The Impact of Search Costs on Consumer Behavior: A Dynamic Approach

May 6, 2012

Abstract

Prices for grocery items differ across stores and time because of promotion periods. Consumers therefore have an incentive to search for the lowest price. However, when a product is purchased infrequently, the effort of checking the price on every shopping trip might outweigh the benefit of spending less. I propose a structural model for storable goods, that takes inventory holdings and search into account. The model is estimated using data on laundry detergent purchases. I find that search costs play a large role in explaining purchase behavior, with consumers not being aware of the price of detergent on 70 percent of of their shopping trips. Therefore, from the retailer's point of view it is important to raise awareness of a promotion through advertising, displays, etc. I also find that a promotion for a particular product increases the consumer's incentives to search. This leads to an increase in category traffic, which is a desirable side-effect of the promotion from the store manager's perspective.

JEL Classi cation: D12, D83, C61, L81

Keywords: Dynamic Demand Estimation, Search Costs, Imperfect Information, Storable Goods, Stockpiling

1 Introduction

Temporary price reductions are used very frequently for consumer packaged goods and represent a large fraction of the marketing mix budget of supermarkets and convenience stores. These promotion periods create an incentive for consumers to strategically time their purchases in order to benefit from the lower promotional price. At the same time, checking for the price of a particular product in order to find out when it goes on promotion requires some effort from the consumer. If search for price information is costly, the consumer has to trade-off the benefits from finding a lower price against the cost of searching. Therefore, it is a part of his utility maximization to decide how informed he wants to be.

In order to capture this, a structural model with imperfect information where consumers engage in costly search is proposed. A pre-purchase stage is modeled in which the consumer decides whether to search based on his expected utility from purchasing and the cost of search. The novelty of this paper is to integrate the search decision into a dynamic demand framework for a storable product. Search is modeled jointly with the purchase decision in order to fully capture consumer behavior in a structural way. This approach makes it possible to quantify search costs and assess their importance in the consumer's decisions. I apply the model to laundry detergent data, but the proposed approach can be used to analyze demand for any storable product. The search aspect of the model will be especially relevant for products with a relatively long inter-purchase duration.

I find that search costs are quantitatively important. The estimated parameters imply that consumers do not search on approximately 70 percent of their shopping trips. With consumers not being aware of prices on most of their shopping trips, marketing tools other than pricing, such as advertising and preferential display, become very important. A counterfactual exercise shows that lowering the search cost by 50 percent when running a promotion leads to a more than three-fold increase in the elasticity of demand. Furthermore, pricing also has an impact on search behavior due to the fact that lower prices make searching more attractive. Therefore, more promotional activity does not only benefit the promoted product, but also leads to an overall increase in category traffic.

This paper contributes to an emerging literature which demonstrates that imperfect information due to search frictions is an important component in the inference of consumer demand. Papers like Kim, Albuquerque, and Bronnenberg (2011), Goeree (2008), Moraga-Gonzalez, Sandor, and Wildenbeest (2009) and Koulayev (2009) show that including imperfect information has an impact on the estimates of consumer preferences.[1] In terms of the estimation strategy, this paper is most closely related to dynamic models of demand for storable products such as Erdem, Imai, and Keane (2003), Sun, Neslin, and Srinivasan (2003) and Hendel and Nevo (2006). The model in this paper also shares some features with the "price consideration model" presented in Ching, Erdem, and Keane (2009), which does incorporate consumer search as well. However, in Ching, Erdem, and Keane (2009) search is used as a computationally-friendly substitute for modeling consumers' inventories in a structural way. This paper

[1]Other papers that estimate the magnitude of search costs include: Hortaçsu and Syverson (2004) for the mutual fund industry or Hong and Shum (2006) and Santos, Hortacsu, and Wildenbeest (2009) for online book purchases. Mehta, Rajiv, and Srinivasan (2003) estimate search costs for grocery shopping items but do not allow consumers to keep an inventory.

instead argues that new insights can be gained from including both search *and* inventory holdings into a structural framework. The empirical model presented here indeed brings together the search aspect modeled in Ching, Erdem, and Keane (2009) with a dynamic inventory model (as in Erdem, Imai, and Keane (2003), Hendel and Nevo (2006)) into one unified structural framework.

The paper is organized in the following way: The next section describes the data. Section three presents some reduced-form results to motivate the structural model. Section four presents the empirical model followed by a discussion of identification in section five. Section six presents the results from the estimation. In section seven I run a validation test and section eight uses the estimates of the model for two counterfactual simulations. Finally, some concluding remarks are made.

2 Data

I use data from the "Kantar Worldpanel UK", a consumer-level panel dataset provided by the Kantar Marketing Research Institute. Each household in the panel is given a scanning device which it uses to scan all products that were purchased. Receipts are then sent to Kantar in order to correctly record the price paid for a particular product. An observation is the purchase of a particular product at a particular store on a particular day. Therefore, it is also known when a household went shopping without buying any laundry detergent, as long as at least one item was purchased on the trip. The data reports purchases for the period from 10/2001 to 10/2007. I therefore have up to 6 years of data for each household. The panel is unbalanced with most households staying in the sample for less than the full duration of the data.

Detergent is chosen for the empirical exercise as it is storable and purchased infrequently. Consumer search behavior is therefore likely to be important for this product. Furthermore, a promotion is unlikely to lead to an increase in consumption of detergent (see Bell, Iyer, and Padmanabhan (2002))). Being able to ignore such a reaction simplifies the analysis. Within the category, there are three main types of detergent: powder, liquid and tablets. The types vary in their effectiveness, i.e. how many loads can be washed with a certain quantity of detergent. In order to compare pack-sizes across products I therefore have to concentrate on one of these types. As consumers rarely switch between these three different types of detergent, it is unproblematic to look at one type in isolation. I choose to concentrate on the market of detergent tablets as there are less brands available than for other types. This facilitates the construction of price series (see next section). Furthermore, as I have to integrate over the consumer's price expectations in the model, having fewer brands makes this part of the algorithm less computationally burdensome.

2.1 Constructing Price Series

As the Kantar Worldpanel has data at the household level, no store-level dataset of prices is readily available. But in order to analyze the consumer's optimal choice I need to know the prices of other products that the consumer could have purchased on his shopping trips, but decided not to. To this end, I have to infer the price of non-purchased products in the consumer's choice set from purchases of other consumers. As households in the panel are distributed over the whole of the UK, I rarely have several observations for the same store in the same week. In order to infer prices, I therefore rely on national pricing policies of the big supermarket chains.[2] However, the construction of a reliable price series is only possible if I observe enough purchases in order to confidently infer the weekly price. Luckily, the market for detergent tablets is very concentrated and 7 brands (Fairy, Daz, Ariel, Persil, Bold, Surf and Tesco's private label brand) make up about 80 percent of purchases. All other brands have substantially lower market shares than these 7. I am able to construct price series for all brands except for Surf and Persil. Both of these brands offer many different pack-sizes and I observe only a small number of purchases for each pack-size, which makes it impossible to construct a reliable price series. I encountered no such problems for the other 5 brands.

The various brands are available in 3-5 different sizes and I allow consumers to buy two packs of the same size. For each brand and pack-size, I construct price series for each of the four big supermarket chains (Asda, Morrisons, Sainsburys and Tesco), plus a residual category for all other stores. This yields a total of 220 price series (44 brand/pack-size combinations at 5 supermarkets). For the four major supermarkets, the prices are identical across stores due national pricing. For all other stores, a simple average of prices across all purchases for a particular brand and pack-size in a particular week is taken. Prices for this residual category will therefore be measured less precisely. Since about 90 percent of purchases occur at the four big chains, this is not problematic.

2.2 Selection of Relevant Households

I restrict the sample to households that regularly consumed detergent. Specifically I drop households that consumed very little detergent (less than 6kg per year) throughout their time in the sample as well as households with gaps of at least 16 weeks between their purchases. I also drop households that are in the sample for less than 20 weeks. Finally, I use only households which bought one of the 5 brands for which I construct price series at least 75 percent of the time. Any detergent purchase of a brand other then these 5 brands is modeled as a residual category. I assume that this "outside good" has a pack-size of 1.3kg and a price of 3 pounds. This corresponds to the average pack-size and price for detergent tablets.[3] Only about 12 percent of purchases fall into the residual category. As a result, there are 686 households that fulfill all criteria. Overall, I observe 113498 shopping trips and 18210 purchases

[2] Supermarkets do engage to some extent in price flexing, i.e. adjusting prices to local conditions, but this is only done for a small subset of products (according to the UK Competition Commission) and does not seem to include laundry detergents.

[3] I also allow liquid and powder detergent purchases within this residual category, but code them at the same price and pack-size for simplicity.

for this sample of households. This corresponds to 165 trips per household and 26.5 purchases per household. I also checked whether the selected households are observably different from other households in the sample and find that they are not. Further details on the selection of households can be found in Section (A.1) of the appendix.

2.3 Descriptive Statistics: Shopping Behavior

It is central for the model presented in this paper that consumers do not observe prices in a particular product category on every shopping trip. This is in part based on the observation that there is great heterogeneity in purchase behavior across shopping trips for the same consumer. Many consumers in the sample visit several supermarkets regularly and have great variation in both the overall size and the composition of their shopping basket. Most importantly, consumers make purchases only in a very limited number of categories on each trip. They are therefore unlikely to observe prices in all categories on each trip.

The whole shopping basket on each trip is observed in the data and it is therefore possible to look directly at the variation across shopping trips. Table (1) presents some evidence of this variation over time. It reports the number of categories in which the consumer made a purchase on a given trip. I use a system of assigning products into various categories that is provided in the dataset. Results for three different levels of aggregation are presented, ranging from the most aggregate level with 5 categories to the most disaggregate one with 226 categories. A category at the aggregate level would be "Toiletries" or "Fresh Food" for example. At the most disaggregate level "Chocolate Biscuit Bars", "Christmas Puddings" or "Fresh Poultry" are examples of product categories. Comparing the first two columns of the table shows that on a given shopping trip, only a small fraction of product categories is covered by the consumer. For example on the most disaggregated level, the fraction is as low as 10 out of 226 product categories, i.e. about 4.4 percent. Of course it might be the case that some consumers never make a purchase in a certain category on any of their shopping trips. The remaining columns therefore show how many categories are covered on a particular trip relative to the number of categories that is covered in a 4-week period for each household. Even in this case, the fraction of categories in which a purchase is made on a particular trip lies only between 20 and 60 percent across different levels of aggregation. In other words, most categories from which the consumer makes a purchase within a 4-week period are unlikely to see a purchase on any particular trip. It is therefore very unlikely that consumers are aware of prices in all categories on all of their shopping trips.

2.4 Descriptive Statistics: Prices

In the context of detergent, there is variation in prices across brands, pack-sizes, across supermarkets chains and over time. This is the case for most storable products, as they are usually available in multiple pack-sizes and are frequently sold at a temporarily lower price. As the multi-dimensional price variation is crucial for the identification of the model, some descriptive statistics along the various dimensions are presented in this section. Note that prices for a particular brand and pack-size do not vary across stores within the same supermarket chain due to national

pricing policies. When refereing to price variation across supermarkets in the remainder of the paper, this will refer to variation across chains.

Table (2) shows the variation in prices across different brands. The price variation is reported for pack-sizes of 900g and 1.9kg, which are the 2 most popular pack-sizes.[4] The mean price is very similar across brands except for Tesco's own brand which is cheaper. There is substantial variation in prices for a given brand and pack-size across supermarkets and time. The larger pack-sizes show a higher standard deviation of prices as they are promoted more often. The quantity discount on the larger pack-sizes is very similar across brands, ranging between a 8 and 12 percent reduction in the per volume price. More detailed descriptive statistics on prices and market-share variation across pack-sizes, brands and supermarkets are presented in the appendix in Tables (B2) and (B3).

One of the most important dimensions of price variation is the variation over time caused by temporary price reductions. These promotion periods entail drastic price changes over a short period of time. For the purpose of illustrating the patterns, the regular price for a particular product is defined as the 75th percentile of the product's price distribution. As promotions are quite infrequent, the 75th percentile will always lie outside of the promotion range of the price distribution. A promotion is defined as a price that is at least 20 percent lower than the regular price. Across all brands, supermarkets, and all 6 years of weekly price data, about 8.5 percent of prices constitute promotions according to this definition. Table (3) disaggregates the occurrence of promotions across brands, supermarkets, and pack-sizes. There is a substantial amount of variation across supermarkets and brands with Tesco promoting its own brand very heavily. Also, the larger pack-sizes are more frequently promoted than the smallest size. The very large pack-sizes are excluded from the analysis as they are only available for a short period of time (see Tables (B2) and (B3)). They are always offered at a discounted price and it is not possible to distinguish a regular and a promotional price according to the definition used here.

There is also substantial variation in the length of promotions and the duration of time between two consecutive promotions. As shown in Table (4) the average promotion period lasts for about 4 weeks, but there is substantial variation in this, with the standard deviation being 2.7 weeks. Similarly, the length of time between promotions is subject to an even higher variation. The mean time between two promotion periods for the same product is 28 weeks with a standard deviation of 36 weeks. Most of the variation in both variables is due to within product variation. Only a small fraction of the overall variation can be explained by variation across brands and supermarkets. This can be seen in the last column, which reports the r-square of a regression of promotion-length (length of a regular price period) on a set of brand and supermarket dummies, which is equal to 0.20 (0.28) respectively. I also analyzed whether the start of a promotion period for a particular brand can be predicted by the time elapsed since the last promotion and found that this is not the case. Overall, the large variation in the way temporary price drops occur makes it difficult to predict promotions. This feature of the data will inform the way price expectations are included into the structural model later on.

[4] The pack-sizes of the different brands are not exactly equal to 900g and 1.9kg. There are only small differences though and I will therefore ignore these differences when looking at price variation across brands.

3 Some Reduced-form Evidence

This section presents some patterns in the data that suggest the presence of search costs. Besides providing out-of-model evidence for the importance of search costs, this section also highlights some of the variation in the data that will allow the identification of search costs. The empirical patterns presented here therefore tie in directly with the section on identification presented later.

3.1 The Impact of Other Items in the Shopping Basket

Products such as laundry detergent are usually bought together with many other items on a shopping trip. Therefore, one has to be careful when analyzing demand for a particular product as it might depend in some way on other products in the shopping basket. Analyzing demand for one product in isolation is a valid procedure only as long as two assumptions are fulfilled: (1) there are no interactions in consumption between the various products,[5] (2) consumers do not have to incur any search costs in order to obtain price information on each trip. For a category such as detergent, there are no other products that are used together with detergent, i.e. there are no complementary products.[6] Also, detergent is a fairly "isolated" category in the sense that there are no close substitutes. As long as a household has to wash clothes, detergent will be the only product that can be used to this end. The first assumption is therefore very likely to hold in the case of detergent.

If this is the case, search costs are an alternative mechanism through which other items in the shopping basket can influence demand for detergent. Specifically, consumers might not make the effort to check the price of detergent on every shopping trip. In the presence of search costs, they are more likely to do so if they also buy products that are located nearby in the store. If they are already near the detergent aisle, the marginal effort to check the price of detergent is relatively smaller. Put differently, the overall search cost can be spread across more products that are located together in the store (see Warner and Barsky (1995) for a similar argument). As was documented in Section (2.3), there is considerable variation in the composition of the shopping basket across trips for the same household. In the presence of search costs which vary as a function of the shopping basket composition, this variation will translate into variation in purchase probabilities.

The impact of other items in the shopping basket on detergent purchases is illustrated in Table (5). In the first column and first row, the unconditional purchase probability for laundry detergent is displayed. On any given trip to a supermarket, the average probability of picking up any brand of detergent is 16.04 percent. The following rows display conditional purchase probabilities for trips that are characterized by a different size and composition of the shopping basket. For simplicity, only binary classifications are used in the table.

[5] To be precise, another condition is the absence of income effects at the trip-level. A budget-constrained household might delay his detergent purchase because there are more urgent products that he needs to buy. Despite the fact that these products are completely unrelated in consumption, demand would be correlated between these products. I looked at whether low income households behave differently and checked for changes in behavior over the duration of a month (due to monthly wage payments). Doing so, I found no evidence of the presence of budget constraints at the trip-level.

[6] The only clearly complementary product is softener, which is therefore excluded from the analysis both in this section and in the structural estimation.

The second and third row of the table show the difference between shopping trips with differently sized shopping baskets. I define a large trip as a trip with above average expenditure (excluding expenditure on detergent) at the household level. The classification of trips therefore only relies on within-household variation in expenditure over time. The conditional probabilities demonstrate the considerable impact of shopping basket size. The purchase probability is only 7.72 percent on a small shopping trip, whereas it increases to 27.49 percent on a large shopping trip. On large trips, the consumer was most likely under less time pressure and also went through more supermarket aisles on his trip, i.e. he was closer to the detergent aisle. Both things lower his search cost for detergent.

The lower two panels in the table show the impact of basket composition on purchase probabilities. In the context of search, it is primarily relevant whether the consumer was close to the detergent aisle in the supermarket. In order to capture this idea I look at whether he bought products that would usually be located in the same part of the store. To this end, the number of other cleaning products and the number of other household goods in the consumer's shopping basket is used. Both constitute product categories of which detergent is part of, the latter being broader than the former. As before, a binary classification is used: supermarket visits are split into trips with above and below (or equal to) the median number of products in the category. Comparing the different rows in the table shows that the purchase probability is considerably higher if other cleaning or household products are purchased on a particular shopping trip. This pattern is consistent with the existence of consumer search costs.

One concern with this type of analysis might be that consumers will use different modes of transport for different shopping trips. For instance, consumers might be more likely to use their car on a large shopping trip. The usage of a car could in turn lead to a higher purchase probability because of transport costs, i.e. the consumer does not want to carry detergent on the trips without the car. This would lead to a correlation between basket size and the purchase probability that is unrelated to search. As pack-sizes are not particularly large in the UK, usually either 900g or 1.9kg, this is most likely not an issue but it is also possible to test this directly. I know from a survey conducted among all consumers in the sample, whether a household always / sometimes / never uses a car for shopping. The second column of Table (5) therefore conditions on households that always use a car. If variation in the mode of transport mattered as described above, one would expect no difference in the purchase probabilities across different types of trips for this subsample of consumers. Instead, the results show that the probabilities vary in a very similar way across the different trips as they do for the full sample. If anything, the probabilities for the different types of shopping trips lie even slightly further apart in the second column.[7]

In summary, in the absence of interactions in consumption between products and if consumers are not subject to any search costs, there is no reason why the composition of the shopping basket should determine whether detergent is purchased. Price and other product characteristics alone should be a sufficient statistic. However, the results in Table (5) confirm that the basket size and composition do matter substantially for the purchase probability.

[7]There is no information on the mode of transport for each shopping trip in the data, it is therefore not possible to test this hypothesis at the trip level. Some further analysis using the distance to each supermarket was also conducted. Distance did not seem to matter for the purchase probability which also speaks against the importance of transport costs for the purchase decision regarding laundry detergent.

This kind of correlation is suggestive of the presence of search costs.[8] In the model presented in this paper, a higher probability of searching due to variation in the shopping basket size and composition will lead to a higher probability of purchasing.[9] More descriptive statistics on the variables used here are provided in Table (B1) of the appendix. All three variables used in this section are also incorporated as "search cost shifters" in the structural model.

3.2 Consumers Missing Promotions

Another interesting pattern in the data is the fact that consumers frequently buy a product at the regular price when they could have bought it at a discounted price on one of their previous shopping trips. Even in a model without search costs, but with inventory holdings, it is not impossible to observe this. As the consumer's inventory decreases over time, his reserve price will go up and he might ignore lower prices due to a high inventory on trips prior to his purchase. But as the time between the missed promotion and the purchase at the regular price becomes very small, it becomes difficult to rationalize this only through inventory holdings. I calculate the occurence of such purchases within two different time windows. Promotions are defined as described in Section (2.4). I find that out of all purchases in the sample at a non-promotional price, in 10.63 percent of cases the consumer could have purchased the same product at a promotional price on another trip earlier in the same week or the previous week. When the time windowed is shortened to only include purchases earlier in the same week this percentage drops to 3.38 percent. This provides another piece of evidence suggesting that consumer are unlikely to be always perfectly informed about prices. If consumers were perfectly informed, one would expect them to (almost) never ignore a promotion and buy at the regular price shortly afterwards. This is due to the fact that the price difference is very large between regular and promotional price and inventory is unlikely to play a large role within such a short time window. Together with the patterns presented in the descriptive statistics, the reduced-form evidence strongly points to the presence of search costs. A crucial part of the structural model presented in the following section will therefore be the inclusion of costly search in the consumer's decision process.

4 The Structural Model

An important contribution of this paper is to include consumer search into a dynamic demand framework for a storable product. In order to do this, one has to be careful to model search behavior for this type of product appropriately. In particular, the search process for a storable, low cost consumer good like detergent works quite

[8]However, this type of correlation could also arise if consumer are perfectly informed about prices but they have to incur a cost to visiting certain parts of the store. This situation is unlikely to be relevant for the data I am using as in the UK little price information is available prior to visiting the store.

[9]Note, that the composition of the whole shopping basket is itself part of a larger optimization problem that the consumer has to solve. In other words, the decision to search and purchase *any* product is part of the consumers decision making process subject to certain constraints. Therefore one could trace back the reasons for why search costs for detergent vary as a function of the shopping basket composition to more basic underlying primitives. I outline such a model verbally in Section (A.2) of the appendix. In the empirical model I will treat the basket composition as exogenous to the search and purchase decision regarding detergent. This is a necessary simplification in order to make the model tractable.

differently from many other goods. When buying a durable and high value product, say a car or TV, consumer search will entail visiting several stores *only* for the purpose of finding information about the desired product. For a low-cost, repeat-purchase product like detergent, no consumer will visit several stores only to search for the best price for this product. Instead, the consumer will go shopping for various other reasons and each shopping trip represents an opportunity to also search for the price of detergent. This is exactly the way in which search is modeled in this paper. Specifically, the timing is as follows: At the beginning of each time period the consumer enters a store and has to decide whether to search, i.e. to go down to the detergent aisle in the particular supermarket. If he decides not to search, he will not have the opportunity to buy any detergent. If he searches, prices are revealed, and the consumer then makes his purchase decision among the available brands.

In more technical terms, the consumer's decision process on each trip is modeled in two stages. In the first

knows the current period prices. The utility in the purchase stage (ps) is indexed $u_{ps,t}$ (The utility in the search stage (ss) is represented by $u_{ss,t}$). If the consumer does not purchase any product ($j_t = 0$) he gets utility

$$u_{ps,t}(j_t = 0) = v(c(i_t)) - T(i_t) + \varepsilon_{0,t}$$

Where $v(\cdot)$ is utility from consumption $c(\cdot)$. The consumer's inventory is denoted by i_t. $T(\cdot)$ represents the storage cost of inventory holdings.

If he decides to purchase any product out of the set of available products ($j_t \in J$) he receives

$$u_{ps,t}(j_t \in J) = -\alpha p_{j,t} + \xi_j + v(c(i_t)) - T(i_t) + \varepsilon_{j,t}$$

This is equivalent to the expression above, but for the inclusion of a negative utility term from having to pay price $p_{j,t}$ and an unobserved (by the econometrician) product quality term ξ_j. A product in this case is defined as a brand / pack-size combination. All the variables in the flow utilities are known to the consumer after searching. Prior to the search he does not have full information about these variables. Most importantly, the realization of prices $p_{j,t}$ ($j \in J$) is unknown prior to search. The exact information structure will be explained in more detail later. $\widetilde{\epsilon}_{ps,t} = (\varepsilon_{0,t}, \varepsilon_{1,t}, ..., \varepsilon_{J,t})$ are product specific taste shocks and are unobserved by the econometrician. Note that in principle, the inventory is also unobserved by the econometrician. However, I do estimate the rate of consumption within the model and treat the inventory as observable. In other words, the rate of consumption inferred from the data is assumed to be measured without noise.

The one-period utility prior to search, i.e. in the search stage, is defined in a similar way. If the consumer does not search he will not be able to make a purchase. He therefore receives the following utility

$$u_{ss,t}(d_t = ns) = v(c(i_t)) - T(i_t) + \varepsilon_{ns,t}$$

This is the same utility he receives if he searches but does not buy anything, except for a different error term $\varepsilon_{ns,t}$. If he searches he receives

$$u_{ss,t}(d_t = s) = -s_t + \varepsilon_{s,t}$$

where $d_t \in \{s, ns\}$ denotes whether the consumer searches or doesn't search. Note, that this expression only captures the flow utility in the search stage. When deciding to engage in search the consumer does anticipate that he will have to take another decision within the same time period. He will therefore receive more *non-discounted* utility as defined by the purchase stage flow utility. The full (expected) flow utility in the current time period if the consumer decides to search (i.e. from both choice stages) can be expressed as follows

$$E[u_{BothStages,t}(d_t = s)] = E\{max_j[u_{ps,t}(j_t)]\} - s_t + \varepsilon_{s,t}$$

A feature of the model is that prior to searching the current period flow utility in the purchase stage is unknown. Therefore, the consumer has to form an expectation over the flow utility in the purchase stage knowing that he will take the optimal purchase decision conditional on the information that he will obtain in the purchase stage.[10] This information is not yet available to him in the search stage and he cannot perfectly predict his decision in the purchase stage.

s_t denotes the search cost which is unobserved (by the econometrician). The search cost has a time-subscript as it will be allowed to vary across trips as a function of certain variables. This will be laid out in detail in the estimation section. $\varepsilon_{ns,t}$ and $\varepsilon_{s,t}$ are unobserved (by the econometrician) iid extreme value error terms. Let $\widetilde{\epsilon}_{ss,t} = (\varepsilon_{ns,t}, \varepsilon_{s,t})$ be the vector of idiosyncratic shocks for time period t in the search stage. Note that different sets of error terms enter the model before and after search. It is assumed that search is a category level decision (see earlier discussion). The consumer decides whether he wants to incur the cost of searching for detergent by making the effort of walking to the detergent aisle. $\widetilde{\epsilon}_{ss,t}$ is therefore a vector of category level taste shocks, whereas $\widetilde{\epsilon}_{ps,t}$ is a vector of shocks that influence the consumer's choice across products.

4.2 The Dynamic Optimization Problem

Formally, a consumer chooses an infinite sequence of decision rules μ in order to maximize the expected, present discounted sum of future utility

$$max_{\{\mu_t\}_{t=0}^{\infty}} E\{\sum_{t=0}^{\infty} \beta^t g_t(\mu_t) \mid x_t, \widetilde{\epsilon}_t\}$$

where $\mu_t = \{(d_t), (j_t \mid d_t = s)\}$. d_t denotes the consumer's search decision with $d_t = s$ if he decides to search and $d_t = ns$ if he does not search. $j_t \in J \cup \{0\}$ denotes the purchase decision conditional on having searched ($d_t = s$). $\widetilde{\epsilon}_t = (\widetilde{\epsilon}_{ps,t}, \widetilde{\epsilon}_{ss,t})$ is the vector of error terms in both decision stages. x_t is a vector of state variables. Finally,

$$g_t(\mu_t) = \begin{cases} \overline{u}_{ps,t}(j_t) + \varepsilon_{j,t} + \overline{u}_{ss,t}(d_t = s) + \varepsilon_{s,t} & \text{if } d_t = s \quad \text{and } j_t \in J \cup \{0\} \\ \overline{u}_{ss,t}(d_t = ns) + \varepsilon_{ns,t} & \text{if } d_t = ns \end{cases}$$

where $\overline{u}_{ps,t}$ ($\overline{u}_{ss,t}$) represents the flow utility $u_{ps,t}$ ($u_{ss,t}$) excluding the error term.

[10] Note that the use of the max-operator in the above equation constitutes a slight abuse of notation. The consumer will make a choice in the purchase stage that maximizes the present discounted value and not the flow utility. The maximization is therefore with respect to the choice-specific value function $v_{ps,t}$ rather than $u_{ps,t}$. Despite this, the expression captures only utility flows in the current time period. $u_{ps,t}$ is therefore the appropriate term inside the max-operator.

In the model the consumer has potentially two consecutive decisions to take in one time period. He first has to decide whether to search or not ($d_t \in \{s, ns\}$). If he does not search, he does not have to take another decision in the current time period. If he decides to search, he then has to decide which product to purchase (or not to purchase anything): $j_t \in J \cup \{0\}$. I will therefore define two different value functions depending on whether the consumer has searched or not: V_{ss} (the value function in the search stage) and V_{ps} (the value function in the purchase stage). They depend on one another and must be solved simultaneously.

4.2.1 Choice-Specific Value Functions

The choice-specific value function in the purchase stage can be written as follows. The value function is specific to a choice of product j

$$
\begin{aligned}
v_{ps,j}(x_{ps,t}, \widetilde{\epsilon}_{ps,t}) &= \overline{u}_{ps,t}(j_t) + \varepsilon_{j,t} + \beta E\{max_d[v_{ss,d}(x_{ss,t+1}, \widetilde{\epsilon}_{ss,t+1} \mid x_{ps,t}, \widetilde{\epsilon}_{ps,t}, j_t)]\} \\
&= \overline{v}_{ps,j}(x_{ps,t}, \widetilde{\epsilon}_{ps,t}) + \varepsilon_{j,t}
\end{aligned}
$$

where $\overline{u}_{ps,t}$ represents utility excluding the error term and $v_{ps,j}$ and $v_{ss,d}$ are the choice specific value functions for the purchase and search stage respectively. $\overline{v}_{ps,j}$ denotes the choice-specific value function in the purchase stage excluding the error term. Furthermore, $j_t \in J \cup \{0\}$, i.e. the option of no purchase is included. Note that the state variables $x_{ps,t}$ and $x_{ss,t}$ are also specific to the search / purchase stage value function, as different factors will be driving the search and purchase decisions. The value function in the search stage next period is a function of the state variables $x_{ss,t+1}$ and error terms $\widetilde{\epsilon}_{ss,t+1}$. The consumer forms expectations about these variables based on current states $x_{ps,t}$, error terms $\widetilde{\epsilon}_{ps,t}$ and action j_t in the purchase stage.

When entering the store, the consumer has to decide whether to search or not. If he doesn't search he receives utility

$$
\begin{aligned}
v_{ss,d=ns}(x_{ss,t}, \widetilde{\epsilon}_{ss,t}) &= \overline{u}_{ss,t}(d_t = ns) + \varepsilon_{ns,t} + \beta E\{max_d[v_{ss,d}(x_{ss,t+1}, \widetilde{\epsilon}_{ss,t+1} \mid x_{ss,t}, \widetilde{\epsilon}_{ss,t}, d_t = ns)]\} \\
&= \overline{v}_{ss,d=ns}(x_{ss,t}, \widetilde{\epsilon}_{ss,t}) + \varepsilon_{ns,t}
\end{aligned}
$$

If he does search he receives

$$
\begin{aligned}
v_{ss,d=s}(x_{ss,t}, \widetilde{\epsilon}_{ss,t}) &= \overline{u}_{ss,t}(d_t = s) + \varepsilon_{s,t} + E\{max_j[v_{ps,j}(x_{ps,t}, \widetilde{\epsilon}_{ps,t} \mid x_{ss,t}, \widetilde{\epsilon}_{ss,t}, d_t = s)]\} \\
&= \overline{v}_{ss,d=s}(x_{ss,t}, \widetilde{\epsilon}_{ss,t}) + \varepsilon_{s,t}
\end{aligned}
$$

Note, that the expectations regarding the purchase stage choice-specific value functions is not multiplied by the discount factor. Despite the purchase decision happening shortly after the search decision, more information will be available to the consumer. This makes it necessary to include the *expected* utility from the purchase stage in the current time period into the search stage value function.

The value functions can be expressed in terms of their choice-specific counterparts

$$\begin{aligned} V_{ss}(x_{ss,t},\widetilde{\epsilon}_{ss,t}) &= max\Big\{\overline{v}_{ss,d=s}(x_{ss,t},\widetilde{\epsilon}_{ss,t})+\varepsilon_{s,t} \quad , \quad \overline{v}_{ss,d=ns}(x_{ss,t},\widetilde{\epsilon}_{ss,t})+\varepsilon_{ns,t}\Big\} \\ V_{ps}(x_{ps,t},\widetilde{\epsilon}_{ps,t}) &= max_{j\in J\cup\{0\}}\Big\{\overline{v}_{ps}(x_{ps,t},\widetilde{\epsilon}_{ps,t})+\varepsilon_{j,t}\Big\} \end{aligned}$$

4.2.2 Ex-Ante Value Functions

As in Rust (1987), let EV_{ps} (EV_{ss}) denote the expectation of the value function, integrated over the realization of $\widetilde{\epsilon}_{ps,t}$ ($\widetilde{\epsilon}_{ss,t}$).

$$\begin{aligned} EV_{ps}(x_{ps,t}) &= \int_{\widetilde{\epsilon}_{ps,t}} V_{ps}(x_{ps,t},\widetilde{\epsilon}_{ps,t})dP_{\widetilde{\epsilon}_{ps}} \\ EV_{ss}(x_{ss,t}) &= \int_{\widetilde{\epsilon}_{ss,t}} V_{ss}(x_{ss,t},\widetilde{\epsilon}_{ss,t})dP_{\widetilde{\epsilon}_{ss}} \end{aligned}$$

Due to the structure of the model, there is no unique sequence of decision stages. If the consumer decides not to search, then the next decision stage after the search stage in (t) is the search stage in $(t+1)$. If he decides to search instead, the search stage in (t) is followed by the purchase stage in (t). Finally, the time period (t) purchase stage is always followed by the search stage in $(t+1)$. The sequence of decision is illustrated graphically in figure (1). Due to there being three possible sequences of decisions, the following conditional independence assumptions (in the spirit of Rust (1987)) are made

$$\begin{aligned} p(x_{ss,t+1},\widetilde{\epsilon}_{ss,t+1}|x_{ss,t},\widetilde{\epsilon}_{ss,t},d_t=ns) &= p(\widetilde{\epsilon}_{ss,t+1}|x_{ss,t+1})*p(x_{ss,t+1}|x_{ss,t},d_t=ns) \\ p(x_{ps,t},\widetilde{\epsilon}_{ps,t}|x_{ss,t},\widetilde{\epsilon}_{ss,t},d_t=s) &= p(\widetilde{\epsilon}_{ps,t}|x_{ps,t})*p(x_{ps,t}|x_{ss,t},d_t=s) \\ p(x_{ss,t+1},\widetilde{\epsilon}_{ss,t+1}|x_{ps,t},\widetilde{\epsilon}_{ps,t},j_t) &= p(\widetilde{\epsilon}_{ss,t+1}|x_{ss,t+1})*p(x_{ss,t+1}|x_{ps,t},j_t) \end{aligned}$$

Note that in order to apply this simplification, I need to assume that the consumer in the search stage does not know the realization of the purchase stage error terms. Assuming that the error terms $(\widetilde{\epsilon}_{ps,t},\widetilde{\epsilon}_{ss,t})$ are iid extreme value distributed, the integration over the error terms yields the following expressions

$$
\begin{aligned}
EV_{ps}(x_{ps,t}) &= \log\left\{\sum_j \exp(\overline{u}_{ps,t}(j_t) + \beta E\{EV_{ss}(x_{ss,t+1} \mid x_{ps,t}, j_t)\})\right\} \\
EV_{ss}(x_{ss,t}) &= \log\Big\{ \exp\Big(\overline{u}_{ss,t}(d_t = ns) + \beta E\{EV_{ss}(x_{ss,t+1} \mid x_{ss,t}, d_t = ns)\}\Big) \\
&\quad + \exp\Big(\overline{u}_{ss,t}(d_t = s) + E\{EV_{ps}(x_{ps,t} \mid x_{ss,t}, d_t = s)\}\Big)\Big\}
\end{aligned}
$$

The term $EV_{ps}(x_{ps,t})$ can be interpreted as the inclusive value of searching on the shopping trip in time period t, excluding the search cost. In the search stage, the consumer will compare the expected utility of this inclusive value minus the search cost with the utility of not purchasing. This is very similar to the optimal stopping problem in replacement models for durable goods (for example Melnikov (2001), Gowrisankaran and Rysman (2007) or Schiraldi (2011)). The difference is that the consumer has to form an expectation about the current period's inclusive value, whereas in the replacement models the current value is known, but the consumer's expectations about the future evolution of the inclusive value influence the purchase decision.

As mentioned before, the consumer might have to take two consecutive decisions within one time period. If he decides to search, new information will be obtained and he has to make a decision which product to buy. Because of the arrival of new information, the purchase decision will be based on different state variables than the search decision. In the search stage the consumer is therefore forming expectations about the *current period* state variables in the purchase stage. If the consumer decides not to search, he will not have to make a second decision in the same time period and the next decision will be the search decision in time period $(t + 1)$. In terms of state transitions it is therefore necessary to specify $\Pr(x_{ps,t} \mid x_{ss,t}, d_t = s)$, $\Pr(x_{ss,t+1} \mid x_{ss,t}, d_t = ns)$ and $\Pr(x_{ss,t+1} \mid x_{ps,t}, j_t)$. As described above, these are the three possible sequences of decision stage / time-period combinations.

4.2.3 Simplifications and Assumptions

In order to make the problem tractable, the dimensionality of the state space has to be reduced. To achieve this, certain assumptions about the relevant state variables and the formation of expectations regarding future realizations of the state variables are made.

In the search stage, it is assumed that the relevant state variables $x_{ss,t}$ are the consumer's inventory i_t, the identity of the store he is visiting k_t, and the search cost s_t. The consumer does not know the prices p_t of the different brands and pack-sizes at this point, but forms expectations based on the identity of the store k_t.[11] Therefore k_t is part of the state space in the search stage. Both the inventory and search costs directly influence the flow utility in the search stage. Once the consumer has searched, he obtains information about the actual prices p_t. Besides

[11]This assumption relies on an institutional feature of the supermarket sector in the UK: the almost absolute absence of feature advertising. Consumers therefore cannot gather price information before going to the supermarket, and instead have to engage in search within the store.

prices, the consumer's inventory i_t and the vector of product quality terms ξ_j are also part of the state variables in the purchase stage $x_{ps,t}$. The current search cost s_t has already been incurred and is not relevant anymore for the purchase decision. Neither is the store identity as the consumer now knows the actual realizations of prices at the store. s_t and k_t are therefore not part of the purchase stage state space $x_{ps,t}$.

When computing expectations, I assume that the consumer knows the empirical distribution of prices at each store. In other words, expectations regarding prices are simply based on the probability density functions of product/store-specific prices.

$$Pr(p_{jk,t} < \widetilde{p}) \quad = \quad F^p_{jk}(\widetilde{p})$$

Where F^p_{jk} denotes the empirical cumulative density function of prices for product j at supermarket k. The empirical cdf is computed from weekly prices for each product over the whole sample period.

Independent of whether the consumer has searched or not, he forms expectations about future search costs and the identity of the store being visited in the next time period. In other words, these expectations are formed in the same way irrespective of whether the consumer is currently in the search or the purchase stage. The expectations over both the search costs and the store identity in $(t+1)$ do not depend on past realizations of any variable. This is done in order to reduce the computational burden. In principle, a Markov-process could be estimated for the search cost and also the store-visit probability. The expectations regarding the store identity in the next time period are formed by computing the discrete probability distribution of store visits based on all shopping trips (across all consumers) in the sample.

$$Pr(k_{t+1} = \widetilde{k}) \quad = \quad \frac{1}{\#\ Shopping\ Trips} \sum_i \sum_{\tau=1}^{T_i} \mathbf{1}(k_{i\tau} = \widetilde{k}) \quad \text{for} \quad \widetilde{k} \in \{1, 2,\ \ldots\ , 5\}$$

Where T_i denotes the number of observations for each consumer i. Expectations are based on this empirical probability distribution for each of the five supermarket chains ($\widetilde{k} \in \{1, 2,\ \ldots\ , 5\}$) in the data.

As it will be explained in more detail later, search costs are allowed to vary as a function of some shopping trip specific variables and estimated parameters. They therefore vary across consumers and shopping trips. This makes it necessary to define expectations over the future realizations of the search cost. As with prices, expectations regarding the future realization of search costs are based on the probability density function of the search cost which is computed using the search cost realizations across all shopping trips in the sample.

$$Pr(s_{t+1} < \widetilde{s}) \quad = \quad F^s(\widetilde{s})$$

Where F^s denotes the empirical cumulative density function of the search cost. A complication in the case of search costs is the fact that the density function has to be recomputed within each iteration of the optimization algorithm. This is necessary because search costs are a function of parameters that the estimation algorithm is searching over. Therefore, the search costs on each trip in the sample are changing within each iteration of the optimization. This in turn leads to a different distribution of search costs.

Finally, the consumer does not have any uncertainty with respect to the product quality terms ξ_j as they are assumed not to vary over time.

Note that I do not model a Markov-process for prices. Price expectations are therefore not influenced by past realizations of prices. This is necessary as I cannot observe whether the consumer decided not to search or if he searched and did not purchase anything. Within the framework of the model, the only time I know for certain that the consumer obtained price information is when he made a purchase. For any period after the most recent purchase, I can only compute the probability of search, i.e. of obtaining price information, based on the estimates of the model. In other words, I do not know exactly when the consumer's information set was "updated" with new prices. It is in principle still possible to make current prices depend on the last observed price by integrating out over all prices the consumer might have seen since his last purchase. But this implies that all prices since the last purchase will become part of the state space, which would lead to an enormous increase in the computational burden. Therefore, in order to keep the model tractable, the simplification of basing price expectations purely on the identity of the store visited is made. Also, the descriptive statistics presented in Section (2.4) show that promotions do not happen in regular intervals. On the contrary, looking at the numbers presented in Table (4), it seems very difficult for a consumer to predict both the start of the next promotion and the end of a promotion that is currently happening. This suggests that it is difficult for the consumer to keep track of prices over time in order to predict when a product will go on promotion. Ultimately, as in most dynamic demand models, consumer beliefs are imposed as an identification assumption that cannot be directly tested (see Manski (2004) for a discussion of the role of expectations). Assuming that consumers do not condition their expectations on past prices might not be a perfect solution. However, it is not obvious in the context of this model (and the pricing patterns of the product analyzed) whether there is a superior way of modeling expectations. Unfortunately, this way modeling expectation does make it more difficult to apply the empirical framework to other products for which promotions are more predictable. This is a weakness with regards to the generalizability of the estimation approach to other product categories.

I also assume that the transition process for the inventory is known to the consumer. It is determined by the consumption rate and the pack-size of the detergent that was purchased in the previous period (if any was purchased):

$$i_t = max(i_{t-1} - \tau + \Delta i_{t-1}, 0)$$

The pack-size purchased in $t-1$ is denoted by Δi_{t-1}. The only component that is unobserved by the econometrician is the rate of consumption τ, which will be estimated from the data. There is no noise in the transition process. Therefore the consumer perfectly predicts the inventory next period. In case the inventory falls below the consumption rate, the inventory in the next time period will be set equal to zero. For simplicity of exposition, the expectations about the inventory are not going to be explicitly written down in the value functions.

The value function can now be written only in terms of the relevant information at each stage.

$$\begin{aligned} EV_{ss}(i_t, k_t, s_t) \quad &= \quad \log \Big\{ \exp \Big(\overline{u}_{ss,t}(d_t = ns) + \beta E_{k_{t+1}, s_{t+1}} \{ EV_{ss}(i_{t+1}, k_{t+1}, s_{t+1}) \} \Big) \\ &\qquad + \exp \Big(\overline{u}_{ss,t}(d_t = s) + E_{p_t} \{ EV_{ps}(i_t, p_t, \xi) \mid k_t \} \Big) \Big\} \end{aligned}$$

$$EV_{ps}(i_t, p_t, \xi) = \log \left\{ \sum_j \exp(\overline{u}_{ps,t}(j_t) + \beta E_{k_{t+1}, s_{t+1}} \{ EV_{ss}(i_{t+1}, k_{t+1}, s_{t+1}) \}) \right\}$$

Note that in the search stage the expectations regarding current prices are conditional on the store identity k_t, whereas the expectations regarding k_{t+1} and s_{t+1} are unconditional ones (as was explained above). The state variables k_t and s_t are not decision-relevant anymore in the purchase stage as the search cost is sunk and the actual price realizations at store k_t are known.

4.3 Choice Probabilities and Likelihood Function

The probability $P_{s,t}$ that a consumer searches in time period t (the consumer index i is omitted) can now be determined from the following equation

$$P_{s,t} \quad = \quad \int_{\widetilde{\epsilon}_{ss,t}} Pr \Big(\overline{v}_{ss,d=s}(i_t, k_t, s_t) + \varepsilon_{s,t} \quad \geq \quad \overline{v}_{ss,d=ns}(i_t, k_t, s_t) + \varepsilon_{ns,t} \Big)$$

The probability of not searching $P_{ns,t}$ is determined similarly. The same logic can be applied to the probability of purchasing product $\widetilde{j}$ conditional on having searched $P_{\widetilde{j}|s,t}$.

$$P_{\widetilde{j}|s,t} \quad = \quad \int_{\widetilde{\epsilon}_{ps,t}} Pr \Big(\overline{v}_{ps,j=\widetilde{j}}(i_t, p_t, \xi) + \varepsilon_{\widetilde{j},t} \quad \geq \quad \overline{v}_{ps,j}(i_t, p_t, \xi) + \varepsilon_{j,t} \quad , \quad \forall j \in J \cup \{0\} \Big)$$

Given the iid extreme value distribution of the error terms the probabilities can be expressed in the following way

$$
\begin{aligned}
P_{s,t} &= \frac{\exp(\overline{v}_{ss,d=s}(i_t,k_t,s_t))}{\exp(\overline{v}_{ss,d=ns}(i_t,k_t,s_t))+\exp(\overline{v}_{ss,d=s}(i_t,k_t,s_t))} \\
&= \frac{\exp(\overline{u}_{ss,t}(d_t=s)+E_{p_t}[EV_{ps}(i_t,p_t,\xi)\mid k_t])}{\exp(\overline{u}_{ss,t}(d_t=ns)+\beta E_{k_{t+1},s_{t+1}}[EV_{ss}(i_{t+1},k_{t+1},s_{t+1})])+\exp(\overline{u}_{ss,t}(d_t=s)+E_{p_t}[EV_{ps}(i_t,p_t,\xi)\mid k_t])} \\
P_{ns,t} &= 1-P_{s,t}
\end{aligned}
$$

$$
\begin{aligned}
P_{\widetilde{j}|s,t} &= \frac{\exp(\overline{v}_{ps,j=\widetilde{j}}(i_t,p_t,\xi))}{\sum_{j\in J\cup\{0\}}\exp(\overline{v}_{ps,j}(i_t,p_t,\xi))} \\
&= \frac{\exp(\overline{u}_{ps,t}(\widetilde{j}_t)+\beta E_{k_{t+1},s_{t+1}}[EV_{ss}(i_{t+1},k_{t+1},s_{t+1})])}{\sum_{j\in J\cup\{0\}}\exp(\overline{u}_{ps,t}(j_t)+\beta E_{k_{t+1},s_{t+1}}[EV_{ss}(i_{t+1},k_{t+1},s_{t+1})])}
\end{aligned}
$$

J denotes the set of all available pack-sizes. The consumer also has the option $\widetilde{j}=0$ of not purchasing anything.

As the search decision is not observable, it cannot be directly used in the estimation. Instead, only the probabilities that are associated with observable actions are used in the estimation:

$$P_{\widetilde{j},t} = P_{\widetilde{j}|s,t} * P_{s,t}$$

$$P_{no-purchase,t} = P_{ns} + (P_{0|s,t} * P_{s,t})$$

With $\widetilde{j}\in J$ (i.e. the option of not purchasing is not included).

In order to allow for heterogeneity in tastes, I specify a finite mixture model (see for example Heckman and Singer (1984) or Kamakura and Russell (1989)). A different set of preference parameters is estimated for various types of consumers. The probability of belonging to a particular group is estimated jointly with the type-specific parameters. In the estimation I allow for 2 types of consumers. Letting $k\in\{1,2\}$ denote the type and θ_k the type-specific vector of preference parameters, the probabilities can be defined for a specific type.

$$P'_{k,\widetilde{j},t} = P_{\widetilde{j},t}(\theta_k)$$

$$P'_{k,no-purchase,t} = P_{no-purchase,t}(\theta_k)$$

The theoretical probabilities derived above can now be used in order to form the likelihood function.

$$L = \sum_{k\in\{1,2\}}\left(\left[\prod_t\prod_j (P'_{k,j,t})^{y_{j,t}}\right]\left[Pr(type=k)\right]\right)$$

$y_{j,t}$ with $j\in\{no-purchase, purchase-\widetilde{j}\in J\}$ is a variable that takes the value one for the decision actually taken in a particular period and zero otherwise. The probabilities of belonging to a particular type are restricted

to sum up to one. One parameter is therefore estimated in order to pin down the weight assigned to each of the two types. Within the estimation routine, I solve the type-specific value functions using a fixed point algorithm.

4.4 Summary of the Structural Model

In summary the relevant ingredients for the estimation are:

1) the flow utilities:

$$\begin{aligned}
u_{ps,t}(j_t = 0) &= v(c(i_t)) - T(i_t) + \varepsilon_{0,t} \\
u_{ps,t}(j_t \in J) &= -\alpha p_{j,t} + \xi_j + v(c(i_t)) - T(i_t) + \varepsilon_{j,t} \\
u_{ss,t}(d_t = ns) &= v(c(i_t)) - T(i_t) + \varepsilon_{ns,t} \\
u_{ss,t}(d_t = s) &= -s_t + \varepsilon_{s,t}
\end{aligned}$$

2) the expected value functions (in both decision stages):

$$\begin{aligned}
EV_{ss}(i_t, k_t, s_t) &= \log\Big\{ \exp\Big(\overline{u}_{ss,t}(d_t = ns) + \beta E_{k_{t+1}, s_{t+1}}\{EV_{ss}(i_{t+1}, k_{t+1}, s_{t+1})\}\Big) \\
&\quad + \exp\Big(\overline{u}_{ss,t}(d_t = s) + E_{p_t}\{EV_{ps}(i_t, p_t, \xi) \mid k_t\}\Big)\Big\} \\
EV_{ps}(i_t, p_t) &= \log\left\{ \sum_j \exp(\overline{u}_{ps,t}(j_t) + \beta E_{k_{t+1}, s_{t+1}}\{EV_{ss}(i_{t+1}, k_{t+1}, s_{t+1})\}) \right\}
\end{aligned}$$

3) the law of motion for the inventory process:

$$i_t = max(i_{t-1} - \tau + \Delta i_{t-1}, 0)$$

4) the probability density functions used to form expectations over observable state variables:

$$\begin{aligned}
Pr(p_{jk,t} < \widetilde{p}) &= F^{p}_{jk}(\widetilde{p}) \\
Pr(s_{t+1} < \widetilde{s}) &= F^{s}(\widetilde{s}) \\
Pr(k_{t+1} = \widetilde{k}) &= \frac{1}{\#\ Shopping\ Trips} \sum_i \sum_{\tau=1}^{T_i} \mathbf{1}(k_{i\tau} = \widetilde{k}) \quad \text{for} \quad \widetilde{k} \in \{1, 2, \ \dots\ , 5\}
\end{aligned}$$

5) the choice probabilities:

$$
\begin{aligned}
P_{s,t} &= \frac{\exp(\overline{u}_{ss,t}(d_t = s) + E_{p_t}[EV_{ps}(i_t, p_t, \xi) \mid k_t])}{\exp(\overline{u}_{ss,t}(d_t = ns) + \beta E_{k_{t+1},s_{t+1}}[EV_{ss}(i_{t+1}, k_{t+1}, s_{t+1})]) + \exp(\overline{u}_{ss,t}(d_t = s) + E_{p_t}[EV_{ps}(i_t, p_t, \xi) \mid k_t])} \\
P_{ns,t} &= 1 - P_{s,t} \\
P_{\widetilde{j}\in J\cup\{0\}|s,t} &= \frac{\exp(\overline{u}_{ps,t}(\widetilde{j}_t) + \beta E_{k_{t+1},s_{t+1}}[EV_{ss}(i_{t+1}, k_{t+1}, s_{t+1})])}{\sum_{j\in J\cup\{0\}} \exp(\overline{u}_{ps,t}(j_t) + \beta E_{k_{t+1},s_{t+1}}[EV_{ss}(i_{t+1}, k_{t+1}, s_{t+1})])}
\end{aligned}
$$

The flow utilities define the consumer's preferences over the states in any particular time period and decision stage. Using the assumptions regarding expectations and the transitions of observable and unobservable states the value functions can be computed, which capture the expectations about future streams of utility. Both of these ingredients enter the choice probabilities defined above, which are used to form the likelihood function.

5 Estimation and Identification

In order to apply the model to the data, a specific functional form for the utility from consumption and the storage cost has to be chosen. I assume storage costs to be linear in inventory: $T(i_t) = c_{storage} * i_t$ and define consumption to be determined by $c(i_t) = min(\tau, i_t)$, where τ is a parameter to be estimated. The min-operator implies that when the inventory falls below the consumption rate (τ), all the remaining inventory is consumed. Consumption is equal to zero once the inventory is depleted. The utility from consumption $v(c(i_t))$ is not well identified in in my model. This is due to the fact that households in my sample are always consuming detergent (except for stock-outs). I do not model an outside option of choosing not to use any detergent which would allow me to estimate $v(c(i_t))$.[12] Formally, I define $v(c(i_t)) = \nu * c(i_t)$ and set $\nu = 1$.

I also restrict the vector of product qualities ξ_j. A product j is a brand / pack-size combination in the model, but product quality does not vary across different pack-sizes. Therefore, I estimate a set of brand specific, but not pack-size specific, fixed effects and multiply the fixed effect with the pack-size purchased. This is done because the choice of pack-size will be informative about some parameters of the model such as storage and search costs. Estimating a full brand / pack-size specific set of fixed effects would eliminate useful variation and make it difficult to identify those parameters. I normalize one of the brand fixed effects to be zero as identification is only possible in relative terms between brands. Due to the scaling, the estimated fixed effects can be interpreted as the per unit utility of buying a certain brand (relative to the normalized brand).

Finally, the search cost s_t has to be parameterized. I choose the following functional form:

$$s_t = \widetilde{s} * \frac{\exp(x_t'\gamma)}{1 + \exp(x_t'\gamma)}$$

[12]This is similar to the situation of a static demand model without an outside option. In that kind of model it is impossible to estimate a constant term in the utility function. For the same reasons $v(c(i_t))$ cannot be identified.

Where x_t is a vector of variables that reflect the size and composition of the consumer's shopping basket. The three covariates that were introduced earlier in Section (3.1) are used here. Namely, I include total expenditure,[13] the number of other cleaning products and the number of other household products purchased. As discussed earlier, one would expect the variables to have an effect on search costs as they are proxies for how close to the detergent aisle the consumer is on a specific trip.[14] For more details on these variables, see Section (3.1). Note that one might want to include store-level variables such as display and feature as variables that shift the search costs as well. These are used for example in Ching, Erdem, and Keane (2009) and Hendel and Nevo (2006) (although they influence utility rather than search costs), and are shown to have a considerable impact. I do not have data on these store characteristics and cannot therefore include them. However, as explained earlier in Section (4), the use of features and displays is rare in the UK. Therefore the omission of these variables should not be problematic. $\widetilde{s}$ and γ are parameters to be estimated. The functional form makes sure that the second term in the expression has the same sign for all values of x. As the whole expression reflects the search cost on a specific trip, it is expected to be non-negative. This condition is fulfilled as long as $\widetilde{s} \geq 0$. In the case of $\widetilde{s} = 0$, there is an absence of any search costs. A simple t-test on this coefficient will therefore allow to test for the relevance of search costs.

The parameters to be estimated are the price coefficient α, product quality ξ_j, the parameterized functions $c(i_t)$ and $T(i_t)$ and the search cost s_t as a function of the vector x_t. The discount factor is difficult to identify (see Rust (1994)) and is set equal to 0.998, which corresponds roughly to a 10 percent annual interest rate. All parameters except for the search cost shifters are allowed to be different for the two types of consumers. The identification of every one of the parameters will be discussed in the following paragraphs. The main focus will lie on the identification of search costs, as including them in the model constitutes the main contribution of this paper.

Finally, in order to implement the estimation, I need to discretize the state-space that is used when computing the value functions. In particular, I have to discretize the price distribution as well as the search cost distribution and the inventory variable. The probability distribution regarding the store visited next time period is already discrete by construction. I also have to take into account that the initial inventory is unobserved. Therefore, I exclude the first 10 trips for each household from the estimation. Some more technical details regarding the implementation of the estimation are provided in Sections (A.4) and (A.5) of the appendix. Furthermore, various sensitivity checks on the model specification are also reported in the appendix. In the remainder of this section I discuss on an intuitive level which variation in the data helps to identify the parameters in the structural model.[15]

[13] I divide expenditure on each trip by the average household-specific expenditure level. The variable therefore only captures household-specific expenditure variation over time.

[14] I am working under the assumption that the three search cost shifters do indeed represent different levels of search costs rather than transport costs. This issues was addressed in Section (3), I am not able to directly test it within the structural model.

[15]

5.1 Short-Term Fluctuation Due to Variation in the Shopping Basket

This section relies on some of the insights presented in both parts of Section (3). Following the discussion in that section, the size and composition of the shopping basket (captured in the model by the vector x_t and set of coefficients γ) will affect the purchase probability by shifting search costs. The identification of γ comes from an "exclusion restriction": in the absence of any interaction in consumption, there is no reason why the variables in x_t would enter into the purchase decision for detergent of a perfectly informed consumer. Therefore, they cannot influence the choice after the consumer has searched and only enter in the search stage. In other words, once the consumer is in front of the detergent shelf and has looked at the prices, the number and types of other products in his shopping cart will not influence his purchase decision. But the shopping basket will have an influence on whether he goes down to the detergent aisle in the first place, i.e. whether he engages in search. Through their influence on search costs, the variables in x_t will lead to a higher likelihood of buying detergent on particular shopping trips, something that cannot be explained without a search decision. The impact of x_t, i.e. the vector γ, can therefore be identified. $\widetilde{s}$ determines how much *all of the variables* matter relative to other components of the utility function. If x_t explains a large part of the variation in purchase behavior relative to say the price coefficient and other factors, the magnitude of $\widetilde{s}$ has to be relatively high.

Furthermore, the search costs can vary at a very high frequency because of variation in the shopping basket between trips. Because of this, the likelihood of searching for price information might differ dramatically across shopping trips, even if the trips happen within a short period of time. In the most extreme case, a consumer might ignore a promotion on a particular shopping trip and buy the product at the regular (i.e. much higher) price on the next trip. This does indeed happen quite frequently as was shown in Section (3.2). In a model without search (but with inventory holdings), the consumer's price sensitivity will increase gradually as the consumer's inventory decreases. This makes it difficult to rationalize consumers missing promotion and buying at the regular price shortly afterwards. If search is included in the model, the gradual depletion of inventory is combined with the high frequency variation in search costs. As inventory decreases, this will translate into a gradual increase in price sensitivity overlaid with sharp drops at a high frequency. These drops are due to the fact that on a trip where the consumer does not search, he will not react to price changes in any way simply because he does not know what the current price level is. Consumers missing promotions therefore provides an additional source of variation to identify $\widetilde{s}$ and γ.

5.2 Purchase-Timing and the Dynamics of Choice

There is also very rich price variation that can be exploited for the purpose of identification, as prices vary across brands, pack-sizes and over time. Roughly speaking, these three dimensions of price variation allow me to separately identify the price coefficient, search and storage costs. Despite the fact that variation in the composition of the shopping basket can identify search costs, the dynamics of consumers' purchase behavior offer an independent source

of identification. Therefore, it would be possible to estimate search costs (together with other relevant parameters) even in the absence of information on the shopping basket composition.

Firstly, brand choice conditional on pack-size choice and the timing of the purchase allows me to identify the price coefficient. The choice of a particular brand conditional on the pack-size and time of purchase does not provide any information regarding storage costs as the increase in inventory is unaffected by the brand chosen. Similarly, the consumer finds out about prices of *all* products after searching. Because of this assumption, the identity of the brand purchased does not help to make any inference about search costs. (Conditional) brand choice therefore only affects the estimate of the price coefficient. Storage costs have an impact on consumer behavior along two dimensions: 1) Larger storage costs will make consumers less willing to purchase at a lower price when they still have detergent left in their stock. In other words they might ignore a promotion in order to avoid an increase in inventory. 2) Larger storage costs give consumers an incentive to purchase smaller pack-sizes and forego the quantity discount associated with larger packs. Both dimensions of price variation, across pack-sizes and over time therefore have explanatory power for the identification of storage costs. In general, a given level of storage costs will not be able to rationalize consumer choice in reaction to both types of price variation. The inclusion of search costs in the model however allows to match the data better in both respects. Similar to storage costs, search costs also affect consumer behavior along the same two dimensions, but in a slightly different way: 1) Higher search costs make consumers less likely to purchase at a lower price when they still have inventory left. If search costs are high, consumer will not engage in search unless the need to purchase becomes more urgent, i.e. their inventory runs low. Note, that this reason for not reacting to a promotion is caused by a lack of information about the price drop rather than by a desire to keep the inventory low. 2) Higher search costs also provide an incentive to purchase larger pack-sizes due to the fact that larger packs imply less frequent purchases. A lower frequency of purchase incidences in turn implies having to search less often as well. Buying larger pack-sizes therefore leads to savings in future search costs.

Note, that higher search and storage costs both lead to less reaction to promotions when inventory is high. However higher search costs imply purchases of larger pack-sizes, whereas higher storage costs have the opposite effect. Therefore, some combination of search and storage costs will be able to fit the data both in terms of pack-size choice as well as the responsiveness of demand with respect to promotions. A validation exercise conducted later on in Section (7) provides supporting evidence to this end.

5.3 Other Parameters

The rate of consumption is defined by $c(i_t) = min(\tau, i_t)$ in the model. As consumers are observed over a long period of time in the sample, the average rate of consumption is well observed. The parameter τ that represents the speed of consumption is therefore identified. Note that this rate could in principle be calculated manually for each consumer, by dividing the total amount of detergent purchased by the overall time in the sample. However,

the manual calculation is problematic as it would assume an absence of consumer stock-outs. With unobserved stock-outs, the actual consumption rate will be larger than the manually calculated one. Therefore, the rate of consumption is estimated together with the other parameters as part of the consumer's optimization problem rather than calculated outside of the estimation.

The product quality terms ξ_j are identified from across product variation in market shares after controlling for price and other factors. They are essentially brand fixed effects and are identified by the brand-specific mean market shares over time.

5.4 Monte-Carlo Simulation

In order to test whether the model is identified in an alternative way, I also conducted a Monte-Carlo simulation.

5.5 Discussion of Assumptions

Having presented the empirical model and the identification strategy fully, I now shortly discuss the key assumptions made in the estimation. In particular, this sections highlights the similarities and differences relative to the two most prominent stockpiling models in the demand estimation literature, Erdem, Imai, and Keane (2003) and Hendel and Nevo (2006). Secondly, I discuss the additional assumptions necessary to incorporate search into the model in detail.

In order to reduce the state space, I follow Hendel and Nevo (2006) in allowing brand-specific tastes only to enter at the moment of purchase but not at the moment of consumption. This allows me to only track one inventory variable, whereas Erdem, Imai, and Keane (2003) have to include a second state variable. In the model of Hendel and Nevo (2006) this particular assumption also allows them to estimate some parameters of the model statically, which further reduces the computational burden. This is not possible in my setup as the presence of the search decision does not allow to factor the likelihood into a dynamic and a static part. On the upside, not having this separation does loosen up another constraint that Hendel and Nevo (2006) face: they cannot allow for unobserved heterogeneity in the brand intercepts. Although I do allow for only a limited amount of flexibility due to computational constraints, this constitutes an advantage of the modeling framework in this paper. However, if the 2-type finite mixture is not sufficient to capture the true preference heterogeneity in the sample, this might lead to biased parameter estimates. Erdem, Imai, and Keane (2003) describe in detail a "self-selection" bias that could arise when unobserved heterogeneity is not appropriately controlled for in a dynamic model with inventory holdings.[16] I also follow Hendel and Nevo (2006) by not including a stock-out cost parameter. This is done primarily because the cost of stocking out should already be captured by the drop of consumption to zero, which is part of the model. It is not obvious why the consumer would occur any cost in addition to being deprived of the consumption utility.

[16] In order to fully avoid any concerns regarding this type of bias one would also have to model inventory as a function of the products in stock, which is not done in the model presented here. Erdem, Imai, and Keane (2003) deal with this issue by including quality-adjusted inventory as a second state variable.

In order to include the search decision, I split up the consumer's decision into two stages. This modeling assumption is, as argued in this paper, an intuitively plausible way to think about consumer behavior. However, with this type of structure, one has to make stronger assumptions than in a dynamic model with only one decision per time period in order to obtain an analytical solution for the two value functions. In particular, stronger assumptions need to be imposed on the error term structure over time and over decision stages as was mentioned in Section (4.2.2). A separate set of error terms enters in each decision stage of the model, which are unknown prior to entering the particular stage. This implies that the purchase stage taste shocks get revealed after making the search decision, but before making a purchase. These taste shocks capture factors such as product placement on the shelf or an unanticipated liking for a particular packing or product design when arriving in front of the shelf. One issue is that the option to consume an additional error term in the purchase stage gives the consumer an extra incentive to incur the search cost. This is problematic if an additional set of errors does not indeed exist and constitutes a mis-specification of the model. In that case the search cost coefficient would be overestimated as a higher search cost is needed to compensate for the increased attractiveness of search due to presence of the purchase stage error terms. On the other hand, if the purchase stage error terms constitute the correct specification, the search cost term correctly picks-up the fact that without search the consumer forgoes these additional errors. This effect can be interpreted as part of the opportunity cost of searching because the consumer will be able to consume a different set of error terms in another product category if he decides not the search for detergent. While the presence of separate purchase stage error terms is not implausible, this assumption can unfortunately not be directly tested. Due to computational considerations it is difficult to avoid this assumption. Without it, one would not be able to obtain a closed form solution for both value functions.

6 Results

Table 6 presents the results for the main regression. The price coefficients as well as the consumption rate and storage cost terms all have the expected sign and are significant at conventional levels. As for the variables influencing the search costs, the coefficients have the expected sign and are precisely estimated. Search costs are lower when the expenditure is high, which makes intuitive sense as more time is spent in the store. If other products in the same product category are purchased, this also lowers the search costs. This reflects that the consumer spends time in the correct aisle, which reduces the effort to engage in search. Also in terms of relative importance of the search cost shifters, the results seem reasonable as the number of products in the narrower product category lowers the search cost by a larger amount. In terms of magnitude, the search cost varies on the interval $[0, 3.8088]$ for type-1 households and $[0, 6.0530]$ for type-2 households. The upper bound of the interval can be obtained by multiplying the estimated $\widetilde{s}$ term by 0.5, which is the maximum value the logit-term in the search cost term can

take.[17] A shopping trip where each of the trip characteristics takes on its respective average value in the sample has a search cost of 1.9071 (3.0308) for a type-1 (type-2) household associated with it. Using the price coefficient one can evaluate the search cost in monetary terms. It is equal to 2.61 (3.16) pounds, which corresponds roughly to the price of a 900g pack-size. Note, that this is the average search cost over all trips. As the consumer optimally decides when to search, the actual incurred cost will be lower. An additional cleaning product (other than detergent) in the shopping basket lowers the search cost by 0.50 (0.38). An additional household product lowers search costs by a much smaller amount of 0.06 (0.04). A one standard deviation change in the third search cost shifter, the relative expenditure, leads to a reduction in search costs of 0.18 (0.13). When comparing a one standard deviation shift in all the three search cost shifters, the number of other cleaning products has the largest impact. This is therefore the most important aspect of the shopping basket composition with respect to the relevance for consumer search behavior. Finally, the storage cost is relatively low in monetary terms with 0.318 (0.022) pounds per time period.

In order to get a sense for the importance of search costs, it is most informative to look at the fraction of consumers that engage in search. To this end, I need to obtain the market-shares, including the fraction of consumers that search, predicted by the estimated parameters. As the market-shares cannot be obtained analytically, I have to simulate consumer behavior. I implement the simulation by taking draws of the taste shocks in the search and purchase stage for a set of simulated consumers. I also randomly draw the identity of the store and the prices that a consumer faces in a particular time period from the empirical distribution of the variables. Finally, after computing the optimal choice for each consumer, the inventory is updated and the same process is repeated for the next time period. I simulate choices for 5000 consumers of each type over 1000 time periods in this fashion.[18] Aggregation over time periods and consumers and types (using appropriate weights), yields simulated market-shares for all products. More importantly, the simulation also allows me to back out the fraction of consumers that engaged in search in a given time period. This is something that is not observed in the actual data. But based on the estimates and the structure of the model, this fraction can be computed when doing the simulation. In Section (A.7) of the appendix more details on how the simulation is implemented are provided.

When simulating the households' behavior in the way just described, I find that the estimated search cost translates into about 70 percent of consumers not searching in a given time period. In other words, out of about 85 percent of consumers that did not make a purchase in any given time period, only 15 percent knew about prices and decided not to purchase. The much larger fraction of consumers, 70 percent, did not engage in search and therefore did not obtain any price information. This number is important as only the 15 percent of consumers that searched would be able to react if the supermarket lowered the price for any product in a given week. There is therefore a large pool of consumers that is untapped for the purpose of the promotion as long as the supermarket cannot change the incentives to search for price information.

[17] Note, that the three search cost shifters only take positive values and all enter the logit-term with a negative sign. The maximum value the logit-term (not scaled by $\widetilde{s}$) can take is therefore 0.5.

[18] In order to find the steady state distribution for the inventory I use an initialization period of 100 weeks. See appendix, Section (A.7) for more details.

7 Validation

In this section, it will be shown that the model is able to predict consumer behavior along various dimensions that were not directly used in the estimation. The results from the model are compared with the actual data as well as with the predictions from a model without search. The model without search is implemented similarly to the baseline model with search. The only difference is that consumers start each time period in the purchase stage. This removes the search decision. All other elements of the model are the same.[19] The parameter estimates of the model without search are reported in Table (B5) of the appendix. The validation exercise is very closely linked to the way that search costs are identified, in particular the mechanisms described in Section (5.2). Besides lending credibility to the structure of the empirical model, the validation therefore also backs up the identification strategy described before.

Table (7) shows the fit of the two competing models in terms of the choice of pack-size. The market-shares reported here are obtained from the simulation described in the previous section. They are aggregated over brands and supermarkets for each pack-size. As there are no pack-size specific fixed effects included in the model, the pattern of purchases over different pack-sizes is purely driven by storage and search cost. The first column shows that in the data, consumers mainly purchase pack-sizes of 900g and 1900g. The smaller market-share of the 1300g pack and the non-linearity in the shares is mainly due to the fact that 1300g packs were often not available.[20] The model *without search* overpredicts the two smaller pack-sizes greatly, whereas the shares for the two larger ones are underpredicted. The model *with search* presented in this paper replicates the general pattern of purchases better and predicts market-shares better across all pack-sizes. However, even when search is included into the model, the market-shares of the smaller pack-sizes relative to the larger ones is still somewhat overpredicted, but less so relative to a dynamic model without search.

Table (8) compares the elasticities of demand from the two models with the raw data. This is done using promotions for various different products. The pricing patterns used in the simulation are chosen in such a way that they replicate as close as possible the actual pricing patterns.[21] Specifically, the depth of the promotions and their frequency are set equal to the depth and frequency in the actual data in each case. The elasticities are computed by comparing market-shares in promotion periods with those in regular price weeks: The percentage change in market-shares is divided by the percentage price drop in the promotion period. In the case of a promotion for small pack-sizes, presented in the first two rows, the model *with search* predicts an elasticity that is somewhat

[19]The only other difference is that the model without search is estimated without allowing for heterogeneity in the price coefficient. I allow for different coefficients in all other terms for two types of consumers (as in the baseline model with search). This is done for the following reason: When allowing for heterogeneity in the price coefficient, I obtain an extremely large price coefficient for one type. This type of consumer is predicted to almost never make a purchase (on less than 1 percent of his shopping trips). As this simple validation exercise predicts such implausible behavior for one type (the predictions for the second type do not exhibit such patterns) I decided to restrict the price coefficient. The restriction imposed onto the model without search is a drawback for the validation exercise. But, as the predictions from model without search and with heterogeneity in the price coefficient are even worse, I am stacking the cards in favor of the model without search by restricting heterogeneity.

[20]See Table (B2) for more details.

[21]Section (A.7) of the appendix provides details about the depth and frequency of the promotions for the products considered.

smaller than the one in the actual data. The model *without search* instead predicts an elasticity that is almost twice as large as the correct one in both cases. In the case of a promotion for a larger pack-size a similar pattern emerges. The prediction from the model with search are in all cases closer to the true elasticity. However, the search model does lower the elasticity by too much and consistently underpredicts the reaction to promotions across all products considered in the table.

Taken together, the two tables confirm the intuition presented in the identification section. Without search, the storage cost term cannot match both the frequent purchases of large pack-sizes and the lack of reaction to promotions. When including search, this tension can be resolved and both a high demand for large pack-sizes and less responsiveness to promotions can be rationalized. The validation exercise shows that incorporating search helps to achieve more accurate price elasticity predictions and improves the fit of the model regarding the pack-size purchase shares. The overall fit along the two dimension is, albeit an improvement over the model without search, still not very tight. This is a weakness of the empirical model presented here.[22]

Finally, I also report the model fit in terms of the time elapsed between two consecutive purchases. The results are represented graphically in Figure (2). For each purchase occasion, the number of trips since the last purchase are reported on the y-axis. This dimension of the model fit is one that is of key importance for any demand model that allows for consumer stockpiling and the inventory dynamics associated with this behavior. I find that a model *without search* fits the data reasonably well along this dimension, which is consistent with the findings of Erdem, Imai, and Keane (2003), Hendel and Nevo (2006) as well as Ching, Erdem, and Keane (2009). The model *with search* performs better and achieves a closer fit relative to the model without search regarding this aspect of the data.

8 Counterfactuals

I further explore the interaction of promotions and consumer search behavior in this section. In particular, I investigate two changes in the marketing mix used by the store. Firstly, I vary the depth of promotions for a particular product. This is of particular interest, as a change in promotional depth does not only change purchase behavior, but it also alters the consumer's incentive to engage in search. Secondly, I allow supermarkets to accompany promotions with a reduction of search costs. The reduction in search costs will make more consumers aware of the promotion, which changes the elasticity of demand. As in the main simulation, I track the behavior of 5000 simulated households over 1000 time periods for each of the two types when running the counterfactuals. The change in marketing strategy is implemented for one of the most popular products: a 900g pack of Ariel at Morrisons,

[22]In principle, one way of fixing the poor fit along the pack-size dimension would be to add a set of pack-size dummies into the utility function. The stance taken in this paper is that the choice of pack-size should in principle be fully explainable by underlying structural parameters such as search and storage costs. Employing pack-size dummies would therefore constitute a somewhat reduced-form way of fixing the poor pack-size fit of the model. For this reason, I prefer not to go down this route. On the upside, the model with search does make some progress towards improving the predictions of the model while relying on a fully structural specification of consumer utility.

one of the four major supermarket chains. All market-shares reported in the table only refer to consumers visiting Morrisons in a given time period and the market-share definition includes the outside option of not purchasing. In other words, the market-share is not conditional on the consumer having purchased any product on a particular trip, but only conditional on the store visit.

8.1 Change in Promotion Depth

In Table (9) the results from a change in promotion depth are reported. The price of a 900g pack of Ariel at Morrisons is lowered from a regular price of 2.5 pounds by 20 to 50 percent in the different scenarios presented in the table. When implementing the simulation not only the depth of the actual promotion is changed, but also consumers' expectations about the price distribution. For each of these cases, prices are drawn in each time period from the empirical distribution. Only prices for the one product for which promotion depth is changed are drawn from a different distribution for each case. The counterfactuals are implemented in a way that the promotions still occur in exactly the same time periods. Therefore only the promotional price varies depending on the depth. The frequency of the promotion is set equal to 10 percent, a typical frequency in the data.

The average market-share of consumers purchasing the product in any of the promotion periods is reported in the first column of Table (9). Similarly, the second column reports the market-share averaged over all periods in which the product was offered at the regular price. Not surprisingly, the market-share in the promotion periods increases when the price reduction is larger. Interestingly, the market-share in non-promotion periods remains almost unchanged and even increases slightly. In principle, there are opposite forces that drive this change. On the one hand, the larger price reduction makes it more attractive to wait for a promotion rather than to purchase at the regular price. The second channel that operates in this model is the following: the incentive to search increases with depth and this can lead to more purchases even in weeks when the product is sold at the regular price. The two effects turn out to be of roughly equal strength, leading to almost no change in sales in weeks without a promotion. Due to this effect on search behavior, the sales in promotion periods do not cannibalize sales in other time periods and there is therefore a strong effect on the market share averaged over weeks with the two different price levels. This average effect is documented in the third column, and the increase in search activity is reported in column (4). The fraction of consumers searching increases by about 0.2 percent when increasing the depth of promotions from 20 to 50 percent. This might not seem to be a very large effect, but one has to bear in mind that these changes in search behavior are caused by the change in promotion depth of only *one* particular product.

In summary, this counterfactual demonstrates the spill-over effect of promotions in the presence of search costs. Similar to a loss-leader type of strategy, the promotion leads to greater incentives to search and therefore increases category traffic. Interestingly, the increase in category traffic is caused not by *actual* lower prices in any particular time period, but by the expectations about *potentially* lower prices. A greater discount in a promotion period can therefore affect purchases even in non-promotion periods. Exactly this effect is illustrated by the counterfactual:

the spill-over affects positively the sales of the promoted product in time periods when it is not on promotion. This effect could not have been captured in a model without search and is of great relevance for marketing strategy. In principle, the store manager might worry about the fact that consumers simply substitute their purchases over time when they encounter a promotion. In that case, the higher sales in promotion periods would (at least partly) be offset by lower sales in weeks without a promotion. The results from the counterfactuals show that this type of mechanism is not strong enough to reduce the impact of promotions. Rather than losing sales in weeks when the product is sold at the regular price, the store's sales remain essentially unaltered.

8.2 Promotions with Lower Search Costs

The second counterfactual demonstrates the importance for the store to lower search costs when running a promotion. Three different scenarios are presented: the baseline case of a promotion without any change in the search cost and two counterfactual scenarios. In the first one search costs are decreased by 25 percent, in the second one they are lowered by 50 percent. These changes in the search cost are only applied to consumers visiting Morrisons in a week where a 900g pack of Ariel was on promotion. The idea is that the lower search costs are achieved by stores accompanying the promotion with other marketing tools that inform consumers about the promotion. The promotion frequency is set to 10 percent, the depth is equal to 20 percent. These values correspond closely to the ones in the actual data and make the baseline case the same as the baseline in the first counterfactual. As in the previous counterfactual, consumer expectations, in this case with respect to search costs, are adjusted.

In the first three columns of Table (10) the market-share in promotion periods, in weeks with a regular price and an average of the two are reported. Column 4 reports the elasticity and the final column shows the fraction of consumers engaging in search at Morrisons in a period where the product is on promotion. The table shows that lowering the search cost whilst running a promotion hugely enhances the impact of the promotion. The fraction of consumers that engage in search increases from 28 to about 45 percent across the three scenarios. This translates into much higher sales in weeks when the product is on promotion. A reduction of the search costs by 50 percent increases the elasticity of demand almost three-fold. This demonstrates the huge impact of a change in search costs and the possibility of enhancing the effectiveness of promotions by informing consumers.

9 Conclusion

A structural model that incorporates search and inventory holdings was presented in this paper. The framework was applied to detergent purchases, but can be used for the analysis of any storable product. On the methodological side, the paper has shown how to integrate a search decision into a dynamic demand model for a storable product. To the best of the author's knowledge, it is the first model to incorporate both search and inventory holdings into a structural demand framework. However, the complexity of the structural model does not come without a cost.

In order to reduce the computational burden, I make assumptions that are quite strong in (at least) two respects: I include only a limited amount of preference heterogeneity in the model and I model price expectations as being unaffected by past realizations of prices. The latter might be less of an issue since promotions do seem to be difficult to predict in the context of my data. However, one major downside of the assumption is that it is more difficult to apply the model to other product categories where promotions are predictable to a larger extent.

Apart from the methodological contribution, the paper also yields several substantive insights for marketing strategy. When estimating the model, I find that search costs are statistically and economically significant. On around 70 percent of shopping trips, consumers are not aware of prices. Given this lack of knowledge about prices, marketing tools other than pricing, such as advertising and preferential display, become very important. In a counterfactual exercise, I find that lowering the search costs by 50 percent leads to an almost three-fold increase in the elasticity of demand. This shows that there is scope for making promotions more effective by accompanying them with other marketing tools that lower search costs. In a second counterfactual, I find that a change in the depth of promotions can lead to an increase in category traffic. Without knowing when the promotion is happening, the knowledge of the fact that promotions are "deeper" will lead to more consumers engaging in search. This in turn has an impact on sales not only in promotion periods, but also on sales for the same product in regular price weeks. Rather than cannibalizing sales in other time periods when running a promotion, the market-share of the product remains constant in weeks when it is not promoted. This finding runs against the conventional wisdom that the strategic purchase timing of consumers will make promotions less effective.

Based on these findings a natural next step would be to explore optimal marketing strategies when consumer are subject to search costs. The firm can both choose its pricing strategy as well as influence consumer search behavior through other marketing tools. Of particular interest is how informed a firm wants its consumers to be. On the one hand informedness can increase the effectiveness of promotions. On the other hand, uncertainty about promotions increases category traffic even in non-promotion periods. It is therefore theoretically ambiguous how much information the firm would want to provide. The framework presented here opens the door to analyzing these type of supply-side issues.

References

Bell, D. R., G. Iyer, and V. Padmanabhan (2002): "Price Competition Under Stockpiling and Flexible Consumption," *Journal of Marketing Research*, 39(3), 292–303.

Ching, A., T. Erdem, and M. Keane (2009): "The Price Consideration Model of Brand Choice," *Journal of Applied Econometrics*, 24, 393–420.

Erdem, T., S. Imai, and M. P. Keane (2003): "Brand and Quantity Choice Dynamics Under Price Uncertainty," *Quantitative Marketing and Economics*, 1, 5–64.

Goeree, M. S. (2008): "Limited Information and Advertising in the US Personal Computer Industry," *Econometrica*, 76(5), 1017–1074.

Gowrisankaran, G., and M. Rysman (2007): "Dynamics of Consumer Demand for New Durable Goods," unpublished manuscript.

Heckman, J., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2), 271–320.

Hendel, I., and A. Nevo (2006): "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 74(6), 1637–1673.

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

Hortaçsu, A., and C. Syverson (2004): "Search Costs, Product Differentiation, and Welfare Effects of Entry: A Case Study of SP 500 Index Funds," *The Quarterly Journal of Economics*, 119(4), 403–456.

Kamakura, W. A., and G. J. Russell (1989): "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26(4), 379–390.

Kim, J. B., P. Albuquerque, and B. J. Bronnenberg (2011): "Online Demand under Limited Consumer Search," *Marketing Science*, forthcoming.

Koulayev, S. (2009): "Estimating Demand in Search Markets: The Case of Online Hotel Bookings," *FRB working paper*.

Manski, C. (2004): "Measuring Expectations," *Econometrica*, 72(5), 1329–1376.

Mehta, N., S. Rajiv, and K. Srinivasan (2003): "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science*, 22(1), 58–84.

MELNIKOV, O. (2001): "Demand for Differentiated Durable Products: The Case of the U.S. Computer Printer Market," unpublished manuscript.

MORAGA-GONZALEZ, J. L., Z. SANDOR, AND M. R. WILDENBEEST (2009): "Consumer Search and Prices in the Automobile Market," unpublished manuscript.

OSBORNE, M. (2011): "Consumer Learning, Switching Costs and Heterogeneity: A Structural Examination," *Quantitative Marketing and Economics*, forthcoming.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

——— (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. L. McFadden, vol. 4, chap. 51, pp. 3082–3143. Elsevier Science.

SANTOS, B. I. D. L., A. HORTACSU, AND M. WILDENBEEST (2009): "Testing Models of Consumer Search using Data on Web Browsing and Purchasing Behavior," *NET Institute Working Paper*, (09-23).

SCHIRALDI, P. (2011): "Automobile Replacement: A Dynamic Structural Approach," *The Rand Journal of Ecomomics*, forthcoming.

SUN, B., S. A. NESLIN, AND K. SRINIVASAN (2003): "Measuring the Impact of Promotions on Brand Switching When Consumers Are Forward Looking," *Journal of Marketing Research*, 11, 389–405.

WARNER, E. J., AND R. B. BARSKY (1995): "The Timing and Magnitude of Retail Store Markdown: Evidence from Weekends and Holidays," *Quarterly Journal of Economics*, 110(2), 321–352.

Sample	Number of Available Categories	Median Number of Categories per Trip All Trips	Fraction of Categories (Relative to 4 Week Period) All Trips	Fraction of Categories (Relative to 4 Week Period) Trips with Above Average Expenditure	Fraction of Categories (Relative to 4 Week Period) Trips with Below Average Expenditure
Most aggregate	5	2	60.00%	75.00%	50.00%
Medium level of aggregation	27	7	36.00%	61.90%	18.18%
Most disaggregate	226	10	19.57%	39.62%	8.47%

Table 1: **Variation in Shopping Behavior.** The number / fraction of categories in which a purchase is made is reported.

900 g	Mean	Std . Dev.	Min	Max	Weeks Available	Market Share
Ariel	2.59	0.203	2.09	2.94	267	25.43
Bold	2.50	0.194	2.25	3.00	235	23.20
Daz	2.38	0.114	1.88	2.60	267	16.89
Fairy	2.62	0.203	2.42	3.02	267	23.96
Tesco	1.86	0.248	1.00	3.00	312	10.51
1.9 kg	Mean	Std . Dev.	Min	Max	Weeks Available	Market Share
Ariel	4.56	0.484	2.48	5.15	312	35.72
Bold	4.57	0.435	3.29	5.15	310	11.13
Daz	4.31	0.381	2.99	4.79	270	12.38
Fairy	4.72	0.500	3.00	5.15	296	20.55
Tesco	3.43	0.511	2.23	3.98	263	20.22

Table 2: **Descriptive Statistics, Prices and Market Shares (Across Brands).** Note: The maximum number of available weeks is 312 (6 years).

Supermarkets	Asda	Morrisons	Sainsbury	Tesco	
Promotion Frequency	3.60%	8.31%	7.01%	12.92%	
Brands	Ariel	Bold	Daz	Fairy	Tesco
Promotion Frequency	7.55%	6.92%	7.59%	7.00%	24.08%
Pack Size	900g	1320g	1920g	Larger Packs	
Promotion Frequency	5.91%	13.89%	10.45%	n/a	

Table 3: **Frequency of Promotion Periods**

	Mean	Std	Min	Max	R-Square of Regression on Set of Brand and Supermarket Fixed Effect
Length of Promotion Periods (in Weeks)	4.37	2.73	1	14	0.20
Length of Regular Price Periods (in Weeks)	27.68	36.43	1	157	0.28

Table 4: **Length of Promotions / Regular Price Periods.**

Sample (Type of Shopping Trips)	Probability of Purchasing any Detergent on a Shopping Trip All Consumers	Probability of Purchasing any Detergent on a Shopping Trip Only Consumers that Always Used a Car
All Shopping Trips	16.04%	16.51%
Large Shopping Trip	27.49%	28.60%
Small Shopping Trip	7.72%	7.78%
At Least One Other Cleaning Product Was Purchased	28.93%	29.58%
No Other Cleaning Product Was Purchased	4.92%	4.79%
Above Median Number of Other Houesholds Products Was Purchased	28.35%	28.52%
Below (or Equal to) Median Number of Other Houesholds Products Was Purchased	4.75%	4.50%
Number of Households	686	527

Table 5: **The Impact of the Shopping Basket on the Purchase Probability for Detergent.**

		Type 1	Type 2
Price Coefficient		0.7281***	0.9580***
(α)		(0.0137)	(0.0182)
Consumption Rate		0.6446***	0.9207***
(τ)		(0.0030)	(0.0048)
Storage Cost		0.2312***	0.0212***
($c_{storage}$)		(0.0030)	(0.0041)
Search Cost Magnitude		7.6176***	12.1060***
($\widetilde{s}$)		(0.0982)	(0.1950)
Search Cost Shifters	Relative	-0.1450***	
(Not Type-Specific)	Expenditure	(0.0134)	
	Number of Other	-0.3655***	
	Cleaning Products	(0.0100)	
	Number of Other	-0.0413***	
	Household Products	(0.0012)	
Probability Type 1		0.4301	
Observations	Households	686	
	Purchases	18210	
	Shopping Trips	113498	
Log-Likelihood		-80526	
AIC		161096	
BIC		161163	

Table 6: **Estimation Results from the Dynamic Model.** *** denotes significance at the 1 percent level, ** at the 5 percent level and * at the 10 percent level.

	Market Shares		
Source	Raw Data (No Simulation)	Simulation	
Type of Model		With Search	Without Search
900g	50.23	57.12	64.75
1300g	6.18	11.94	13.62
1900g	38.77	27.82	19.18
2100g and larger	3.37	3.13	2.45

Table 7: **Comparison of the Raw Data with a Model with Search and a Model without Search: Market-Shares Across Pack-sizes.** Market shares for different pack-sizes are computed by aggregating over all brands. The market share is computed conditional on the consumer making a purchase.

	Elasticities		
Source	Raw Data (No Simulation)	Simulation	
Type of Model		With Search	Without Search
Promotion of a 900g Pack of Ariel (at Morrisons)	4.21	2.72	7.38
Promotion of a 900g Pack of Tesco's Private Label Brand (at Tesco)	2.80	1.80	4.42
Promotion of a 1900g Pack of Ariel (at Morrisons)	7.94	4.10	14.26

Table 8: **Comparison of the Raw Data with a Model with and a Model without Search: Elasticities of Promotions.**

	Market Share During a Promotion Week	Market Share During a Week With Regular Price	Average Market Share Over All Weeks	Fraction of Consumers Searching (in All Weeks)
20 percent	4.01	2.78	2.90	28.35
30 percent	4.82	2.79	2.98	28.40
40 percent	5.78	2.80	3.07	28.45
50 percent	6.81	2.80	3.17	28.52

Table 9: **Counterfactual: The Change in Market-Share from Deeper Promotions.** The analysis is conducted for a 900g pack of Ariel at Morrisons.

	Market Share During a Promotion Week	Market Share During a Week With Regular Price	Average Market Share Over All Weeks	Elasticity	Fraction of Consumers Searching (During a Promotion Week)
Baseline	4.01	2.78	2.90	2.21	28.46
Search Costs Lowered by 25 Percent	4.99	2.78	2.99	3.96	35.16
Search Costs Lowered by 50 Percent	6.44	2.79	3.13	6.56	45.28

Table 10: **Counterfactual: The Effect of Lower Search Costs on Price Elasticities.** The analysis is conducted for a 900g pack of Ariel at Morrisons.

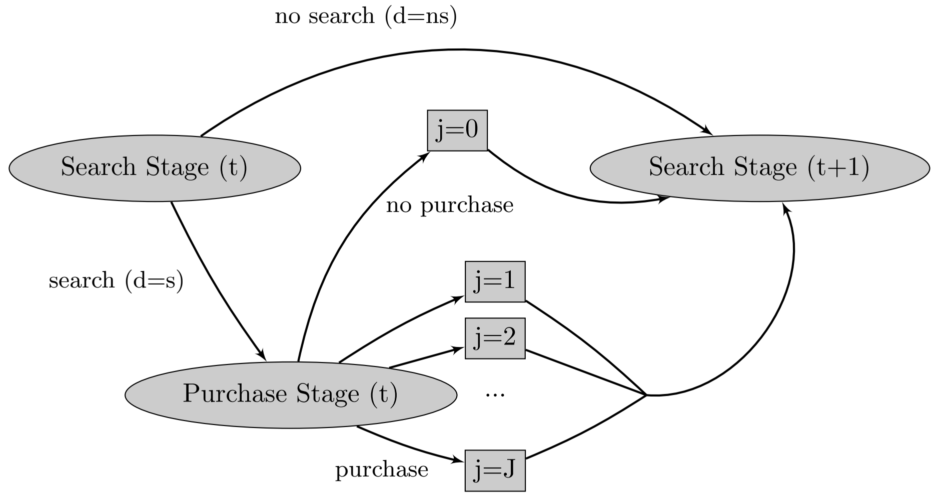

Figure 1: **Timing in the Structural Model.**

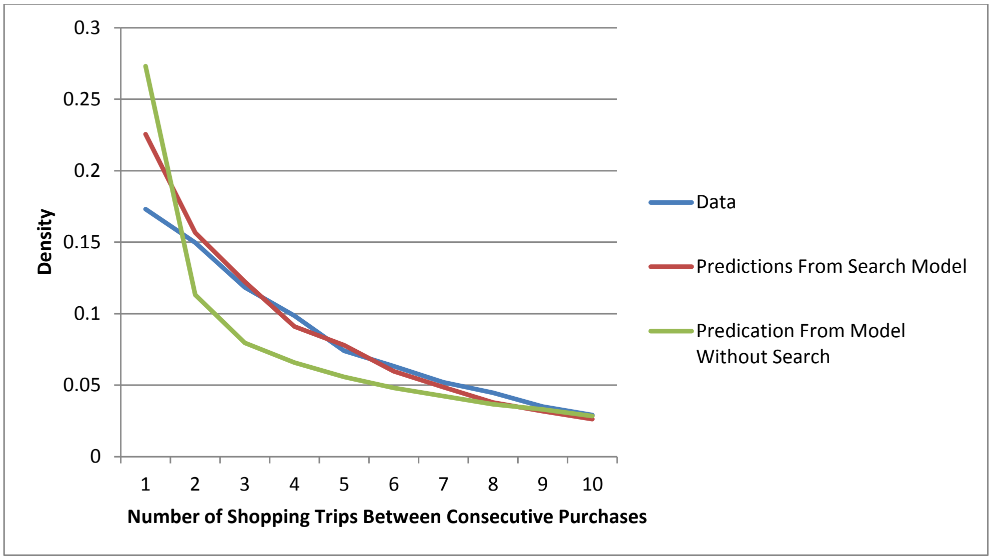

Figure 2: **Predictions of Interpurchase Times**

A Appendix: Estimation

A.1 Household Selection

When selecting the households that are included in the estimation, several criteria (all described in the main text) are applied. This section provides some further justification for the selection criteria and provides details about how the sample size was affected. The full dataset contains about 40000 households; the final sample used in the estimation comprises 686 households.

In a first step all households that were in the sample for less than 20 weeks are eliminated. This is done as information from "un-committed" consumers that spend only a short period of time in the panel might be less reliable. This reduces the sample to roughly 31000 households. I then eliminate all households that bought less than 6 kilograms of detergent per year and households that did not purchase any detergent for a period of at least 16 weeks. This eliminates households with extremely low consumption rates, that possibly visit a launderette some of the time. In the sample 90 percent of households buy between 10 and 35 kilograms the mean being 20 kilograms. This suggests that the 6 kilograms constitute an unusual behavior that the model will not be able to capture. Similarly, a large gap in purchases might be due to the household going on holiday, etc. This also constitutes an unusual behavior that the model cannot capture. These 2 criteria decrease the sample size to 12000. Next, I eliminate all households that bought detergent tablets less than 75 percent of the time. This removes households that primarily purchased other types of detergent such as powder or liquid detergent. I focus on one type of detergent for the purpose of making pack-sizes comparable across products. Tablets are chosen because there are fewer brands in this market. This has the advantage of allowing me to obtain reliable price series and reduces the computational burden as expectation over future prices of each brand have to be formed. This step leaves me with about 2000 households. Finally, I use only households which bought one of the 5 brands for which I construct price series at least 75 percent of the time. As a result, there are 686 households that fulfill all criteria.

Arguably, some of the criteria applied are quite conservative. For example one might try to eliminate longer periods without purchases, but still keep most of the time series for a particular household. This is not done here as I end up with a fairly large sample of households compared to other papers in the literature (for example Hendel and Nevo (2006)). I also have a much longer time series of purchases (6 years) than what is available in other datasets used to analyze demand for similar products. Therefore, very little is lost by eliminating households in a conservative way if one is doubtful about irregularities in their behavior. Also, other papers such as Osborne (2011) take a random sample of all households in order to reduce the computational burden in the estimation. Instead of doing this I prefer to apply the conservative criteria of elimination outlined above.

A.2 Outline of a Model of Shopping Basket Size and Composition Choice

Section (3.1) of the paper shows that the size and composition of the shopping basket has an influence on the purchase probability for detergent. Of course, the size and composition of the whole shopping basket is not truly exogenous, but itself part of a larger optimization problem that the consumer has to solve. In other words, the decision to search and purchase *any* product is part of the consumers decision making process subject to certain constraints. Therefore, one could trace back the reasons for why search costs for detergent vary as a function of the shopping basket composition to more basic underlying primitives. In order to do this, one can think of a model in which the consumer decides whether to search for price information across *all* products in the supermarket (and possibly decides to make a purchase). Although I do not attempt to formally derive (or estimate) such a model, it is still instructive to sketch out the trade-offs inherent in such a model. Assume that the consumer enters the store with a certain need for various products. The immediacy of the need is determined by the inventory of the product he holds at home and his future consumption needs. Also, time spent in the store is costly for the consumer. Even without any formal derivations, one can intuitively think of the type of predictions that can be obtained: 1) When the opportunity cost of time is high, a consumer will only want to stock up on the most necessary items. He will therefore purchase less items on such a trip and engage in less search. Furthermore, he is more likely to purchase perishable goods which need to be stocked up more frequently rather than durable products such as laundry detergent. 2) Assume the consumer has already searched for a product in a particular product category. The marginal cost of searching for a further product within the category is lower than for another product in a different category, due to products typically being arranged by category. This type of cost saving in the time spent searching gives consumers an incentive to lump purchases within a category together on a particular trip. The mechanisms just described will create the type of correlation between shopping basket size and composition with the purchase probabilities for detergent reported earlier in this section. The variation in the shopping basket composition, although being an outcome of the consumer's decision problem, will therefore reflect differences in the search cost for detergent caused by the underlying variation in the time constraint and the consumption needs across various products. In the empirical model, I will treat the basket composition as exogenous to the search and purchase decision regarding detergent. This is a necessary simplification in order to make the model tractable.

A.3 Selection of Trips in Section 3.2 (Consumers Missing Promotions)

In order to compute the percentages presented in the table I first have to define a promotion. As described in Section (2.4) this is done as follows: The 75th percentile of the price distribution of each brand at a particular supermarket is assumed to be the regular price. As promotions are very infrequent and because the regular price varies very little over time, this is an appropriate definition. The 75th percentile will always lie outside of the promotion range of the price distribution. A promotion is defined as a price that is at least 20 percent below the regular price.

I then compute the identity of the product purchased and the price of the product for every purchase made. In the next step, I look up the price for the purchased product on every shopping trip of the same consumer that happened before the actual purchase and after the previous purchase in the detergent category. Any previous trip to a store where the particular product was not available is dropped. This allows me to find out whether the product purchased had been on promotion on any previous trip of the same consumer for any arbitrary time window. No matter which time window is chosen, it will always go no further back than the first trip after the previous purchase. As detergent is purchased very infrequently, this constraint is usually not binding for the time windows used in the table. I also eliminate all trips to supermarkets in the "Other" category as I do not have reliable price information for those trips. They are not considered both in terms of purchases and in terms of possible purchases on previous trips.

A.4 Discretizing the State Variables

In order to implement the dynamic programming problem, I discretize the price distribution as well as the search cost distribution and the inventory variable. The probability distribution regarding the store visited next time period is already a discrete distribution by construction.

In order to cover all the possible price realizations of any of the available products, I construct a grid for prices between 0 and 16 British pounds. A price of zero never occurs, therefore I use this grid-point in order to deal with temporarily unavailable products. Products are made effectively unavailable by assigning an extremely high price (99999 instead of 0) to them, which reduces utility from this option to minus infinity. This will enter the consumer's expectations about futures prices together with the price distribution conditional on availability. I use 25 gridpoints, this makes the grid fine enough in order to capture the typical promotion depth for any pack-size and brand available during the sample period.

In order to discretize the distribution of future expected search costs I do the following: Based on the estimates of $\widetilde{s}$ and the vector γ, I calculate the search costs for all the shopping trips and compute the distribution of the search costs over a set of grid-points. Defining the grid is made particularly easy by the functional form chosen, as the search cost has to be an element of the compact interval $\widetilde{s} * [0, 1]$. I therefore use a grid of values between zero and one and compute the distribution of $\exp(x_t'\beta)/[1 + \exp(x_t'\beta)]$. This term is then multiplied by $\widetilde{s}$. I use 11 grid-points for the search cost distribution in the estimation.

Finally, inventory is discretized using a grid ranging from 0 to 15 kilograms of detergent inventory with 30 gridpoints in the dynamic problem. When constructing the inventory variable for each consumer, I allow the transition (as a function of the consumption rate and the pack-size of a purchase) to be continuous. For every choice available, the expected value function is computed by linearly interpolating between the value functions defined for the closest grid-point to the left and to the right of the actual inventory value (which resulted from the continuous transition process). As the grid is relatively fine, the method of interpolation presumably does not have

a large effect.

A.5 Initial Inventory

When estimating the model I have to deal with the problem of an unknown initial inventory. Note that the consumption rate is convex by construction. Once the inventory falls below the rate (τ), consumption is reduced until it becomes zero when the inventory is completely depleted. Because of this, the impact of the initial inventory will fade over time. I start with the first observed purchase for each household and assume that no inventory was held before that time period. I then calculate the evolution of the inventory implied by the estimated consumption parameter τ and the observed purchases. Only after the first ten trips is the observed behavior used in order to form the likelihood function. This helps to mitigate the initial inventory problem. As a sensitivity check (see main text), I also tried excluding the first 20 trips instead of only 10 from the estimation. This had little impact on the results.

A.6 Sensitivity Checks

I ran several tests in order to check the sensitivity of the estimates. First of all, one might be worried that there exists some reverse causality for the search cost shifters. This type of mechanism is particularly worrisome in the case of the "number of other cleaning products purchased" search cost shifter, as it might be the case that the consumer decides to buy detergent and because of this he also buys other cleaning products. It is presumably less likely that buying detergent will have an influence on the overall trip expenditure or the number of household products, which is a very wide category. Therefore, I re-estimated the model dropping the "number of other cleaning products purchased" as a shifter of the search costs. The results of this regression are reported together with the baseline results in Table (B4) of the appendix. There is relatively little change in the parameter estimates of the maintained parameters. Furthermore, I try to include a quadratic storage cost term for both types of households. Both coefficients are positive but not significantly different from zero.

A.7 Implementing the Simulation

In order to analyze the predictions from the model I need to simulate consumer behavior based on the parameter estimates of the model. To this end, I take draws from the distribution of error terms, determine the optimal choice for each consumer, and aggregate over the choices of all simulated consumers in order to obtain the market shares for each choice. I also randomly draw the identity of the store and the prices that a consumer faces in a particular time period from the empirical distribution of the variables. The probability distributions of store-visits and prices can easily be computed from the raw data. In both cases, the distribution is computed from sample frequencies (of store visits / store-specific prices) and is therefore independent of any parameters of the estimation. I discretize the price distribution for this purpose; the store-visit distribution is discrete by construction. I use a total of 5000

simulated consumers for each one of the 2 types and simulate the behavior over 1000 weeks. Total market shares are calculated by weighing the market share of each type of consumer with the estimated weight of the respective type.

Finally, consumers also differ by the inventory they hold in a particular time period. In order to get sensible results from the simulation I need to know the inventory distribution implied by the model. I find this distribution by starting at an arbitrary distribution. I then simulate consumer behavior for enough time periods such that the distribution reaches a steady state. Specifically, I start by assigning an inventory of zero to each consumer. I then simulate the consumer's behavior over 100 time periods and update the inventory each period according to the rate of depreciation derived in the estimation and the simulated purchases. The inventory changes very little from period to period at the end of the 100 simulated time periods. Therefore, it seems that the impact of the initial inventory should have faded completely after this time span. I then use this "steady state" inventory distribution as the initial inventory for the various simulation exercises in the results section, the validation, and the counterfactuals.

When looking at market-shares in promotion periods and regular price weeks (this is done for the validation and the two counterfactuals), I aggregate the market-shares separately for all promotion and regular price weeks. The elasticity is computed by comparing the percentage difference in market-share with the price change implied by the promotion. I therefore do not use one particular promotion in order to assess the effect on demand. Instead I look at all promotions (which are by construction randomly timed) and average over all promotional weeks. The prices of all other products are drawn from their empirical distribution in every week. The promotion of a particular product can therefore sometimes coincide with promotions for other products as implied by the price distributions. This way of constructing price elasticities is as close as possible to the choice situations consumers face in reality. This makes the elasticities comparable to the ones calculated from the raw data.

Furthermore, in the case of all 3 products reported in Table (8), the depth and frequency of the promotion was set as close as possible to the actual price patterns. A 900g pack of Ariel at Morrisons was promoted 3.4 percent of the time with a 27 percent discount. A 900g pack of Tesco's Private Label at Tesco was promoted 19 percent of the time with a 26 percent discount. A 1900g pack of Ariel at Morrisons was promoted 4.7 percent of the time with a 21 percent discount.

B Appendix: Tables

	Mean	Median	10th Percentile	90th Percentile	Std. Dev.
Expenditure	29.73	18.22	2.73	73.92	30.30
Relative Expenditure	1	0.83	0.13	2.07	0.88
Other Cleaning Products	1.17	0	0	4	1.78
Other Household Products	12.69	8	1	31	12.88

Table B1: **Variation in Expenditure and Shopping Basket Composition.** *Relative Expenditure* is computed by dividing trip-level expenditure by the average expenditure at the household-level.

		Store Weeks	Market Share
Brand			
	Ariel		30.79
	Bold		17.17
	Daz		16.73
	Fairy		21.35
	Tesco		13.96
Pack-Size			
	900 g	1381	50.97
	1300 g	243	6.27
	1900 g	1331	39.34
	2340 g	38	2.27
	2496 g	80	1.15
Store			
	ASDA		19.26
	Morrisons		15.91
	Other		10.44
	Sainsbury's		16.08
	Tesco		38.32

Table B2: **Descriptive Statistics: Aggregate Market Shares.** The maximum number of available weeks is 312 (6 years) for each of 5 supermarket chains. A product could therefore be available for a maximum of 1560 "Store-Weeks".

		Mean	Std. Dev.	Min	Max	Weeks Available	Market Share
ASDA							
	900 g	2.59	0.189	2.42	2.88	265	13.05
	1300 g	3.4	0.272	2.39	3.48	49	2.10
	1900 g	4.76	0.223	3.18	4.98	170	4.11
	2340 g	n.a.					
	2496 g	n.a.					
Morrisons							
	900 g	2.64	0.213	1.99	2.99	269	10.64
	1300 g	3.48	0.259	2.89	3.89	34	0.43
	1900 g	4.73	0.474	3.49	5.39	252	4.17
	2340 g	4.88	0.004	4.88	4.89	24	0.51
	2496 g	5.14	0.092	4.99	5.19	15	0.17
Other Stores							
	900 g	2.77	0.261	1.99	3.47	268	4.76
	1300 g	3.45	0.397	2.39	4.09	64	0.82
	1900 g	4.78	0.609	2.91	8.46	311	3.32
	2340 g	n.a.					
	2496 g	5.25	0.224	4.99	5.82	27	0.09
Sainsbury's							
	900 g	2.75	0.198	2.49	2.99	312	5.02
	1300 g	3.39	0.425	2.39	3.59	51	1.11
	1900 g	4.96	0.485	3.18	5.39	302	9.73
	2340 g	n.a.					
	2496 g	5.35	0.085	5.19	5.39	14	0.23
Tesco							
	900 g	2.62	0.203	2.42	3.02	267	16.76
	1300 g	3.48	0.000	3.48	3.48	45	1.73
	1900 g	4.72	0.500	3.00	5.15	296	17.44
	2340 g	4.88	0.000	4.88	4.88	14	1.73
	2496 g	5.15	0.008	5.15	5.19	24	0.65

Table B3: **Descriptive Statistics: Prices and Market Shares of the Brand Fairy.** The maximum number of available weeks is 312 (6 years).

		Baseline Model		Model with Fewer Search Cost Shifters	
		Type 1	Type 2	Type 1	Type 2
Price Coefficient		0.7281***	0.9580***	0.6962***	0.9933***
		(0.0137)	(0.0182)	(0.0213)	(0.0214)
Consumption Rate		0.6446***	0.9207***	0.6915***	0.9035***
		(0.0030)	(0.0048)	(0.0033)	(0.0076)
Storage Cost		0.2312***	0.0212***	0.2497***	0.0141***
		(0.0030)	(0.0041)	(0.0041)	(0.0043)
Search Cost		7.6176***	12.1060***	7.7921***	10.7468***
		(0.0982)	(0.1950)	(0.1187)	(0.1789)
Search Cost Shifters (Not Type-Specific)	Relative Expenditure	-0.1450***		-0.1736***	
		(0.0134)		(0.0132)	
	Number of Other Cleaning Products	-0.3655***		n/a	
		(0.0100)			
	Number of Other Household Products	-0.0413***		-0.0635***	
		(0.0012)		(0.0015)	
Probability of Type 1		0.4301		0.4379	
Observations	Households	686		686	
	Purchases	18210		18210	
	Shopping Trips	113498		113498	
Log-Likelihood		-80526		-81499	
AIC		161096		163040	
BIC		161163		163104	

Table B4: **Sensitivity Check: Comparison of the Baseline Model with a Model with fewer Search Cost Shifters.**

		Baseline Model		Model without Search Costs	
		Type 1	Type 2	Type 1	Type 2
Price Coefficient		0.7281*** (0.0137)	0.9580*** (0.0182)	1.5955*** (0.0069)	
Consumption Rate		0.6446*** (0.0030)	0.9207*** (0.0048)	1.4023*** (0.0058)	1.4911*** (0.0201)
Storage Cost		0.2312*** (0.0030)	0.0212*** (0.0041)	0.0015 (0.0047)	0.5609*** (0.0523)
Search Cost		7.6176*** (0.0982)	12.1060*** (0.1950)	n/a	
Search Cost Shifters (Not Type-Specific)	Relative Expenditure	-0.1450*** (0.0134)		n/a	
	Number of Other Cleaning Products	-0.3655*** (0.0100)		n/a	
	Number of Other Household Products	-0.0413*** (0.0012)		n/a	
Probability of Type 1		0.4301		0.2711	
Observations	Households	686		686	
	Purchases	18210		18210	
	Shopping Trips	113498		113498	
Log-Likelihood		-80526		-91426	
AIC		161096		182894	
BIC		161163		182958	

Table B5: **Comparison of the Baseline Model with a Model without Search Costs.**