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

Stephan Seiler *

December 3, 2011

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 creates a positive spill-over for other products in the same category, which is a desirable side-effect of the promotion from the store manager's perspective.

JEL Classification: D12, D83, C61, L81

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

*London School of Economics, Centre for Economic Performance and Institute for Fiscal Studies. I would like to thank my advisors John Van Reenen and Pasquale Schiraldi for their invaluable guidance and advice. I am also grateful to Michaela Draganska, Alan Sorenson and Tat Chan who discussed the paper for great feedback as well as participants at EARIE (Ljubljana), QME (Chicago), Workshop on Search and Switching Costs (Groningen), RES (University of Surrey), IIOC (Vancouver), CEPR Applied IO (Toulouse), Marketing Science (Cologne), SED (Montreal), MIO (New York) and seminar participants at the LSE, the Institute for Fiscal Studies, Frankfurt, CREST (Paris), Stanford, UCLA, Rochester, Washington University in St. Louis, Carnegie Mellon, Chicago, San Diego, Zurich and Tilburg. I would also like to thank Rachel Griffith at the Institute for Fiscal Studies for great help with understanding the data and detailed discussions as well as Pedro Gardete, Joachim Groeger, Wes Hartmann, Guenter Hitsch, Claire LeLarge, Fabio Pinna, Peter Rossi, Thomas Schelkle and Philipp Schmidt-Dengler for helpful discussions. Any remaining errors are my own. A previous version of this paper was circulated under the title "A Dynamic Model with Consideration Set Formation".

1 Introduction

Temporary price reductions are used very frequently for grocery items 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. A counterfactual exercise shows that an increase in the price drop during promotion periods increases the fraction of consumers searching. This in turn leads to spill-over effects of the price decrease for one product to sales of other products and also to sales of the same product in non-promotion periods. Put differently, the greater "depth" of the promotion increases category traffic, which in turn increases sales for other products. In a second counterfactual I alter the search cost during a promotion period. Lowering the search cost by 50 percent when running a promotion leads to a more than three-fold increase in the elasticity of demand. Marketing tools that make search less costly could therefore greatly enhance the effectiveness of a promotion by making more consumers aware of the price reduction.

The findings of this paper are important for two reasons. First, the model gives insight into the importance of imperfect information in explaining consumer behavior. 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. According to the findings of this paper, retailers should raise awareness of price whenever they run a promotion. Second, retailers might worry that higher sales during promotion periods are at least partly offset by lower sales in regular price weeks. This can happen due to the fact that consumers strategically time their purchases in the presence of price promotions. The predictions from the model show that this concern is unsubstantiated. I find that deeper promotions do not cannibalize sales in weeks when the product is sold at the regular price. Instead, sales for the promoted product in non-promotion weeks are almost unaffected due to the fact that expectations about more attractive promotions increase category traffic.

This paper's general theme fits into 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 which is also the case here. 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.[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, the model presented here is different in some important ways. 1) In Ching, Erdem, and Keane (2009), search is modeled in a more reduced-form way whereas in this paper the consumer's search decision is explicitly included as part of the structural model. 2) In the "price consideration model", search is used as a computationally-friendly substitute for modeling consumers' inventories in a structural way. This paper instead argues that new insights can be gained from including both search *and* inventory holdings into the structural model. 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.

This paper is also closely related to the literature on consideration sets as this literature also analyzes the impact of imperfect information on consumer choice. Examples include Andrews and Srinivasan (1995), Roberts and Lattin (1991) or Bronnenberg and Vanhonacker (1996). Conditional on purchasing a particular type of product, it is assumed in this literature that consumers do not take all available brands into consideration when making their purchase decision. Due to cognitive limitations, they only compare a smaller set of brands that form their consideration set. The differences in awareness across brands can originate from different sources such as previous consumption of the product or exposure to advertising for a particular brand. Imperfect information therefore has an impact on the choice of product, but not on purchase incidence. This paper instead looks at imperfect information at the purchase incidence level.

The paper is organized in the following way: The next section describes the data. Section three presents some reduced-form results to motivate the use of a particular structural model. Section four presents the empirical

[1]There is a large theoretical literature on search in general and also on search in the context of temporary promotion periods with seminal contributions by Stigler (1961) and Varian (1980). Other papers on sales like Salop and Stiglitz (1982), Assuncao and Meyer (1993) or Pesendorfer (2002) explicitly incorporate stockpiling into the theoretical model, a feature that is also modeled in this paper. Empirical studies such as Bell and Hilber (2006), Sorensen (2000) or Lach (2006) show that some the qualitative predictions from the theoretical models can indeed be found in the data.

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 Worldpanel UK 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. Being able to ignore such a reaction simplifies the analysis. For other products such as food items, one would have to control for this aspect. 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

[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.

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. The latter is important as "second pack half price" promotions are frequent in the UK. The per-unit price therefore changes when several units are purchased. 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 for each store and within each week, except for a few deviations.[3] This confirms that prices from other stores of the same chain can indeed be used to infer prices. For all other stores, a simple average of prices 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

Many households in the sample buy only very small quantities of laundry detergent or none at all. Households with a low volume purchased per year are therefore dropped from the sample. Also households for which I observe no purchases for a very long period of time are dropped. Specifically, I drop households with purchases of less than 6 kilograms of detergent per year (90 percent of households buy between 10 and 35 kilograms, the mean being 20 kilograms) and households who did not buy detergent for a period of at least 16 weeks. The latter might be due to the household going on holiday, etc. In any case, it constitutes an unusual behavior that the model cannot capture. 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.[4] 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 for this sample of households.

[3]Specifically I define a price to be the "correct" price if I have at least 2 observations for a week/supermarket combination and if strictly more than 50 percent of price observations are identical. If I cannot define a price for a given week in this way, the value is interpolated from prices in adjacent weeks.

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

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. Finally, the last 2 columns split trips into above and below average expenditure at the household level. Unsurprisingly, the fraction of categories covered is much lower on shopping trips with low overall expenditure. The difference in the fraction between low and high expenditure trips is about 25 to 35 percentage points depending on the level of aggregation.

Table (1) provides some initial evidence suggesting that consumers are not perfectly informed about prices. Most categories from which the consumer makes a purchase within a 4-week period are unlikely to see a purchase on any particular trip. Even at the most aggregate level, consumers do not cover all categories on each trip. It is therefore very unlikely that they are aware of prices in all categories on all of their shopping trips.

Table (2) is more focused on factors that influence search behavior specifically with respect to laundry detergent. Apart from providing some evidence that consumers might not be perfectly informed about detergent prices, the variables presented will also be used directly in the structural estimation. The first row presents the overall expenditure on a given shopping trip. Consumers spend around 30 pounds on average, the standard deviation is high and the distribution is very skewed with a large mass in the left tail, i.e. a lot of low expenditure trips. The second row presents a transformation of the expenditure variable. The trip-specific expenditure is divided by the average household-specific expenditure. This takes out the household-level variation in expenditure. The variable is not interesting in itself at this point, but is presented as it will be used in the structural model. The final two rows present descriptive statistics on the number of other cleaning products and the number of other household products. Both categories include detergent, the latter one being broader than the former. On about 55 percent of trips consumers do not buy any cleaning product other than detergent and there is also great variation in the number of products in both categories. Overall, these descriptive statistics are the detergent-specific equivalent of Table (1). As detergent tends to be located near other cleaning products (and other households products) in the store, Table (2) gives some indication that on many trips the consumer will not have been in the part of the store where detergent was located. He was therefore most likely unaware of prices within this product category. The information presented in Table (2) will be used in some reduced-form evidence regarding the presence of search costs in Section (3.1). Also, except for the first row, all variables are used as "search cost shifters" in the structural model. A more detailed description of this is provided later on.

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 (3) 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.[5] 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 (B1) and (B2).

One of the most important dimensions of price variation is the variation over time caused by temporary price

[5] 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.

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 (4) 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 (B1) and (B2)). 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 (5) 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.

Finally, a typical pattern of prices over time is shown in Figure (1). The graph shows the evolution of prices for three different pack-sizes of Fairy at Tesco. In this case, only the 1.9kg pack-size was ever promoted. Besides illustrating the pricing patterns, the graph also shows a typical pattern of change in product assortment. For the first 46 weeks of the sample, Tesco sold a 1.3kg pack of Fairy. The supermarket chain stopped selling this product when they included a 900g pack-size in their assortment.

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,[6] (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.[7] 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 and size 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 (6). In the first column and first row, the unconditional purchase probability for laundry detergent is displayed. Given that a consumer is undertaking a trip to the 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.

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

[6] 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.

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

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. As mentioned before, 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 (see Table (2)). 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. In the case of "cleaning products", the median is equal to zero, with no purchase of any cleaning products happening on 55 percent of trips. In the case of "household products", the median number is equal to eight. 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 (6) 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.[8]

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 (6) confirm that the basket size and composition do matter substantially for the purchase probability. This kind of correlation can be explained with a model of consumer search, but not within a static or dynamic model

[8] 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.

that does not allow for search costs. 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]

3.2 Consumers Missing Promotions

Another interesting pattern in the data is the fact that consumers "miss out" on promotions. Quite frequently consumers buy a product at the regular price when they could have bought the product 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 (see for example Erdem, Imai, and Keane (2003) or Hendel and Nevo (2006)). 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. If consumers are price sensitive, then a sharp drop in inventory is needed in order to lead to a reserve price increase that is larger than the gap between the regular and promotional price. An alternative explanation for why we might see consumers ignoring promotions is that they care very little about price. In that case, inventory holdings alone would determine purchases and the timing would be completely independent of prices. The patterns presented in this section show that inventory holdings and / or price (in)sensitivity cannot rationalize consumer behavior. At the same time, the observed purchase patterns are consistent with the presence of search costs.

In Table (7) the pattern of consumer purchases around promotion periods are documented. The table is constructed based on all purchases which happened either (a) at a promotional price or (b) at the regular price, but the product had been available at a promotional price on one of the consumer's previous trips within a certain time period.[10] All purchases that fulfill these conditions can be grouped into three categories: (1) *"Purchase Acceleration"*, the product was purchased at a promotional price but has not previously been available at a promotional price; (2) *"Missing-out"*, the product was purchased at the regular price, but has previously been available at a promotional price; (3) *"No Reaction"*, the product was purchased at the promotional price, but could also have been purchased at a discounted price on a previous trip. The three terms are used only to distinguish the different categories and are therefore used in a loose way. For instance, case (1) might not necessarily represent an acceleration if the promotion simply coincides with the time period in which the consumer would have purchased even in the absence of the promotion. As the following discussion will show, the distinction is still helpful in order to get a sense of the extent of strategic timing of purchases. Promotions are defined as described in Section (2.4).

The first panel of the table shows the percentage of purchases that fall into each one of these categories for two

[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.

[10] See Section (A.3) in the appendix for more details on how the percentages in the table were constructed.

different time windows. The first row only considers previous trips that happened in the same (calendar) week, the second row widens the window to include shopping trips in the previous week. The percentages show that in most cases where a product is purchased on promotion, the consumer could have bought it at a promotional price also on one of his previous trips. A small fraction constitutes purchase accelerations. Most interestingly, there is a substantial fraction of "missed promotion" purchases. In particular with the narrow time window this percentage should be very close to zero if consumers were perfectly informed.

It could also be the case, that the consumer cares little about price. In this situation, the behavior of a perfectly informed consumer could give rise to the purchase pattern found in the data. A price insensitive (but informed) consumer would disregard price in any of his decisions and therefore occasionally make a purchase that falls into the "Missing-Out" category. This alternative explanation has nothing to do with search costs. In order to assess the validity of this hypothesis, I compare the data with a random sequence of purchases. This is helpful as the purchase patterns of a price insensitive consumer should not be different from a randomly generated sequence. If instead consumers care about price, they are expected to accelerate their purchases more often and miss promotions less often relative to the random sequence. The comparison therefore allows to test whether consumers react to intertemporal price differences, i.e. whether they engage in strategic timing of their purchases. The lower panel of Table (7) reports percentages for the three categories used above for a random sequence of purchases. In order to obtain these numbers, I use the actual store visits of the consumers but randomly assign the timing of purchase as well as the identity of the purchased product. This assignment of purchases is based on the purchase probabilities in the actual data. I compute 10 random sequences of purchases and report the average percentage across those 10 sequences in the table. The comparison between the two panels provides a measure for how strongly the consumer's purchase behavior differs from a random timing with respect to the reaction to promotions. In the case of a one week window, the percentage of missed promotion is 12.37 percent in the data compared to 17.89 percent for the random sequence. The percentage of accelerated purchases is higher in the actual data with 13.85 percent of accelerations relative to 10.59 percent. This provides evidence that consumers do engage to some extent in strategic timing of their purchases, i.e. they are not completely price insensitive. At the same time, a price sensitive consumer that has perfect information about prices would be expected to almost never miss a promotion. In terms of benefitting from lower prices, the missed promotions are "low-hanging fruit". Therefore, it would be expected that a consumer not only outperforms the random sequence to some extent on this dimension, but reduces these incidences to zero.[11]

Taken together, the two panels provide evidence that consumers do engage in strategic timing of their purchases, as they "outperform" the random sequence. At the same time they do not react to a significant number of promotions that had been available to them. In a model of perfect information, it is difficult to rationalize why consumers

[11] Instead of the comparison with the random sequence one might be tempted to look at the duration since the last purchase whenever a consumer buys at a promotional price. As Keane (1997) points out, strategic timing might lead to consumers bringing their purchase forward when encountering a promotion. At the same time they might also wait longer than usual in the hope of encountering a promotion in the future. There is therefore no clear prediction as to how the inter-purchase duration is influenced by promotions. For this reason no analysis along those lines is presented here.

only sometimes adjust the timing of their purchase. Search costs imply a discrete jump in the reserve price: Up to a certain point the reserve price is effectively minus infinity as the consumer is not engaging in search and will therefore not react to any level of price. Once he starts searching it will jump up to a strictly positive level. Search costs can therefore create the discrete jump in the reserve price that is needed to rationalize the fact that consumers sometimes miss promotions and buy shortly afterwards at the regular price. The findings are also in line with evidence presented in Murthi and Srinivasan (1999). The authors find that on a large fraction of shopping trips consumers do not react to price changes or other marketing inputs. It is exactly this type of price insensitivity that leads to consumers missing promotions on some trips.

Together with the patterns presented in the descriptive statistics, the reduced-form evidence points strongly 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 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. Note, that more information is available in the purchase stage. The search and purchase decisions therefore have to be modeled as two consecutive decisions. This way of modeling 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. If they had the possibility to obtain price information prior to going to the store, one would have to allow for this in the search model. Luckily, this type of mechanism can be ignored in the context of the UK.

In more technical terms, the consumer's decision process is modeled in two stages. In the first stage (the search stage), the consumer has to decide whether to find out the price of detergent, i.e. to search for the product. In order to do this, he will compare the expected utility from buying at a particular store minus the search cost, with the utility from not purchasing any laundry detergent in the current time period. This optimal search decision, which is embedded in a dynamic structural model, has not been modeled in the previous literature. I am assuming that the consumer will include either all brands in his choice set or none. Therefore, there is no variation in the number of products in the choice set, but only a binary decision whether to consider detergent as a product category or not. If the consumer decides to search, he then has the option of buying one of the available products or nothing at all. This second stage (the purchase stage) is modeled like the purchase decision in a standard dynamic demand model. The timing is illustrated in Figure 2. As in other dynamic models with inventory holdings, I assume that consumers are able to store the good and receive utility from consuming part of their inventory every period.

It is further assumed that the intention to purchase detergent never causes the household to go shopping. Instead, shopping trips are undertaken for reasons exogenous to the decision of buying detergent. This can be justified as detergent makes up only a small fraction (on average about 2 percent in my data) of total expenditure on the average shopping trip. The search process is therefore modeled as a decision to search or not *within* each store for an exogenously given sequence of shopping trips. This is similar to the way Hartmann and Nair (2010) model consumers visiting different stores over time. In order to model consumer choice recursively, I assume that the consumer is uncertain about the identity of the store he will visit the next time period. Therefore, he will form expectations about the identity of the store he will visit on his next shopping trip.

I do not explicitly model optimal consumption behavior. Instead, I assume that households are consuming detergent at a constant rate until they run out. This disregards any effect of purchase behavior on consumption, i.e. households are not allowed to react to promotions by increasing their consumption. This absence of a category expansion effect is a valid assumption for a product like detergent. There is presumably little scope and incentive for adjusting consumption (see for example Bell, Iyer, and Padmanabhan (2002)).

Finally, the cost of search will be allowed to vary across shopping trips. To this end I make the search cost a function of the size and composition of the shopping basket using the variables described in Section (2.3) and Table (2). Specifically, I allow the search cost to vary with the overall expenditure, the number of other cleaning products, and the number of other households products purchased. The reasoning for why these variables qualify as search cost shifters was laid out earlier. More details on how they enter the search cost are provided in Section (5). I do not constrain the direction in which they shift the search cost a priori. In terms of available information, I assume that the search decision is taken at a point in time where the identity of the store as well as the composition of shopping basket and therefore the search cost are known. This decision could either happen in the store or at home when a consumer is writing down his shopping list and decides whether to include detergent on the list.

4.1 Flow Utility

When defining the utility within a particular time period one has to distinguish whether the consumer has already searched or not. The flow utility is therefore defined as the utility in a particular time period and decision stage. In the following equations I will refer to the choice of a consumer on whether to engage in search as the "search stage". After having searched, the consumer then has to make a decision on which product to buy. This choice is referred to as the "purchase stage". Note that for simplicity of exposition, the consumer subscript i is omitted from the following equations. The model does allow for preference heterogeneity across various types of consumers using a finite mixture (see for example Heckman and Singer (1984) or Kamakura and Russell (1989)). Despite this, all the derivations are done in the same way for each type. The heterogeneity will therefore be ignored for the moment and only be introduced into the model at the very end.

I will start by defining the flow utility in the purchase stage, i.e. after the consumer has searched and therefore 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.[12] 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). $\varepsilon_{ns,t}$ 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

[12] 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.

$d_t = ns$ 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.

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 (2). 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. 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 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.

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

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 ($\widetilde{k} \in \{1, 2, \ ... \ , 5\}$) 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 (5), 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. Furthermore, it would not be enough to estimate a Markov-process of prices between shopping trips in this context. As most consumers visit several stores regularly, one would instead have to estimate a store-specific Markov-process for each of the five supermarkets in the sample. Tracking weekly prices over all brands of detergent at several supermarkets implies quite a strong assumption about the consumer's ability to process information. It is therefore not clear whether such an approach would be more realistic in terms of the assumptions it is build on. 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. This is captured in the law of motion by the max-operator. 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

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

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 $P_{\widetilde{j}|s,t}$.

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

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 $v(c(i_t))$. 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. 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 more difficult to identify those parameters. By assuming that the brand fixed effects are scaled up by the pack-size, I am also in line with the modeling assumption in Hendel and Nevo (2006). Due to the scaling, the estimated fixed effects can be interpreted as the per unit utility of buying a certain brand. I normalize one of the brand fixed effects to be zero as identification is only possible in relative terms between brands. Furthermore, the fixed effects are restricted to be all non-negative in order to assure that scaling by the pack-size leads to higher utility for larger pack-sizes.

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

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

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 the variables in the bottom three rows of Table (2): relative 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. 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

[13] See Section (2.3) for details on how this variable is computed.

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 functional form also makes the incorporation of expectations regarding future search costs computationally easy, as the search cost varies on the compact set $\widetilde{s} * [0, 1]$. This makes it easy to define an appropriate grid when expectations about the future value of search costs are computed. This has to be done when solving for the value function.

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 cannot be identified 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 restriction is imposed in order to limit the number of additional parameters to be estimated. The baseline model has 22 parameters in total. 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. Because the search process is not observed, the search behavior will be inferred from the patterns of purchases.

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.[14] 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.

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

[14] Some more technical details regarding the implementation of the estimation are provided in Sections (A.4) and (A.5) of the appendix.

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.

Also, note that the term $\widetilde{s}$, which scales the logistic term in the search cost expression, is identified by the same argument. While the vector of coefficients γ identifies which one of the variables in x_t has the largest impact on search costs, it is $\widetilde{s}$ that 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 reserve price will rise 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 the reserve price 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, the reserve price is effectively minus infinity. 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 a very rich price variation that can be exploited for the purpose of identification, as prices vary across brands, pack-sizes and over time. The price coefficient α will influence brand and pack-size choice as well as the timing of purchases. As will be explained in the following paragraphs, pack-size choice and purchase timing provide identifying variation for storage and search costs and the price coefficient. However, brand choice has explanatory power *only* for 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. In other words, the purchase of a 900g pack of brand A or brand B on a particular trip implies the same evolution of inventory and therefore provides the same information regarding the storage cost. Similarly, the model assumes that the consumer knows about the prices of *all* products after searching. Because of this, the identity of the brand purchased does not help to make any inference about search costs. Brand choice conditional on pack-size choice and the timing of the purchase therefore allows to identify the price coefficient separately from storage and search costs.

Having pinned down the price coefficient, I will now explain how search and storage costs can be separately identified. 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 for the magnitude of the search costs together with storage costs. Therefore, it would be possible to estimate search costs even in the absence of information on the shopping basket composition. The basic idea for the identification is that search costs have two effects in the dynamic model: (1) they make consumers miss promotions as a result of not checking prices, (2) search costs work as a transaction cost, i.e. in their presence consumers prefer to purchase less often. For the purpose of exposition, the following paragraphs illustrate the identification of search and storage costs relative to a given level of the price coefficient. As brand choice identifies the price coefficient, this is without loss of generality.

First of all, consumers have a certain willingness to change the timing of their purchases in reaction to variation of prices over time. They might, for example, purchase detergent earlier than they intended because they encounter a promotion. This behavior of "purchase acceleration" has the disadvantage of increasing inventory above the usual level. On the other hand, the consumer benefits from the lower promotional price. If he waited with his purchase he would avoid the higher inventory, but he knows that the price will most likely have returned to the regular level when he makes the purchase. Figure 3 illustrates the dynamics of this decision. If consumers do *not* react to promotions by accelerating their purchase, this could be explained by high storage costs. In this case, the disutility from the larger inventory outweighs the benefit of the lower price. Alternatively, this behavior is consistent with a high search cost. This implies that the consumer was not aware of the promotion and therefore did not react to it. Only with this type of variation it is not possible to separately identify search and storage costs.

Secondly, consumers have preferences over the available pack-sizes. Big pack-sizes lead to a larger increase in inventory and therefore higher disutility from storage costs. At the same time the per-volume price is lower for larger pack-sizes. If consumers buy bigger packs, this might be due to their low storage costs. In this case, the benefit of a lower per volume price is larger than the disutility from the increase in inventory. Another explanation for this type of behavior are high search costs: Buying a large pack-size increases the time-span until the next purchase. As a purchase always entails that search costs have to be incurred, buying a large pack-size leads to savings on future search costs. Again, search and storage costs offer equally valid explanations for this particular type of consumer behavior.

In general, neither search nor storage costs alone can match the choice behavior in terms of both the pack-sizes purchased and the extent of purchase acceleration. However, the two types of costs have different predictions along the two dimensions just described. A higher storage cost implies *less* purchase acceleration and *less* purchases of large pack-sizes. Increasing the search cost implies *less* purchase acceleration and *more* purchases of larger pack-sizes. Therefore, variation along the pack-size and purchase acceleration dimension can jointly identify search and storage costs. To see this, consider the set of storage/search cost combinations that can rationalize the degree of purchase acceleration in the sample. In a graph with storage and search costs measured on each axis, this can be represented as a negatively slopped line, i.e. a higher storage cost is needed in order to keep the degree of purchase acceleration constant when the search cost is lowered. Similarly the set of storage/search cost combinations that can rationalize a certain share of large and small pack-sizes can be represented by an upward-sloping line in the same space. As long as the two lines are strictly monotonically increasing and decreasing respectively, there is a unique point of intersection. Therefore identification is achieved under the weak assumption of strict monotonicity in the relationships just described.[15]

The descriptive statistics and out-of-model evidence presented earlier show that the purchase patterns are consistent with the presence of search costs, but are difficult to explain with a dynamic model with inventory holdings alone. In particular, Table (B1) shows that larger pack-sizes have a large market-share. In a model with inventory holdings and without search, a big market-share for large pack-sizes has be rationalized by storage costs that are low relative to the price coefficient. In the absence of search costs this would imply also, that consumers will accelerate their purchases when they encounter a promotion. As shown in Table (7) this is not the case, as consumers actually miss out on promotions frequently. The magnitudes of consumers' reactions along the two dimensions are at odds with the inventory model with perfect information. In particular, there is only a small quantity discount (8 to 12 percent between a 900g and a 1.9kg pack-size), but large pack-sizes have a considerable market share. On the other hand, there is a significantly larger discount in promotion periods (20 to 30 percent), but consumers frequently do not react to this sharp decrease in price. The different reactions along those two dimensions of choice make it difficult to rationalize the purchase patterns in a model without search. Taking search costs into account allows this paper to explain a lack of purchase acceleration as well as a large market share for bigger pack-sizes. High search costs will make consumers want to purchase less often. At the same time, they will check prices less often and therefore not react to some promotions. I will come back to choice along these two dimensions in the validation exercise in Section (7).

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 known. 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

[15] In order to confirm that the model is (at least) parametrically identified I also ran a Monte-Carlo simulation and was able to recover the parameters of the model.

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 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.

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 my model. It is not obvious why the consumer would occur any cost in addition to being deprived of the consumption utility. Furthermore, I do not feel confident that the parameter is well identified because it will be mainly pinned down by variation in the latent inventory process. As the inventory process is itself a function of estimated parameters, the inference is very indirect and the construction of the latent process presumably creates noise that will make identification difficult. I model consumption very similar to Erdem, Imai, and Keane (2003), but do not allow consumers to switch between different consumption rates over time. This is another simplifying assumption I make in order to reduce computational costs. The disadvantage of this is that it allows for less flexibility with respect to modeling consumption behavior.

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. In terms of the empirical implementation it has several consequences, some of them desirable, others less so. First of all, the split of the decision leads to a reduction in the computational burden of the model. This might seem surprising as the setup using two value functions, one for each decision stage, seems more complex than a standard dynamic model with only one decision per time period. However, splitting the decisions can in fact lead to a reduction in the state space if some state variables enter only into one of the decision stages, which is the case here. Take for example the search costs and the price vector: If these two state variables were to enter the same decision stage, the value function would have to be defined for every search cost / price combination (defined on some appropriate grid). Instead, with price only entering the purchase stage and the search cost only entering the search stage, the dimension of the "joint state space" of the two value functions is equal to the sum of the grid-points for the state variables in each stage. If all state variables were combined into one decision stage instead, their grid-point dimensions would be multiplied instead of added. This reduction in computational burden is a very convenient side effect of the model structure.

Despite this advantage, 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 directly 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 8 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 relative expenditure is high, which makes intuitive sense as less 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.[16] 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.[17] Aggregation over time periods and consumers and types (using appropriate weights), yields simulated market-shares for all products.

[16] 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.

[17] In order to find the steady state distribution for the inventory I use an initialization period of 100 weeks. See appendix, Section (A.6) for more details.

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.6) 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. This finding inspires one of the counterfactuals presented in Section (8.2). Specifically, the counterfactual shows that lowering search costs when running a promotion greatly enhances the demand response. The results are in line with findings from surveys such as those presented in Dickson and Sawyer (1990) or Gauri, Sudhir, and Talukdar (2008). Both papers find that consumers have limited knowledge about prices and only engage to a limited extent in inter-temporal price search. This is exactly what I find as well, using an entirely different methodology. The estimation results are also consistent with the findings of Ching, Erdem, and Keane (2009) which use a more reduced-form approach to incorporating imperfect knowledge about prices. Taken together it is reassuring that a multitude of approaches all point to the fact that (un)informedness about prices does play a large role in understanding consumer choice behavior.

6.1 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 (B3) 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. Finally, with respect to the initial condition problem, I try excluding the first 20 weeks from the estimation (rather than only the first 10). I find that the results remain qualitatively similar.

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.[18] The parameter estimates of the model without search are reported in Table (B4) 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 (9) 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.[19] 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 (10) 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.[20] 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

[18] 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. One possible explanation for why the fully flexible model without search performs badly is related to the identification arguments given earlier. In a model without search, the price coefficient and storage costs have to match three dimensions of price variation in the data. It is entirely possible (but speculative) that part of the variation that helps identify the search cost term is falsely attributed to heterogeneity in the price coefficient when search is removed from the model. 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.

[19] See Table (B1) for more details.

[20] Section (A.6) of the appendix provides details about the depth and frequency of the promotions for the products considered.

change in market-shares is divided by the percentage price drop in the promotion period.[21] 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 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. The fact that consumers react only in a limited way to promotions will lead to larger storage costs, which in turn leads to a lower predicted market-share for large pack-size. This causes the bad fit for pack-sizes of 1900g or larger presented in Table (9). Inversely, in order to fit the pattern of purchases over pack-sizes, the storage costs needs to be relatively low in a model without search. But at the same time, this leads to a larger predicted response to promotions as shown in Table (10). 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.

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. It seems hard to imagine that consumers have an inherent preference for certain pack-sizes after controlling for price, storage and search costs (as well as brand-specific taste parameters). 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. It is left to future research to shed more light on this issue.

Finally, I also report the model fit in terms of the time elapsed between two consecutive purchases. The results are represented graphically in Figure (4). 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

[21] An elasticity of 4 therefore can be therefore be interpreted as a 4 percent increase in market-share for a 1 percent decrease in price. With a typical promotion depth of 20 percent this would imply an 80 percent increase in market-share. More details on this are provided in Section (A.6) of the appendix.

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

A key feature of this paper is to model consumer search behavior for a storable product which is subject to temporary price reductions. In such an environment consumers have an incentive to time their purchases. At the same time, the presence of search costs limits the extent to which they can engage in strategic purchase timing. Consumers have to pay a search cost in order to find out about prices on the current shopping trip. Only if they are willing to incur this cost will they have the option of reacting to a promotion in the store.

In some counterfactual exercises, 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. Therefore, analyzing promotional depth within the context of this model can shed new light on the effects of this aspect of the marketing mix. Secondly, I allow supermarkets to accompany promotions with a reduction of search costs. Marketing tools such as feature advertising or top-of-the-aisle placement could be employed together with a promotion in order to implement such a strategy. The reduction in search costs will make more consumers aware of the promotion and therefore lead to a bigger 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, 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. The shares therefore have to be interpreted as the shares of *all consumers visiting the store* that purchased certain products.

One downside of the counterfactuals presented here is that I am not able to compute profits. The model will therefore only be used to demonstrate changes in consumer behavior. In the absence of cost information and a supply-side model of profit, the demand reaction cannot be translated into profits. This exercise is outside of the scope of this paper. In future research, it would be very interesting to perform this type of analysis in order to analyze how an optimal promotion should be implemented.

8.1 Change in Promotion Depth

In Table (11) 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 changes I remain entirely within the framework of the empirical model. Therefore, 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 of the market-share of consumers purchasing the product in any of the promotion periods is reported in the first column of Table (11). Similarly, the second column reports the market-share averaged over all periods in which the product is 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. This channel is alluded to in the counterfactuals in Osborne (2010) which also involve varying the depth of promotions. 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. Finally, the last column in the table reports the market-share of other products purchased. Comparing these market-shares to the third column, one can see that higher sales of the promoted product only partially come at the expense of lower sales for other products. Only about 20 percent of the increase in market-share of the promoted product is offset by lower sales for other products.

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 are essentially unaltered in those time periods.

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. This is done in order to be consistent with the more frequent lower realization of the search cost in promotion periods.

In the first three columns of Table (12) 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. How exactly such a policy should be implemented cannot be answered within the model unfortunately. Although the change in the elasticity is large, reducing consumer search costs will be costly for the supermarket. In the absence of any information on how much of a decrease is feasible and how costly it would be, the impact on profits cannot be assessed. Despite this caveat, the big change in the elasticity is suggestive that accompanying promotions with other marketing tools that lower search costs might be desirable.

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.

The empirical model presented in this paper is able to explain certain patterns in the data better than previously used models of dynamic demand that assume consumers are perfectly informed. In particular, the model is able to rationalize the relatively low responsiveness of consumers to promotions. Secondly, the performance of the model in terms of matching consumers' pack-size choices also constitutes an improvement over the perfect information demand model. Despite improving the fit relative to a model without search, the fit is still not particularly tight along those two dimensions. A further improvement of the model's predictive power would be desirable, but is left to future research.

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. According to the findings of this paper, retailers should raise awareness of prices when running a promotion. 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.

References

Andrews, R. L., and T. C. Srinivasan (1995): "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32(1), 30–41.

Assuncao, J. L., and R. J. Meyer (1993): "The Rational Effect of Price Promotions on Sales and Consumption," *Management Science*, 39(5), 517–535.

Bell, D. R., and C. A. L. Hilber (2006): "An Empirical Test of the Theory of Sales: Do Household Storage Constraints Affect Consumer and Store Behavior?," *Quantitative Marketing and Economics*, 4(2), 87–117.

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

Bronnenberg, B. J., and W. R. Vanhonacker (1996): "Limited Choice Sets, Local Price Response and Implied Measures of Price Competition," *Journal of Marketing Research*, 33(2), 163–173.

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

Dickson, P. R., and A. G. Sawyer (1990): "The Price Knowledge and Search of Supermarket Shoppers," *The Journal of Marketing*, 54(3), 42–53.

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

Gauri, D. K., K. Sudhir, and D. Talukdar (2008): "The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data," *Journal of Marketing Research*, 45(2), 226–240.

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.

Hartmann, W. R., and H. S. Nair (2010): "Retail Competition and the Dynamics for Tied Goods," *Marketing Science*, 29(2), 366–386.

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.

Keane, M. (1997): "Current Issues in Discrete Choice Modeling," *Marketing Letters*, 8(3), 307–322.

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*.

Lach, S. (2006): "Existence And Persistence Of Price Dispersion: An Empirical Analysis," *The Review of Economics and Statistics*, 84(3), 433–444.

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.

Murthi, B. P. S., and K. Srinivasan (1999): "Consumers' Extent of Evaluation in Brand Choice," *Journal of Business*, 72(2), 229–256.

Osborne, M. (2010): "Frequency versus Depth: How Changing the Temporal Process for Promotions Impacts Demand for a Storable Good," unpublished manuscript.

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

Pesendorfer, M. (2002): "Retail Sales. A Study of Pricing Behavior in Supermarkets," *Journal of Business*, 75(1), 33–66.

Roberts, J. H., and J. M. Lattin (1991): "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28(4), 429–440.

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

Salop, S., and J. E. Stiglitz (1982): "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents," *American Economic Review*, 72(5), 1121–1130.

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.

Sorensen, A. (2000): "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs," *Journal of Political Economy*, 108(4), 833–850.

Stigler, G. J. (1961): "The Economics of Information," *Journal of Political Economy*, 69(3), 213–225.

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.

Varian, H. R. (1980): "A Model of Sales," *American Economic Review*, 70(4), 651–659.

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.

	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 2: **Variation in Expenditure and Shopping Basket Composition.**

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 3: **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 4: **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 5: **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 6: **The Impact of the Shopping Basket on the Purchase Probability for Detergent.**

Type Of Data	Time Window Prior To Purchase	Fraction of Consumers that Fall into each Category: Purchase Acceleration	Missing-Out	No Reaction
Actual Data	same week	7.27%	4.26%	88.46%
	1 week	13.85%	12.37%	73.78%
Randomly Generated Data	same week	4.13%	9.28%	86.59%
	1 week	10.58%	17.89%	71.52%

Table 7: **Consumer Behavior around Promotion Periods.**

		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 8: **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 9: **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 10: **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)	Market Share of Other Products (in All Weeks)
20 percent	4.01	2.78	2.90	28.35	14.63
30 percent	4.82	2.79	2.98	28.40	14.61
40 percent	5.78	2.80	3.07	28.45	14.59
50 percent	6.81	2.80	3.17	28.52	14.57

Table 11: **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 12: **Counterfactual: The Effect of Lower Search Costs on Price Elasticities.** The analysis is conducted for a 900g pack of Ariel at Morrisons.

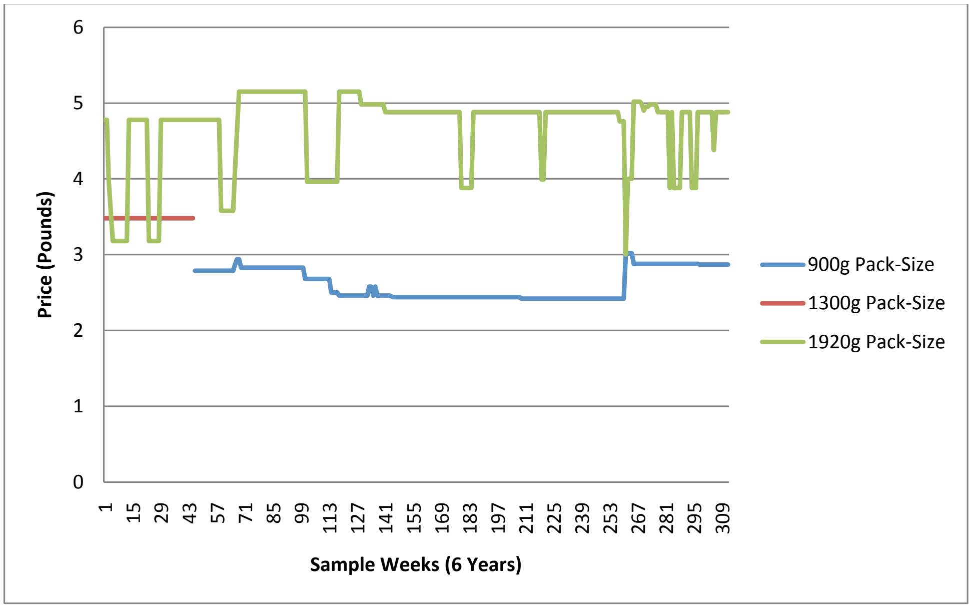

Figure 1: **Price Series Example: Different Pack-Sizes of Fairy at Tesco.**

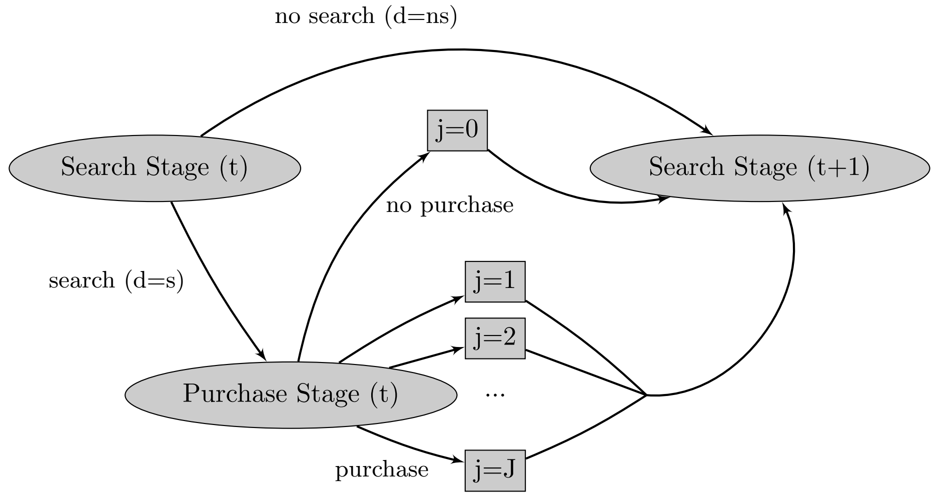

Figure 2: **Timing in the Structural Model.**

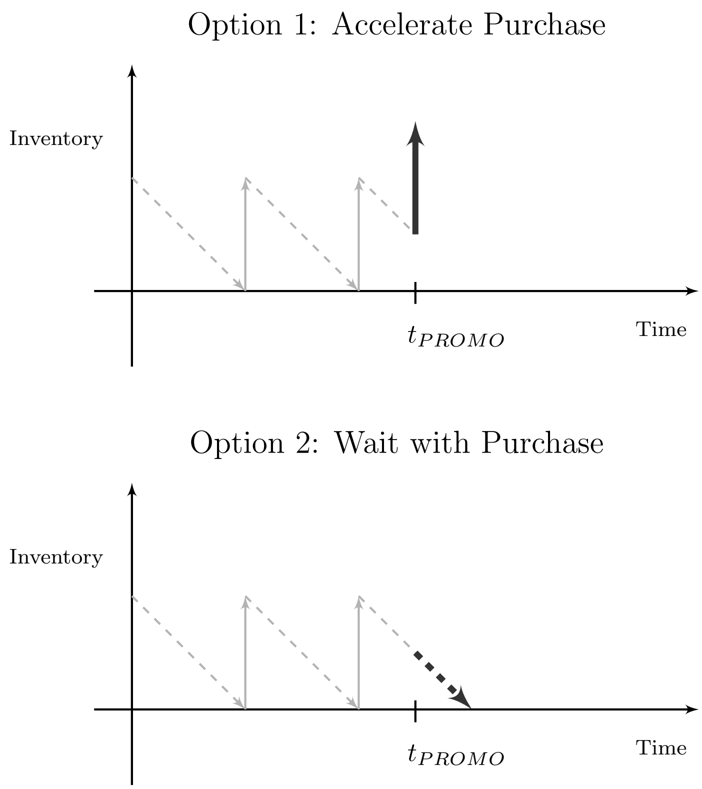

Figure 3: **Inventory Dynamics around a Promotion Period.** The dashed arrows represent a slow decrease in inventory due to consumption. The solid arrows represent a discrete jump in the inventory holding when a new purchase is made. t_{PROMO} denotes the point in time when the promotion occurs.

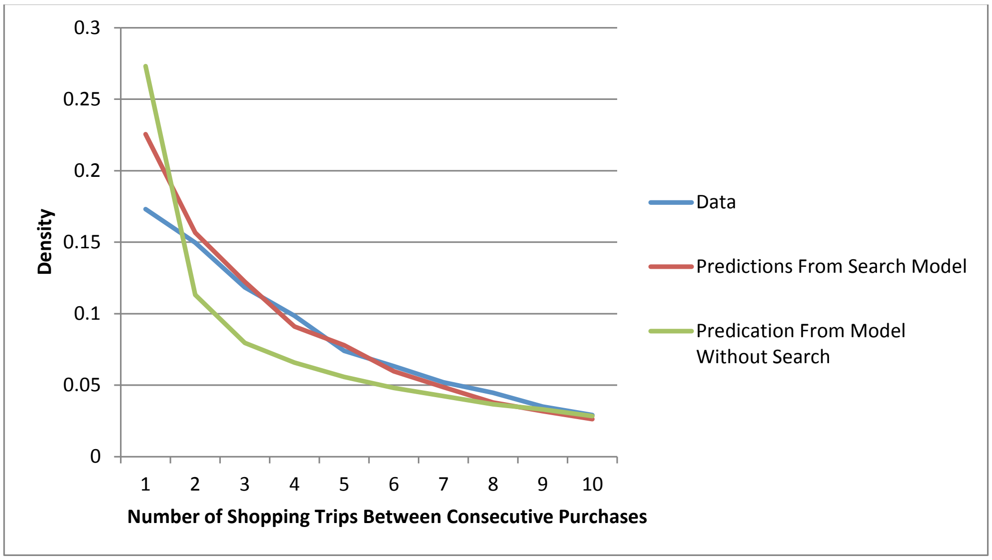

Figure 4: **Predictions of Interpurchase Times**

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. Finally, in many instances the product is purchased at the regular price and was available also at the regular price on previous trips, i.e. there was no possibility in the recent past to purchase the product on promotion. As these situations do not allow for any strategic purchase timing, they are less interesting for the analysis. They are therefore not reported in the table.

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 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 (10), 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

		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 B1: **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 B2: **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*** (0.0137)	0.9580*** (0.0182)	0.6962*** (0.0213)	0.9933*** (0.0214)
Consumption Rate		0.6446*** (0.0030)	0.9207*** (0.0048)	0.6915*** (0.0033)	0.9035*** (0.0076)
Storage Cost		0.2312*** (0.0030)	0.0212*** (0.0041)	0.2497*** (0.0041)	0.0141*** (0.0043)
Search Cost		7.6176*** (0.0982)	12.1060*** (0.1950)	7.7921*** (0.1187)	10.7468*** (0.1789)
Search Cost Shifters (Not Type-Specific)	Relative Expenditure	-0.1450*** (0.0134)		-0.1736*** (0.0132)	
	Number of Other Cleaning Products	-0.3655*** (0.0100)		n/a	
	Number of Other Household Products	-0.0413*** (0.0012)		-0.0635*** (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 B3: **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 B4: **Comparison of the Baseline Model with a Model without Search Costs.**