SPATIOTEMPORAL SEARCH

By

Navid Mojir, K. Sudhir and Ahmed Khwaja

March 2014

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/
Spatiotemporal Search

Navid Mojir, K. Sudhir and Ahmed Khwaja

Yale School of Management

February 2014

1 Navid Mojir is a PhD candidate in Marketing (e-mail: navid.mojir@yale.edu), K. Sudhir is James L. Frank Professor of Private Enterprise and Management (k.sudhir@yale.edu) and Ahmed Khwaja is an Assistant Professor of Marketing (ahmed.khwaja@yale.edu), all at the Yale School of Management, 135 Prospect Street, 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 Tilburg, Utah, Madison, Yale Marketing, and Yale IO prospectus seminars. We thank Steven Berry, Stephan Seiler and Che-Lin Su for comments.
Abstract

Despite evidence that consumers search across both stores (spatial) and time (temporal), the search literature models search in only one dimension. We develop a model of spatiotemporal search that nests a finite horizon model of spatial search 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 for latent class heterogeneity. The empirical analysis is based on 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 importance of stockouts relative to stockpiling in the milk category. Further, contrary to the conventional wisdom that promotions reduce 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 across time is widespread in many 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, consumers may choose different search strategies along the space and time dimensions (Gauri, Sudhir, & Talukdar, 2008). Yet, structural empirical models of consumer price search have focused on search either along the spatial or temporal dimension, but not both. Omitting either dimension can lead to bias in the estimation of both search cost and price elasticity. It can also lead to misleading counterfactual estimates in evaluating the effectiveness of price promotions policies and its impact on store loyalty. We therefore develop and estimate the first dynamic spatio-temporal structural model of price search across stores and across time. The model nests a finite horizon model of search across stores within an infinite horizon model of search over time.

There is a vast literature in economics and marketing on price search, both theoretical and empirical, focused on search across stores, that does not consider the temporal dimension. Two types of search models dominate the search (across stores) literature. The first is the fixed sample size search 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. In marketing, the literature on consideration sets is based on the fixed sample size model (Roberts & Lattin, 1991; Mehta, Rajiv, & Srinivasan, 2003). Honka (2013) also assumes a fixed sample size model. 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. One example of the prediction they test is that
consumers will always purchase at the last store; but in the data there are many cases where the consumer does not purchase at the last store.\textsuperscript{2} Honka and Chintagunta (2013) develop an identification strategy to distinguish sequential versus simultaneous search using only price and consideration set size data.

There is also a literature on price search over time.\textsuperscript{3} 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. 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. Hartmann and Nair (2010) study the problem of inter-temporal demand estimation of tied goods (razors and razor blades) across multiple store formats, but treating store visits as exogenous. Seiler (2011) studies the problem of inter-temporal price search for detergents treating store choice as exogenous, but endogenously modeling 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.

There are a number of modeling issues and challenges that we need to address in developing a model of across store and across time search and applying it to frequently purchased consumer goods. First, this is a unique setting, in which we have to nest a dynamic optimal stopping problem of sequential search and purchase across stores in a time period within another optimal stopping problem 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 last 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. Finally since store visit decisions are not typically driven by needs in only one category, we need to account for the possibility of non-focal category needs impacting store visits. This

\textsuperscript{2} Bell et al. (1999) model store format choice 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) to model consumer choice of EDLP versus high-low formats. However, their paper does not account for forward looking behavior.

\textsuperscript{3} There is a large literature on purchase acceleration in response to price promotions using scanner data (e.g., Neslin, Henderson and Quelch 1985).
issue has never been addressed in extant temporal search models. We estimate the dynamic structural model allowing for discrete heterogeneity by solving the dynamic program as a mathematical program with equilibrium constraints; and nesting this within an EM algorithm similar to Arcidiacono and Jones (2003).

We estimate the structural model using household visit and purchase choices in the milk category. We find that there are three segments of consumers that vary in their level of search costs and price sensitivity. Not accounting for the time dimension of search leads to considerable bias in estimates of search costs and price elasticity; but in a direction opposite to what has been reported in the literature (e.g., Erdem, Imai and Keane 2003; Hendel and Nevo 2006). Based on the estimates of the structural model, we perform a counterfactual that sheds insight on how promotions can induce loyalty even for low search cost consumers to their preferred store. Our results question 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. 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 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. We allow for the possibility that other factors affect a household’s decision to visit stores. As mentioned earlier, we develop a finite

\footnote{We do not model brand choice to focus on the essentials of the across-store and across-time search process. Our empirical application is for the milk category, where brand choice is not the critical dimension of purchase decision. While the modeling framework itself could be extended to accommodate brand choice, including brand choice can create computational challenges due to the explosion in the dimensionality of the state space.}
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.

2.1 The Basic Set Up

A household $h$ can search across a finite consideration set of stores denoted by $\Omega_h$. Let $N_h^{\text{max}}$ be the number of stores in $\Omega_h$; then, there are potentially a maximum of $N_h^{\text{max}}$ stages of store search in any given period $t$ until all stores in the consideration set $\Omega_h$ 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 $\Omega_{htn}$ denote the set of unvisited stores for household $h$ at time period $t$ at spatial search stage $n$.

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

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

(1) Visit Decision $(t,n)$: Household $h$ observes visit-related state variables $x_{ht}^n$ and decides whether (1) or not to visit another store from the set of unvisited stores at stage $n$ in period $t$ ($\Omega_{htn}$) so to maximize the household’s value function across the remaining stages in period $t$ and across future time periods.

---

(1) Visit Decision $(t,n)$: Household $h$ observes visit-related state variables $x_{ht}^n$ and decides whether (1) or not to visit another store from the set of unvisited stores at stage $n$ in period $t$ ($\Omega_{htn}$) 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 moves to the purchase decision at stage \((t,n)\).

b. No visit: A household that decides not to visit an additional store, 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 \(\Omega_{htn}^t\), household observes purchase-related (including 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 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 \(N_{h max}^t\)) 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 \(N_{h max}^t\) of time \(t\) leads to the visit decision in stage 1 at time \(t + 1\), i.e., Visit Decision \((t + 1,1)\).

---

5 Note that here we have \(k\) in the subscript in the purchase-related state variable \(x_{htk}^p\). This represents the identity of the store being visited. While we define visit-impacting variables (i.e., \(x_{ht}^v\)) to include visit-related information of all stores in the household’s consideration set (this will become clearer below where we define flow utilities for visit decisions), we do not do the same for purchase-related variables. This is due to the fact that variables that affect visit decision (e.g. store format) are known for all stores at the beginning of each time period, whereas variables that affect purchase decision (e.g. price) are revealed after visiting each store. Therefore, we do not aggregate them to define a store-independent purchase-related state variable the same way that we do for state variables related to the visit decision to acknowledge that purchase-related information for stores is revealed gradually and sequentially during the spatial search process.
To summarize, a household $h \in \{1,2,...,H\}$ at time period $t \in \{1,2,3,...\}$ and store search stage $n \in \{1,2,...,N^\text{max}_h\}$, observes state variables $x^v_{ht}$ that affect the decision to visit a store. The household makes a decision about whether to visit and which store to visit $y^v_{htn} \in \Omega_{htn} \cup \{0\}$, where $y^v_{htn} = 0$ 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 $\Omega_{htn}$ (i.e., $y^v_{htn} = k > 0$), the household observes purchase-related state variables $x^p_{htk}$ for that store and makes a decision $y^p_{htn} \in \{0,1\}$, where 0 indicates no purchase in the focal category and 1 indicates purchase in the focal category.$^6$

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 $k \in \Omega_h$.$^7$ 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 $x^v_{ht} = \{W_t, i_{ht}, i^\text{stock up}_{ht-1}, d_h\}$, 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 $k \in \Omega_{hn}$ at stage $(t,n)$ is given by:

$^6$ Note that although we call $x^v_{ht}$ and $x^p_{htk}$ state variables, some of the items included in them might be store specific or consumer-store specific characteristics (e.g., distance between a consumer and a store) which do not change over time, while some other items are truly state variables (i.e., change over time via a transition process). We do not separate them here to avoid notational complexity.

$^7$ In the final implementation and estimation of the model we use square root of travel time measured in minutes. The concave transformation of minutes fits the data better than the actual minutes.
\[ u^v_{htnk}(x^v_{ht}) = X_{hk} \beta_h - S_h(d_{hk}, W_t) + \eta I_{n=1}^{stock\ up} + \varepsilon^v_{htnk} \text{ for } k > 0 \]

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 \( k \in \Omega_{htn} \). The third term \( \eta I_{n=1}^{stock\ up} \) is a parsimonious attempt 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), she would be reluctant to make a visit at the next period as she has enough inventory of non-focal items. We capture this idea without the need to keep track of inventory of non-focal items (and considerably expanding the state space)—serving as a parsimonious and computationally convenient modeling device.\(^8\) We account for this reluctance just once for each period; we use \( I_{n=1}^{stock\ up} \) as a dummy variable that is one for first stage and zero for other stages while \( I_{n=1}^{stock\ up} = 1\{SP_{h,t-1} \geq \bar{SP}_h\} \) counts a period as a stock up period if total spending of household \( h \) in that period is higher than average spending per period by that household.\(^9\) Finally, \( \varepsilon^v_{htnk} \) 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: \( S_h(d_{hk}, W_t) = \tau_h + \delta_h d_{hk} + \omega_h W_t \).

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 store characteristics (\( X_{hk} \)) using two variables that account for store differentiation: (1) Whether store \( k \) is EDLP and (2) Whether store \( k \) is the primary grocery store for household \( h \).\(^{10}\)

A household that forgoes search has the following utility,

\[ u^v_{htnk0}(x^v_{ht}) = \varepsilon^v_{htnk0} \]

\(^8\) The idea that consumption outside the category can impact store visits has not been addressed in the dynamic (temporal) structural modeling literature (e.g., Erdem, Imai and Keane 2003; Hendel and Nevo, 2006; Hartmann and Nair 2010). Although, these papers allow for a non-focal outside good in the utility function, this does not affect the search decision directly but rather passively by being the outside option or the residual in the budget constraint. In contrast, the non-focal category in our model has a direct and dynamic impact on the search for the focal category itself by affecting the store visit decision.

\(^9\) A model which allows stockup to cause reluctance to visit additional stores does not fit as well.

\(^{10}\) We define the store with highest share of visit in each household’s consideration set to be the primary store for that household. This is to capture store differentiation due to factors that we do not observe.
**Purchase Decision**

After visiting store \( k \), the household decides whether to make a purchase or not in the focal category. Flow utility if household decides to make a purchase is given by:

\[
u_{htnk1}(x_{htk}^p) = \alpha_h p_{ht} + \varepsilon_{htnk1}^p
\]

\[
= \tau_{htk1}(x_{htk}^p) + \varepsilon_{htnk1}^p
\]

Where \( \alpha_h \) is the price sensitivity of household \( h \), \( p_{ht} \) is price of the focal category in store \( k \) at time period \( t \), and \( \varepsilon_{htnk1}^p \) 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 gets:

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

All the structural error shocks in the above equations are assumed to be independent and identically distributed (i.i.d). The identical distribution of these shocks is assumed to be 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. This consumption utility is represented as,

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

Where \( i_{ht} \) is the inventory level of the focal category, \( c(i_{ht}^c + y_{ht}^p \chi) \) is consumption as a function of inventory level and \( \varphi \) is utility of consuming \( c(i_{ht}^c + y_{ht}^p \chi) \) units.\(^{11}\) Here, \( \chi \) represents the amount that gets added to consumer inventory if she makes a purchase (i.e., milk container size in our application) and \( y_{ht}^p = \sum_{n=1}^{N_{ht}} y_{htn}^p \) (which is equal to one if the consumer makes a purchase in time period \( t \) and zero otherwise). Let \( \rho_h \) be the household \( h \)'s consumption rate of the focal category. Specifically we assume \( c(i_{ht}^c + y_{ht}^p \chi) = \min\{\rho_h, i_{ht} + y_{ht}^p \chi\} \), where household

\(^{11}\) Note that we do not need to have an error shock in this equation as the utility of consumption can easily be included in flow utilities of either the search or purchase stages. It is like a “salvage value” that is revealed after the visit and purchase decisions have been made conditional on inventory at the end of the time period. We define utility of consumption separately as it makes the notation clearer and easier to follow.
consumes an amount equal to consumption rate if there is more than one serving left in inventory and consumes what is left in inventory otherwise.  

We assume a linear form for utility from consumption. Specifically, \( \varphi(c(i_{ht} + y_{ht}^p \chi)) = \sigma c(i_{ht} + y_{ht}^p \chi) + \tau \), where \( \sigma \) and \( \tau \) are parameters to be estimated.

### 2.3 State Transitions

Here we define appropriate state transitions and expectations associated with variables.

**Inventory**

Inventory held by household evolves as follows:

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

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

**Price Distribution**

We assume that prices follow an exogenous discrete distribution with \( m \) different levels of possible prices. \(^{13}\) Prices are also assumed to have different distributions for different stores and distributed independently over time. \(^{14}\) More formally: \( p_{ht} \sim \text{Multinomial}(1, p_k) \).

**Stock Up Dummy**

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:

\[
I_{ht}^{stock up} \sim \text{Bernoulli}(\pi_{h}^{S\rightarrow S}I_{ht-1}^{stock up} + \pi_{h}^{N\rightarrow S}(1 - I_{ht-1}^{stock up}))
\]

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.

---

\(^{12}\) We define serving as the amount that household consumes in one time period.

\(^{13}\) We use five price levels for estimation purposes.

\(^{14}\) 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. See Khan et al (2013) for a discussion of institutional reasons like state and federal pricing regulations that make milk prices plausibly immune to demand shocks and more a function of supply and cost shocks. The search literature typically assumes a first order Markov process, but does not model the decision to visit the store. In our setting where we model store visits, the assumption of a first order Markov process is problematic because if a household does not visit a store at time period \( t \), the household cannot form expectations of prices for that store at time period \( t + 1 \) using a Markov process since the price at time period \( t \) at that store would be unknown to the household. We find the independence assumption empirically more appealing for our analysis.
Weekend

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

Store consideration set evolves as follows, where the store visited in stage \( n-1 \), it is removed from the consideration set at stage \( n \).

\[ \Omega_{ht0} = \Omega_h \quad \text{and} \quad \Omega_{htn} = \Omega_{htn-1} \setminus y_{htn-1}^v \]

2.4 The 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 \( h \),

\[
\max_{\{\Delta_H\}_{t=1}^{\infty}} \mathbb{E}\left( \sum_{t=0}^{\infty} \beta^t \varpi_{ht}(\Delta_{ht}) \mid \mathcal{F}_{ht} \right),
\]

where \( \Delta_{ht} = \{y_{htn}^v, y_{htn}^p\} \) represents the vector of a household’s visit (\( y_{htn}^v = \{y_{htn}^v\}_{n=1}^{N_h(t)} \)) and purchase (\( y_{htn}^p = \{y_{htn}^p\}_{n=1}^{N_p(t)} \)) decisions. These decisions in each time period are conditional on visit and purchase-related observed and unobserved state variables: \( \mathcal{F}_{ht} = \{x_{htn}, x_{ht}^p, \varepsilon_{ht}^v, \varepsilon_{ht}^p\} \). Here \( x_{htn}^p = \{\{x_{htnk}^p\}_{k \in \Omega_n} \}_{n=1}^{N_h(t)} \) includes all the relevant observed state variables for the purchase state, while \( \varepsilon_{htn}^v = \{\varepsilon_{htnk0}^v, \varepsilon_{htnk}^v\}_{k \in \Omega_n} \}_{n=1}^{N_h(t)} \), and \( \varepsilon_{ht}^p = \{\{\varepsilon_{htnk1}^p, \varepsilon_{htnk0}^p\}_{k \in \Omega_n} \}_{n=1}^{N_p(t)} \) 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 visit and purchase stages up to the \( N_h(t) \) stage\(^{15} \) plus consumption utility:

\[
\varpi(\Delta_{ht}) = \sum_{n=1}^{N_h(t)} \prod_{l=0}^{N_p(t)-1} (u_{htnl}^p)_{y_{htnl}^p}^{y_{htnl}^p} + \sum_{n=1}^{N_p(t)-1} \prod_{l=0}^{N_p(t)-1} (u_{htnl}^p)_{y_{htnl}^p}^{y_{htnl}^p} + u_{ht}^c
\]

\(^{15}\) Recall that \( N_h(t) \) denotes the stage \( n \) at which household \( h \) stops search in period \( t \).
2.5 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 $\Omega_{htn}$. 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 $N_{htn}^{max}$ stores in its consideration set, at any stage before the last stage (i.e. $n < N_{htn}^{max}$) 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 $\Omega_{htn+1}$. With a purchase, the household gets the corresponding flow utility plus discounted value of utility (across time) that she will get starting next period.

$$v_{htnk1}^p(x_{htk}^p, x_{ht}^v) = \overline{w}_{htnk1}^p + u_{ht}^v + \beta E_{\bar{z}_{t+1}} \Delta w_{ht} \Delta \pi_{ht}^v EV_{htnk+1}^v(x_{ht}^v, \Omega_h) + \varepsilon_{htnk1}^p$$

$$= \overline{v}_{htk1}^p + \varepsilon_{htnk1}^p$$

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.

$$v_{htnk0}^p(x_{htk}^p, x_{ht}^v) = EV_{ht,n+1}^v(x_{ht}^v, \Omega_{ht,n+1}) + \varepsilon_{htnk0}^p$$

$$= \overline{v}_{ht0}^p + \varepsilon_{htnk0}^p$$

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

$$v_{htnk}^v(x_{ht}^v) = v_{htnk}^c + EV_{htnk}^p + \varepsilon_{htnk}^v$$

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.,

$$v_{ht0}^v(x_{ht}^v) = u_{ht}^c + \beta.E_{x_{ht+1}^v|x_{ht}^v,\Delta_t,v_{ht}}EV_{ht+1,1}(x_{ht+1}^v,\Omega_h) + \varepsilon_{ht0}^v$$

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

$$v_{htN_h^{max}k0}^p(x_{htk}^p, x_{ht}^v) = u_{ht}^c + \beta.E_{x_{ht+1}^v|x_{ht}^v,\Delta_t,v_{ht}}EV_{ht+1,1}(x_{ht+1}^v,\Omega_h) + \varepsilon_{htN_h^{max}k0}^p$$

This completes definitions of all the necessary choice-specific value functions.

### 2.6 Ex-Ante Value Functions

Now we can 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,
\[
EV_{htn}^v(x_{ht}^v) = E_{\epsilon_{htn}^v | x_{ht}, \Omega_{htn}, \epsilon_{htn-1}^v} \max_{k \in \Omega_{htn} \cup \{0\}} \left[ v_{htn}^v(x_{ht}^v) \right]
= \log \left( \sum_{k \in \Omega_{htn} \cup \{0\}} \exp (\bar{v}_{htn}^v) \right),
\]

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,
\[
EV_{htnk}^p = E_{x_{htnk}^v, \epsilon_{htnk1}^v, \epsilon_{htnk0}^v | x_{htnk-1}^v, \epsilon_{htnk-1}^v} \max\{ v_{htnk1}^p, v_{htnk0}^p \}
= \int_{x_{htnk}^v} \log [\exp (\bar{v}_{htnk1}^p) + \exp (\bar{v}_{htnk0}^p)] \cdot dP(x_{htnk}^p)
\]
Again, the second equality is based on the extreme value distribution and the conditional independence assumption.

2.7 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:

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

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,

\[
P_{htnk1}^p = \frac{\exp (\bar{v}_{htnk1}^p)}{\exp (\bar{v}_{htnk1}^p) + \exp (\bar{v}_{htnk0}^p)}.
\]

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,

$$
L_{h|g} = \prod_{t=1}^{T_h} \prod_{n=1}^{N_{h}} (P_{h\text{ continues} \ g} | g) \prod_{k=0}^{K} (P_{h}^{p} | g) \prod_{k=1}^{K} (P_{h}^{p} \ o f f = k) \prod_{k=0}^{K} (1 - P_{h}^{p} | g) \prod_{k=0}^{K} (1 - P_{h}^{p} = k).
$$

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

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

3 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 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 other papers 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. 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.

16 Note that we do not need to assume that probability of being a member of each group (interpreted here as segment size) is the same for all households. In fact we will relax this later.

17 Note that in Equation (1) $L_{h|g}$ is the product of probabilities of the sequence of decisions for all the time periods during which household $h$ is observed. This sequence can include between one to $2N_h^{max}$ probability terms (a visit and a purchase decision for each store) depending on actions that household $h$ takes.
### 3.1 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

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

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 method proves to be computationally intensive considering the size of the state space and structure of the problem. Therefore, instead of using NFXP, we re-formulate the problem as a constrained optimization problem. To that end, we re-write likelihood function 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.

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

subject to:

$$EV_h^{v}(x_h^{v}, \Omega_{h|v}) = \log \left\{ \sum_{k \in \Omega_{h|v} \cup \{0\}} \exp(\bar{\pi}_{h|v}^{v}(x_h^{v})) \right\}, \quad \forall \tau \in \{1, \ldots, T_h\}, \forall h \in \{1, \ldots, N_h\}$$

$$EV_h^{p} = \int_{x_h^{p}} \log[\exp(\bar{\pi}_{h|p}^{p}) + \exp(\bar{\pi}_{h|p}^{0})] \, dP(x_h^{p})$$

$$\forall \tau \in \{1, \ldots, T_h\}, \forall h \in \{1, \ldots, N_h\}, \forall k \in \{1, \ldots, N_h\}$$

Where $EV_h^{v} = \{\{EV_h^{v}\}_{t=1}^{T_h} \}_{n=1}^{N_h}$ and $EV_h^{p} = \{\{EV_h^{p}\}_{t=1}^{T_h} \}_{n=1}^{N_h} \}_{k=0}^{N_h}$ are set of ex-ante value functions for the search and purchase stages respectively. Similarly, $\bar{\pi}_h^v = \{\{\bar{\pi}_h^v\}_{t=1}^{T_h} \}_{n=1}^{N_h} \}_{k=0}^{N_h}$ and $\bar{\pi}_h^p = \{\{\bar{\pi}_h^p\}_{t=1}^{T_h} \}_{n=1}^{N_h} \}_{k=0}^{N_h}$ represent set of deterministic parts of the choice-specific value functions for the search and purchase stages.

---

18 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.
Now 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 $\Pr(g \mid x_h^v, x_h^p, \Delta_h, p; \Theta)$ 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

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

### 3.2 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 $p = (p_1, ..., p_G)$. 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

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

Where $L_{h|g}$ is individual likelihood for household $h$ conditional on being of type $g$, and

$$x_h = \{x_h^v \cup \{x_h^p \}_{k=1}^{N_h} \}_{t=1}^{T_h}$$

represents set of all observed state variables for household $h$. The maximum likelihood estimate of $\hat{p}_g$ is given by

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

### 3.3 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 $\Theta$.

**Step 1:** Calculate $p_g^\hat{h}$, using equation (3), conditional on $p_g$ and $\Theta$.  \hspace{1cm} (19)

Consider that to calculate $p_g^\hat{h}$ we need to calculate likelihoods conditional on $\Theta$. We can do it using contraction mapping, or instead, we can formulate the problem as a constrained optimization problem. Note that this optimization
Step 2: Given the estimates of $p^h_g$, use equation (4) to update $p_g$.

Step 3: Using estimates of $p^h_g$, maximize equation (2) subject to Bellman equations as constraints to update $\Theta$.

Step 4: Iterate over steps 1 to 3 till convergence on $\Theta$.

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).

3.4 Identification

Identification of different parameters of the model is straightforward. Price coefficient is identified using variation in prices and also consumers’ purchase decisions. Parameters of consumption utility function ($\sigma$ and $\tau$) are identified from the observed variation in households consumption rate and the imputed stockouts. Utility from consumption of non-focal categories ($\eta$) 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.

4 Data

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.

---

20 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.

21 Typically, weekly discount factor is assumed to be 0.995 in empirical research. Our assumption of 0.993 for half-week time period results in a smaller weekly discount factor, but this is consistent with findings of more recent empirical stream of research that estimates discount factor (Song, Mela, Chiang, & Chen, 2012; Chung, Steenburg, & Sudhir, forthcoming). For a review of the literature on discounting look at Frederick, Loewenstein, & O’Donoghue (2002).
We use milk as our focal category. Milk is an ideal category for our purposes, because (i) due to its perishable nature, there can be only limited stockpiling and therefore it is frequently purchased by a large share of households; this provides us multiple purchase occasions within a yearly sample; (ii) the product is frequently promoted making price search sensible; and (iii) brand is not a major consideration in household purchases allowing us to focus on category choice.

To avoid outliers in terms of distance, we drop households that shop for milk at stores more than 15 kilometers away. We believe it is likely that these households are purchasing from these stores due to their proximity from work. Since their work location is unobserved, we decided to drop such households from the analysis. 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. We focus on households with two or fewer stores in their consideration set; hence $N_h^{\text{max}} = 2$ or below for all households. This allows us to gain computational tractability in the finite horizon model; even with this assumption, we have 4 potential stages in the finite horizon model, because each of the two stores has a visit and purchase stage. Third, we omitted households who do not shop frequently (less than 20 shopping trips over the period of data collection) and do not purchase milk frequently (purchase milk in less than 5% of their shipping trips). Finally, to avoid the issue of size choice, we focused our analysis on households loyal to the most common size (one gallon) over the term of data collection. In all we use 373 households who shop from 690 stores.

5 Results

We begin by providing some model free evidence of search across stores and across time to warrant a dynamic structural model of search across stores and across time. Next we report the results of the full structural model that have the spatial and time 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 Model Free Evidence

We present some model-free evidence to show that there is price search spatially across stores and across time.

---

22 Roughly 70% of households make more than 90% of their milk purchases from one or two stores (regardless of size of the consideration set). Therefore, limiting number of stores in consideration of households in the sample should not impact final results dramatically.
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 distribution of share of time periods in which a household visits both stores within a time period. A large number of households visit two stores within the same weekday or weekend period.

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

To assess whether there is spatial search for milk or it is consumers have strong store-category loyalty, Figure 3 presents distribution of purchases of milk from their “favorite store” (store from which consumer has purchased the item from most often) for milk purchases. In fact, milk is purchased from both stores by multiple households.

Figure 3. Store-category loyalty for milk
Finally, we test whether milk purchases at the two stores are not simply due to the sequence in which the store is visited but are likely due to search. Figure 4 shows the probability distribution of purchasing milk from the second store conditional on visiting two stores in the same time period. Many households purchase milk at the second store during the same period. These suggest evidence of cross-store search.

**Figure 4: Milk purchases at second store visited**

![Figure 4: Milk purchases at second store visited](image)

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 two stores as a function of the inventory level of milk controlling for heterogeneity by including household fixed effects in the model.\(^{23}\) In this regression, the coefficient of inventory of milk is negative and significant \((p < .01)\) showing that increase in inventory of milk decreases probability of visiting two stores in the same period.

**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

---

\(^{23}\) 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.
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 7.51 periods (half-weeks) across households when purchases were made when there was no promotion, and 6.58 (half-weeks) across the same households when purchases were made on promotion. The difference of 0.92 periods is statistically significant at \( p = 0.06 \), suggesting that even for the milk category there is evidence of purchase acceleration.

5.2 Estimates of the structural model

The result of estimation of the model with three segments is presented in Table 1. All coefficients are highly significant \( (p < 0.01) \) and have expected signs, except for the weekend coefficient for the first segment, travel time (we used square root of travel time to account for diminishing marginal effects) of the second segment, and preference for EDLP stores for the second segment.

Segment 1 comprises 55% of the sample households, while second and third segments represent 24% and 21% of the sample, respectively. Segment 1 has the highest search cost and lowest price sensitivity; therefore they do not place much value on price search; Hence, they should perform the least amount of search across time and across stores. Segment 3 has the lowest search cost and the highest price sensitivity; hence, they value gains from search, but also have low cost of search, therefore, for a given level of price dispersion, they will search more intensely on both the store and time dimensions. Segment 2 is in between the other two segments on both search cost and price sensitivity. However, during weekends, their search cost is comparable to segment 3; to the extent they do grocery shopping only on weekends, one can expect them to search across stores similarly. But given their lower price sensitivity, they do not value deals as much. Hence overall, this segment will have “moderate search”.

---

24 We estimated model with one, two, and also four segments. Although the four segment model has slightly better fit based on the Bayesian Information Criterion (BIC), the three segment model is considerably superior in terms of segment interpretability. We therefore focus our discussion in the paper based on results of the three segment model.

25 Note that while second segment prefers to go shopping during weekends, the third segment prefers weekdays. Search cost for weekends for the second segment would be \( 2.250 - 0.736 = 1.514 \), whereas search cost for weekdays for the third segment is 0.905 (only based on the intercept and ignoring the effect of travel time for now). Although, these two segments would behave very differently in terms of which day they prefer to go shopping, the frequency of store visits for these two segments would be to some extent similar.
Table 1. Search model with both store and time dimensions

<table>
<thead>
<tr>
<th></th>
<th>Segment 1</th>
<th>Segment 2</th>
<th>Segment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Sensitivity (α)</td>
<td>-0.1709***</td>
<td>-0.2559***</td>
<td>-0.4424***</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0151)</td>
<td>(0.0166)</td>
</tr>
<tr>
<td>Marginal Consumption Utility (σ)</td>
<td>4.3329***</td>
<td>3.8487***</td>
<td>4.2917***</td>
</tr>
<tr>
<td></td>
<td>(0.1995)</td>
<td>(0.2577)</td>
<td>(0.3078)</td>
</tr>
<tr>
<td>Intercept of Consumption Utility (τ)</td>
<td>-0.735***</td>
<td>-0.756***</td>
<td>-0.755***</td>
</tr>
<tr>
<td></td>
<td>(0.0505)</td>
<td>(0.0671)</td>
<td>(0.0769)</td>
</tr>
<tr>
<td>Stock Up Previous Period (η)</td>
<td>-0.505***</td>
<td>-0.457***</td>
<td>-0.258***</td>
</tr>
<tr>
<td></td>
<td>(0.0744)</td>
<td>(0.0697)</td>
<td>(0.0635)</td>
</tr>
<tr>
<td>Search Cost Intercept (λ)</td>
<td>2.3555***</td>
<td>2.2496***</td>
<td>0.9048***</td>
</tr>
<tr>
<td></td>
<td>(0.0572)</td>
<td>(0.0778)</td>
<td>(0.0774)</td>
</tr>
<tr>
<td>Travel Time (δ)</td>
<td>0.0830***</td>
<td>-0.040</td>
<td>0.1032***</td>
</tr>
<tr>
<td></td>
<td>(0.0158)</td>
<td>(0.0217)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Preferred store (ψ₁)</td>
<td>-0.961***</td>
<td>-0.819***</td>
<td>-0.965***</td>
</tr>
<tr>
<td></td>
<td>(0.0257)</td>
<td>(0.0290)</td>
<td>(0.0295)</td>
</tr>
<tr>
<td>EDLP (ψ₂)</td>
<td>-0.106***</td>
<td>-0.111***</td>
<td>-0.283***</td>
</tr>
<tr>
<td></td>
<td>(0.0313)</td>
<td>(0.0363)</td>
<td>(0.0401)</td>
</tr>
<tr>
<td>Weekend (ω)</td>
<td>0.1135***</td>
<td>-0.736***</td>
<td>0.7696***</td>
</tr>
<tr>
<td></td>
<td>(0.0288)</td>
<td>(0.0381)</td>
<td>(0.0399)</td>
</tr>
<tr>
<td>Segment Size</td>
<td>0.55</td>
<td>0.24</td>
<td>0.21</td>
</tr>
</tbody>
</table>

To test if the intuition presented above is valid, we compare the observed behavior across three segments. Table 2 presents metrics on the visit and purchase behavior for each segment. Segment 1 visits stores least often. Given that two periods constitute a week, as predicted based on structural estimates, the first segment has the minimum percentage of store visits, followed by second and third segments. In fact, the first segment does very little spatial search considering the fact that a consumer in this segment on average visits both stores in the consideration set only 2.2% of the time. The second segment does perform some spatial search, but not as much as the third segment. This was also predicted considering the lower search cost of the third segment.
Table 2. Observed search behavior for each of the three segments

<table>
<thead>
<tr>
<th></th>
<th>Segment 1</th>
<th>Segment 2</th>
<th>Segment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of Shopping Periods in Which at Least One Store Has Been Visited</td>
<td>33.6%</td>
<td>51.6%</td>
<td>59.3%</td>
</tr>
<tr>
<td>Percentage of Periods with Both Stores Visited</td>
<td>2.2%</td>
<td>10.3%</td>
<td>13.5%</td>
</tr>
<tr>
<td>Percentage of Periods with Both Stores Visited Conditional on Visiting at Least One Store</td>
<td>6.7%</td>
<td>19.5%</td>
<td>22.4%</td>
</tr>
<tr>
<td>Average Price Paid ($)</td>
<td>2.83</td>
<td>2.78</td>
<td>2.63</td>
</tr>
</tbody>
</table>

Table 3 reports the search costs in dollar terms for the three segments during weekdays and weekends at the primary and secondary stores based on the estimated parameters and price sensitivity. As expected from the parameters, the search cost for all segments at the primary store is low, relative to the secondary store. For segment 1, the search costs are roughly the same over weekdays and weekends. Segment 2 has higher weekend cost, but segment 3 has higher weekend costs. Segment 3 has very low weekday costs at their primary store, allowing such households to search extensively during weekdays.

Table 3: Search costs estimates

<table>
<thead>
<tr>
<th></th>
<th>Segment 1</th>
<th>Segment 2</th>
<th>Segment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekend secondary store</td>
<td>$15.42</td>
<td>$5.60</td>
<td>$4.25</td>
</tr>
<tr>
<td>Weekday secondary store</td>
<td>$14.76</td>
<td>$8.48</td>
<td>$2.51</td>
</tr>
<tr>
<td>Weekend primary store</td>
<td>$9.76</td>
<td>$2.41</td>
<td>$2.05</td>
</tr>
<tr>
<td>Weekday primary store</td>
<td>$9.09</td>
<td>$5.29</td>
<td>$0.31</td>
</tr>
</tbody>
</table>

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 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 for the secondary store. To calculate effective search cost for the primary store, we also included in the sum the coefficient on the preferred store dummy variable before dividing the sum by the price sensitivity coefficient.
5.3 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 dimensions, but not both. As we argued, this can lead to biased estimates of search costs and price sensitivity. We now assess the extent of bias by omitting the temporal dimension. By setting the discount factor to zero, the full model reduces to a pure store search only model; i.e., the pure store dimension model is the myopic version of the full model. 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 over-estimated. The bias is identical in sign across all three segments, though greatest for segment 1 and least for segment 3. 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.

We discuss the intuition for the three biases in our analysis. 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, 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 stock-outs, 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 stock-outs 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 would be 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 would rationalize it as due to the need to avoid a future stock-out.
Table 4. Search model with only store dimension

<table>
<thead>
<tr>
<th></th>
<th>Segment 1</th>
<th>Segment 2</th>
<th>Segment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Sensitivity (α)</td>
<td>-0.0335***</td>
<td>-0.1846***</td>
<td>-0.3019***</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0150)</td>
<td>(0.0163)</td>
</tr>
<tr>
<td>Marginal Consumption Utility (σ)</td>
<td>7.8523***</td>
<td>9.8897***</td>
<td>9.7317***</td>
</tr>
<tr>
<td></td>
<td>(0.2178)</td>
<td>(0.2747)</td>
<td>(0.3283)</td>
</tr>
<tr>
<td>Intercept of Consumption Utility (τ)</td>
<td>-1.779***</td>
<td>-2.324***</td>
<td>-2.157***</td>
</tr>
<tr>
<td></td>
<td>(0.0525)</td>
<td>(0.0702)</td>
<td>(0.0809)</td>
</tr>
<tr>
<td>Stock Up Previous Period (η)</td>
<td>-0.545***</td>
<td>-0.542***</td>
<td>-0.408***</td>
</tr>
<tr>
<td></td>
<td>(0.0720)</td>
<td>(0.0696)</td>
<td>(0.0631)</td>
</tr>
<tr>
<td>Search Cost Intercept (λ)</td>
<td>2.3010***</td>
<td>2.2485***</td>
<td>0.8426***</td>
</tr>
<tr>
<td></td>
<td>(0.0569)</td>
<td>(0.0791)</td>
<td>(0.0766)</td>
</tr>
<tr>
<td>Travel Time (δ)</td>
<td>0.0983***</td>
<td>-0.042</td>
<td>0.1250***</td>
</tr>
<tr>
<td></td>
<td>(0.0158)</td>
<td>(0.0219)</td>
<td>(0.0217)</td>
</tr>
<tr>
<td>Preferred store (ψ₁)</td>
<td>-0.951***</td>
<td>-0.807***</td>
<td>-0.991***</td>
</tr>
<tr>
<td></td>
<td>(0.0254)</td>
<td>(0.0291)</td>
<td>(0.0297)</td>
</tr>
<tr>
<td>EDLP (ψ₂)</td>
<td>-0.109***</td>
<td>-0.106***</td>
<td>-0.273***</td>
</tr>
<tr>
<td></td>
<td>(0.0311)</td>
<td>(0.0367)</td>
<td>(0.0401)</td>
</tr>
<tr>
<td>Weekend (ω)</td>
<td>0.1150***</td>
<td>-0.732***</td>
<td>0.7535***</td>
</tr>
<tr>
<td></td>
<td>(0.0284)</td>
<td>(0.0385)</td>
<td>(0.0402)</td>
</tr>
<tr>
<td>Segment Size</td>
<td>0.56</td>
<td>0.24</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Why is the direction of bias in our paper different relative to all of past research on temporal search? Past research has 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 stock-outs, as consumer can store goods for longer time-periods, giving them more flexibility to perform temporal search without fear of stockouts.\footnote{Also note that in a category like detergents, consumer can adjust her consumption due to the level of inventory to some extent and that intensifies the effect of the difference explained here.} Hence, households purchase less frequently at high
prices, because there are enough opportunities to buy at low prices. Hence 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

One of the substantive goals of the paper is to understand how price promotions impact store loyalty. Conventional wisdom suggests that as promotions become more frequent, cherry-picking behavior will increase, leading to reduced loyalty. However when there are search costs, and these costs are different for different stores, and a consumer can choose between searching spatially across stores or across time, a household might choose to time purchases at its preferred store, rather than shopping across stores, if promotions occur frequently enough. This might lead to potentially increased store loyalty for a household. The nature of the tradeoff between search across time and across stores is complicated: the tradeoff stems from differential search cost of primary and secondary stores, how that compares to price sensitivity, and the frequency with which purchases need to be made. If average search cost for a consumer is low enough compared to her price sensitivity, an increase in promotion frequency would cause her to search more. But there is another factor in choosing between spatial and temporal search and that is relative search cost of the primary store to the secondary store. If search cost for the primary store is very low compared to that of secondary store, then it makes sense for the consumer to go to the primary store more often to take advantage of more frequent promotions, instead of checking both stores since the latter would be more costly. This is what we expect to observe in the third segment since it has relatively low search cost and much lower effective search cost for the primary store compared to the secondary store. Following a similar reasoning, we expect to observe least change in behavior (i.e. lowest increase in search) for the first segment, since it has highest search cost and lowest price sensitivity. Note that we also expect less switching from spatial to temporal search in the limited amount of search that is done by the first segment since difference between search cost of primary and secondary stores is relatively low compared to the third segment. The second segment naturally would fall between the first and third segment based on the estimates.
for this segment. It is worth mentioning that the link between store loyalty and promotional frequency has never been addressed in the theoretical or empirical literature. Our structural model with both spatial (store) and time dimensions provides us an opportunity to investigate this link.

To evaluate the link between store loyalty and promotional frequency, we vary promotional frequency symmetrically at two stores following a HILO pricing strategy, keeping average and regular price at the stores constant. HILO stores have a regular price and a promotional price occurring at the chosen promotional frequency. This implies that when promotional frequency increases, a consumer can have more opportunities to obtain discounts, but the discount levels will be smaller. For our analysis 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%.\(^{28}\) We set the travel times to the primary and secondary 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. For loyalty, we report household level share of visits. To gain a better understanding of efficiency of consumer search, we also report average price paid by segment. For profits, we report annual profit per segment and total profits.\(^{29}\)

Figure 5 shows the share of visits to the primary store for the three segments. We find that for each segment an increase in promotional frequency increases the store visit share to the primary store. Note that the greatest change in share of primary store visits happens for the third segment, followed by second and first segments. This was expected since the third segment has the highest price sensitivity and lowest search cost with highest relative difference between effective search cost of primary store and secondary store.

\(^{28}\) That translates into a change in promotion depth from roughly 64% to 18%.

\(^{29}\) To obtain stationary estimates with minimal simulation error, we forward simulate 1000 households over a large number of periods (10,000) and average the metrics across households. We also define the cost to be 60% of the average price (set to be $4.3 in our simulations).
Figure 5. Share of visits to the primary store

<table>
<thead>
<tr>
<th>Promo. Prob.</th>
<th>Seg1</th>
<th>Seg2</th>
<th>Seg3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.25%</td>
<td>$59.11</td>
<td>$51.82</td>
<td>$49.82</td>
<td>$55.41</td>
</tr>
<tr>
<td>7.14%</td>
<td>$59.06</td>
<td>$51.84</td>
<td>$50.06</td>
<td>$55.43</td>
</tr>
<tr>
<td>8.33%</td>
<td>$59.11</td>
<td>$51.98</td>
<td>$50.45</td>
<td>$55.58</td>
</tr>
<tr>
<td>10.00%</td>
<td>$59.29</td>
<td>$52.26</td>
<td>$50.98</td>
<td>$55.86</td>
</tr>
<tr>
<td>12.50%</td>
<td>$59.34</td>
<td>$52.43</td>
<td>$51.41</td>
<td>$56.02</td>
</tr>
<tr>
<td>16.67%</td>
<td>$59.55</td>
<td>$52.72</td>
<td>$52.01</td>
<td>$56.33</td>
</tr>
<tr>
<td>25.00%</td>
<td>$59.51</td>
<td>$52.83</td>
<td>$52.39</td>
<td>$56.41</td>
</tr>
</tbody>
</table>

We next explore how an increase in promotional frequency impacts store profitability. Table 5 reports annual store profit per household as a function of store frequency by segment and in the aggregate. Figure 6 shows the profit per household from each segment, and the aggregate average profit per household across all segments. We find that profit per household increases for all three segments and in the aggregate. Thus, increasing promotional frequency (with correspondingly shallower promotions) leads to increases in profitability in the presence of spatial and temporal promotions.

Interestingly, the profit per household also increases for both stores as promotional frequency increases. Note that we do allow for category expansion in spend per household.
7 Conclusion

This paper introduces a dynamic structural model of search along both the spatial (store) and temporal dimensions allowing for discrete 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 demonstrate that the large literature on search which does not accommodate search on the temporal dimension 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 models temporal search using highly stockpilable 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 elasticities 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 extant research 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. We believe our dynamic structural model of spatiotemporal search would 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.
References


Honka, E and P. Chintagunta (2013). Simultaneous or Sequential? Search Strategies in the U.S. Auto