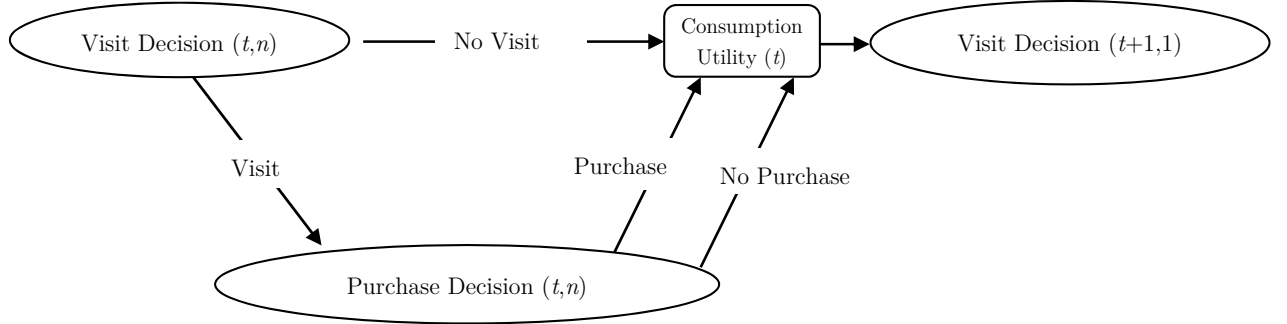
- (1) Visit Decision (t,n) : Household h observes visit-related state variables x_{ht}^v and decides whether or not to visit another store k from the set of unvisited stores at stage n in period t (Ω_{htn}) so to maximize the household's value function across the remaining stages in period t and across future time periods.
- a. Visit: A household that decides to visit another store k moves to the purchase decision at stage (t,n) .
 - b. No visit: A household that decides not to visit an additional store k , concludes its store search for period t and moves to stage 1 of store search at time $t + 1$ i.e., $(t + 1,1)$.
- (2) Purchase Decision (t,n) : When at store k from the set of unvisited stores Ω_{htn} , household observes purchase-related (including the k^{th} store specific) state variables x_{htk}^p and decides whether to purchase or not at that store to maximize the household's value function across the remaining stages in period t and across future time periods.
- a. Purchase: Upon purchasing the product, period t activities conclude and household will move to the search decision at time $t + 1$ in stage 1, i.e., Visit Decision $(t + 1,1)$.
 - b. No Purchase: If household does not purchase at stage n , household moves to the next stage of store search $(n + 1)$ at time period t ; i.e., Visit Decision $(t, n + 1)$.

Note that each household gets the utility from consumption at each time period only once. We assume that consumption occurs after the household is done with the search process and right before moving to the next time period. Thus, we ensure that changes in the level of inventory are taken into account when the household gets utility from consumption.

Figure 1b represents the final stage of store search (i.e. stage N_h^{max}) for time period t . The process is identical to Figure 1a, except that given the finite horizon nature of the store search process, not purchasing at the final stage N_h^{max} of time t leads to the visit decision in stage 1 at time $t + 1$, i.e., Visit Decision $(t + 1,1)$.

To summarize, a household $h \in \{1,2,\dots,H\}$ at time period $t \in \{1,2,3,\dots\}$ and store search stage $n \in \{1,2,\dots,N_h^{\text{max}}\}$, observes state variables x_{ht}^v that affect the decision to visit a store. The household makes a decision about whether to visit and which store to visit $y_{htn}^v \in \Omega_{htn} \cup \{0\}$, where $y_{htn}^v = 0$ represents a decision to stop search for period t at stage n . Let $N_h(t)$ denote the stage n at which household h stops search in period t . Conditional on visiting store k from the set of unvisited stores Ω_{htn} (i.e., $y_{htn}^v = k > 0$), the household observes purchase-related state variables x_{htk}^p for that store and makes a decision $y_{htn}^p \in \{0,1\}$, where 0 indicates no purchase in the focal category and 1 indicates purchase in the focal category.

Figure 1b: Schematic of model at period t and final store search stage $n = N_h^{max}$



2.2 Flow Utilities

Visit Decision

We begin with the flow utility (i.e., the immediate utility) from visit and purchase at stage n . Define d_{hk} as the travel time of household h to store $k \in \Omega_h$. Let d_h be the vector that includes the travel times to all the stores in the consideration set of household h . Let the variables that the household observes prior to visit be denoted by the set $x_{ht}^v = \{W_t, i_{ht}, I_{h,t-1}^{\text{stockup}}, d_h\}$, where W_t is a dummy variable coded as 1 if time period t is a weekend, 0 otherwise, and i_{ht} is the inventory held by the household at beginning of time period t . The immediate flow utility for household h from visiting store $k \in \Omega_{hm}$ at stage (t, n) is given by:

$$\begin{aligned} u_{htnk}^v(x_{ht}^v) &= X_{hk} \beta_h - S_h(d_{hk}, W_t) + \eta I_{n=1} \cdot I_{h,t-1}^{\text{stockup}} + \varepsilon_{htnk}^v \quad \text{for } k > 0 \\ &= \bar{u}_{htnk}^v(x_{ht}^v) + \varepsilon_{htnk}^v \end{aligned}$$

Where the first term indicates preferences for store characteristics, and the second term $S_h(d_{hk}, W_t)$ is the travel cost incurred by household h to visit store $k \in \Omega_{hm}$. The third term $\eta I_{n=1} \cdot I_{h,t-1}^{\text{stockup}}$ is a parsimonious approach to capture the role of the non-focal categories in search. If a consumer spends a lot on non-focal categories in any period (i.e., has a stock up period), there is likely less need to make a visit at the next period as she has enough inventory of non-focal items. To the extent that non-focal categories play a more important role in store visit decisions, the estimate of η would be larger. This provides a flexible framework to account for the fact that the focal category could only partially be responsible for store visit decisions. We note that this issue has been abstracted away from, in the extant literature on dynamic structural models of temporal search (e.g., Hendel and Nevo 2006; Hartmann and Nair, 2009). We account for the effect of stockup

only on the first stage in a period ($I_{n=1}$) as the fit of the model becomes poorer when we include it in later stages. We define a period as a stockup period as if total spending of household h in that period is higher than average per period spend for that household. Finally, ε_{htnk}^v is a visit-choice specific structural error shock that represents factors observed by the consumer but unobserved by the researcher that affect the decision to visit store k at stage n at time t for household h .

The search cost function is specified as a linear function: $S_h(d_{hk}, W_t) = \iota_h + \delta_h d_{hk} + \omega_h W_t$. While the effect of travel time (d_{hk}) on travel/search cost is obvious, the weekend dummy variable allows us to account for the fact that working households can have a higher opportunity cost of search during weekdays, while households with retired seniors or an adult non-working member may have higher opportunity costs of search on weekends. We include two store characteristics in X_{hk} to account for store differentiation: (1) Whether store k is EDLP and (2) Whether store k is the primary grocery store for household h , where we operationalize the primary store as that which has the highest share of visits in its consideration set.

A household that forgoes search obtains the following utility:

$$u_{htn0}^v(x_{ht}^v) = \varepsilon_{htn0}^v$$

Purchase Decision

After visiting store k , the household decides whether to make a purchase or not in the focal category. The flow utility for a household making a purchase is given by:

$$\begin{aligned} u_{htnk1}^p(x_{htk}^p) &= \alpha_h p_{kt} + \varepsilon_{htnk1}^p \\ &= \bar{u}_{htk1}^p(x_{htk}^p) + \varepsilon_{htnk1}^p \end{aligned}$$

Where α_h is the price sensitivity of household h , p_{kt} is price of the focal category in store k at time period t , and ε_{htnk1}^p is a purchase-choice specific structural error shock representing factors that affect the purchase decision and are observed by the household but not the researcher.

A household that does not purchase obtains:

$$u_{htnk0}^p(x_{htk}^p) = \varepsilon_{htnk0}^p$$

The structural error shocks in the above equations are all assumed to be independent and identically distributed (i.i.d) type I extreme value.

Consumption Utility

Before moving to the next time period, the consumer gets utility from consumption of the focal category, which is a function of the inventory that includes purchases in the current period. We represent the consumption utility as:

$$u_{ht}^c(i_{ht}, y_{ht}^p) = \varphi(c(i_{ht} + y_{ht}^p \chi))$$

where i_{ht} is the inventory level of the focal category, $c(i_{ht} + y_{ht}^p \chi)$ is consumption as a function of inventory level and φ is utility of consuming $c(i_{ht} + y_{ht}^p \chi)$ units. We do not include an error shock on consumption utility as it is non-separable from the flow utilities from the search and purchase stages. Here, χ represents the amount that gets added to consumer inventory if she makes a purchase (i.e., milk container size in our application) and $y_{ht}^p = \sum_{n=1}^{N_h^{max}} y_{htn}^p$ (which is equal to one if the consumer makes a purchase in time period t and zero otherwise).

Let ρ_h be the household h 's consumption rate of the focal category. Specifically we assume $c(i_{ht} + y_{ht}^p \chi) = \rho_h \cdot \min\{1, \lfloor (i_{ht} + y_{ht}^p \chi) / \rho_h \rfloor\}$, where $\lfloor x \rfloor$ represents the floor of x . This means the household consumes an amount equal to consumption rate if there is more than one serving left in inventory and consumes zero otherwise. This specification allows us to capture a drop in utility when a household does not have adequate inventory and thus captures the cost of household stockouts. We assume a linear form for utility from consumption. Specifically, $\varphi(c(i_{ht} + y_{ht}^p \chi)) = \sigma \cdot c(i_{ht} + y_{ht}^p \chi) + \tau$, where σ and τ are parameters to be estimated.

State Transitions

Here we define appropriate state transitions and expectations associated with inventory, prices, stockup, weekday/weekend and store consideration sets.

Inventory held by household evolves as follows:

$$i_{h(t+1)} = i_{ht} - c(i_{ht} + y_{ht}^p \chi) + y_{ht}^p \chi$$

Where χ is increase in inventory after purchase (i.e. milk container size in our application) and $y_{ht}^p = \sum_{n=1}^{N_h^{max}} y_{htn}^p$.

We assume that prices follow an exogenous discrete distribution with m different levels of possible prices. We allow for prices to have different distributions for different stores. We assume a

store specific multinomial distribution of prices over the m price levels over time. Formally:
 $p_{kt} \sim \text{Multinomial}(1, \vec{p}_k)$.²

We assume that decision on how much to spend at each period is exogenous to the model and form a first order Markov process for transition of dummy variable on stock up periods. More formally: $I_{ht}^{\text{stockup}} \sim \text{Bernoulli}(\pi_h^{S \rightarrow S} I_{h,t-1}^{\text{stockup}} + \pi_h^{N \rightarrow S} (1 - I_{h,t-1}^{\text{stockup}}))$, where $\pi_h^{S \rightarrow S}$ and $\pi_h^{N \rightarrow S}$ are transition probabilities from stock up to stock up and from non-stock up to stock up period for household h respectively.

Weekends and weekdays alternate. We initialize the first period to be Weekend or Weekday as appropriate. In our case, the first period falls on weekdays, so we initialize the variable to zero $W_1 = 0 \quad W_t = 1 - W_{t-1}$.

Store consideration set evolves as follows, where the store visited in stage $n - 1$ is removed from the consideration set at stage n . $\Omega_{ht0} = \Omega_h$ and $\Omega_{htn} = \Omega_{htn-1} \setminus y_{htn-1}^v$

2.3 The Visit and Purchase Sequence Problem

Each consumer makes a sequence of visit and purchase decisions to maximize utility from the current time period plus discounted utility from future periods. Based on flow utilities defined in previous section, we can write the optimization problem as a sequence problem of visit and purchase decisions for each household h ,

$$\max_{\{\Delta_{ht}\}_{t=1}^{\infty}} E \left(\sum_{t=0}^{\infty} \beta^t \cdot \varpi_{ht}(\Delta_{ht}) \mid \tilde{x}_{ht} \right),$$

Where $\Delta_{ht} = \{y_{ht}^v, y_{ht}^p\}$ represents the vector of a household's visit ($y_{ht}^v = \{y_{htn}^v\}_{n=1}^{N_h(t)}$) and purchase ($y_{ht}^p = \{y_{htn}^p\}_{n=1}^{N_h(t)-1}$) decisions. These decisions in each time period are conditional on visit and purchase-related observed and unobserved state variables: $\tilde{x}_{ht} = \{x_{ht}^v, x_{ht}^p, \varepsilon_{ht}^v, \varepsilon_{ht}^p\}$. Here $x_{ht}^p = \{\{x_{htk}^p\}_{k \in \Omega_h \setminus \Omega_{htn}}\}_{n=1}^{n=N_h(t)-1}$ includes all the relevant observed state variables for the purchase state, while $\varepsilon_{ht}^v = \{\varepsilon_{htn0}^v, \{\varepsilon_{htnk}^v\}_{k \in \Omega \setminus \Omega_{htn}}\}_{n=1}^{n=N_h(t)}$, and $\varepsilon_{ht}^p = \{\{\varepsilon_{htnk1}^p, \varepsilon_{htnk0}^p\}_{k=y_{htn}^v}\}_{n=1}^{N_h(t)-1}$ represent all the relevant unobserved state variables for visit and purchase stages, respectively. The total utility that the household gets across all stages within time period t is the sum of flow utilities from the

² The exogeneity assumption is common in the dynamic structural modeling literature; see Erdem, Imai and Keane 2003, for a detailed discussion on the plausibility of the price exogeneity assumptions in modeling choice of frequently purchased consumer goods. In particular Khan et al (2013) discuss institutional reasons like state and federal pricing regulations that make milk prices plausibly exogenous to demand shocks and more a function of supply and cost shocks. Without store search, the literature typically assumes a first order Markov process. With store search the assumption of a Markov process is questionable, as the household does not know prices for stores that have not been visited. Hence we assume independence across time for all stores.

visit and purchase stages up to the $N_h(t)$ (the stage at which household stops search at period t) plus consumption utility:

$$\varpi(\Delta_{ht}) = \underbrace{\sum_{n=1}^{N_h(t)} \prod_{l=0}^{N_h^{\max}} (u_{htnl}^v)^{\mathbb{1}\{y_{htn}^v=l\}}}_{\text{visit utility}} + \underbrace{\sum_{n=1}^{N_h(t)-1} \prod_{l=0}^1 (u_{htnkl}^p)^{\mathbb{1}\{y_{htn}^p=l\}}}_{\text{purchase utility}} + \underbrace{u_{ht}^c}_{\text{consumption utility}}$$

2.4 Choice-Specific Value Functions

Within the finite horizon spatial search model, a household has to make two consecutive decisions in each stage of each time period (i.e. a decision to visit a store, potentially followed by a decision to make a purchase in the focal category). We therefore define *two sets of value functions*, one for visit decisions and the other for purchase decisions. To keep notation simple, we use the ex-ante value functions of search and purchase to write the choice-specific value functions. Precise definition of these value functions is presented in the next subsection. Let $EV_{htn}^v(x_{ht}^v, \Omega_{htn})$ represent the ex-ante value function of search at stage n of time period t for household h ; i.e., the highest expected value of utility that the household can get starting at search stage n if the set of unvisited stores is Ω_{htn} . Similarly, let EV_{htnk}^p represent the ex-ante value function at purchase stage n of period t if household h is visiting store k .

Consider household h with N_h^{\max} stores in its consideration set, at any stage before the last stage (i.e. $n < N_h^{\max}$) visiting store k , making a purchase decision at time t . After observing purchase-related variables, the household has two options; (1) to make a purchase and end store search for the current period t and presumably start at $t+1$ with a higher inventory level, or (2) to wait for stage $(n+1)$ and consider visiting an unvisited store from its store choice set Ω_{htn+1} . With a purchase, the household gets the corresponding flow utility plus discounted value of utility (across time) that she will get starting next period.

$$\begin{aligned} v_{htnk1}^p(x_{htk}^p, x_{ht}^v) &= \bar{u}_{htnk1}^p + u_{ht}^c + \beta \mathbb{E}_{x_{h,t+1}^v | x_{ht}^v, \Delta_{ht}, \epsilon_{ht}} EV_{h,t+1,1}^v(x_{h,t+1}^v, \Omega_h) + \varepsilon_{htnk1}^p \\ &= \bar{v}_{htk1}^p + \varepsilon_{htnk1}^p \end{aligned}$$

If household does not purchase, the household receives the corresponding flow utility plus expected value of utility that she gets starting next search stage. Note that expected value of the next search stage is not discounted as it happens in the same time period.

$$\begin{aligned}
v_{htnk0}^p(x_{htk}^p, x_{ht}^v) &= EV_{ht,n+1}^v(x_{ht}^v, \Omega_{ht,n+1}) + \varepsilon_{htnk0}^p \\
&= \bar{v}_{htn0}^p + \varepsilon_{htnk0}^p
\end{aligned}$$

Moving one step back, household faces a decision of whether to visit a store and which store to visit. At this point, household knows the realizations of random shocks for the visit stage but not for the purchase stage. The household also has not observed purchase-related state variables for that store yet (e.g., does not know prices before visiting the store). Therefore, the household should use the expected value of the utility for the purchase stage in making the decision whether to visit the store or not. As this expected value is represented by EV_{htnk}^p , the choice-specific value function for search stage n can be written as

$$\begin{aligned}
v_{htnk}^v(x_{ht}^v) &= \bar{u}_{htnk}^v + EV_{htnk}^p + \varepsilon_{htnk}^v \\
&= \bar{v}_{htnk}^v + \varepsilon_{htnk}^v,
\end{aligned}$$

where $k \in \Omega_{htn}$, implying that at this stage household can choose a store from the set of unvisited stores in the current time period. If household decides to stop search (i.e., $k=0$), instead of expected value of the next purchase stage in the current time period, the household will get the discounted expected value of utility starting from the first visit stage of next time period, i.e.,

$$\begin{aligned}
v_{htn0}^v(x_{ht}^v) &= u_{ht}^c + \beta.E_{x_{h,t+1}^v|x_{ht}^v, \Delta_{ht}, \varepsilon_{ht}} EV_{h,t+1,1}^v(x_{h,t+1}^v, \Omega_h) + \varepsilon_{htn0}^v \\
&= \bar{v}_{htn0}^v + \varepsilon_{htn0}^v
\end{aligned}$$

So far, we have presented choice-specific value functions for search and purchase at an arbitrary stage $n < N_h^{\max}$. We present the value functions separately for $n = N_h^{\max}$ because the value function of the purchase stage at the last remaining store will not include the expected value of the next search stage, if consumer decides not to make a purchase. In that case as there are not any stores left unvisited for the current time period, upon a decision not to make a purchase, the consumer will move on to the next time period

$$\begin{aligned}
v_{htN_h^{\max}k0}^p(x_{htk}^p, x_{ht}^v) &= u_{ht}^c + \beta.E_{x_{h,t+1}^v|x_{ht}^v, \Delta_{ht}, \varepsilon_{ht}} EV_{h,t+1,1}^v(x_{h,t+1}^v, \Omega_h) + \varepsilon_{htN_h^{\max}k0}^p \\
&= \bar{v}_{htN_h^{\max}0}^p + \varepsilon_{htN_h^{\max}k0}^p
\end{aligned}$$

2.5 Ex-Ante Value Functions

Next, we define value functions and ex-ante value functions based on choice-specific value functions defined in the previous subsection. Denoting $V_{htn}^v(x_{ht}^v) = \max_{k \in \Omega_{htn} \cup \{0\}} \{v_{htnk}^v(x_{ht}^v)\}$ as value function of search stage, the ex-ante value function at the visit stage is given by,

$$\begin{aligned} EV_{htn}^v(x_{ht}^v) &= E_{\epsilon_{htn}^v | x_{ht}^v, \Omega_{htn}, \epsilon_{ht, n-1}^v} \left\{ \max_{k \in \Omega_{htn} \cup \{0\}} [v_{htnk}^v(x_{ht}^v)] \right\} \\ &= \log \left\{ \sum_{k \in \Omega_{htn} \cup \{0\}} \exp(\bar{v}_{htnk}^v) \right\}, \end{aligned}$$

where $\epsilon_{htn}^v = \{\epsilon_{htnk}^v\}_{k \in \Omega_{htn} \cup \{0\}}$. The second equality follows from the properties of extreme value distribution and the conditional independence assumption. Similarly, let the value function of the purchase stage be denoted by $V_{htnk}^p = \max\{v_{htnk1}^p, v_{htnk0}^p\}$, then we can write ex-ante value function at the purchase stage as,

$$\begin{aligned} EV_{htnk}^p &= E_{x_{htk}^p, \epsilon_{htnk1}^p, \epsilon_{htnk0}^p | \epsilon_{ht, n-1, k1}^p, \epsilon_{ht, n-1, k0}^p} \{ \max[v_{htnk1}^p, v_{htnk0}^p] \} \\ &= \int_{x_{htk}^p} \log[\exp(\bar{v}_{htk1}^p) + \exp(\bar{v}_{htn0}^p)] . dP(x_{htk}^p) \end{aligned}$$

Again, the second equality is based on the extreme value distribution and the conditional independence assumption.

2.6 Choice Probabilities and Likelihood Function

Based on the choice-specific value functions presented in the previous section, we can write the choice specific probabilities at each stage in any given time period, given the distribution of error shocks. As the error shocks are drawn from a Type I extreme value distribution, the choice specific probabilities can be represented as follows:

$$P_{htnk}^v = \frac{\exp(\bar{v}_{htnk}^v)}{\sum_{j \in \Omega_{htn} \cup \{0\}} \exp(\bar{v}_{htnj}^v)}$$

where P_{htnk}^v is the probability that household h at time period t and stage n chooses to search store $k \in \Omega_{htn}$ from the set of unvisited stores or chooses to stop search in the current period $k=0$. The probability of the same household making a purchase, while visiting store k can be written as,

$$P_{htnk1}^p = \frac{\exp(\bar{v}_{htk1}^p)}{\exp(\bar{v}_{htk1}^p) + \exp(\bar{v}_{htn0}^p)}$$

We allow for discrete heterogeneity among households, i.e., a household h can belong to one of G segments denoted by g . Using the representation of probabilities above and the household's observed decision, the likelihood for household h conditional on being from segment g can be written as,

$$L_{h|g} = \prod_{t=1}^{T_h} \prod_{n=1}^{N_h(t)} \prod_{k=0}^{N_h^{\max}} (P_{htnk}^v | g)^{1_{\{y_{htn}^v=k\}}} \cdot (P_{htnk1}^p | g)^{1_{\{y_{htnk}^p=1 \& y_{htn}^v=k\}}} \cdot (1 - P_{htnk1}^p | g)^{1_{\{y_{htnk}^p=0 \& y_{htn}^v=k\}}}.$$

The unconditional likelihood for the sample of size N can be written as follows where p_g denotes the size of group g .

$$L = \prod_{h=1}^H \left(\sum_{g=1}^G p_g \cdot L_{h|g} \right) \quad (1)$$

3 Data and Model-Free Evidence

We use a Nielsen household-level panel data set of all grocery purchases by a sample of households across the United States from January to December 2006. We observe every shopping trip and all grocery items purchased and price paid for each item by each household. We also observe store zip code and household census tract county code which allows us to calculate (an approximate) distance between each household and each store in their consideration set. We complement this data with Retail Scanner Data from Nielsen, to construct the weekly prices for relevant stores. Appendix B provides details on the price construction.

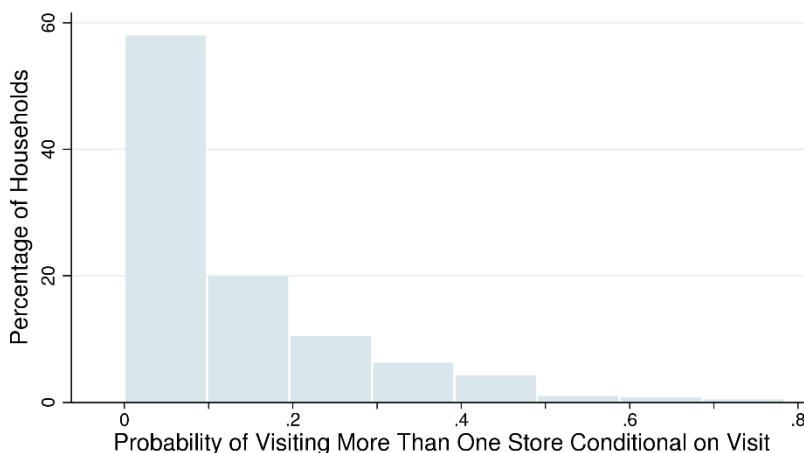
We use milk as our focal category. Milk is an ideal category for our purposes. It has the second highest spend (\$80 with a 3.39% basket share) after soda (\$117 with a basket share of 4.81%), but the highest level of penetration (88%) among the top ten of the high-spend categories. See Appendix A for a category analysis of spend and penetration. It is also purchased frequently as it is perishable and therefore there is limited stockpiling. Thus it is frequently a relevant category in a household's shopping basket and given the high level of spend, the price effect and value of price search is plausibly high. One advantage of milk relative to soda is that brand choice is not a central issue in the category as the vast majority of purchases are private labels at the store. This allows us to abstract away from brand choice and to keep the decision and state space of the dynamic programming problem manageable. Given that we model store visits and category purchases, our decision and state space is much larger than most models estimated in the literature.

We construct our sample of households from the panel data as follows. First we drop households who do not shop frequently (< 20 shopping trips over the year across all stores) or do not purchase milk frequently ($< 5\%$ of their shipping trips). Second, we consider a store to be in a household consideration set only if the household spends greater than or equal to 10% of its annual spending in grocery in that store. Based on this cutoff, 94% of households shop at three stores or less. Hence we set $N_h^{\max} = 3$. Third, we abstract away the issue of size choice again to keep the state space manageable, by focusing on single-unit buyer households who purchase the most common size (one gallon) over the term of data collection. Fortunately, the size loyalty in the category is high with the median share of the favorite size being 93%. Finally, we dropped a small number of households shopped at store for which we do not have price data. In all, we use 948 households.³

3.1 Model-Free Evidence

We begin our analysis with model-free evidence to show that there is price search spatially across stores and across time.

Figure 2. Share of periods that a household visits both stores



Spatial Search across Stores

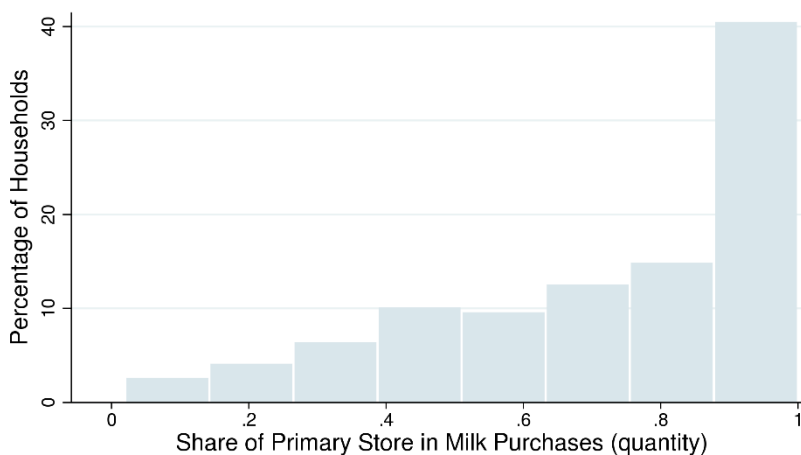
To separate weekday and weekend behaviors, we treat Friday-Sunday as the weekend period and Monday-Thursday as weekday period. Figure 2 shows the distribution of share of time periods

³ In a previous version of this paper, we estimated the model with a smaller sample of households who shop at no more than two stores. (i.e., $N_h^{\max} = 2$). Our key results and insights remain virtually identical, lending us confidence that the sample selection has little impact on our conclusions. The two store version of the paper is available on request.

in which a household visits more than one store within a time period. A large number of households visit multiple stores within the same weekday or weekend period.

A possibility is that people visit multiple stores, but always buy milk at one store, suggesting there is no search for milk across stores. Figure 3 presents distribution of purchases of milk from a household’s “favorite store” (store from which consumer has purchased the item most often). We find that milk indeed is purchased from different stores by multiple households.

Figure 3. Store-category loyalty for milk

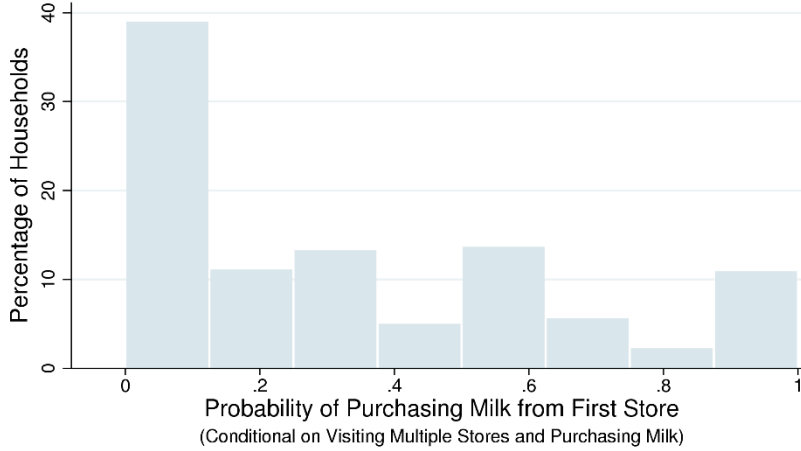


A related concern is whether milk may be bought at different stores, but always chosen at the first store visited. Figure 4 shows the probability distribution of purchasing milk from the first store conditional on visiting multiple stores in the same time period. Many households purchase milk at the second or third store during the same period. These suggest evidence of cross-store search.

To explore the consumer search among stores and checking for the fact that milk could have an effect on consumer’s decision to perform spatial search, we estimated a logistic regression where we model the probability of visiting multiple stores as a function of the inventory level of milk controlling for heterogeneity by including household fixed effects in the model.⁴ In this regression, the coefficient of inventory of milk is negative and significant ($p < .01$) showing that an increase in inventory of milk decreases the probability of visiting multiple stores in the same period.

⁴ The inventory level is not observed, so we construct inventory levels by tracking purchases and adjusting for consumption rates. We initialize the inventory level for households with a random value.

Figure 4: Milk purchases at second store visited



Search across Time

To study whether consumers adjust purchase timing in response to milk promotions we test the differences in inter-purchase times between milk purchases as a function of whether milk is purchased on promotion or not. The idea is that consumers accelerate their purchases when there is a promotion before consuming their current inventory as demonstrated in the early work of Neslin, Henderson and Quelch (1985) and Hendel and Nevo (2006). Given that milk is a perishable item, that can only be stockpiled for short periods, it is an empirical question as to whether purchase acceleration is likely in the milk category. To answer this question, we performed a paired sample t-test comparing average inter-purchase time for purchases that are made on promotion versus those that are made on regular price. We found that the average inter-purchase time was 4.47 periods (half-weeks) across households when purchases were made when there was no promotion, and 4.88 (half-weeks) across the same households when purchases were made on promotion. The difference of 0.41 periods is statistically significant at $p = 0.01$, suggesting that there is evidence of purchase acceleration in the milk category.⁵

4 Estimation

We formulate the estimation problem of the dynamic programming model as a Mathematical Program with Equilibrium Constraints (Su & Judd, 2012). However, instead of estimating the

⁵ For this test, we dropped households that never bought milk on promotion as we could not do a paired test for this group. We also dropped households that had lapses between purchases of more than 12 periods (one and a half months) as these few outlier households disproportionately impact the duration between purchases relative to the large number of households making regular purchases.

heterogeneous model using nonlinear constrained optimization as suggested in Su and Judd (2012), we combine the MPEC approach with an iterative EM algorithm procedure (Arcidiacono and Jones 2003). We use a finite mixture of types to capture heterogeneity. Although we can technically use the nonlinear constrained optimization approach even with finite heterogeneity, a practical challenge arises in our setting, where we model choices of store and purchase visits in each time period, compared to the case where only purchase choices are modeled conditional on store visits. With such a large number of choice probabilities the likelihood of each household’s purchase string becomes smaller than numerical precision of the computer.⁶ With heterogeneity, the log likelihood function with heterogeneity cannot be written simply as a summation of log of choice probabilities. By nesting the constrained optimization within an EM algorithm procedure, at any stage of the optimization process, the objective functions only enter in the form of summations of log of choice probabilities with the probability of membership in each segment set at the value of the previous iteration, thus bypassing the numerical precision problem.

4.1 The Mathematical Programming with Equilibrium Constraints

In the unconditional likelihood function, presented in Equation (1), $L_{h|g}$ is a function of choice-specific value functions of the model. In fact, this equation could be re-written as

$$L = \prod_{h=1}^H \underbrace{\left(\sum_{g=1}^G p_g \cdot L_{h|g}(\bar{v}_h^p, \bar{v}_h^v; x_h^p, \Delta_h, \Theta) \right)}_{L_h} .$$

While traditional nested fixed point approach (NFXP) suggests application of an unconstrained optimization algorithm and calculation of value functions outside the optimization loop using contraction mapping, this methods proves to be computationally intensive considering the size of the state space and structure of the problem.⁷ Therefore, instead of using NFXP, we re-formulate the problem as a constrained optimization problem. To that end, we re-write the likelihood function

⁶ This happens for two reasons; first, long panel structure, which is not unique to our model. Note that $L_{h|g}$ is the product of probabilities of the sequence of decisions for all the time periods during which household h is observed. Second, due to the nested structure of the model (i.e. a finite horizon cross-store model nested in an infinite time horizon model). The sequence of probabilities can include between one to $2N_h^{max}$ probability terms for each time period (a visit and a purchase decision for each store) depending on actions that household h takes. This exacerbates the long panel issue by up to 3 depending on the maximum number of stores in households’ consideration sets.

⁷ The specific nested structure of the problem in this case results in a system of Bellman equations which adds to the computational burden in each iteration of the contraction mapping.

as a function of choice-specific and ex-ante value functions and replace the contraction mapping with a set of constraints, each of which representing a Bellman equation.

$$\max_{\Theta} \prod_{h=1}^H \left(\sum_{g=1}^G p_g \cdot L_{h|g}(EV_h^v, EV_h^p, \bar{v}_h^p, \bar{v}_h^v; x_h^v, x_h^p, \Delta_h, p, \Theta) \right)$$

subject to:

$$EV_{htn}^v(x_{ht}^v, \Omega_{htn}) = \log \left\{ \sum_{k \in \Omega_{htn} \cup \{0\}} \exp(\bar{v}_{htnk}^v) \right\}, \quad \forall t \in \{1, \dots, T_h\}, \forall n \in \{1, \dots, N_h\}$$

$$EV_{htnk}^p = \int_{x_{htnk}^p} \log[\exp(\bar{v}_{htk1}^p) + \exp(\bar{v}_{htn0}^p)] \cdot dP(x_{htk}^p) \\ \forall t \in \{1, \dots, T_h\}, \forall n \in \{1, \dots, N_h\}, \forall k \in \{1, \dots, N_h\}$$

where $EV_h^v = \{\{EV_{htn}^v\}_{t=1}^{T_h}\}_{n=1}^{N_h}$ and $EV_h^p = \{\{\{EV_{htnk}^p\}_{t=1}^{T_h}\}_{n=1}^{N_h}\}_{k=0}^{N_k}$ are set of ex-ante value functions for the search and purchase stages respectively. Similarly, $\bar{v}_h^v = \{\{\{\bar{v}_{htnk}^v\}_{t=1}^{T_h}\}_{n=1}^{N_h(t)}\}_{k=0}^{N_h^{\max}}$ and $\bar{v}_h^p = \{\{\{\bar{v}_{htk1}^p, \bar{v}_{htn0}^p\}_{t=1}^{T_h}\}_{n=1}^{N_h(t)}\}_{k=1}^{N_h^{\max}}$ represent set of deterministic parts of the choice-specific value functions for the search and purchase stages.

To address the issue of small numbers arising from the fact that taking the log of the above objective would not transform multiplication of numerous probability terms inside $L_{h|g}$, we adopt the EM approach presented in Arcidiacono and Jones (2003). Assuming that $\Pr(g | x_h^v, x_h^p, \Delta_h, p; \hat{\Theta})$ represents conditional probability that household h belongs to group g conditional on observed state variables, decisions, group sizes, and set of parameters, the objective function of the above constrained optimization problem could be replaced with

$$\max_{\Theta} \sum_{h=1}^H \sum_{g=1}^G \Pr(g | x_h^v, x_h^p, \Delta_h, p; \hat{\Theta}) \ln(L_{h|g}(EV_h^v, EV_h^p, \bar{v}_h^p, \bar{v}_h^v; x_h^v, x_h^p, \Delta_h, p, \Theta)) \quad (2)$$

4.2 Segment Sizes and Household Probability of Membership

Allowing for a finite number of groups, let p_g denote the unconditional probability that a consumer belongs to group g and $p = (p_1, \dots, p_G)$. Following Bayes' theorem, we can write the probability that household h is from group g , conditional on household's observed behavior and a set of parameters

$$\Pr(g | x_h, \Delta_h, p; \Theta) = \frac{p_g L_{h|g}(x_h, \Delta_h, p; \Theta)}{\sum_{g=1}^G p_g L_{h|g}(x_h, \Delta_h, p; \Theta)} \quad (3)$$

Where $L_{h|g}$ is individual likelihood for household h conditional on being of type g , and $x_h = \{x_{ht}^v \cup \{x_{htk}^p\}_{k=1}^{N_h}\}_{t=1}^{T_h}$ represents the set of all observed state variables for household h . The maximum likelihood estimate of \hat{p}_g is given by

$$\hat{p}_g = \frac{1}{H} \sum_{h=1}^H \Pr(g | x_h, \Delta_h, p; \Theta) \quad (4)$$

4.3 The Estimation Algorithm

We combine the procedure presented for estimating models with discrete heterogeneity in (Arcidiacono & Jones, 2003) with MPEC approach (Su and Judd 2012). Equations (2), (3) and (4) suggest an iterative algorithm for estimation.

Step 0: Assume starting values of p_g and Θ .

Step 1: Calculate p_g^h , using equation (3), conditional on p and Θ .⁸

Step 2: Given the estimates of p_g^h , use equation (4) to update p_g .

Step 3: Using estimates of p_g^h , maximize equation (2) subject to Bellman equations as constraints to update Θ .

Step 4: Iterate over steps 1 to 3 till convergence on Θ

The above iterative algorithm is an adaptation of the EM algorithm presented in (Arcidiacono & Jones, 2003), in that instead of using the Rust (1987) nested fixed point algorithm to solve the dynamic programming problem, we solve the DP problem using a mathematical program with equilibrium constraints (Su & Judd, 2012).

4.4 Identification

We present an informal discussion of identification in this section. The two most critical parameters for a search model across stores and time are search cost and price parameters. Intuitively, the purchase/no purchase decision in response to price variation as a function of state

⁸ To calculate p_g^h we need to calculate likelihoods conditional on Θ . We obtain the likelihood not through a contraction mapping, but through constrained optimization. The optimization problem has a constant objective function as we are solving conditional on Θ ; hence the optimizer minimizes feasibility error of constraints (Bellman equations) rather than minimizing optimality error (which is zero with a constant objective function).

variables such as inventory identifies price sensitivity, while the frequency of store visits identifies search cost. However, the sequence of store search in a model coupled with information about the store at which the purchase is made provides even tighter identification.

Imagine two consumers who have similar frequency of store visits and purchase at a high price. Even though both customers visit roughly the same number of stores in any given time period, the first one usually makes a purchase at the first store that she visits, while the second one makes a purchase at the first store only if the price is low. Since the frequency of store visits is the same, the model would estimate their search costs to be similar, whereas the fact that the first consumer always buys from the first store regardless of price, would identify her lower price sensitivity.

Identification of other parameters is fairly straightforward. Parameters of consumption utility function (σ and τ) are identified from the observed variation in households consumption rate and the imputed stockouts.⁹ Utility from consumption of non-focal categories (η) is identified from observations where households visit stores without making a purchase in the focal category. We can identify preference for store formats based on household share of visits to different store formats. As is typical in the dynamic structural modeling literature, the discount factor is not identified in this model and we assume it to be 0.993 for each period.¹⁰

5 Results

We first report the results of the full structural model with both spatial and temporal dimensions. We then report the extent and nature of bias in estimates when the time dimension is omitted. We provide intuition for the bias.

5.1 Estimates of the structural model

⁹ We estimate consumption rate for each household separately using each household's purchase decisions. For each household the consumption rate would be simply total amount purchased over number of time periods that the household is observed in our data.

¹⁰ Typically, weekly discount factor is assumed to be 0.995 in empirical research. Our assumption of 0.993 for half-week time period is slightly smaller than the standard assumption, consistent with recent empirical estimates of the discount factor (Song, Mela, Chiang, & Chen, 2012; Chung, Steenburgh, & Sudhir, 2014). For a review of the literature on discounting look at Frederick, Loewenstein, & O'Donoghue (2002).

Table 1. Search model with both store and time dimensions

	Segment 1	Segment 2	Segment 3
Price Sensitivity(α)	-0.1075*** (0.0072)	-0.3038*** (0.0075)	-0.3456*** (0.0072)
Marginal Consumption Utility (σ)	3.3346*** (0.1186)	2.9469*** (0.0924)	2.1048*** (0.0752)
Intercept of Consumption Utility (τ)	-0.6301*** (0.032)	-0.5599*** (0.0498)	-0.4593*** (0.0273)
Stock Up Previous Period (η)	-0.4026*** (0.0474)	-0.1811*** (0.0549)	0.0172 (0.0295)
Search Cost Intercept (ι)	2.594*** (0.0386)	1.702*** (0.0154)	1.5991*** (0.0366)
Travel Time (δ)	0.0769*** (0.011)	0.1097*** (0.0267)	0.0195* (0.0101)
Preferred store (ψ_1)	0.9773*** (0.018)	2.0481*** (0.0387)	0.7554*** (0.0158)
EDLP (ψ_2)	-0.0256 (0.0295)	-0.3026*** (0.0263)	0.0329 (0.0265)
Weekend (ω)	-0.1187*** (0.0194)	0.6057*** (0.0263)	-0.1532*** (0.0205)
Segment Size	0.49	0.22	0.29

All coefficients are highly significant ($p < 0.01$) and have expected signs, except for the coefficient on stock-up of previous period for the third segment, and preference for EDLP stores for the first and third segments. Segment 1 comprises 49% of the sample households, while second and third segments represent 22% and 29% of the sample, respectively. Segment 1 has the highest search cost and lowest price sensitivity; therefore it does not place much value on price search; Hence, it should perform the least amount of search across time and across stores. Segments 2 and 3 have lower search costs and higher price sensitivities. Hence, they value gains from search. But as they have low search cost, they will search more intensely for a given level of price dispersion. Segments 2 and

3 differ in their preferences for spatial and temporal dimensions of search. Segment 2 has a strong preference for their primary store (2.048 vs. 0.755); therefore households in this segment would focus on intense temporal search at their primary store. In contrast, households in segment 3 will perform more spatial search, as their relative preference for the primary store is not that high.

Segments 2 and 3 also differ in their preference for shopping during weekends, and their preference for EDLP stores. While the second segment seems to prefer weekdays, the third segment has a preference for shopping over the weekends.

Table 2. Observed search behavior and demographics by segment

	Segment 1	Segment 2	Segment 3
Percentage of Shopping Periods in Which at Least One Store Has Been Visited	32.2%	60.0%	55.5%
Percentage of Periods with Both Stores Visited	2.7%	5.0%	13.3%
Percentage of Periods with Both Stores Visited Conditional on Visiting at Least One Store	7.6%	8.0%	20.2%
Average Price Paid (\$)	2.98	2.96	2.87
Median Household Income (\$1000)*	50-60	45-50	40-45

* Based on a quantile regression of income category on probability of being a member of each segment.

To test if our predictions based on the structural model estimates above is valid, we compare the observed behavior across three segments. Table 2 presents metrics on the visit and purchase behavior for each segment. As expected, segment 1 visits stores least often. In fact, the first segment does very little spatial search considering the fact that a consumer in this segment on average visits more than one store in the consideration set only 2.7% of the time. The second segment does perform some spatial search, but not as much as the third segment. This is consistent with the larger estimated differential preference for the primary store for the second segment, which pushes this segment to perform more temporal search, as explained before. Not surprisingly, there is a negative relationship between median income of the segment and its price sensitivity. The third segment,

which is the most price sensitive has the lowest median income and the first segment, which is the lowest price sensitive, has the highest median income.

Table 3 reports the search costs in dollar terms for the three segments during weekdays and weekends based on the estimated parameters and price sensitivity.¹¹ Note that households decide about their search strategy based on the search cost and also their preference for primary store compared to other stores in their consideration set. Monetary values of preference for the primary store are estimated to be \$9.1, \$6.7, and \$2.2 for segments one, two, and three respectively. The fact that this preference is the largest for the second segment relative to its search cost explains why the second segment is more inclined towards temporal search.

Table 3. Search costs estimates

	Segment 1	Segment 2	Segment 3
Weekend	\$25.20	\$8.67	\$4.36
Weekday	\$26.30	\$6.68	\$4.80

5.2 Bias from omission of the temporal dimension

As discussed in the introduction, the search cost literature thus far has focused on either the spatial or temporal dimension, but not both. As we argued, this can lead to biased estimates of search costs and price sensitivity, especially in a frequently purchased category. We now assess the extent of bias by omitting the temporal dimension. To make sure that the results are comparable, we keep the spatial dimension of search and turn off the temporal dimension (i.e. consumer forward-looking behavior). Parameter estimates are presented in Table 4. We observe three important biases in these results. Search cost and price sensitivity parameters are underestimated in the store-search-only model relative to the full model with both store and time search. In contrast, the utility from consumption is overestimated. We discuss the intuition for the three biases in our analysis.

¹¹ To calculate search cost for each segment we sum the estimate of the search cost intercept, the product of coefficient on travel time and square root of average travel time for each segment. For weekends, we also include in the sum the estimate of the coefficient on weekend dummy. We then divided the sum of coefficients by the estimate of price sensitivity to get dollar value equivalent of search cost.

Table 4. Search model with only store dimension

	Segment 1	Segment 2	Segment 3
Price Sensitivity(α)	-0.0635*** (0.0072)	-0.0346*** (0.0076)	-0.2806*** (0.0067)
Marginal Consumption Utility (σ)	8.6573*** (0.1482)	2.1872*** (0.0947)	2.6494*** (0.0725)
Intercept of Consumption Utility (τ)	-2.0259*** (0.0359)	-0.6636*** (0.0561)	-1.0271*** (0.0252)
Stock Up Previous Period (η)	-0.4196*** (0.0449)	-0.5013*** (0.0645)	-0.2706*** (0.0287)
Search Cost Intercept (ι)	2.5392*** (0.0376)	1.7868*** (0.0171)	1.4945*** (0.0349)
Travel Time (δ)	0.0661*** (0.0108)	0.1626*** (0.0329)	0.0226** (0.0098)
Preferred store (ψ_1)	-0.9345*** (0.0174)	-2.2016*** (0.0481)	-0.8636*** (0.0152)
EDLP (ψ_2)	0.0227 (0.0276)	-0.3273*** (0.0298)	0.0391 (0.0257)
Weekend (ω)	-0.0815*** (0.0191)	0.4736*** (0.0298)	-0.0426** (0.0194)
Segment Size	0.50	0.17	0.33

First, utility from consumption in the myopic case is inflated because what was previously attributed to future utility in the dynamic model is now all attributed to the current period. Second, search cost is underestimated because the value that accrues in the future from gaining a lower price due to current search is not accounted for in the myopic model; so the observed level of search cannot be rationalized by the potential future value from the search in the model, and therefore the model rationalizes it as due to low search cost.

Third, price sensitivity is underestimated in the myopic model. The direction of the bias on price sensitivity is at first blush surprising given that previous research that has focused on the

temporal dimension (e.g., Hendel and Nevo 2006) find that price sensitivities are over-estimated in a myopic model. To make it easier to interpret, Table 5 presents price elasticities for both the forward-looking and myopic cases.¹² Considering size of different segments, the myopic model underestimates price elasticity across the whole sample by roughly 30%.

To understand the underestimation of price sensitivity, one should consider three main factors that control the household’s current decision to purchase; current inventory/current consumption, utility from future consumption/cost of future stockouts, and expectation over future prices (getting a better deal in future). In a perishable frequently purchased category like milk, where the consumer cannot stockpile much, when she is low on inventory, the cost of future stockouts can overwhelm potential gains from getting a better price in the future. When we turn off the forward-looking dimension of the model, observing a consumer with a low level of inventory who makes a purchase at a high price (which is fairly common due to limited time span that consumer has to perform temporal search) the myopic model rationalizes it as low price sensitivity, while a forward-looking model rationalizes it as due to the need to avoid a future stockout.

Table 5. Myopic and long-term price elasticities.

Segment	Full Model	Myopic Model
1	-0.28	-0.20
2	-0.58	-0.07
3	-0.91	-0.78

Why is the direction of bias different relative to the previous literature on temporal search? Past research analyzed categories like detergents, razors etc., which have large inter-purchase times due to ease of stockpiling. In such categories, the effect of expectations over future prices (desire to get a better deal in future) is more powerful than that of avoiding stockouts, as consumer can store goods for longer time-periods, giving them more flexibility to perform temporal search without fear of stockouts. Further, in categories like detergents, consumers can more flexibly adjust consumption

¹² To calculate “long-term” price elasticities, we followed a procedure similar to that in Hendel and Nevo (2006). Specifically, we used the following procedure; first we solved the model for choice probabilities of each household using observed price distribution of each store. Then we modified price distributions by increasing prices by one percent (i.e. shifting price distribution, keeping its shape intact). Then we re-solved the model for choice probabilities using new prices. Finally, we simulated households’ visit and purchase decisions and measured percentage change in purchases with the modified prices.

by shifting wash cycles to after purchase or reducing the amount of detergent they use to reduce the cost of stockouts much more easily than with milk. This can further mute the effect of stockouts. Hence, households purchase less frequently at high prices, because there are enough opportunities to buy at low prices. In that case, a myopic model overestimates price sensitivity. In contrast, in a perishable category like milk, the frequency of purchase is relatively high at high prices due to fear of a stockout, which leads to underestimation of price sensitivity. Thus, by analyzing a truly “frequently purchased category” such as milk in contrast to detergents, we gain the insight that the direction of the bias is driven by the ratio of purchase to promotional frequency.

6 Impact of Promotional Frequency on Store Loyalty and Profits

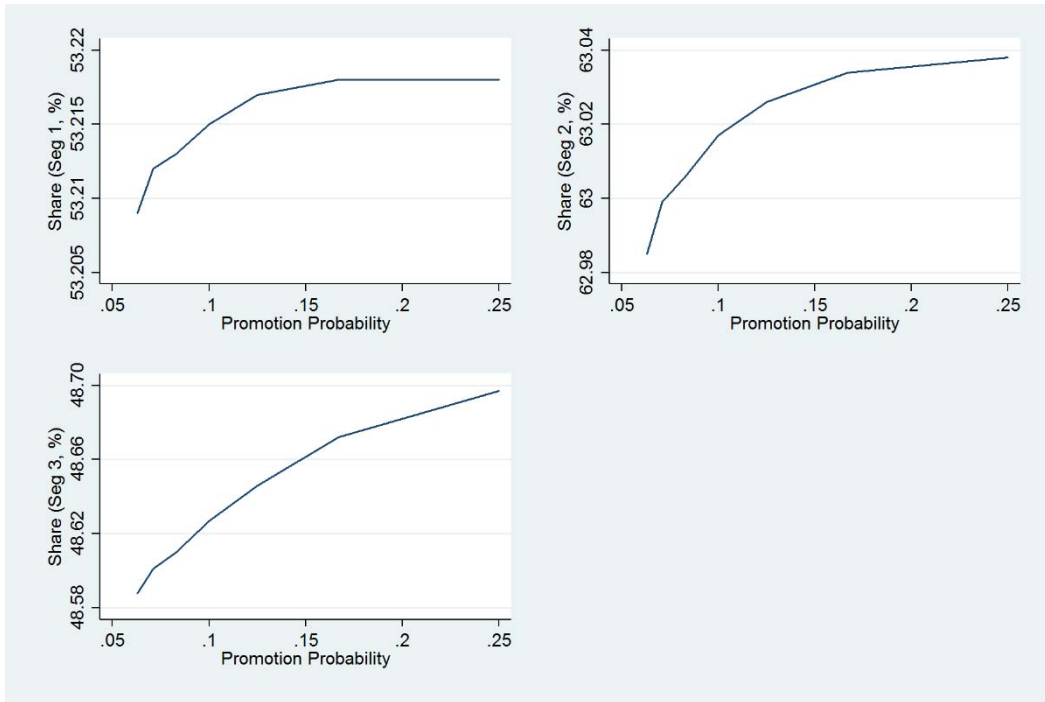
We now seek to understand how price promotions impact store loyalty in the presence of spatial search across stores and temporal search. Conventional wisdom suggests that as promotions become more frequent, cherry-picking behavior will increase, leading to reduced loyalty. However, Gauri et al. (2008) conjecture that in the presence of search costs, households may respond to more frequent promotions by choosing to shop more at their preferred store as they can take advantage of the periodic promotions without incurring the search costs of spatial store search.

We perform a counterfactual using our structural model of search across stores and time to test the Gauri et al. (2008) conjecture. We vary promotional frequency symmetrically at all stores, keeping average and regular price at the stores constant. This implies that when promotional frequency increases, a consumer can have more opportunities to obtain discounts, but the discount levels will be smaller. For the counterfactual, we vary frequency of promotion occurrence from once every eight weeks to once every two weeks in one week steps. This translates into an increase in promotion probability from 6.25% to 25% with corresponding promotional depth changing from 64% to 18%. We set the travel time to the primary and non-primary stores to be the average observed in the data.

Given this promotional environment, we forward simulate the behavior of households to compute a number of relevant metrics of loyalty and profits. To obtain stationary estimates with minimal simulation error, we forward simulate 1000 households for 20,000 periods and average the metrics across households. For loyalty, we report household level share of visits. Assuming a gross margin of 40%, we compute the annual profit per segment and total profits.

Figure 5 shows the share of visits to the primary store for the three segments. Consistent with the conjecture in Gauri et al. (2008), we find that the store visit share increases with promotional frequency across all three segments.

Figure 5. Share of visits to the primary store



We next explore how the increase in promotional frequency impacts store profitability. Table 6 reports annual store profit per household for the primary store as a function of promotion probability by segment and in the aggregate. Figure 6 shows that profit per household increases for each segment for the primary stores; the increase in overall gross margin is almost 4%. This is a very significant increase in profits for the grocery sector where net margins tend to be around 1-2% of revenues. Thus, increasing promotional frequency (with correspondingly shallower promotions) not only leads to greater store loyalty to the primary store, but also greater profitability.

Finally, we check how the average paid prices change in response to promotional frequency. Figure 7 shows the average paid price in the primary store for each segment increases as promotion probability increases. This increase in average paid price arises from two sources; first, when customers run into a promotion, the promotion depth is smaller. Second, when they don't face a promotion they still might buy the item without waiting for the next store visit since they will not

save as much under promotion anyway. In a sense, the increase in promotion probability with shallower promotions decreases the value of spatial price search, but increases the value of temporal search leading to the greater loyalty towards primary stores.

Table 6. Annual profit per household for primary store

Promo. Prob.	Seg1	Seg2	Seg3	Total
6.25%	\$42.98	\$45.18	\$18.63	\$36.40
7.14%	\$42.94	\$45.40	\$18.95	\$36.52
8.33%	\$43.12	\$45.92	\$19.39	\$36.86
10.00%	\$43.23	\$46.33	\$19.76	\$37.10
12.50%	\$43.21	\$46.60	\$20.07	\$37.24
16.67%	\$43.38	\$47.09	\$20.48	\$37.56
25.00%	\$43.45	\$47.47	\$20.80	\$37.76

Figure 6. Change in primary store profit versus change in promotion probability

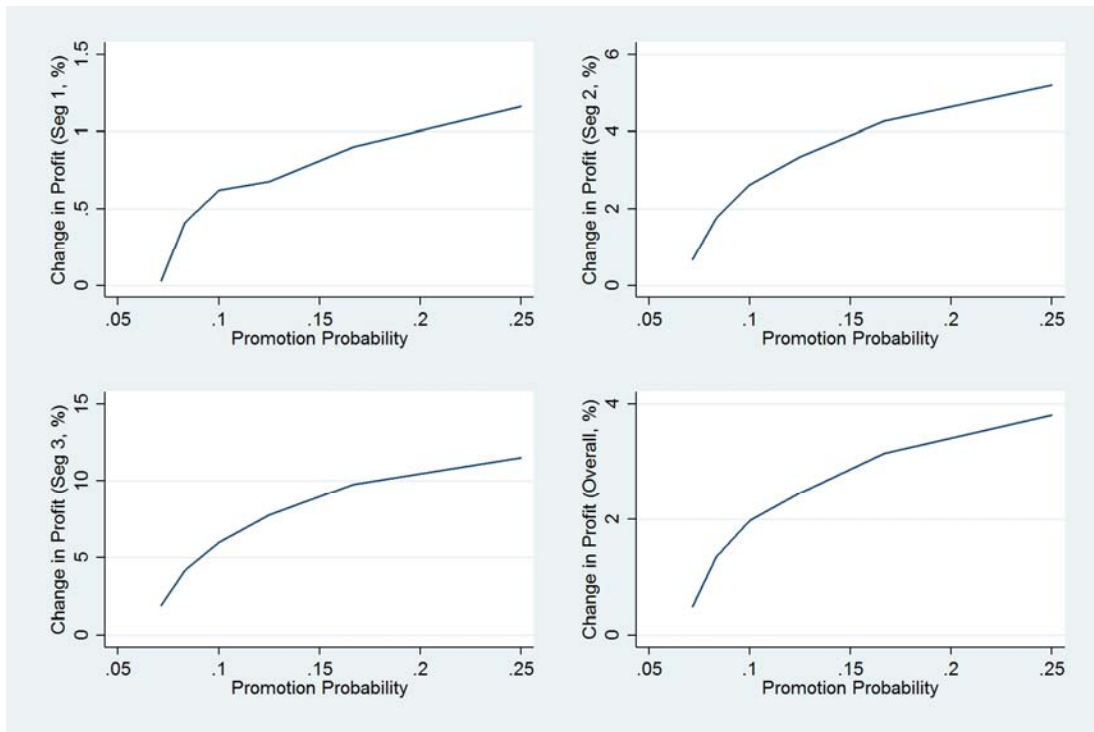
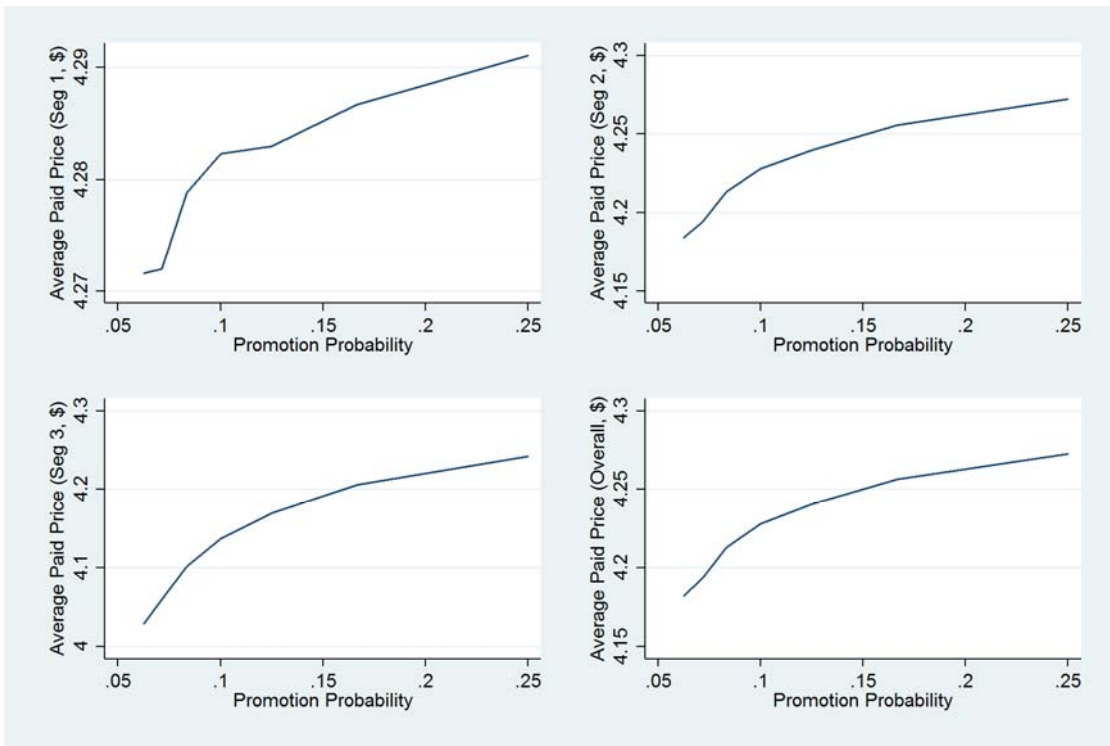


Figure 7. Change in average paid prices by segment



7 Conclusion

This paper introduces a dynamic structural model of search along both the spatial (store) and temporal dimensions allowing for discrete unobserved heterogeneity. The model nests a finite horizon model of spatial search across stores within an infinite horizon model of search across time. We use an iterative EM-algorithm based approach in combination with an MPEC formulation of the dynamic model to obtain estimates of the structural model accommodating discrete heterogeneity.

We calibrate the model using household purchases in the milk category—where consumers purchase often and there is limited stockpiling due to the perishable nature of the good even if there are promotions. We find different search strategies along the spatial and temporal dimensions by different segments as a function of their search costs, price sensitivity and relative preference for the primary store. We demonstrate that not accounting for temporal search can have substantial bias in the estimates. Our analysis on the milk category helps to provide a more nuanced sense on the direction of the bias relative to the existing literature which focused on temporal search using highly stockpiling categories such as detergents. We find that the direction of the bias by omitting the temporal dimension is determined by the relative frequency of purchase and frequency of promotions.

When frequency of promotions is much greater than the frequency of purchases as in laundry detergents, omitting the temporal dimension leads to overestimation of price elasticities. However, when the frequency of promotions is comparable to the frequency of purchases (due to inability to stockpile) as in the milk category, omission of the temporal dimension leads to underestimation of price sensitivities because the stockout avoidance motivation is stronger. Further, search costs are also underestimated.

Finally, we evaluate the substantive question of how price promotions impact store loyalty. We find that in the presence of search costs, price sensitive shoppers respond to price promotions by reducing cross-store price search and increasing temporal price search at their preferred store, thus increasing the level of store loyalty to their preferred store. Thus, in contrast to conventional wisdom which suggests that price promotions reduce loyalty among price sensitive shoppers, we find that the presence of even small search costs in combination with small levels of store differentiation can increase the level of store loyalty in the market.

Our analysis is an initial foray in the search literature into developing a simultaneous model of search along the spatial and temporal dimensions. We believe there is more opportunity for both theoretical and empirical work in a joint model of search along both dimensions. A theoretical model that characterizes equilibrium pricing when both dimensions of search are present can help gain more insight into how the two dimensions interact to generate marketplace outcomes both on the consumer and firm side. Our analysis demonstrates that the nature of biases in omitting time dimension of search can be category specific; for example, we discovered that the relative frequency of price promotions and purchase can impact the nature of bias in estimated price sensitivities. A systematic investigation of factors that drive the bias can be valuable for retailers and academics seeking to understand the role of retail promotions and consumer behavior. Finally, we found that store differentiation, search cost and temporal search interact to impact household search strategies and outcomes such as store loyalty. Further, while our analysis has been for a frequently purchased category, it should be valuable to apply our framework to one-time purchases of durable goods to gain insight into the nature of spatial and temporal search in such categories. Overall, our dynamic structural model of spatiotemporal search should provide the impetus to ask additional questions about how market outcomes change as a function of category characteristics, store promotional strategies and store locational configurations.

Appendix

A. Choice of Milk Category for Analysis

To select a category for the analysis, we wanted a category with high penetration, high levels of spend and share of customer basket. Table A1 shows the top ten product modules ranked based on average share of household spending. Soft drinks had the highest share of basket and average spending, but the large number of brands and varieties in this category made it a difficult category to study category choice. Milk had the second highest share of total household spending at 3.3% and the highest penetration level. Also as most consumers choose private labels, we can abstract away from brand choice, keeping the state space manageable given the large state space needed to model store visit and purchase choice across time.¹³

Table A1. Average Share of Top Ten Most-Spent Product Modules across Household*

Rank	Product Module	Avg. Share	Avg. Spending(\$)	Penetration
1	SOFT DRINKS(CARBONATED & LOW CALORIE)	4.81%	117.42	87%
2	DAIRY-MILK-REFRIGERATED	3.39%	79.63	88%
3	CIGARETTES	2.70%	88.66	17%
4	CEREAL - READY TO EAT	2.60%	60.01	81%
5	BAKERY - BREAD – FRESH	2.10%	49.00	87%
6	COOKIES	1.54%	35.72	73%
7	ICE CREAM – BULK	1.44%	32.67	66%
8	SOUP-CANNED	1.29%	30.14	69%
9	CANDY-CHOCOLATE	1.26%	28.76	63%
10	WATER-BOTTLED	1.23%	30.45	48%

* A few households spend large amounts (some in excess of \$10,000) on cigarettes. If we drop such outlier households, cigarettes drop out and fruit drinks enters the list at No. 10. Penetration figures are based on at least a \$10 spend in the category during the year of data.

¹³ Nielsen categorizes products into 10 departments, about 125 product groups, and about 1100 product modules. The product group level include products from a variety of modules. Promotions usually happen at product module level and hence is the right level for our purposes. Note that milk would still be among top ten groups even if the analysis was at the product group level.

B. Constructing Price Data

We supplement our panel data with retail scanner data from Nielsen, provided by Kilts Center for Marketing at the University of Chicago for milk prices at the stores. There are two challenges in using this data: 1) the unique store identifier code in the Chicago data is different from that in our panel data.¹⁴ 2) some stores in our panel data are not available in the Chicago dataset.

To address the first issue, we match stores in our panel data set with that in the retail scanner data using retailer code, first three digits of zip code, and unique identification number of the household in the sample who buys from each store. In cases where we find more than one store matching the same retailer name and area (zip3), we take average of price across those stores. This is equivalent to assuming that pricing is set at region level rather than individual store level, for stores with the same chain name—a reasonable assumption. To address the second issue, we take the following steps. First, if we do not observe price data for any of the stores in a household’s consideration set, we drop that household. Second, if we have data on price in the retail scanner data for at least one of the stores in a household’s consideration set, we keep the household in the sample, but impute price data for stores that are not observed in the scanner data, using observed prices (and their distribution) in the panel data, when a panelist in the sample makes a purchase from the store. Such stores tend to be second or third stores for the household with limited purchases from them; hence this does not affect many observations. Further, due to overlaps among household consideration set of stores, we often observe store prices even for periods even when a household in the sample has not paid a visit to a store, as another household has bought the item from the store.

¹⁴ As the panel data provided by Kilts Center does not provide address or full zip code of each store, which is needed calculate travel time between each household and corresponding stores, we need to use the panel dataset provided by Nielsen for this paper.

References

- Arcidiacono, P., & Jones, J. B. (2003). Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm. *Econometrica*, 933-946.
- Besanko, D., & Winston, W. L. (1990). Optimal Price Skimming by a Monopolist Facing Rational Consumers. *Management Science*, 555-567.
- Chung, D. J., Steenburgh, T., & Sudhir, K. (2014). Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans. *Marketing Science*, 165-187.
- Conlisk, J., Gerstner, E., & Sobel, J. (1984). Cyclic Pricing by a Durable Goods Monopolist. *Quarterly Journal of Economics*, 489-505.
- De Los Santos, B., Hortaçsu, A., & Wildenbeest, M. R. (2012). Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior. *American Economic Review*, 2955-2980.
- Erdem, T., Imai, S., & Keane, M. (2003). Consumer Price and Promotion Expectations: Capturing Consumer Brand and Quantity Choice Dynamics Under Uncertainty. *Quantitative Marketing and Economics*, 5-64.
- Fotheringham, A. S. (1988). Consumer Store Choice and Choice Set Definition. *Marketing Science*, 299-310.
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, 351-401.
- Gauri, K. D., Sudhir, K., & Talukdar, D. (2008). The temporal and spatial dimensions of price search: Insights from matching household survey and purchase data. *Journal of marketing research*, 226-240.
- Gentry, M. (2012). Displays, Sales, and In-Store Search in Retail Markets. *Working Paper*.
- Goeree, M. S. (2008). Limited Information and Advertising in the US Personal Computer Industry. *Econometrica*, 1017-1074.
- Hartmann, W. R., & Nair, H. S. (2010). Retail Competition and the Dynamics of Demand for Tied Goods. *Marketing Science*, 366-386.
- Hendel, I., & Nevo, A. (2006). Measuring the Implications of Sales and Consumer Inventory Behavior. *Econometrica*, 1637-1673.
- Hong, H., & Shum, M. (2006). Using price distributions to estimate search costs. *Rand Journal of Economics*, 257-275.
- Honka, E. (2013). Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry. *Working Paper*.
- Kim, J. B., Albuquerque, P., & Bronnenberg, B. J. (2010). Online Demand Under Limited Consumer Search. *Marketing Science*, 1001-1023.
- Koulayev, S. (2009). Estimating Demand in Search Markets: The Case of Online Hotel Bookings. *Working Paper*.

- McCall, J. J. (1970). Economics of Information and Job Search. *Quarterly Journal of Economics*, 113-126.
- Mehta, N., Rajiv, S., & Srinivasan, K. (2003). Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation. *Marketing Science*, 58-84.
- Meyer, J. I., & Eagle, T. C. (1982). Context-Induced Parameter Instability in a Disaggregate-Stochastic Model of Store Choice. *Journal of Marketing Research*, 62-71.
- Mortensen, D. T. (1970). Job Search, the Duration of Unemployment, and the Phillips Curve. *The American Economic Review*, 847-862.
- Neslin, S. A., Henderson, C., & Quelch, j. (1985). Consumer Promotions and the Acceleration of Product Purchases. *Marketing Science*, 147-165.
- Popkowski Leszczyc, P. T., Sinha, A., & Timmermans, H. J. (2000). Consumer store choice dynamics: an analysis of the competitive market structure for grocery stores. *Journal of Retailing*, 323-345.
- Roberts, J. H., & Lattin, J. M. (1991). Development and Testing of a Model of Consideration Set Composition. *Journal of Marketing Research*, 429-440.
- Schuler, H. J. (1979). A Disaggregate Store-Choice Model of Spatial Decision-Making. *The Professional Geographer*, 146-156.
- Seiler, S. (2013). The Impact of Search Costs on Consumer Behavior: A Dynamic Approach. *Quantitative Marketing and Economics*, 155-203.
- Solgaard, H. S., & Hansen, T. (2003). A hierarchical Bayes model of choice between supermarket formats. *Journal of retailing and consumer services*, 169-180.
- Song, Y., Mela, C. F., Chiang, J., & Chen, Y. (2012). Determining Consumers' Discount Rates With Field Studies. *Journal of Marketing Research*, 822-841.
- Stigler, G. J. (1961). The Economics of Information. *The Journal of Political Economy*, 213-225.
- Stiglitz, J. E., & Salop, S. C. (1982). The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents. *American Economic Review*, 1121-1130.
- Su, C., & Judd, K. (2012). Constrained Optimization Approaches to Estimation of Structural Models. *Econometrica*, 2213-2230.
- Weitzman, M. L. (1979). Optimal Search for the Best Alternative. *Econometrica*, 641-654.