PRICE SEARCH ACROSS TIME AND ACROSS STORES

By

Navid Mojir and K. Sudhir

March 2014
Revised July 2019

COWLES FOUNDATION DISCUSSION PAPER NO. 1942R

Price Search across Time and across Stores

Navid Mojir

Harvard Business School, nmojir@hbs.edu

K. Sudhir

Yale School of Management, k.sudhir@yale.edu

In response to price dispersion across stores and price promotions over time, consumers search across both stores and time, 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 search across stores and across time that nests a finite horizon model of 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 stockouts and purchase frequency in the milk category. Interestingly, we find that for a given mean and variance in price, increasing the frequency while reducing depth of price promotions across all stores can increase share of visits and profits for consumers' preferred store for the segment with high price sensitivity and low cross-store search cost; consumers who are most prone to search across stores and across time.

Key words: Consumer Search, Price Search, Store search, Spatial search, Temporal search, Spatiotemporal Search, Dynamic Structural Models, MPEC

1. Introduction

Price dispersion across stores and price promotions across time is widespread in retail settings. In response, consumers can search across stores and across time 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 store and time dimensions (Gauri et al. 2008).[1] While the analysis by Gauri et al. was for grocery products involving repeat purchases, search across stores and time is widespread even for one-time purchases. For instance, a potential car or household appliance buyer

---

[1] In Gauri et al. (2008), search across stores is considered as spatial search and search across stores and time is referred to as spatio-temporal search, but the general ideas are relevant even if search is online and consumers have heterogeneous preferences over stores due to retail characteristics or household-specific store preferences which can make certain stores more attractive to search first.

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. With online consumer journeys becoming more visible, researchers can now observe store search and revisits over time. But though search across stores and time are widely pervasive in the real world, there are no theoretical or empirical models that endogenize search on both the store 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 and Lattin 1991, Mehta et al. 2003). Honka (2014) 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) find 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 only price and consideration sets are available, Honka and Chintagunta (2017) develop an identification strategy to distinguish between sequential and simultaneous search.

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 et al. (1984) and Besanko and Winston (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 et al. 1985). For example, Erdem et al. (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 et al. 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 within store price search for various brands, conditional on visiting the store. The model provides a structural "search cost" based framework for the "price consideration" model in Ching et al. (2009). Our paper extends the literature by endogenizing search along both the store and temporal dimensions, but we abstract away from within-store price search.

This paper also contributes to the literature on store choice, particularly in the context of groceries. A key characteristic of grocery choice is that consumers typically shop for a basket of categories, rather than just one category. Like in other contexts outside of groceries, stores may also vary in their product assortments and prices. Bell et al. (1998) consider the role of shopping basket on store choice by breaking down shopping costs into fixed costs (i.e. independent of shopping basket) and variable costs (i.e. dependent on the shopping basket) and develop a model of store visits based on consumers choosing a store with the lowest total shopping cost. They find that consumers prefer Everyday Low Price (EDLP) stores when shopping for larger baskets. Leszczyc et al. (2000) estimate a dynamic

hazard model that takes into account choice of store based on factors such as basket size (stockup or fill-in), and time between trips. Fox et al. (2005) explore shopping behavior across different retail formats and Fox and Hoch (2004) explore cherry-picking behavior among consumers with different shopping and store visit behavior. They conclude that cherry-pickers, on average, save more per item while having larger shopping baskets.

We contribute to this literature by developing a forward-looking dynamic structural model that includes store choice–accounting for household-specific variables (e.g., distance to store), trip-specific variables (e.g., as type of shopping trip—stockup or fill-in, time since last stockup shopping trip) and store characteristics (e.g., EDLP) from the past literature as well as expectations about price promotions. But our focus still is on search within a particular category.

The key intertemporal tradeoff behind the forward-looking structural model is that given (1) price promotions within the focal milk category and (2) factors impacting store choice (including basket-related factors), households not only consider current states of milk inventory, prices, and time since stockup, but also (1) future expectations of milk prices at different stores in the household's consideration set and (2) likelihood of visiting various stores in the future for the overall basket needs. This justifies a forward-looking dynamic structural model of search across stores and across time. The structural model then allows us to answer novel questions about how consumers trade off search across time versus across stores within a category (especially regularly purchased categories with significant spend such as milk and soda). This helps generate insights on how promotion policies may enhance or detract from the behavioral loyalty and profitability of a household's preferred store.

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 prices are high. As we noted earlier, 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. In settings where durable goods prices are negotiated (rather than posted) after visiting stores, prices upon revisits may have temporal dependence (e.g., it may not exceed last negotiated price), but these may be easily incorporated within the modeling framework by appropriately modifying the distribution of price expectations. Finally, we 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. Table 1 summarizes characteristics of structural search models to highlight and clarify relevant features that are incorporated into the models.

Table 1   Selected Empirical Structural Models of Search

Paper	Product Category	Online/ Offline	Endogenous Store Search	Store Visit Basket Effect	Within Store Search	Temporal Search
One-time purchase sequences						
Kim et al. (2010)	Camcorders	Online	Sequential	NA	Yes	No
Honka (2014)	Auto Insurance	Offline	Simultaneous	NA	NA	No
Ursu (2018)	Hotels	Online	Sequential	NA	Yes	No
Repeat Purchase Sequence						
Erdem et al. (2003)	Ketchup	Offline	NA	No	No	Yes
Hendel and Nevo (2006)	Detergent	Offline	NA	No	No	Yes
Hartmann and Nair (2010)	Razor & Blades	Offline	No	Exogenous	No	Yes
Seiler (2013)	Detergent	Offline	NA	No	Yes	Yes
This paper	Milk	Offline	Sequential	Yes	No	Yes

We estimate the dynamic structural model allowing for discrete heterogeneity. Given that we model visits and purchases (not just purchases as in extant models), the number of events included in each households 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 households' 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 spending among groceries and high frequency of purchases, milk is an ideal category for studying price search across stores and time. We also observe sufficient variation across weeks in stores visited, factors outside the milk category that impact store choice (e.g., basket size) and milk purchases at a store conditional on visiting a store to aid model identification.

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 (41%) has high cost of store 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 (40%) has relatively low search cost for its primary store, hence visits the preferred store often, and prefers to shop during weekdays. The third segment (19%) has the lowest search cost; this segment searches both temporally and across stores and obtains the lowest prices. The implicit search costs for a visit to a store varies from $3.23 for the low search cost segment to $29.08 for the high search cost segment.

Second, not accounting for either the store or time dimension of search leads to underestimated price elasticities. The direction of the bias in estimates from omitting the temporal dimension is opposite to what has been reported in the literature (e.g., Erdem et al. 2003; Hendel and Nevo 2006). We provide an explanation for this difference based on the fact that the previous literature has 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 context of our model.

Finally, we use our model to study the effect of increasing promotion frequency on consumer store visit and purchase behavior. We find that while one segment of consumers do not respond to more frequent promotions (the first segment with low price sensitivity

and high search cost), a second segment of consumers who search across time respond to such an increase by timing their purchases (cherry-picking) and taking advantage of price promotions more. This results in reduced paid prices and profits from this segment. An interesting phenomenon happens for a third segment of consumers who search both across time and across stores. For this segment of consumers, increased promotion frequency results in less search across stores (hence increased share of visits to consumer's favorite store). From the perspective of the primary store, the increased share of store visits overpowers the timing of purchases over time and results in higher paid prices and profits from this segment of consumers. Therefore we conclude that increased promotion frequency does not necessarily result in increased cherry-picking and reduced profits. In fact, the result of such policy would depend on consumers' search strategy and varies across different consumer segments.

The rest of the paper is organized as follows: Section 2 describes the data and the model-free evidence in support of the modeling assumptions. Section 3 describes the model and Section 4 describes the estimation. 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 (behavioral) store loyalty. Section 7 concludes.

2. Data and Model-Free Evidence

We begin with a description of the data and descriptive statistics. We then provide model-free evidence in support of the primary features of the model—search across stores and search across time.

2.1. 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.[2] 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 and travel time between each household

---

[2] Researchers own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

and each store in its consideration set. We complement this data with Retail Scanner Data from Nielsen, to construct the weekly prices at the relevant stores.

We use milk as our focal category. Milk is an ideal category for modeling search in the grocery category due to its high spend and purchase frequency. It has the second highest spend ($80 with a 3.39% basket share) after soda ($117 with a basket share of 4.81%), and 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 households that purchase in the category typically consume it daily, and it has limited shelf-life.

We construct our sample of households from the panel data as follows. First we drop households who do not shop frequently (fewer than 20 shopping trips over the year across all stores) or do not purchase milk frequently (less than 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 or fewer stores across the entire period of the data. We exclude the 6% of households who shop at more than 3 stores.

We discuss three other features of the data that inform our modeling choices. First, the vast majority (84%) of purchases are private labels at the store and the median share of private brands across all households is 83% with many households buying only private labels. Further, most of the households who make non-private label purchases could be considered in one of two groups; One group consistently buys the same national brand–which means there is no brand choice. A second group mixes private labels and national brands, but they buy the national brand almost always when it is on promotion, and thus when it is priced lower than the private label. A Pearson chi-square test rejects the independence of promotion and choice of non-private label (p-value, 0.000). Collectively, all of these groups represent 91.5% of households in the data.[3]

Given these features of the data, we abstract away from modeling brand choice, as it adds limited empirical value in the category. For those households who exclusively buy private labels across stores, we simply track category purchases and use prices of the private label

---

[3] As a point of comparison, we note that the 91.5% cumulative share of included households is comparable to the cumulative shares accounted in Hendel and Nevo (2006) who endogenize brand choice, but not store choice in their study of laundry detergents (liquid: 93% and powder: 90%).

at each corresponding store. For those households who buy exclusively a national brand, we simply track category purchases and use the prices of the relevant national brand. Finally, for those who mix between national and private brands based on the lowest price, the relevant price is that of the lowest priced product in their consideration set for each period. The relevant price distribution for each is also constructed—for the relevant loyal or cheapest brand—as appropriate.[4]

Second, households in the category are largely size-loyal with median share of the preferred size at 93%. Given the extent of size loyalty among households in the category, it makes little sense to model size choice at the household level–because the i.i.d. error assumption across sizes in the logit would produce worse fit with the data than by simply not modeling size choice. Among the high frequency and high volume households (over 20 store visits with at least 5% of trips including milk purchase), the preferred size was 1 gallon for most of households, with its median share at 81%. We restrict our analysis to the 1 gallon households.[5] We also dropped 5% of the high frequency and high volume households that occasionally purchased more than 1 unit. While our model can conceptually accommodate multi-unit purchases, the restriction to single unit purchases helps to mitigate computational complexity from an expanded state space.

While it is easy to obtain the price for the product that is purchased from the database, we need to impute the right price considered by a household at the store when the household does not purchase milk. Appendix D provides details on construction of price data. We dropped a small number of households who shopped at stores for which we do not have any price data. In all, we use panel data on 948 households in our estimation. Given that we expect search behavior to differ across weekdays and weekends, we split the week into a weekend period (Friday-Sunday) and weekday period (Monday-Thursday). Thus the temporal unit of analysis is $\frac{1}{2}$ week.

---

[4] For the remaining 8.5% of households, we use the price of their most-bought brand to form the price distribution.

[5] Seiler (2013) endogenizes search within the store and uses loyalty to brands and sizes even in the presence of lower prices to identify within store "price search" cost across households; however he notes that the low share products market share are consistently over-predicted, even if the search cost is modeled. Such overprediction would be even greater at the household level for those very loyal to a particular size. Given the trade-offs in fitting the data, and our research focus on across store and across time search, and the high levels of size loyalty, we abstract away from endogenizing within store price search for sizes and restrict our analysis to category choice on the preferred size and brand.

2.2.   Model-Free Evidence

We begin by demonstrating model-free evidence for the key premise of search across stores and across time.

Price Variation across Stores and Across Time: We begin with some metrics about the extent to which prices vary across stores and across time to justify price search.

To measure the extent of cross-store price variance observed by households in a given period, we compute the coefficient of variation (CV) for each household across stores within their consideration set at each time period The average of these CV is 19%.

To measure the cross-time price variance within stores experienced by households, we compute the CV of prices at each store. The average CV is 5.3%. This number is 5.6% for HiLo stores and 2.3% for EDLP stores.

Finally, we note that there is little seasonality in demand for milk; hence the price variation reported above is not driven by seasonality. Overall, the reported metrics of CVs of prices indicate potential gain for households from performing price search across stores or time. Further, note that there is variance around the average CVs, so individual households differ in their incentives for search—which of course provides identifying price variation for estimation of the search model.

Price Search across Stores: We build up the evidence for price search across stores for milk through the following series of five analysis: (1) households shop at multiple stores within a time period, (2) households buy milk at different stores, (3) milk is not always bought at the first store visited in a period, (4) milk inventory impacts probability of visiting multiple stores, and (5) milk inventory impacts probability of visiting the store with lowest average milk price.

Figure 1 shows the household level distribution of share of periods in which multiple stores are visited, conditional on visiting at least one store in the period. Clearly, a large number of households visit multiple stores within the same weekday or weekend period.

However, multi-store visits need not reflect store search for milk. Even if a household visited multiple stores to satisfy their basket needs, they could always have bought milk at one store–indicating no search across stores. Figure 2 presents distribution of purchases of milk from household's "favorite store" (i.e. store from which consumer has purchased the item most often). Many households buy milk from multiple stores and are not tied into a "preferred" store for milk.

Figure 1   Distribution of share of multi-store periods across households

(i.e. periods during which a household visits more than one store)
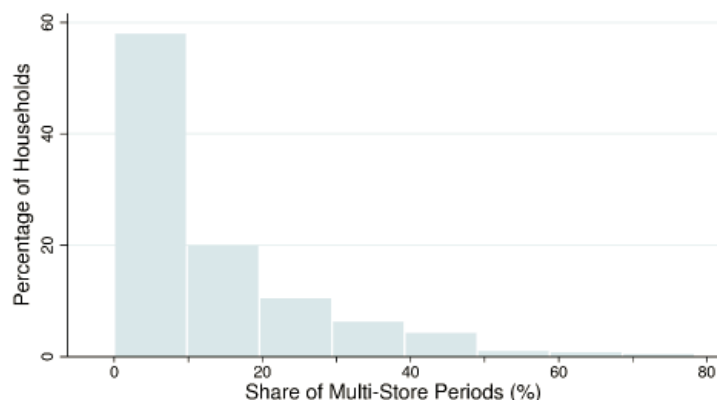
Figure 2   Distribution of share of "favorite store" in the milk category

from total category purchases across households

One possibility is that even if milk may be bought at different stores, it may always be at the first store visited in that period. Figure 3 shows the distribution of probability of purchasing milk from the first store visited during any period conditional on visiting multiple stores in the same time period. As the figure suggests, only 10% of households have a probability of 90% or more to make their purchase at the first store that they visit during any given period. Most households purchase milk at the second or third store during the same period. This is suggestive evidence of cross-store search.

To specifically assess whether milk drives multiple store visits (and thus cross-store search), we estimated a logistic regression where we model the probability of visiting multiple stores as a function of milk inventory level, controlling for inventory levels of other categories by using "time since last stockup trip" as a proxy and household heterogeneity

by including household fixed effects.[6] The coefficient of milk inventory is negative and significant ($p < 0.01$); suggesting that an increase in milk inventory decreases the probability of visiting multiple stores in the same period.

Finally, to check whether milk price distribution across stores affects search, we estimate a similar logistic regression but with visiting the store with "cheapest milk" as the dependent variable. Again, the coefficient on milk inventory is negative and significant ($p < 0.01$); suggesting that the higher the milk inventory, the less likely the household will visit the store with the lowest milk prices, controlling for the influence of other categories.

The results of both regressions are presented in Appendix B. While the results presented are for regressions across all time periods, we get qualitatively identical results if we condition on periods when there is at least one store visit. We therefore conclude that milk prices influence store search decisions.

Figure 3   Distribution of probability of purchasing milk from the first store visited in multi-store periods conditional on purchasing milk
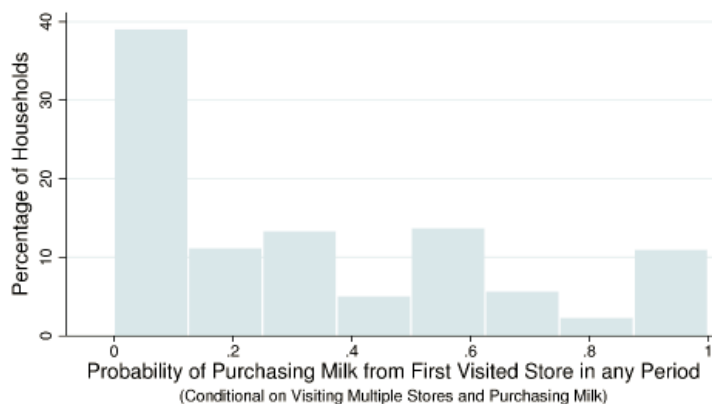
Search across Time: To study whether consumers adjust purchase timing in response to milk promotions we test the differences in inter-purchase times between milk purchases as a function of whether milk is purchased on promotion or not. The idea is that consumers accelerate their purchases when there is a promotion before consuming their current inventory as demonstrated in the early work of Neslin et al. (1985) and Hendel and Nevo (2006). Given that milk is a perishable item that can be stockpiled only for short periods, it is an empirical question as to whether purchase acceleration is likely in the milk category.

---

[6] 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. The result is robust to different initial random values.

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 at the regular price, 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.[7]

3. The Model

We develop a household model of repeat purchase for a non-durable good. We allow the household to hold inventory and consume over time, thus decoupling purchase and consumption. The model embeds endogenous price search across stores (taking into account the household's benefit relative to search cost from visiting a store) and across time (by timing the purchase when prices are low so as to balance stockpiling and stockout costs) to take advantage of periodic price promotions.[8] Given that the model is applied to a grocery category, we model the fact that store visits are driven not only by the focal category, but also by the basket of groceries that needs to be purchased.

As discussed in the introduction, the intertemporal tradeoff underlying the forward-looking model is that it takes into account not just the current states (e.g., price, inventory, time since stockup), but also (1) future expectations of milk prices at different stores in the household's consideration set and (2) likelihood of visiting various stores in the future for overall basket needs. We next present the details of the model of store visits, purchases and consumption over time.

3.1. The Basic Set Up

At each time period t, a household h can search across a finite consideration set of stores denoted by Ω_h.[9] The choice of time period can be flexible depending on the model setting.

---

[7] 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 of more than 12 periods (one and a half months) between purchases, as these few outlier households disproportionately impact the duration between purchases relative to the large number of households making regular purchases.

[8] Given the features of the milk category that we described in the last section, we model choice at the category level and abstract away from brand choice. For categories where brand choice is important, the model can be easily extended to incorporate brand choice. However, the model would be computationally more difficult to estimate.

[9] We assume that the consideration set for each household is constant and does not change over time. Furthermore, as we will see in the estimation section, in our implementation, all the model parameters are assumed to be heterogeneous up to a finite number of segments.

We use a half week period in our empirical analysis due to the grocery setting, as very few people revisit the same store within the same half week, but we could model shorter periods as appropriate if we need to model online purchase journeys. Let N_h^{max} be the maximum number of stores in Ω_h; then, there is potentially a maximum of N_h^{max} stages of store search in any given period t until all the stores in the consideration set Ω_h are exhausted. Note however, that a household may stop search within a period after visiting $0 \leq n \leq N_h^{max}$ stores. 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 4 represents one time period of store search. Each time period consists of potentially N_h^{max} different stages of search. Each stage involves two decisions by the household: a store visit decision and a category purchase decision.

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 as to maximize the household's value function across the remaining stages in period t and across future time periods.[10] A household that decides to visit another store k moves to the purchase decision at stage (t, n). 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., $(t + 1, 1)$.

Purchase Decision (t, n): When at store k from the set of unvisited stores Ω_{htn}, the household observes purchase-related state variables x_{htk}^p (this includes store-specific variables for the kth store) 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. After the decision of whether to make a purchase or not in the focal category, the household moves to the next visit decision which would include remaining, unvisited, stores in the consideration set, unless all the stores in the consideration set have been visited, in which case, the household moves to the next time period.

---

[10] While our framework does not allow the same store to be revisited in the same time period, a household can go back to the same store in the next time period. If we make the time period small enough, this can practically accommodate store revisits. In our empirical analysis, we treat time period as 1/2 week. The median probability of revisiting the same store conditional on visiting at least one store in the half week across the households in our sample is 6.1%. Therefore, we consider that our time period is short enough that store revisits in the same time period should not have significant effect on the estimates.

Figure 4   Schematic of model at period t
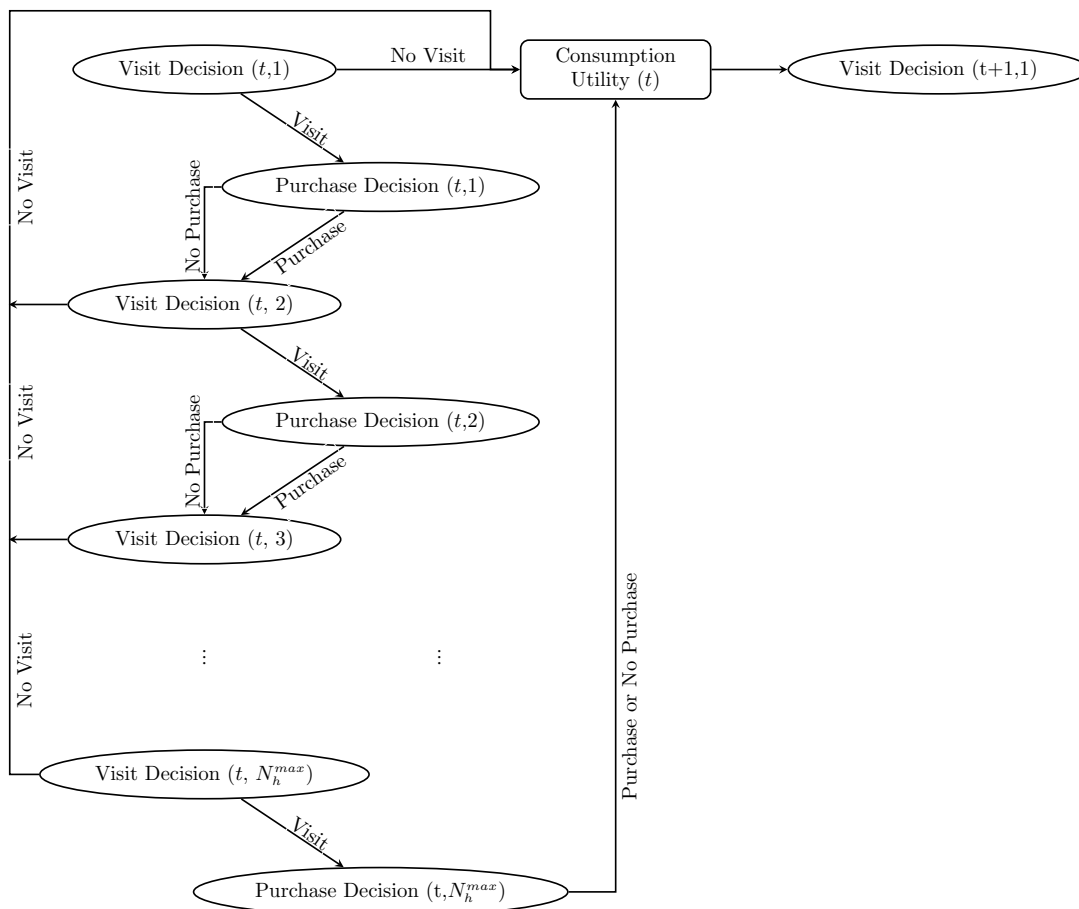
Notes: At each visit stage, consumers have already observed visit-related state variables and are aware of price distribution at all stores in their consideration set, but they do not know realized prices. At each purchase stage, consumers know the realized price at the store that is being visited and also the observed state variables for the next visit stage. More details about the specific state variables for visit and purchase are discussed below in the visit and purchase flow utilities.

Note that each household gets the utility from consumption of the focal category 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.

To summarize, a household $h \in \{1, 2, \ldots, H\}$ at time period $t \in \{1, 2, 3, \ldots\}$ and store search stage $n \in \{1, 2, \ldots, N_h^{max}\}$, 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

$y^v_{htn} \in \Omega_{htn} \cup \{0\}$, where $y^v_{htn} = 0$ represents a decision to stop search for period t at stage n. Conditional on visiting store k from the set of unvisited stores Ω_{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.

3.2. Flow Utilities

Visit Decision

We begin with the flow utility (i.e., the immediate utility) from visit and purchase at stage n. There are three sets of factors that influence the decision of which store to visit and when to make the visit. First, there are store-specific characteristics that could make a store more or less attractive. These include observed factors (e.g. store format) and unobserved factors (e.g. proximity to work). Second, travel costs which would depend on geographical distance or travel time for a store and also opportunity cost of consumer's time at the time of visit. For example, 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. Third, consumer shopping needs or equivalently, level of inventory of different categories in the consumer's shopping basket. We capture these factors in our specification of the utility of visiting store $k \in \Omega_{htn}$ at stage (t, n) as follows:

$$u^v_{htnk}(x^v_{ht}) = X_{hk}\beta_h - S_h(d_{hk}, W_t) + \eta T^{stockup}_{htn} + \varepsilon^v_{htnk} \qquad for \ \ k > 0 \qquad (1)$$
$$= \bar{u}^v_{htnk}(x^v_{ht}) + \varepsilon^v_{htnk}$$

Where the first term indicates preferences for store characteristics. In our implementation, we include two store characteristics in X_{hk} to account for store differentiation: (1) Whether store k is EDLP[11] and (2) Whether store k is the primary grocery store for household h, where we operationalize the primary store as the one which has the highest share of visits in the household's consideration set.[12]

---

[11] Note that the effect of store pricing policy is structurally captured (partially) by price expectations for the focal category. The coefficient on EDLP captures other aspects of preference that includes the influence of pricing policy of other, non-focal, categories among other factors.

[12] Given share of visits will be affected by price promotions, one concern is whether the definition of primary store is endogenous. In practice, visits to primary stores are substantially larger than for secondary stores; hence the classification is very robust and not very sensitive to pricing variation.

The second term $S_h(d_{hk}, W_t)$ is the travel cost incurred by household h to visit store $k \in \Omega_{htn}$. We interchangeably refer to this cost as store visit cost or search cost (in line with the search literature). While this cost needs to be incurred to learn about milk prices, it also allows consumer to search for prices and buy in other categories. This is a critical insight in modeling category level search behavior in the grocery sector, where purchases often are for a basket of goods. Here d_{hk} is defined as the travel time to store k for household h.[13] We consider search cost to be a linear function: $S_h(d_{hk}, W_t) = \iota_h + \delta_h d_{hk} + \omega_h W_t$, where W_t is a dummy variable coded as 1 if time period t is a weekend, 0 otherwise. This weekend dummy variable allows us to account for difference in opportunity cost of consumers' time during weekends and weekdays.

The third term, $\eta T_{htn}^{stockup}$ the number of time periods since last stockup, summarizes the effect of a household's basket of grocery needs on store visits in a parsimonious manner. The intuition is that the longer it has been since a stockup visit by the household, the more likely the household would need to visit stores due to depletion of inventory in the non-focal categories independent of the inventory in the focal category.[14] We define a trip as a stockup trip if total spending (excluding milk) of household h in that trip is higher than average spend (excluding milk) for that household across all their trips. We exclude milk from the definition of stockup so that it captures only the effect of categories other than milk on store visits. Note that in this case we are assuming the base utility for visiting a store is zero right after a stockup visit and that utility increases with the number of periods since last stockup trip. This is to capture the increase in household's willingness to make a visit due to depletion of inventory of the basket of non-focal categories.[15] Finally,

---

[13] The travel time is from the center of the zip code area for each household's residence to each store as we do not observe the workplace address for households' members). The center of zipcode approximation is common in the literature as household addresses are withheld for privacy reasons. (e.g. Gauri et al. (2008), Bell et al. (1998), and Briesch et al. (2009)).We measured travel time using Google maps Application Programming Interface (API). Final estimates presented in the paper are based on the square root of travel times as these results are robust to distance outliers.

[14] Note that we have not included the effects of focal category inventory in the flow utility for visits directly. This effect will impact visits indirectly through its effect on the choice-specific value function by impacting consumption utility (when there is not adequate stock) in the dynamic structural framework.

[15] A household is more likely to visit its primary store for stockup trips. In our data, the share of primary store for stockup trips is 63%. Given our subsequent assumption that the unobserved error shocks are distributed type I extreme value, the resulting logit functional form of visit probability can potentially accommodate a positive interaction effect between time since last stockup and visit to the primary store. If we have positive coefficients for "time since last stockup visit" and primary store, then in the convex region of the curve, the increase in the probability of visiting the primary store would be higher than that of non-primary stores, the longer it has been since last stockup visit. We verified this later based on the estimates of the structural model.

ε^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.

A household that forgoes search obtains the following utility:

$$u^v_{htn0}(x^v_{ht}) = \varepsilon^v_{htn0} \tag{2}$$

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:

$$\begin{aligned} u^p_{htnk1}(x^p_{htk}) &= \alpha_h p_{kt} + \varepsilon^p_{htnk1} \\ &= \bar{u}^p_{htk1}(x^p_{htk}) + \varepsilon^p_{htnk1} \end{aligned} \tag{3}$$

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^p_{htnk1} 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:

$$u^p_{htnk0}(x^p_{htk}) = \varepsilon^p_{htnk0} \tag{4}$$

The structural error shocks in the above equations are all assumed to be independent and identically distributed (i.i.d) type I extreme value.[16]

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:

$$u^c_{ht}(i) = \varphi_h(c(i)) \tag{5}$$

where i represents the inventory level of the focal category at the end of all visits at time t (i.e. after taking into account potential increases in the inventory because of purchases).

---

[16] We note that purchase and visit error shocks across different stores for the same household at the same time period can be potentially correlated. We assume the standard conditional independence for computational reasons.

$c(i)$ is consumption as a function of inventory level and φ_h is the utility of consuming $c(i)$ 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.

Let ρ_h be the household h's per-period consumption rate of the focal category. Specifically, we assume consumption to be either the per period consumption if there is sufficient inventory or zero (we set fractional consumption to zero limit the state space):

$$c(i) = \begin{cases} 0 & i < \rho_h \\ \rho_h & i \geq \rho_h \end{cases}$$

This means the household consumes an amount equal to the consumption rate if there is more than one serving left in the inventory and consumes zero otherwise. This specification allows us to capture a drop in the consumption utility when a household does not have adequate inventory. The assumption that the consumption utility from a fractional serving is zero (rather than linear) is an approximation that helps us model the inventory state space in terms of number of servings, which is an integer, rather than as a continuous variable.

Further, as described in the inventory transition equation below, that inventory is perishable after T periods; this accommodates a nonlinear stockpiling (inventory) cost—the entire value of inventory not consumed within T periods is lost to the household.[17]

We assume that utility of consumption is $\varphi_h(c(i)) = \sigma_h c(i) + \tau_h$, where σ_h and τ_h are parameters to be estimated. Thus when there is not enough inventory for a household serving, i.e., $i < \rho_h$, the utility will be τ_h, the disutility from stockout.

State Transitions

Here we define the state transitions associated with inventory, prices, time since last stockup visit, weekday/weekend, and store consideration sets.

We define the inventory for each household h, at each time period t, and at the end of each stage of search n to be represented by i_{htn}. We need to define the transition of inventory across both time periods and also across different stages of search within the same time period.

First, transition of inventory across time periods; At the end of the first stage of each time period, the level of inventory equals that of the end of the last stage of the previous time

---

[17] A linear in time inventory cost over the short time period of three weeks is not identified; hence our approach to model a highly nonlinear inventory cost is more appropriate for perishable products.

period, represented by $N_h(t-1)$, plus any additions that might come from any purchase that might have occurred in the first stage of search. That is:

$$i_{ht1} = \min(i_{ht-1N_t(t-1)} + y^p_{ht1}q, \ T \times \rho_h)$$

Here, q represents the quantity purchased (i.e. milk container size in our application). The min operator is used to account for perishability of milk and impose a non-linear inventory cost (i.e. we assume that the inventory is perishable after T periods).

Second, transition of inventory across search stages within the same time period; In this case, the inventory at the end of each stage equals the inventory at the end of the previous stage plus any changes due to a purchase at the current stage. If the current stage is the last stage of search for the period, we need to account for the reduction in the inventory level due to consumption. That is:

$$i_{htn} = \min(\underbrace{i_{htn-1}}_{\text{carryover}} + \underbrace{y^p_{htn}q}_{\substack{\text{purchased} \\ \text{amount}}} - \underbrace{I\{n = N_h(t)\}c(i_{htn-1} + y^p_{htn}q)}_{\text{consumed amount}}, \quad \underbrace{T \times \rho_h}_{\substack{\text{maximum storable} \\ \text{inventory}}} \quad)$$

Here, $I\{n = N_h(t)\}$ is an indicator function that equals one if the current stage is the last stage of search in the period.

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.[18] We assume a store-specific multinomial distribution of prices over the m price levels over time. While many of the papers in the literature that do not model store search assume that prices, while exogenous to demand shocks, follow a Markov process, we assume consumer price expectations are independent across time. Formally: $p_{kt} \sim Multinomial(1, \vec{p_k})$.[19] This temporal independence assumption is driven by both conceptual and practical issues. At a conceptual level, consumers have to observe prices at each store in order for a Markov

---

[18] To be more precise, we have a different distribution for each household-store combination. That is because the relevant price for different households shopping from the same store might be different (e.g. one household might always buy a national brand while another always buys the store brand). Note that the price distribution assumption can be used flexibly for different settings. For instance, when we consider a negotiated price situation like cars, we could model price distribution to be at or below the negotiated price for a fixed period of time, that the dealer might honor the price.

[19] The price exogeneity assumption is common in the dynamic structural modeling literature; see Erdem et al. (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. (2015) 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. We follow the literature in assuming price exogeneity. Note that while consumers know the parameters of the distribution, $\vec{p_k}$, before visiting a store, they do not know the realized price before making a visit.

assumption on price expectations to work. Since consumers only visit a subset of stores (or none at all) in any given period, in practice a Markovian assumption on the evolution of price expectations will not be a clearly superior assumption compared to the independence assumption. In theory, we can consider consumers to have price expectations that are independent from past prices, when the past price for a store is not observed, and have Markovian expectations, when the past price is observed. In practice, however, that would lead to an explosion of the state space, since we would need to keep track of not only past prices for each of the stores, but also, stores that have or have not been visited by the household in the previous period. We also abstract away from potential correlations of prices of different stores in a household's consideration set since we find it to be generally small and statistically insignificant for more than half of the households in our sample.

Let $T_{htn}^{stockup}$ represent time periods since last stockup visit at the beginning of stage n. We assume that probability of stockup visit (i.e. θ_h) is a function of time since last stockup visit[20] Specifically, across search stages within the same time period:

$$T_{htn+1}^{stockup} = \begin{cases} 0 & \text{probability: } \theta_h(T_{htn}^{stockup}) \\ T_{htn}^{stockup} & \text{probability: } 1 - \theta_h(T_{htn}^{stockup}) \end{cases}$$

And across time periods ($N_h(t)$ represents the latest stage that a visit decision is considered at time t):

$$T_{h(t+1)1}^{stockup} = \begin{cases} 1 & \text{probability: } \theta_h(T_{htN_h(t)}^{stockup}) \\ T_{htN_h(t)}^{stockup} + 1 & \text{probability: } 1 - \theta_h(T_{htN_h(t)}^{stockup}) \end{cases}$$

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: $W_1 = 0$ and $W_t = 1 - W_{t-1}$.

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

3.3. 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

---

[20] The function θ is empirically calibrated nonparametrically over the finite range of positive integer values in the data. It turns out to be monotonically nondecreasing function of $T_{htn}^{stockup}$. We set it to zero outside the range observed in the data.

defined in previous section, we can write the optimization problem as a sequence problem of visit and purchase decisions for each household h,

$$\max_{\{\Delta_{ht}\}_{t=0}^{\infty}} \mathbb{E}\left(\sum_{t=0}^{\infty} \beta^t \varpi_{ht}(\Delta_{ht})|\tilde{x}_{ht}\right) \tag{6}$$

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 ($y_{ht}^p = \{y_{htn}^p\}_{n=1}^{N_h(t)}$) decisions ($N_h(t)$ represents the latest stage that a visit decision is considered at time period t). 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 includes all the relevant observed state variables for the purchase state, while $\varepsilon_{ht}^v = \{\varepsilon_{htn0}^v, \{\varepsilon_{htnk}^v\}_{k\in\Omega_h\backslash\Omega_{htn}}\}_{n=1}^{n=N_h(t)}$ and $\varepsilon_{ht}^p = \{\{\varepsilon_{htnk1}^p, \varepsilon_{htnk0}^p\}_{k=y_{htn}^v}\}_{n=1}^{N_h(t)}$ 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 stage $N_h(t)$ plus consumption utility:

$$\varpi(\Delta_{ht}) = \underbrace{\sum_{n=1}^{N_h(t)}\prod_{l=0}^{N_h^{max}}(u_{htnl}^v)^{\mathbb{I}(y_{htn}^v=l)}}_{\text{visit utility}} + \underbrace{\sum_{n=1}^{N_h(t)}\prod_{l=0}^{1}(u_{htnkl}^p)^{\mathbb{I}(y_{htn}^p=l)}}_{\text{purchase utility}} + \underbrace{u_{ht}^c}_{\substack{\text{consumption}\\\text{utility}}} \tag{7}$$

3.4. Choice-Specific Value Functions

Within the finite horizon store 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 $\mathbb{E}V_{htn}^v(x_{ht}^v, \Omega_{htn})$ 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 $\mathbb{E}V_{htnk}^p$ 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 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, or (2) to wait for stage $(n+1)$ and consider visiting an

unvisited store from the set of remaining stores $\Omega_{htn} - \{k\}$, or to wait till next time period if the set of unvisited stores is empty. If it is the last stage of search (i.e. $n = N_h^{max}$), with a purchase, the household gets the corresponding flow utility of purchase plus the utility from consumption of the focal category and the discounted value of utility (across time) that she will get starting next period. On the other hand, if there are still more stores left to be visited (i.e. $n \neq N_h^{max}$), the household gets the purchase utility plus the expected value of utility that she will get starting next search stage. Since this is in the same time period, there will be no discounting. Also, note that to avoid double-counting, we do not add consumption utility here.

$$
v_{htnk1}^p(x_{htk}^p, x_{ht}^v, \Omega_{htn}) = \begin{cases} \bar{u}_{htnk1}^p + \mathbb{E}V_{ht,n+1}^v(x_{ht}^v, \Omega_{htn} - \{k\}) + \varepsilon_{htnk1}^p & \text{if } n \neq N_h^{max} \\ \bar{u}_{htnk1}^p + u_{ht}^c + \\ \beta\mathbb{E}_{x_{h,t+1}^v | x_{ht}^v, \Delta_{ht}, \varepsilon_{ht}}\mathbb{E}V_{h,t+1,1}^v(x_{h,t+1}^v, \Omega_h) + \varepsilon_{htnk1}^p & \text{if } n = N_h^{max} \end{cases}
$$
$$
= \bar{v}_{htk1}^p + \varepsilon_{htnk1}^p \tag{8}
$$

If the household does not purchase, and it is the last stage of search in the current time period, the household receives the utility from consumption plus the discounted value of expected utilities, starting from next time period. If there are still unvisited stores in the current time period, the household would get the expected value of stream of utilities from potential visits and purchases from those stores and future time periods. That would be captured in $\mathbb{E}V_{htn+1}^v$.

$$
v_{htnk0}^p(x_{htk}^p, x_{ht}^v, \Omega_{htn}) = \begin{cases} \mathbb{E}V_{htn+1}^v(x_{ht}^v, \Omega_{htn} - \{k\}) + \varepsilon_{htnk0}^p & \text{if } n \neq N_h^{max} \\ u_{ht}^c + \beta\mathbb{E}_{x_{h,t+1}^v | x_{ht}^v, \Delta_{ht}, \varepsilon_{ht}}\mathbb{E}V_{h,t+1,1}^v(x_{h,t+1}^v, \Omega_h) + \varepsilon_{htnk0}^p & \text{if } n = N_h^{max} \end{cases}
$$
$$
\tag{9}
$$

Moving one step back, the household faces a decision of whether to visit a store and which store to visit. At this point, the 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 unvisited stores yet (e.g., does not know prices before visiting the store). Therefore, the household should use the expected value of the utility for what comes next in making the decision whether and which store to visit. If the household

decides to visit one of the unvisited stores, the next step would be making a purchase decision. Therefore, we can write the choice-specific value function for the visit stage as:

$$v_{htnk}^v(x_{ht}^v, \Omega_{htn}) = \bar{u}_{htnk}^v + \mathbb{E}V_{htnk}^p(x_{ht}^v, \Omega_{htn} - \{k\}) + \varepsilon_{htnk}^v \tag{10}$$

$$= \bar{v}_{htnk}^v + \varepsilon_{htnk}^v$$

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

$$v_{htn0}^v(x_{ht}^v) = \bar{u}_{ht}^c + \beta \mathbb{E}_{x_{h,t+1}^v | x_{ht}^v, \Delta_{ht}, \varepsilon_{ht}} \mathbb{E}V_{h,t+1,1}^v(x_{h,t+1}^v, \Omega_h) + \varepsilon_{htn0}^v \tag{11}$$

$$= \bar{v}_{htn0}^v + \varepsilon_{htn0}^v$$

3.5.   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 $V_{htn}^v(x_{ht}^v, \Omega_{htn}) = \max_{k \in \Omega_{htn} \cup \{0\}} \{v_{htnk}^v(x_{ht}^v, \Omega_{htn})\}$ as value function of search stage, the ex-ante value function at the visit stage is given by,

$$\mathbb{E}V_{htn}^v(x_{ht}^v, \Omega_{htn}) = \mathbb{E}_{\epsilon_{htn}^v | x_{ht}^v, \Omega_{htn}, \epsilon_{htn-1}^v} \left[\max_{k \in \Omega_{htn} \cup \{0\}} \{v_{htnk}^v(x_{ht}^v, \Omega_{htn})\} \right] \tag{12}$$

$$= log \left[\sum_{k \in \Omega_{htn} \cup \{0\}} exp(\bar{v}_{htnk}^v(x_{ht}^v, \Omega_{htn})) \right]$$

where $\epsilon_{htn}^v = \{\varepsilon_{htnk}^v\}_{k \in \Omega_{htn} \cup \{0\}}$. 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,

$$\mathbb{E}V_{htnk}^p(x_{ht}^v, \Omega_{htn}) = \mathbb{E}_{x_{htk}^p, \varepsilon_{htnk1}^p, \varepsilon_{htnk0}^p | \varepsilon_{ht,n-1,k1}^p, \varepsilon_{ht,n-1,k0}^p} \{max[v_{htnk1}^p(x_{ht}^v, \Omega_{htn}), v_{htnk0}^p(x_{ht}^v, \Omega_{htn})]\}$$

$$\tag{13}$$

$$= \int_{x_{htk}^p} log[exp(\bar{v}_{htk1}^p) + exp(\bar{v}_{htn0}^p)]dP(x_{htk}^p)$$

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

3.6. Choice Probabilities and the 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:

$$P_{htnk}^v = \frac{exp(\bar{v}_{htnk}^v)}{\sum\limits_{j \in \Omega_{htn} \cup \{0\}} exp(\bar{v}_{htnj}^v)} \tag{14}$$

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:

$$P_{htnk1}^p = \frac{exp(\bar{v}_{htk1}^p)}{exp(\bar{v}_{htk1}^p) + exp(\bar{v}_{htn0}^p)} \tag{15}$$

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 households observed decision, the likelihood for household h conditional on being from segment g can be written as,

$$L_{h|g} = \prod_{t=1}^{T_h} \prod_{n=1}^{N_h(t)} \prod_{k=0}^{N_h^{max}} (P_{htnk}^v|g)^{\mathbb{I}(y_{htn}^v=k)} (P_{htnk1}^p|g)^{\mathbb{I}(y_{htnk}^p=1 \& y_{htn}^v=k)} (1 - P_{htnk1}^p|g)^{\mathbb{I}(y_{htnk}^p=0 \& y_{htn}^v=k)} \tag{16}$$

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

$$L = \prod_{h=1}^{H} \left(\sum_{g=1}^{G} p_g L_{h|g} \right) \tag{17}$$

4. Estimation

We formulate the estimation problem of the dynamic programming model as a Mathematical Program with Equilibrium Constraints (Su and 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 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.[21] With heterogeneity, the log likelihood function 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 17, $L_{h|g}$ is a function of choice-specific value functions of the model. In fact, this equation could be re-written as:

$$L = \prod_{h=1}^{H} \underbrace{\left(\sum_{g=1}^{G} p_g L_{h|g}(\bar{v}_h^p, \bar{v}_h^v; x_h^v, x_h^p, \Delta_h, \Theta) \right)}_{L_h} \tag{18}$$

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. Specifically, the finite horizon nested within an infinite horizon structure of our problem results in a system of Bellman equations that is computationally challenging in each iteration of the contraction mapping. Therefore, instead of using NFXP, we re-formulate the problem as a constrained optimization problem. To that end, we re-write the likelihood function as a function of

---

[21] 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 a factor of up to 6 depending on the maximum number of stores in households' consideration sets.

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

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

subject to:

$$\mathbb{E}V_{htn}^v(x_{ht}^v, \Omega_{htn}) = \log \left(\sum_{k \in \Omega_{htn} \cup \{0\}} \exp(\bar{\upsilon}_{htnk}^v) \right) \qquad \forall t \in \{1, \ldots, T_h\}, \forall n \in \{1, \ldots, N_h\}$$

$$\mathbb{E}V_{htnk}^p = \int_{x_{htnk}^p} \log[\exp(\bar{\upsilon}_{htk1}^p) + \exp(\bar{\upsilon}_{htn0}^p)]dP(x_{htk}^p) \qquad \forall t \in \{1, \ldots, T_h\}, \forall n, k \in \{1, \ldots, N_h\}$$

$$(19)$$

where $\mathbb{E}V_h^v = \{\{\mathbb{E}V_{htn}^v\}_{t=1}^{T_h}\}_{n=1}^{N_h^{max}}$ and $\mathbb{E}V_h^p = \{\{\{\mathbb{E}V_{htnk}^p\}_{t=1}^{T_h}\}_{n=1}^{N_h^{max}}\}_{k=0}^{N_h}$ are set of ex-ante value functions for the search and purchase stages respectively. Similarly, $\bar{\upsilon}_h^v = \{\{\{\bar{\upsilon}_{htnk}^v\}_{t=1}^{T_h}\}_{n=1}^{N_h^{max}}\}_{k=0}^{N_h^{max}}$ and $\bar{\upsilon}_h^p = \{\{\{\bar{\upsilon}_{htk1}^p, \bar{\upsilon}_{htn0}^p\}_{t=1}^{T_h}\}_{n=1}^{N_h^{max}}\}_{k=1}^{N_h^{max}}$ 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 (derivation presented in Appendix E):

$$\max_{\Theta} \sum_{h=1}^{H} \sum_{g=1}^{G} Pr(g|x_h^v, x_h^p, \Delta_h, p; \hat{\Theta}) \log(L_{h|g}(\mathbb{E}V_h^v, \mathbb{E}V_h^p, \bar{\upsilon}_h^p, \bar{\upsilon}_h^v; x_h^v, x_h^p, \Delta_h, p, \Theta)) \qquad (20)$$

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 $p = (p_1, \ldots, p_G)$. 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

$$p_g^h = Pr(g|x_h, \Delta_h, p; \Theta) = \frac{p_g L_{h|g}(x_h, \Delta_h, p; \Theta)}{\sum_{g=1}^{G} p_g L_{h|g}(x_h, \Delta_h, p; \Theta)} \qquad (21)$$

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

$$\hat{p}_g = \frac{1}{H} \sum_{h=1}^{H} Pr(g|x_h, \Delta_h, p; \Theta) \tag{22}$$

4.3.   The Estimation Algorithm

We combine the procedure presented for estimating models with discrete heterogeneity in Arcidiacono and Jones (2003) with MPEC approach (Su and Judd 2012). Equations 20, 21 and 22 suggest an iterative algorithm for estimation.

Step 0: Assume starting values of p_g and Θ.

Step 1: Calculate p_g^h, using equation 21, conditional on p and Θ.[22]

Step 2: Given the estimates of p_g^h, use equation 22 to update p_g.

Step 3: Using estimates of p_g^h, maximize equation 20 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 and 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 as a mathematical program with equilibrium constraints (Su and 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 price and search cost parameters. Intuitively, the purchase/no purchase decision in response to price variation as a function of state variables such as inventory identifies price sensitivity, while the frequency of store visits identifies search cost. We will elaborate more on this in the following paragraphs.

Identification of price parameter in our model might initially seem a bit tricky since the decision to purchase not only depends on the current price at the store that is being visited,

---

[22] 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). We also assessed sensitivity to starting values of inventory, by estimating the model with random starting values of inventory and find our estimates to be robust.

but it also depends on the flow of future utilities from potential store visits in the future. Since the utility that comes from potential visits in the future depends on how high the search cost is, one might imagine that search cost parameters are not separately identified from price sensitivity. However, note that even though search cost parameters enter the purchase decision indirectly through the potential future utility of visiting another store, conditional on the store that is being visited and whether the period is a weekend or weekday, they are constant across different visits for the same household, whereas price is not and it is this variation in price conditional on store that gives us separate identification of the price coefficient.

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.[23] The effect of non-focal categories (η) on visit decision 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 $\frac{1}{2}$ period.[24] While variance in travel time across household-stores identifies coefficient on travel time, and the overall frequency of visits identifies intercept of search cost, the difference between frequency of visits to primary and non-primary stores identifies preference for the primary store.

Note that we do not estimate the inventory cost since it is not identified. This is due to the lack of meaningful variation in choice of quantity, and also the short (i.e. three weeks) time horizon for inventory because of the perishable nature of milk. However, we are able to accommodate the cost of stockpiling in the model because the entire value of purchase is lost if not consumed within the "expiration" period.

---

[23] 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.

[24] 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 (Yao et al. 2012, Chung et al. 2013). For a review of the literature on discounting look at Frederick et al. (2002).

5.   Results

We begin with the estimates of our full structural model with three latent segments.[25] We then discuss how price elasticities are biased when the store or time dimensions are omitted.

5.1.   Estimates of the Full Structural Model

Table 2 reports the estimates of the structural model. All coefficients are significant and have expected signs, except for the coefficient on preference for weekends for the third segment, and preference for EDLP stores for the second segment, which are insignificant.

Segment 1 comprises 41% of the sample households, while second and third segments represent 40% and 19% 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 obtain greater value from search.

With the lowest level of search cost and a strong preference for the primary store, the third segment is expected to perform search across both dimensions. Compared to the third segment, the second segment has a higher search cost, which should result in lower levels of search across stores. Further, their preference for the primary store is high, relative to their search cost; hence, we expect to observe greater temporal search in this segment. The second segment has a preference for shopping during weekdays; the third segment has a preference for shopping at EDLP stores.

To test if our predictions for segment behavior based on the structural model estimates above are valid, we compare the observed behavior across three segments. Table 3 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 search across stores considering the fact that a consumer in this segment on average visits more than one store in the consideration set only 2.3% of the time. Consistent with their less frequent store vists, the average basket size per trip is highest. The second segment visits stores more frequently, and performs some search across stores. But the level of spatial search for this segment is

---

[25] We tested models with other numbers of segments. Even though a model with four segments does slightly better in terms of AIC and BIC criteria, two of the segments in the four segment model were qualitatively very similar, making the more parsimonious three-segment model easier to interpret and more practical.

Table 2   Search model with both store and time dimensions

	Segment 1	Segment 2	Segment 3
Price Sensitivity (α)	-0.121***	-0.289***	-0.494***
	(0.006)	(0.005)	(0.006)
Marginal Consumption Utility (σ)	4.204***	2.953***	2.240***
	(0.155)	(0.097)	(0.110)
Intercept of Consumption Utility (τ)	-0.788***	-0.612***	-0.437***
	(0.041)	(0.031)	(0.040)
Time Since Last Stock Up Period (η)	0.259***	0.316***	0.196***
	(0.058)	(0.043)	(0.054)
Search Cost Intercept (ι)	3.350***	2.263***	1.486***
	(0.012)	(0.010)	(0.014)
Travel Time (δ)	0.056***	0.056***	0.035*
	(0.021)	(0.016)	(0.020)
Preferred Store (ψ_1)	1.015***	1.102***	1.312***
	(0.035)	(0.026)	(0.033)
EDLP (ψ_2)	0.083***	-0.010	0.198***
	(0.022)	(0.019)	(0.027)
Weekend (ω)	-0.338***	0.255***	0.029
	(0.022)	(0.015)	(0.018)
Segment Size	0.41	0.40	0.19

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

pretty low compared to the third segment. Their average basket size is in between segments 1 and 3. This is consistent with larger estimated search cost and high preference for the primary store, relative to the search cost. Households in the third segment seem to visit stores most frequently and perform the highest level of search across stores among all three segments and have the lowest basket size pert trip Not surprisingly, there is a negative relationship between median income of the segment and its price sensitivity and share of private labels The third segment, which is the most price sensitive has the lowest median income and the highest share of private label (which has lower average price) and the first

Table 3   Observed behaviors and demographics by segment
The figures in this table are calculated based on household-specific segment
probabilities (i.e. p_g^h) computed in the EM estimation from equation 21.

	Segment 1	Segment 2	Segment 3
Percentage of shopping periods with at least one store visit	30.3%	49.2%	69.1%
Percentage of periods with more than one store visit	2.3%	6.5%	15.0%
Percentage of periods with more than one store visit conditional on at least one store visit	6.9%	11.9%	20.0%
Average price paid for milk ($)	2.99	2.95	2.82
Share of private labels (%)	79.3	81.4	84.9
Average basket size ($)	36.03	28.05	22.33
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.

segment, which is the lowest price sensitive, has the highest median income and lowest share of private label.

Table 4 reports the search costs in dollar terms for the three segments during weekdays and weekends based on the estimated parameters and price sensitivity.[26] 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.

Table 4   Search cost estimates

	Segment 1	Segment 2	Segment 3
Weekend	$26.29	$9.29	$3.29
Weekday	$29.08	$8.41	$3.23

5.2.   Bias from Omission of Either of the Two Dimensions of Search

As discussed in the introduction, the search cost literature thus far has focused on either the store or temporal dimension, but not both. We investigate how the omission of either of these dimensions can lead to biased estimates of price elasticity.

---

[26] To calculate search cost for each segment we summed 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 included 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.

We begin by turning off the temporal dimension of search. One straightforward approach to turn off temporal search is to set the discount factor, β in the structural model, to zero. However, in such a model the household would not internalize the utility of the purchased product that would be consumed in the future. Further, inventory also would have no impact on the purchase decision. So in addition to setting $\beta = 0$, we (1) include the inventory in the the purchase equation 3 and (2) modify the consumption utility such that consumer obtains the entire consumption utility for the purchase in the period of purchase by having equation 5 be simply a factor of purchase quantity.[27]

To investigate potential biases that arise from not modeling the store search dimension, we need a model that keeps the temporal dimension of the framework intact, while making the store visit exogenous. One straightforward approach would be to assume there is only one "virtual" store in each household's consideration set. In that case, each household would have a probability of making a store visit in each time period (which could be modeled as a Markov process), and price expectation for that store will be based on the distribution of prices pooled across all the stores in her consideration set. But in that case, potential differences across stores would be completely ignored. Hence, to turn off the store search, we consider a model similar to Hartmann and Nair (2010),but instead of modeling probability of visiting each store format (as in their model), we model probability of visiting each store (i.e. most visited, second most visited, and third most visited store). This allows us to have a much richer specification that models store visit behavior more accurately and yet exogenously, while keeping the temporal dimension of the model intact. Details of the store-only and temporal-only search models and their parameter estimates are presented in online Appendix F.

Table 5 presents price elasticities for all three models. We find that omitting either the temporal or spatial dimension underestimates the price elasticity.[28]

We first consider underestimation of price elasticity without temporal search. The direction of the bias is at first blush surprising given that previous research that has focused on

---

[27] We tried two other myopic models; one with only setting $\beta = 0$, and another model in which in addition to setting $\beta = 0$, we modified consumption utility such that consumer gets utility from the whole purchase in the same period, but without letting inventory influence purchase equation (i.e. (1) mentioned above). The direction of bias remains the same in all the three models.

[28] Note that food category elasticities tend to be much smaller than brand elasticities. Overall, our category elasticity estimate for milk is consistent with past research. Reviewing 160 studies on the price elasticity of demand for major food categories Andreyeva et al. (2010) find that price elasticities for foods and nonalcoholic beverages ranged from 0.27 to 0.81.

Table 5	Price elasticities		
Segment	Full Model	No Temporal Search	No Store Search
1	-0.34	-0.002	-0.37
2	-0.80	-0.62	-0.43
3	-1.24	-1.21	-0.47
Overall	-0.70	-0.59	-0.41
(% Change)		(-21%)	(-40%)

the temporal dimension (e.g., Hendel and Nevo (2006)) find that price elasticities are over-estimated in a myopic model without temporal search. To understand the underestimation of price elasticity, 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 the 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) would be rationalized as low price sensitivity by the myopic model. That is while a forward-looking model would rationalize it as due to the need to avoid a future stockout.

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. That is because the frequency of consumption and purchase is lower and consumers can store goods for longer time periods, which gives them more flexibility to perform temporal search without fear of stockouts. Further, in categories like detergents, consumers can more flexibly adjust consumption 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. The fact that we observe the highest level of bias for the first segment with lowest price sensitivity and highest stockout disutility (τ), compared to the third segment with highest price sensitivity and lowest stockout disutility seem to be consistent with our explanation for the bias.

Next we consider the case where we turn off the store dimension of search and retain the temporal dimension. Here the intuition for the underestimation is more straightforward. When we turn off the store search dimension, we rationalize the purchases at the primary store even at higher prices as due to low price sensitivity rather than due to the search costs that have to be incurred to obtain the lower prices by visiting other stores. This leads to the downward bias.

Our results suggest that the existing literature that has not accounted for endogenous store search may have significantly underestimated price elasticities. However, not accounting for the temporal dimension of search can lead to either overestimation or underestimation depending on the extent to which stockpiling (purchase acceleration) or stockout (loss in consumption utility) effects may be stronger in the category. We conclude that it is important to account for both dimensions to obtain unbiased price elasticities, when search along either or both dimensions are prevalent.

6. The Impact of Promotional Frequency on Store Loyalty and Profits

Gauri et al. (2008) documented different combinations of spatial and temporal price search strategies of households in their descriptive analysis. But without a structural model of search across stores and time, they were unable to study the impact of different promotional policies for stores under endogenous search. Using the structural model estimates, we now investigate the impact of changes in promotional frequency on store outcomes.

For the purposes of this counterfactual simulation, we evaluate how *symmetric changes* in promotional frequency lead to *asymmetric effects* on store visits and profit across primary and non-primary stores. To focus on promotional frequency, we keep the average price level and variance of the distribution constant across various levels of promotional frequency.[29]

---

[29] We do so for a given price average $\overline{p}$, variance σ_p^2, and frequency of promotions φ, by solving the following system of equations for regular and promotional price for each price distribution:

$$\begin{cases} \varphi p_{\text{promo}} + (1-\varphi)p_{\text{reg}} = \overline{p} \\ \varphi p_{\text{promo}}^2 + (1-\varphi)p_{\text{reg}}^2 = \sigma_p^2 + \overline{p}^2 \end{cases}$$

Specifically, we vary promotional frequency 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% (given a period is $\frac{1}{2}$ week) with the corresponding promotional depth changing from 71% to 38%. We set the travel time to the primary and non-primary stores to be the average, as observed in the data. Figure 5 presents a schematic representation of changes in price distribution used in simulations.

Figure 5   **Price distributions used in simulations: from low frequency but deep promotions to high frequency and shallow promotions, while keeping the mean and variance of price constant.**
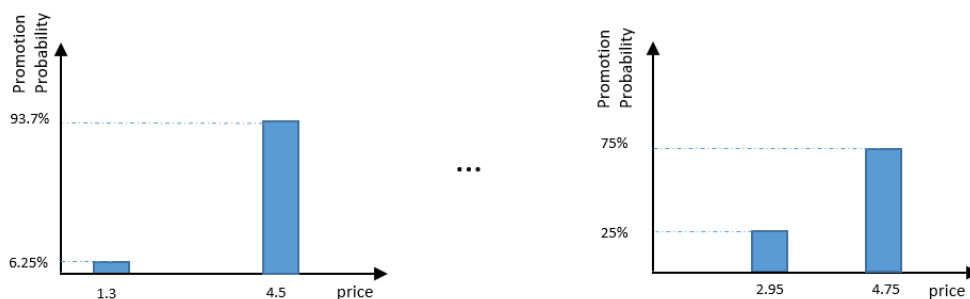

Given this promotional environment, we forward-simulate the behavior of households to compute a number of relevant metrics of loyalty and profits. To obtain stationary estimates with minimal simulation error, we forward-simulate 5000 households for 10,000 periods and average the metrics across households.

Figures 6, 7 and 8 present changes in the share of primary store visits, average price paid, and profit (at gross margin of 40%) for different promotional frequencies relative to the lowest promotion frequency case of 6.25%, respectively.

Increasing promotional frequency has heterogeneous impact on store loyalty and profits from the three segments. Not surprisingly, the incidental search segment 1—price insensitive and high search cost—does not change much its primary store loyalty in response to changes in promotional frequency and there is very little impact on primary store profits. The temporal search segment 2—price sensitive, but high primary store loyalty—which already exhibited high levels of loyalty to their primary store, does not change much in terms of store loyalty, but interestingly has lower levels of profits for the store as these primary store loyal customers are more often exposed to promotional prices.

The most interesting result is related to the response of the spatio-temporal search segment – with the highest price sensitivity and lowest search cost. This segment is impacted
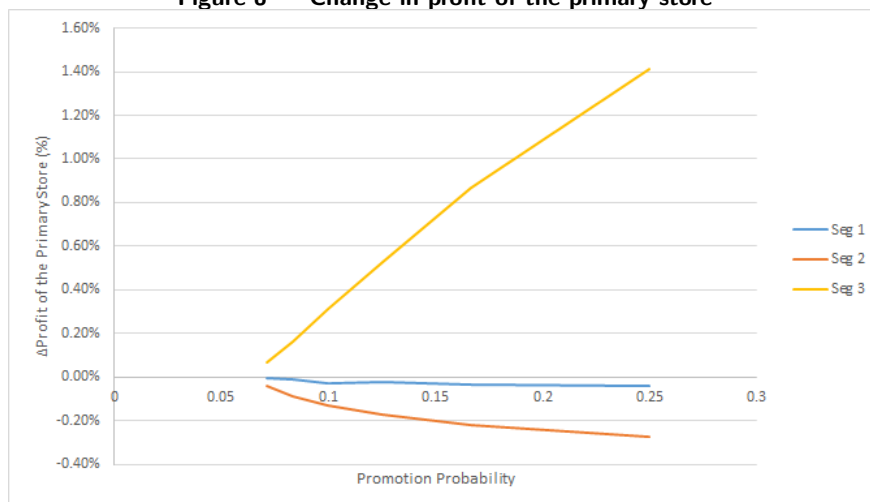
Figure 6   Change in share of the primary store visits
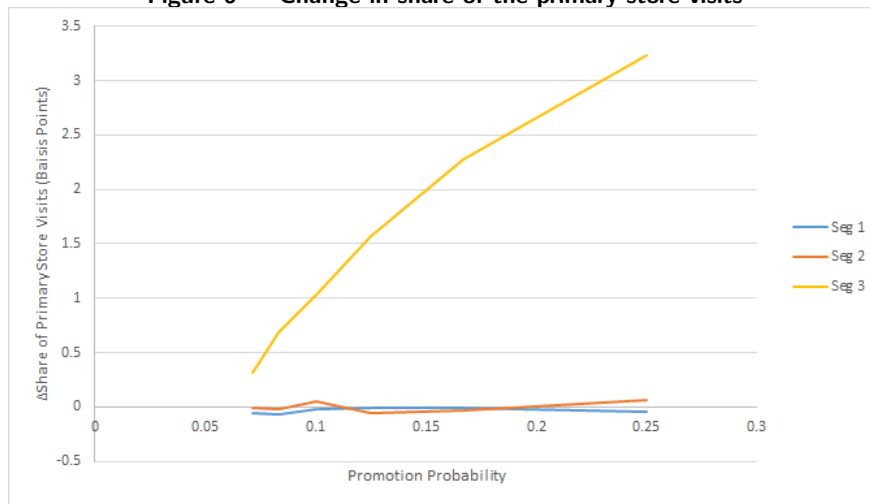
Figure 7   Change in paid prices for different segments
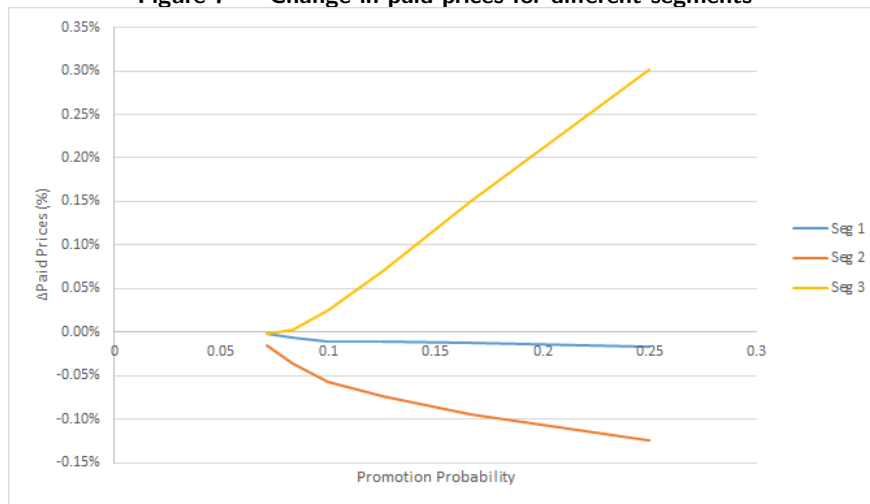
Figure 8   Change in profit of the primary store

the most by higher frequency of price promotions. Not only does their primary store loyalty increase, but overall profits increase as well—as these households have sufficiently high price sensitivity and low search cost for their store visit behavior to be affected by price changes of milk. But the increased frequency (with shallower discounts) causes greater consolidation of milk purchases at the primary store, leading to greater profits for the primary store. We note that though promotions are shallower as promotional frequency increases, the increased primary store loyalty for Segment 3 is not due to lower price variance—which we hold fixed in the simulations.

Thus, the structural model with endogenous search across stores and time is able to generate richer insights on how promotional changes will impact store outcomes. As stores will differ in the composition of households across different segments, our modeling framework can help managers evaluate the impact of alternative price promotion strategies under alternative expectations of competing store reactions.

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 search across stores within an infinite horizon model of search across time. We use an iterative approach based on EM-algorithm 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 either spatial or temporal dimension of search can result in substantial biases in the estimates. Our analysis on the milk category helps to provide a more nuanced sense on the direction of the bias that arises from ignoring consumers' forward-looking behavior relative to the existing literature which focused on temporal search using highly stockpilable categories such as detergents. Our results suggest that the direction of the bias by omitting the temporal dimension is determined by the relative frequency of purchase and

frequency of promotions, which impacts the relative costs of stockpiling and stockouts. When frequency of promotions is 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.

Finally, we evaluate the substantive question of how price promotions impact store loyalty by varying the frequency and depth of price promotions symmetrically across all stores, keeping the mean and variance of prices fixed. The effects vary across segments. Interestingly we found little change in loyalty among the incidental and temporal search segments, but the highest increase in loyalty to the primary store among the spatiotemporal search segment, which had the highest price sensitivity and lowest search costs.

Our analysis is an initial foray in modeling search across stores and across time. 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.

This paper suggests that the nature of biases in omitting the time dimension of search can be category specific; a systematic investigation of how these biases affect different categories 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 search across stores and across time in such categories. Also, given our focus on price variation in one focal category (milk); we only controlled for the store's price positioning strategy (e.g., EDLP/Hi-Lo) in how it impacts store visits and search. Future research can explore how price positioning endogenously impacts consumer store search especially as it relates to store choice as a function of basket level price variation. 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 charac-
teristics, store promotional strategies, price positioning and store locational configurations.

References

Andreyeva T, Long MW, Brownell KD (2010) The impact of food prices on consumption: A systematic review of research on the price elasticity of demand for food. *American Journal of Public Health* .

Arcidiacono P, Jones JB (2003) Finite mixture distributions, sequential likelihood and the em algorithm. *Econometrica* 71(3):933–946.

Bell DR, Ho TH, Tang CS (1998) Determining where to shop: fixed and variable costs of shopping. *Journal of Marketing Research* 35(3):352–369.

Besanko D, Winston WL (1990) Optimal price skimming by a monopolist facing rational consumers. *Management Science* 36(5):555–567.

Briesch RA, Chintagunta PK, Fox EJ (2009) How does assortment affect grocery store choice? *Journal of Marketing research* 46(2):176–189.

Ching A, Erdem T, Keane M (2009) The price consideration model of brand choice. *Journal of Applied Econometrics* 24:393–420.

Chung DJ, Steenburgh T, Sudhir K (2013) Do bonuses enhance sales productivity? a dynamic structural analysis of bonus-based compensation plans. *Marketing Science* 33(2):165–187.

Conlisk J, Gerstner E, Sobel J (1984) Cyclic pricing by a durable goods monopolist. *The Quarterly Journal of Economics* 99(3):489–505.

De los Santos B, Hortaçsu A, Wildenbeest MR (2012) Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review* 102(6):2955–80.

Erdem T, Imai S, Keane MP (2003) Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and economics* 1(1):5–64.

Everitt BS, Hand DJ (1981) *Finite Mixture Distributions* (Springer).

Fox EJ, Hoch SJ (2004) Cherry-picking. *Journal of Marketing* 77(S2):S25–S60.

Fox EJ, Montgomery AL, Lodish LM (2005) Consumer shopping and spending across retail formats. *The Journal of Business* 69(1):46–62.

Frederick S, Loewenstein G, O'donoghue T (2002) Time discounting and time preference: A critical review. *Journal of economic literature* 40(2):351–401.

Gauri DK, 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* 45(2):226–240.

Hartmann WR, Nair HS (2010) Retail competition and the dynamics of demand for tied goods. *Marketing Science* 29(2):366–386.

Hendel I, Nevo A (2006) Measuring the implications of sales and consumer inventory behavior. *Econometrica* 74(6):1637–1673.

Hong H, Shum M (2006) Using price distributions to estimate search costs. *The RAND Journal of Economics* 37(2):257–275.

Honka E (2014) Quantifying search and switching costs in the U.S. auto insurance industry. *The RAND Journal of Economics* 45(4):847–884.

Honka E, Chintagunta P (2017) Simultaneous or sequential? search strategies in the U.S. auto insurance industry. *Marketing Science* 36(1):21–42.

Khan R, Misra K, Singh V (2015) Will a fat tax work? *Marketing science* 35(1):10–26.

Kim JB, Albuquerque P, Bronnenberg BJ (2010) Online demand under limited consumer search. *Marketing science* 29(6):1001–1023.

Koulayev S (2009) Estimating demand in search markets: the case of online hotel bookings .

Leszczyc PTP, Sinha A, Timmermans HJ (2000) Consumer store choice dynamics: an analysis of the competitive market structure for grocery stores. *Journal of Retailing* 76(3):323–345.

McCall JJ (1970) Economics of information and job search. *The 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* 22(1):58–84.

Mortensen DT (1970) Job search, the duration of unemployment, and the phillips curve. *The American Economic Review* 60(5):847–862.

Neslin SA, Henderson C, Quelch J (1985) Consumer promotions and the acceleration of product purchases. *Marketing science* 4(2):147–165.

Roberts JH, Lattin JM (1991) Development and testing of a model of consideration set composition. *Journal of Marketing Research* 429–440.

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

Salop S, Stiglitz JE (1982) The theory of sales: A simple model of equilibrium price dispersion with identical agents. *The American Economic Review* 72(5):1121–1130.

Seiler S (2013) The impact of search costs on consumer behavior: A dynamic approach. *Quantitative Marketing and Economics* 11(2):155–203.

Stigler GJ (1961) The economics of information. *Journal of political economy* 69(3):213–225.

Su CL, Judd KL (2012) Constrained optimization approaches to estimation of structural models. *Econometrica* 80(5):2213–2230.

Ursu RM (2018) The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science* .

Weitzman ML (1979) Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society* 641–654.

Yao S, Mela CF, Chiang J, Chen Y (2012) Determining consumers' discount rates with field studies. *Journal of Marketing Research* 49(6):822–841.

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[30] 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 for studying category choice. Milk had the second highest share of total household spending at 3.3% and the highest penetration level. In terms of the frequency of price promotions, on average, milk has been purchased on promotion 17% of the times. While this number shows an opportunity for consumers to save in this category by searching over time, it is smaller compared to a category like soft drinks, mostly due to its perishability and the higher cost of stockout.

Table A1   Average share of top ten most-spent product modules across households*

Rank	Product Module	Avg. Share	Avg. Spending ($)	Penetration	% Purchased on Promotion
1	Soft drinks (carbonated & low calorie)	4.81%	117.42	87%	39%
2	Dairy - milk - refrigerated	3.39%	79.63	88%	17%
3	Cigarettes	2.70%	88.66	17%	6%
4	Cereal - ready to eat	2.60%	60.01	81%	32%
5	Bakery - bread - fresh	2.10%	49.00	87%	22%
6	Cookies	1.54%	35.72	73%	26%
7	Ice cream - bulk	1.44%	32.67	66%	43%
8	Soup - canned	1.29%	30.14	69%	30%
9	Candy - chocolate	1.26%	28.76	63%	28%
10	Water - bottled	1.23%	30.45	48%	21%

* 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. The last column reports the average share of purchase occasions where a household purchases the item on promotion.

B.   Reduced-Form Regressions

Here we present the results of our reduced-form regressions meant to show the effect of milk inventory on store visit decisions. These regressions are not conditional on visiting a store. The results of similar regressions that condition on store visit practically remain the same, even though the exact value of estimates change.

---

[30] 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.

Table B1 **The effect of milk inventory on the likelihood of visiting store with the lowest milk prices, and also visiting multiple stores in the same time period**

	Visiting Store with Cheapest Milk		Visiting Multiple Stores	
Milk Inventory	-0.037***	(0.006)	-0.038***	(0.009)
Periods Since Last Stockup Visit	0.027***	(0.003)	0.045***	(0.004)
Observations	54805		53787	
FE	Yes		Yes	

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The first regression considers the likelihood of visiting the store with cheapest milk as a function of number of time periods since last stockup visit and inventory of milk. The second regression considers the likelihood of making multiple store visits in the same time period. Both regressions show that inventory of milk can affect store visit decisions (i.e. the lower the inventory of milk, the higher the likelihood of visiting multiple stores in the same time period, or visiting the store with the lowest average milk price).

## C. Store Types: Consumer Choice, Pricing and Promotion

In this section, we describe how consumers choose between EDLP and Hi-Lo Stores. We also describe sample promotional schedules of Hi-Lo stores to give a sense of the price variance.

C.1. Share of EDLP and Hi-Lo Stores

EDLP stores make up 11.3% of all the stores in the consideration set of the households in our sample. Histogram of EDLP percentages for each consumer in the sample is provided below.

C.2. Promotion Schedules of Hi-Lo Stores

We describe the promotion schedules at 3 Hi-Lo Stores to give the reader a sense of the nature of price promotions seen by the households. Since each household's price expectations is driven by the stores in their individual consideration set, these illustrative graphs suggest that there is enough identifying variation in terms of promotion to estimate the model across all the households in our sample.

Figure C1   Histogram of number of EDLP stores in each HH's consideration set
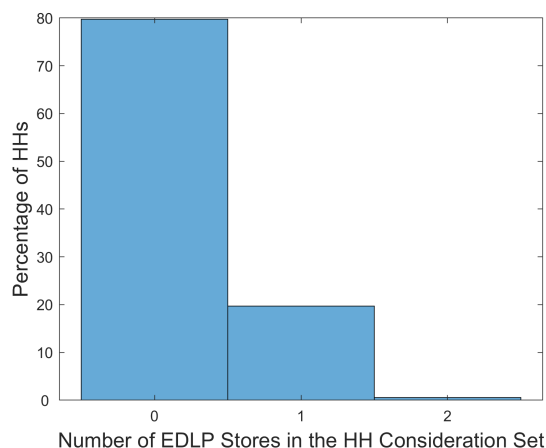
Figure C2   Samples of HiLo price sequences
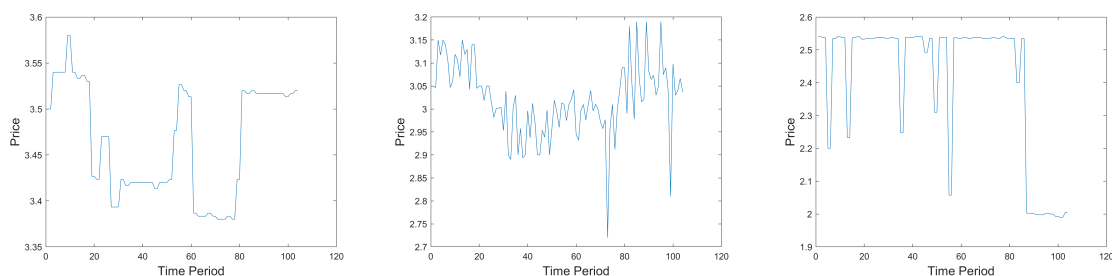

C.3.   Choice between EDLP and Hi-Lo Stores

Conditional on having both EDLP and Hi-Lo in the consideration sets, there seems to be a fair level of heterogeneity among households between how they alternate between HiLo and EDLP stores. Figure C3 shows the histogram of share of EDLP store visits among households that have both EDLP and HiLo stores in their consideration set.

D.   Constructing Price Data

For periods when a household makes a purchase in the milk category, we use the paid price. When the household does not purchase in the category, we consider the price for each household's relevant brand. Depending on the household's buying behavior, described in section 2.1 of the paper, this could be either the store brand, the most-purchased national brand, or the cheapest available brand.

To get data on prices when a household in our panel does not make a purchase, we supplement our panel data with retail scanner data from Nielsen, provided by Kilts Center for Marketing at the University of Chicago. There are two challenges in using the data:

Figure C3 **Histogram of share of EDLP store visits from HH's total visits, conditional on having both EDLP and HiLo stores in their consideration set.**
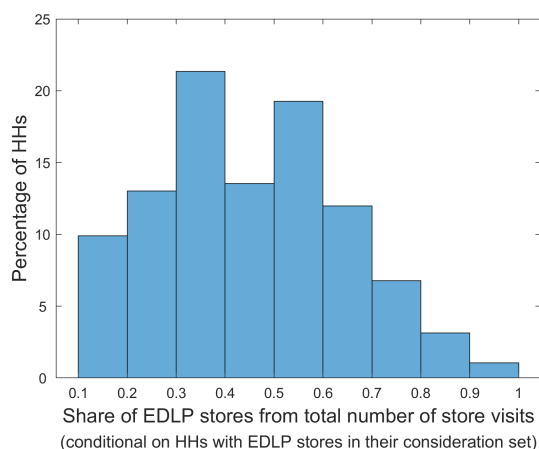
(i) the unique store identifier code in the Chicago data is different from that in our panel data.[31] (ii) some stores in our panel data are not available in the Chicago dataset.

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, this seems to be a reasonable assumption, given institutional practice. 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 households' consideration sets of stores and the fact that most households purchase private labels most of the time, we often observe store prices even for periods when a household in the sample has not paid a visit to a store, as another household has bought the item from the store.

---

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

E.   Derivation of the Conditionally-Group-Averaged Log-Likelihood

In this appendix, we derive equation 20, which we call conditionally-group-averaged log-likelihood from the unconditionally-group-averaged log-likelihood. That is, we want to show that the following two maximization problems are equivalent. This is an adaptation of what is presented in Arcidiacono and Jones (2003), and Everitt and Hand (1981) to our specific problem.

$$\max_{\Theta} \prod_{h=1}^{H} \left(\sum_{g=1}^{G} p_g L_{h|g}(\underbrace{\mathbb{E}V_h^v, \mathbb{E}V_h^p, \bar{v}_h^p, \bar{v}_h^v; x_h^v, x_h^p, \Delta_h, p}_{X}, \Theta) \right) \tag{E.1}$$

is equivalent to

$$\max_{\Theta} \sum_{h=1}^{H} \sum_{g=1}^{G} Pr(g|\underbrace{x_h^v, x_h^p, \Delta_h, p}_{Y}; \hat{\Theta}) ln(L_{h|g}(\underbrace{\mathbb{E}V_h^v, \mathbb{E}V_h^p, \bar{v}_h^p, \bar{v}_h^v; x_h^v, x_h^p, \Delta_h, p}_{X}, \Theta)) \tag{E.2}$$

We start with taking the log of the first equation and writing its first order conditions. To simplify the notation, we denote all the arguments in L with X, and all the arguments in the conditional probability with Y (shown above):

$$\sum_{h=1}^{H} \frac{\partial log(\sum_{g=1}^{G} p_g L_{h|g}(X_h, \Theta))}{\partial \theta_s} = \sum_{h=1}^{H} \frac{p_s \partial L_{h|s}(X_h, \Theta)/\partial \theta_s}{\sum_{g=1}^{G} p_g L_{h|g}(X_h, \Theta)} \qquad \forall s = 1, \ldots, G \tag{E.3}$$

From Bayes's theorem:

$$Pr(s|Y_h; \Theta) = \frac{p_s L_{h|s}(X, \Theta)}{\sum_{g=1}^{G} p_g L_{h|g}(X, \Theta)} \tag{E.4}$$

Substituting p_s from above into the first order condition, we will get:

$$\sum_{h=1}^{H} Pr(s|Y_h; \Theta) \frac{\partial L_{h|s}(X_h, \Theta)/\partial \theta_s}{L_{h|s}(X_h, \Theta)} = \sum_{h=1}^{H} Pr(s|Y_h; \Theta) \frac{\partial log(L_{h|s}(X_h, \Theta))}{\partial \theta_s} = 0 \qquad \forall s = 1, \ldots, G \tag{E.5}$$

This is the first order condition for:

$$\max_{\Theta} \sum_{h=1}^{H} \sum_{g=1}^{G} Pr(g|Y_h; \hat{\Theta}) ln(L_{h|g}(X_h, \Theta)) \tag{E.6}$$

F. One-Dimensional Models

In this appendix, we describe details of our implementation and also present raw estimates for temporal-only and spatial-only models for which we have presented price elasticities in section 5.2.

F.1. The Temporal-Only Model

For the temporal-only model in section 5.2, we follow Hartmann and Nair (2010) in assuming that store visits are exogenous and conditional on visit, consumer decides about whether to make a purchase in the focal category during that visit. We model the probability ϕ_{ht}, that household h visits store k at time t as follows.

$$\phi_{ht}(k) = \frac{\exp(\mu_{hk} + \rho \mathbb{I}_{ht-1}(k))}{\sum\limits_{m \in \{0\} \cup \Omega_{ht}} \exp(\mu_{hm} + \rho \mathbb{I}_{ht-1}(m))} \tag{F.1}$$

Where $\mathbb{I}_{ht}(k)$ is an indicator function that equals one if household h visits store k at time t and zero otherwise.

We can write the choice-specific value function for making a purchase as follows:

$$v_{htk1}^p(x_{htk}^p) = \overline{u}_{ht}^p + u_{ht}^c + \beta \sum_{k \in \{0\} \cup \Omega_h} \phi_{ht+1}(k) \mathbb{E}V_{ht+1k}^p + \varepsilon_{htk1} = \overline{v}_{htk1}^p + \varepsilon_{htk1} \tag{F.2}$$

The first summand captures the flow utility of purchase, while the second summand captures the flow utility of consumption. The third term sums over expected value of maximum value that could be gained in potential purchase in the remaining stores in the consideration set, weighted by probability of visiting each store, ϕ_{htk}, and finally the last term is a structural error shock similar to that in the full model. Similarly, we can write the value of not making a purchase as:

$$v_{htk0}^p(x_{htk}^p) = u_{ht}^c + \beta \sum_{k \in \{0\} \cup \Omega_h} \phi_{ht+1}(k) \mathbb{E}V_{ht+1k}^p = \overline{v}_{htk0}^p + \varepsilon_{htk0} \tag{F.3}$$

Based on the above choice-specific value functions, we can write the expected value of purchase decision as follows:

$$\mathbb{E}V_{htk}^p(x_{ht}^p) = \mathbb{E}_{x_{htk}^p, \varepsilon_{htk1}, \varepsilon_{htk0} | \varepsilon_{htk1}, \varepsilon_{htk0}} \{\max[v_{htk1}^p, v_{htk0}^p]\} \tag{F.4}$$

$$= \int\limits_{x_{htk}^p} \log[\exp(\overline{v}_{htk1}^p) + \exp(\overline{v}_{ht0}^p)] dP(x_{htk}^p)$$

Using the above expressions, we can write choice probabilities and likelihood function similar to the main model. Table F1 presents the parameter estimates.

<div align="center">

Table F1   **Search model with only time dimension**

	Segment 1	Segment 2	Segment 3
Price Sensitivity(α)	-0.1885***	-0.223***	-0.265***
	(0.0096)	(0.0065)	(0.0076)
Marginal Consumption	9.6571***	6.1282***	3.4197***
Utility(σ)	(1.1821)	(0.2641)	(0.1268)
Intercept of Consumption	-1.6602***	-1.2361***	-0.7868***
Utility(τ)	(0.211)	(0.0673)	(0.0676)
First Store Visit	-1.702***	-0.6107***	0.4569***
Intercept(μ_1)	(0.0189)	(0.012)	(0.0185)
Second Store Visit	-1.7359***	-0.7533***	0.4205***
Intercept(μ_2)	(0.021)	(0.0164)	(0.0244)
Third Store Visit	-1.6306***	-0.7169***	0.3818***
Intercept(μ_3)	(0.0348)	(0.0322)	(0.0414)
Store Visit Persistence(ρ)	-0.1112***	-0.3747***	0.2387***
	(0.0171)	(0.0117)	(0.0161)
Segment Size	0.39	0.43	0.18

</div>

F.2.   The Spatial-Only Model

For the myopic model, we made the following changes in the full model:

1. We set the discount factor to zero (i.e. $\beta = 0$).

2. We set the consumption utility to be $u_h^c = \sigma(q \sum_{n=1}^{N_h} y_{hn})$. That is, the household gets the utility from consuming the whole purchased amount during the time period at once.[32]

3. We modified equation 3 to include the effect of the inventory carried over from the past on purchase likelihood. This is the level of inventory either at the end of the previous time period, or at the end of the previous stage in the same time period. Note that in this case, even though we need to keep track of the inventory, and its transition follows the same rules as in the full model, the level of inventory that is

---

[32] Note that compared to the full model, the utility from consumption does not have an intercept in this case. Since most of the households purchase one unit in each period and in this case the whole purchased amount is assumed to be consumed in the same period, there will be very little variation in consumed quantity across households and time. That makes identification of two parameters for the consumption utility extremely difficult. In the full model, however, variation in households' consumption rates helps with the identification.

carried over from the past does not influence consumption utility.

$$u_{hnk1}^p(x_{hk}^p) = \begin{cases} \alpha_h p_k + \gamma_h i_h' + \varepsilon_{hnk1}^p & \text{if} \quad n = 1 \\ \alpha_h p_k + \gamma_h i_{hn-1} + \varepsilon_{hnk1}^p & \text{if} \quad n > 1 \end{cases}$$

Here, i_h' is the level of inventory at the end of the previous time period, and i_{hn-1} is the level of inventory at the end of the previous stage, after taking into acccount potential purchases in the previous stages.

Table F2   Search model with only store dimension

	Segment 1	Segment 2	Segment 3
Price Sensitivity (α)	-0.0015	-0.2214***	-0.4613***
	(0.0074)	(0.0081)	(0.0096)
Past Inventory	-0.2309***	-0.0816***	-0.0636***
Coefficient (γ)	(0.008)	(0.0106)	(0.0105)
Consumption Utility (σ)	0.3702***	0.1985***	-0.0319
	(0.0245)	(0.0461)	(0.038)
Time Since Last Stock Up	0.3067***	0.0513	0.0159
Period (η)	(0.0453)	(0.0521)	(0.0452)
Search Cost Intercept (ι)	2.2673***	2.5977***	1.4755***
	(0.0111)	(0.012)	(0.0117)
Travel Time (δ)	0.1461***	0.0942***	0.0083
	(0.0181)	(0.0196)	(0.0182)
Preferred Store (ψ_1)	1.4148***	0.8843***	0.9144***
	(0.0294)	(0.03)	(0.031)
EDLP (ψ_2)	0.0117	-0.1071***	-0.0001
	(0.0204)	(0.0214)	(0.0233)
Weekend (ω)	0.0154	-0.0318	0.0801***
	(0.0163)	(0.0198)	(0.0161)
Segment Size	0.34	0.42	0.24

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table F2 presents parameter estimates for the myopic model. We also estimated a model with only the first item in the above list, and another model with making the first two changes in the above list (i.e. the whole inventory gets consumed at the end of each period

and its level does not influence the utility from purchase). The direction of bias for all three segments remained the same across all implementations.