

A dynamic demand model with consideration set formation

Stephan Seiler *

December 3, 2009

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. When a product is purchased infrequently though, the hassle 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 a large proportion of consumers not being aware of the price of detergent in a given time period. Trip characteristics such as the amount of money spent on other items and the number of products purchased in the same product category cause the search cost to vary across shopping trips. I also compute between-store price elasticities and find that temporary promotions have little impact on competing stores. There is no post-promotion dip in sales. Permanent price reductions lead to a significant shift in market share towards the store that lowered its price. The adjustment of market shares is almost immediate.

1 Introduction

Temporary price reduction 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 wait until the next reduction in price. As many consumers visit several different stores regularly, they will search both across stores and intertemporally. Many grocery items, such as laundry detergent for example, are bought regularly but with relatively long time intervals between purchases. Despite having an incentive to wait for the next promotion period before buying again, it might be too much hassle to check for the price of detergent on every store visit. Finding out the price of a certain product (in particular in a very large supermarket) might actually be quite costly, as the consumer needs to spend time searching for it in the store. Furthermore shopping trips are very heterogeneous in many respects and on some trips only goods from a certain category (food

*London School of Economics, Centre for Economic Performance and Institute for Fiscal Studies.

for example) are purchased. On other trips the overall expenditure is very small and only a few items are purchased and the consumer spends little time in the store. On such trips it is even less likely that consumers are aware of the price of other products.

Most discrete choice model assume that the consumer is aware of all the prices in the store on each visit. For the reasons just outlined this paper proposes a structural model with imperfect information, where consumers engage in costly search. In order to do this, a pre-purchase stage is modeled in which the consumer decides whether to search based on his expected utility from purchasing and the search cost. The novelty of this paper is to integrate the search decision in a structural way into a dynamic demand framework where consumers hold an inventory. This allows me to quantify the search cost as well as identify its drivers. Although I apply the model to laundry detergent data it can be used to analyze demand for any storable product with a relatively long inter-purchase duration.

Furthermore I will explicitly analyze across-store substitution pattern. In a dynamic model with inventory holdings the interaction between stores becomes relevant as a promotion in one store will lead to an increase in inventory and therefore to less purchases on future shopping trips. There is therefore a potential effect of pricing decisions at one store on its competitors in future time periods. This is true even if consumers do not alter their store choice because of the price change.¹ In order to get a complete picture I will include both the frequency of store visits and the identity of the stores visited in my model. This is necessary as both things vary substantially between households. Using the estimates from the model I then compute the contemporaneous and the long-term impact of price changes at a particular store on the store itself and other stores.

I find that the search costs are quantitatively important and equivalent to half a pack of (standard size) detergent in terms of the consumer's disutility. About 70 percent of consumers that do go shopping in a particular time period do not search. The search costs are a negative function of the overall amount of money spend in the store and the number of products purchased in the same category. Search costs are also lower if one or several pack-sizes are on promotion. When simulating the dynamic adjustment to a temporary price change I find a strong contemporaneous reaction but very little impact on future time periods. This is in line with the lack of a post-promotion dip that is well documented in the literature.² There is hardly any negative effect of store-specific promotions on other stores. For permanent price reductions the adjustment is almost immediate with a substantial negative impact on products and stores that did not lower prices.

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

¹Throughout the paper I will treat the store choice decision as exogenous. Any substitution effect across different stores is therefore purely driven by inter-temporal substitution due to stockpiling.

²Many paper describe this phenomenon. Among the papers that explicitly analyze it are Hendel and Nevo (2003), Heerde, Leeflang, and Wittnik (2000) and Neslin and Stone (1996).

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. There are also papers that do estimate search costs such as Hortasu and Syverson (2004) for the mutual fund industry or Hong and Shum (2006) for online purchases of textbooks.³ Mehta, Rajiv, and Srinivasan (2003) estimate search costs for grocery shopping but do not allow consumers to keep an inventory. To the best of the authors knowledge there is no paper that estimates the magnitude of search costs for grocery shopping whilst also taking stockpiling into account.

The demand estimation will be based on Berry, Levinsohn, and Pakes (1995) who estimate a demand system for the US automobile market. Their paper along with many other applications did not consider any dynamic aspects in the consumers' purchase decision. In the more recent literature many contributions try to explicitly include dynamic aspects. Other papers that (like this one) use dynamic demand models to analyze consumers' demand for storable products include Hendel and Nevo (2006), Erdem, Imai, and Keane (2003) and Sun, Neslin, and Srinivasan (2003).⁴

The modeling of the search process in this paper is based on the concept of a consideration set that can be found mainly in the marketing literature. In this literature demand models are used to explore cognitive limitations of consumers when making a purchase decision. For a lot of products many different varieties are available. Comparing the prices and utilities from using the product for all of them is therefore too complex a task for the consumer. Instead he will only consider purchasing certain brands that he is aware of. This awareness can originate from different sources such as previous consumption of the product or exposure to advertising for a particular brand. The set of products for which the consumer compares the utilities and out of which he chooses the best one is called consideration set. Papers such as Andrews and Srinivasan (1995), Roberts and Lattin (1991) and Bronnenberg and Vanhonacker (1996) show that variables that should not influence the utility from consumption like the the familiarity with the store, advertising and shelf positioning still have an impact on the purchase decision by influencing the consideration set. Results often show that the consideration part of the model explains a much larger part of the variation in purchase behavior than choice conditional on the relevant consideration set does (Hauser (1978)). Other papers have looked at imperfect information in a dynamic context. This literature has concentrated on the effects of

³There is indeed a large literature on search in online markets, which I am not going to review. The search process for online purchases works quite differently from the one for in-store search for a particular product.

⁴Interestingly Hendel and Nevo (2006) and Erdem, Imai, and Keane (2003) both allude to the existence of imperfect perfection in the way it is modeled in this paper. They do not model imperfect information and search explicitly though. Hendel and Nevo (2006) point out the strong effect of including a dummy for whether the product is on display in the utility function. This presumably captures that consumers are more likely to be aware of the product when it is on display. This is exactly the way I am modeling the consumer behavior here.

advertising and past consumption on choice (see Akerberg (2003), Erdem and Keane (1996) or Mehta, Rajiv, and Srinivasan (2003)).⁵ None of them models the dynamic aspects arising from inventory holdings though. Papers using consideration sets for markets other than retail include Ben-Akiva and B.Boccaro (1995) who analyze the choice of transport and Goeree (2008) who looks at the US personal computer industry

The paper is organized in the following way: The next sections describes the data. In the third section some reduced-form results are presented to further motivate the use of a particular structural model. Section four presents the empirical model followed by some discussion of the assumptions and results from the regressions. Section six discusses identification, section seven present results on short- and long-term price elasticities. Finally some concluding remarks are made.

2 Data

The data used is a consumer panel provided by the TNS (Taylor Nelson Sofres) marketing research institute. The data is collected at the household level. Each household in the panel is given a scanning device which he uses to scan all products that were purchased. Receipts are then send in to TNS in order to correctly record the price paid for a particular product. The dataset includes about 19,000 household over the period from 10/2001 to 10/2007. An observation is the purchase of a particular product at a particular store on a particular day. Therefore it is also known when a particular household went to a store without buying any laundry detergent (as long as at least one item was bought on the trip).

2.1 Constructing Price Series

As the TNS worldpanel has data at the household level, no store-level dataset of prices is readily available. For the discrete choice demand model used later it is important though, to know the prices of products that were available but were not purchased by the household. Households in the panel are distributed over the whole of the UK, I therefore rarely have several observations for the same store in the same week. In order to infer prices I rely on national pricing policies of the big supermarket chains.⁶ As there exists a large variety of brands⁷ from different manufacturers I observe only very few purchases in each week and supermarket for most products. The construction of reliable price series is only possible though if I observe enough purchases in order to confidently infer the weekly

⁵Only Mehta, Rajiv, and Srinivasan (2003) explicitly refer to a consideration set. The way of modeling unawareness with respect to certain brands in the other paper does capture the same idea though.

⁶Supermarkets 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.

⁷If different pack-sizes are treated as different goods, I observe 2126 (!) different product in the dataset

price. I therefore construct the price series only for the most popular brand in the market, "Fairy Non-Bio Tabs". This brand is available in 6 different sizes and I allow consumers to buy two packs of the same size.⁸ For each pack-size I construct price series for each of the four big chains (Asda, Morrisons, Sainsburys and Tesco) plus a residual category for all other stores. This yields a total of 60 price series. For these the prices are identical for each store and within each week except for very few deviation.⁹ This confirms that prices from other stores of the same chain can indeed be used to infer prices. This is done for Asda, Morrisons, Sainsburys and Tesco. For all other stores a simple average of prices for a particular pack-size in a particular week is taken. Prices for this residual category will therefore be measured less precisely. As about 90 percent of purchases occur at the four big chains this is not very problematic. Table 1 shows some descriptive statistics on the shopping frequencies and pricing at the supermarkets in our sample. The price information is based on a 936 gram pack, the most popular pack-size. There is some heterogeneity in the pricing at different supermarkets that will make it potentially worthwhile for consumers to search across stores.

For the structural estimation it will be important to know how many of the different pack-sizes of Fairy Non-Bio Tabs were on promotion in a particular week. There is a promotion variable that is provided by TNS in the dataset. I will use this information in order to construct a variable in the following way: For each week and supermarket I sum up the number of purchases where the product is flagged to be on promotion and divide this by the number of total purchases. As there is some measurement error in the data, the variable is not always (but most of the time) equal to zero or one. I do this for every one of the 12 different pack-sizes that I allow for and sum the variables up. This captures the overall extent of promotions for the brand.

2.2 Selection of Relevant Households

Many households in the sample do buy only very small quantities of laundry detergent or non at all. Households with a low volume purchased per year are therefore dropped from the sample. Also households for which I observe no purchase for a very long period of time are dropped.¹⁰ I also drop households that are in the sample for less than 20 weeks.¹¹ Finally I am constraint by the availability of price information. Therefore only households which have never bought any product other than

⁸Treating 2 packs of the same size as a different product is important as "2 for 1" promotions are frequent. The per unit price therefore changes when several units are purchased.

⁹Specifically I define a price to be the "correct" price if I have at least 2 observations for a week/store 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. With this method 10 percent of prices are assumed to be mismeasured and replace by the "correct" price.

¹⁰Specifically I drop households with purchases of less than 6 kilogram of detergent per year (90 percent of households buy between 10 and 35 kilogram, the mean being 20) 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. It constitutes an unusual behavior in any case that the model cannot capture.

¹¹TNS deems that information from "un-committed" consumers that spend only a short period of time in the panel is less reliable.

”Fairy Non-Bio Tabs” all the time are kept. This might seem very restrictive and it definitely is so for any analysis of product choice. As I am more concerned with the intertemporal dimension of the purchase decision this is less of an issue here. I end up with 257 households that fulfill all criteria.

2.3 Timing

For the dynamic problem it is crucial to define the relevant time period. Other papers have simply used each shopping trip as an observation.¹² This is problematic as the frequency of shopping trips varies substantially across households and over time. Applying the same discount factor for each time-period and every household is therefore not valid. Furthermore the depreciation of the inventory also has to take into account how much time elapsed since the last trip.¹³ Because of this issue I opt to use ”real time” time periods. As I cannot allow for several trip in the same period and household often go shopping several times per week,¹⁴ I define half weeks as the relevant time period. This way of modeling avoids the problems mentioned above, but it still leaves some households with more than one trip in a half week period. It is possible to pick one trip at random for each of these time periods in order to solve the problem. If detergent was purchased on both trips though, the inventory transition would be incorrect. I therefore move shopping trips in excess of one in a particular time period forward by one period. Given that the time periods are very short this seemed the cleanest way.¹⁵ Note that, when ”moving” an observation its characteristics such as price and the characteristics of the shopping trip also move. Therefore the prices and all other characteristics remain the ones that the household actually faced. I simply pretend that they happened one period later. As I have daily information on trips, I am usually also able to tell the order of trips within one time period.

3 Some Reduced-form Results

This section presents some reduced-form results that are suggestive of the features that will be included in the structural model. Logit-regressions with a dummy equal to one if a particular pack-size was purchased on a particular shopping trip as dependent variable are reported in Table 2. Each shopping trip is an observation whether the household purchased any detergent or not. The utility from the outside option of not purchasing anything is normalized to zero. This corresponds to a standard static discrete choice estimation. It is referred to as ”reduced-form” as neither the inventory formation nor

¹²This is the case in Hendel and Nevo (2006) for example.

¹³In order to define a recursive problem I need the transition process of the inventory for a particular household to be constant over time. This would not be the case here.

¹⁴On average households visit about 1.5 stores each week.

¹⁵Just over 50 percent of observations are not affected by the shifting. 70 percent are shifted by a maximum of 4 time periods (2 weeks). There are some observations that were shifted by more than this though. Luckily the percentiles from the shopping trips carry over to the household level. I.e. 70 percent of households have a maximum shift of 4 time periods. For a sensitivity check households with a lot of shifts could be excluded.

the consideration set are explicitly modeled. Nevertheless some effects that we would expect to see if these two features are indeed relevant can be illustrated with the regressions. As the per-unit price for different pack-sizes differs, I do include dummies for all 12 pack-sizes. The price coefficient is therefore purely identified from within pack-size variation in price.¹⁶

Column (1) shows the result when only price and the control variables are used as independent variables. The price coefficient has the expected sign and is highly significant. More interestingly, the relative expenditure on a particular shopping trip also has a significant influence on the probability of purchasing detergent. If a household spends more money overall on a particular shopping trip, he is more likely to purchase detergent. This can be interpreted as capturing the time spent in the store.¹⁷ The coefficient in column (2) is highly significant. In column (3) a dummy variable equal to one if a particular pack-size was on promotion is included.¹⁸ This has the expected positive impact and is highly significant. Column (4) include two variable that capture whether the consumer has purchased products from the same product category. The first variable is the number of cleaning products other than detergent that the consumer has purchased. The second variable is the number of products within the category "household goods", which includes cleaning products but is a broader category. Buying other products from the same category has a positive impact on the purchase probability, the narrower category being more important than the broader one.

Column (5) includes two variables into the regression that capture the effect of stockpiling. The more time elapsed since the last purchase, the higher the likelihood of another purchase. The higher the volume bought previously, the less likely is another purchase. This confirms that stockpiling is important.

In column (6) all the variables are used in the same regression. The qualitative results for all the coefficients are the same as for the regressions described above. The results are also robust to including a full set of household dummies, as column (7) shows.¹⁹

In the standard discrete choice model consumers are assumed to have perfect information about the prices of all available products at every point in time where they can make a purchase (i.e. on

¹⁶I also include a set of control variables. The frequency of trips to a particular supermarket defined as (number of shopping trips to a particular supermarket) / (total number of shopping trips) enters positively as visiting a store more often leads to more opportunities to purchase the product there. The overall frequency of shopping trips enters negatively as for a given consumption of detergent more shopping trips will lead to more trips without a purchase. Household size captures the heterogeneity in consumption and enters positively. A set of supermarket dummies is also included in all the regressions. None of these control variables is reported in the table.

¹⁷Relative expenditure is defined as the expenditure on a particular trip divided by the average expenditure over all shopping trips for a certain household. This adjusts the measure for across household variation in average expenditure and is therefore a better measure for time spend in the store. The money spend on laundry detergent is excluded from these measures of expenditure in order to avoid expenditure to be higher because of the detergent purchase itself. The expenditure measure is therefore exogenous with respect to the laundry detergent purchase.

¹⁸Note that this is a pack-size specific variable. It is not the brand-level promotion variable that was defined in the previous section.

¹⁹These enter all the options except for the outside option and therefore capture the household specific average need for detergent. Note that some of the controls used in the other regressions do not vary within each household and therefore have to be dropped.

every shopping trip in the case of a grocery shopping item). The fact that consumers are more likely to purchase detergent when they spend more money on other items is at odds with this assumption. Under perfect information there is no reason why the characteristics of the shopping trip should determine which products are purchased (Bronnenberg and Vanhonacker (1996)). Price and other product characteristics alone should be a sufficient statistic.²⁰ The regression results confirm though, that trip characteristics do matter for the purchase probability. If a consideration stage is included and it is costly to search for the product this result can be rationalized. For example: If the probability of including detergent in the consideration set is higher when the consumers spends more money in the store, this will lead to a higher probability of purchasing. This kind of correlation can only be explained with a model of search and not within a static or dynamic model with perfect information.

One worry might be that consumers do buy detergent only on particular trips because they want to avoid carrying the product home on certain trips. This alternative explanation has nothing to do with search and would still lead to a correlation of trip characteristics with the purchase probability. It is not clear though, why the variables used in the regression would capture whether it is more or less burdensome to carry the product. In order to assess the the validity of this theory I do include the distance to the store into the regression in column (8).²¹ In column (9) I use only shopping trips that occurred on a weekend as distance will be a better measure for those trips.²² The distance to the store has no significant influence in either of the two regressions. The coefficients on the other variables do not change dramatically and remain significant.²³ Although the distance to the store is an imperfect measure for the burden of carrying the product, the results are indicative of this not being much of an issue.

The results from the reduced-form regression suggest that both stockpiling and a consideration decision are relevant for the purchase of laundry detergent. This provides further evidence that these aspects have to be included in a structural model that models consumer behavior in a realistic way.

4 The Structural Model

In order to model the consumers' behavior I assume that they are able to store the good and receive utility from consuming part of their inventory every period. Their decision process is modeled in two-stages. In a first stage the consumer has to decide whether to include laundry detergents in his

²⁰As I use dummy variables for each pack-size, product characteristics are controlled for in a very conservative way.

²¹As there is missing data for both supermarket and consumer location the number of observations drops in this regression.

²²During the week consumers might shop close to their workplace or on the way home. The distance from the household's address to the supermarket is therefore not very informative. For weekends it presumably provides more useful information.

²³Except for the coefficient on relative expenditure in column (9). Presumably this happens because weekend trips are associated with higher expenditure. Using only shopping trips on weekends eliminates some of the identifying variation for this coefficient.

consideration set or not. 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. As I am using data on only one brand of detergent, the consumer will include either all pack-sizes in his consideration set or none. There is therefore 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.²⁴ If the consumer decides to consider purchasing the good, he then has the option of buying one of the available brands or not to buy anything. This second stage looks like the purchase decision in a standard dynamic demand model.

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 on the average shopping trip a basket of products is bought out of which detergent makes up only a small fraction. This also accords with findings that there are (if anything) small effects on store-traffic from promotions on individual items (see Urbany, Dickson, and Sawyer (2000), Walters (1991) or Walters and Rinne (1986)). The search process is therefore modeled as a decision to search or not *within* each store for an exogenously given sequence of shopping trips.²⁵ The cost of visiting a particular store is not something I attempt to model here. With exogenous store choice it is assumed not to have any influence on the search cost.

Furthermore the cost of search will be allowed to vary across shopping trips. Mitra (1995) provides some experimental evidence that consideration sets are not stable over time, but vary across purchase occasions. This fact has been captured in many papers by variables such as whether the product was on promotion or on display as well as the positioning within the shelf or aisle (see for example Bronnenberg and Vanhonacker (1996) or Allenby and Ginter (1995)). These variable will pick up how likely it is for a consumer to consider a particular brand. In this model I am analyzing whether the whole category is considered. I will therefore use data on how costly it is to consider the category as a whole, using the same trip characteristics variables that were also used in the reduced-form regressions.

I will assume that the search decision is taken at a point in time where both the identity of the store as well as the search cost and therefore the trip characteristics are known. In terms of timing this decision could either happen in the store or at home when consumer is writing down his shopping list and decides whether to include detergent in the list. The only constraint that I am imposing is that store identity and trip characteristics are already in the consumer's information set.

²⁴In most of the marketing literature on consideration sets the number of product included in the set was the main focus. Here the consideration decision is modeled purely as a category level decision similar to Ching, Erdem, and Keane (2006)

²⁵For many other (more expensive) products the consumers will choose to go to different stores specifically in order to find information about a certain product. This is very different from the way search is modeled here.

4.1 Flow Utility

The per period utility when no good is purchased is defined in the following way:

$$u_{0,t} = v(c(i_t)) - T(i_t)$$

i_t denotes the inventory in period t and $c(\cdot)$ consumption as a function of the inventory, where $c(\cdot)$ is increasing at a decreasing rate in its argument. This reflects the fact that when the inventory decreases, the consumer will use the remaining inventory for a longer time as he wants to avoid completely running out of it. $v(\cdot)$ is the utility received from consumption. $T(\cdot)$ reflects the cost of inventory holdings. This is also the flow utility a household obtains in a period in which he does not go shopping at all.

The formula for the per period utility when a product is purchased is:

$$u_{j,t} = -\alpha p_{j,t} + v(c(i_t)) - T(i_t)$$

This is equivalent to the expression above, but for the inclusion of a negative utility term from having to pay price $p_{j,t}$. A good is in this case defined as a particular pack-size of Fairy Non-Bio tabs.²⁶ The consumer does not derive any utility directly from the purchase, but the increased inventory will give him utility from consumption in future periods.

4.2 Value Function

In order to model the dynamic decision process of the consumer I define an indicator variable S , which takes the value 0 in the consideration stage and 1 in the purchase decision stage. This variable is part of the state space for the dynamic problem. In the first stage (state $S = 0$) the consumer can decide to either not search for the product or include it in his consideration set. As he does not know the prices before searching, the utility from inclusion in the consideration set has to be expressed in expected value terms. It is equal to the expected value function at $S = 1$ and will differ across consumers i as some of the parameters to be estimated and the expectation formation vary across individuals.

$$V_i(\Omega_{S=0,t}) = \max \begin{cases} u_{0,t} + \beta E[V_i(\Omega_{S=0,t+1}) | \Omega_{S=0,t}] + \varepsilon_{0,t}, \\ E[V_i(\Omega_{S=1,t}) | \Omega_{S=0,t}] - s_t + \varepsilon_{s,t} \end{cases}$$

$\varepsilon_{0,t}$ and $\varepsilon_{s,t}$ represent iid extreme value error terms, β is the discount factor and $\Omega_{S,t}$ denotes the consumer's information set at time t and state $S \in \{0, 1\}$. The variable s_t is the search cost that has

²⁶As in the reduced-form regression I do allow for 12 different pack-sizes.

to be incurred when including the product in the consideration set. This search cost will vary with trip characteristics. It is assumed that the trip characteristics and therefore the search cost is known at this point and no expectation has to be formed. The value function in a time period without a shopping trip is equal to the expression above with the second term inside the max-operator being dropped.

Once the consumer has searched, he is in state $S = 1$, in which the value function looks very similar to a standard dynamic discrete choice model:

$$V_i(\Omega_{S=1,t}) = \max \left\{ \begin{array}{l} u_{0,t} + \beta E[V_i(\Omega_{S=0,t+1}) | \Omega_{S=1,t}] + \varepsilon_{0,t}, \\ \max_{j \in J_t} \{u_{j,t} + \beta E[V_i(\Omega_{S=0,t+1}) | \Omega_{S=1,t}] + \varepsilon_{j,t}\} \end{array} \right.$$

Here the consumer chooses a good j out of the sets of available goods J_t or he decides not to purchase anything.

The ex ante value functions can be calculated using the extreme value distribution of the error terms (following Rust (1987)):

$$EV_i(\Omega_{S=0,t}) = \log \left\{ \begin{array}{l} \exp(u_{0,t} + \beta E[EV_i(\Omega_{S=0,t+1}) | \Omega_{S=0,t}]) \\ + \exp(E[EV_i(\Omega_{S=1,t}) | \Omega_{S=0,t}] - s_t) \end{array} \right\}$$

$$EV_i(\Omega_{S=1,t}) = \log \left\{ \begin{array}{l} \exp(u_{0,t} + \beta E[EV_i(\Omega_{S=0,t+1}) | \Omega_{S=1,t}]) \\ + \sum_{j \in J_t} \exp(u_{j,t} + \beta E[EV_i(\Omega_{S=0,t+1}) | \Omega_{S=1,t}]) \end{array} \right\}$$

The term $EV_i(\Omega_{S=1,t})$ can be interpreted as the inclusive value of going shopping in time period t , excluding the search cost. The consumer will compare the expected utility of this inclusive value with the utility of not purchasing in the consideration stage. This is very similar to the optimal stopping point problem in replacement models for durable goods (for example Melnikov (2001), Gowrisankaran and Rysman (2007) or Schiraldi (2009)).²⁷

In order to make the problem tractable the relevant information sets have to be defined as well as the formation of expectations regarding future realizations of the state variables. In $S = 0$, when the consumer has not searched yet, the relevant information to him is his inventory, the identity of the

²⁷The difference is that the consumer has to form an expectation about the current period inclusive value, whereas in the replacement models the current value is known, but the consumers' expectations about the future evolution of the inclusive value are very important.

store he is visiting and the search cost. He does not know the prices of the different pack-sizes at this point, but has to form expectations based on the identity of the store. I assume that the consumer know the empirical distribution of prices at each store and computes his expected utility from searching based on the relevant price distribution conditional on store choice.²⁸ Once the consumer has searched he obtains information about the actual prices of the available pack-sizes. Independent of whether the consumer has searched or not, he will have to form expectation about future prices and future search costs. As the consumer does not know yet which store he is going to visit, he will therefore form an expectation about the identity of the store. The expectations regarding prices conditional on the store identity will be formed in the same way as it is done for current price prior to the search decision. Finally expectations about future values of the search cost are formed based on the empirical distribution of search costs. As for price expectations, the search cost does not depend on past realizations of any variable.²⁹

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

$$EV_i(S = 0, i_t, k_t) = \log \left\{ \begin{array}{l} \exp(u_{0,t} + \beta E_{k_{t+1}, s_{t+1}}[EV_i(S = 0, i_{t+1}, k_{t+1}, s_{t+1})]) \\ + \exp(E_{p_{k,t}}[EV_i(S = 1, i_t, p_{k,t}) | k_t] - s_t) \end{array} \right\}$$

$$EV_i(S = 1, i_t, p_{k,t}) = \log \left\{ \begin{array}{l} \exp(u_{0,t} + \beta E_{k_{t+1}, s_{t+1}}[EV_i(S = 0, i_{t+1}, k_{t+1}, s_{t+1})]) \\ + \sum_{j \in J_t} \exp(u_{j,t} + \beta E_{k_{t+1}, s_{t+1}}[EV_i(S = 0, i_{t+1}, k_{t+1}, s_{t+1})]) \end{array} \right\}$$

Note that in $S = 0$ the expectation regarding the current prices is conditional on the store identity, whereas the expectations regarding k_{t+1} and s_{t+1} are unconditional ones.

Finally I have to define the transition process for the inventory. This is simply given by the consumption that enters the per-period utility and the pack-size of the detergent that was purchased in the previous period (if any was purchased). The inventory evolves deterministically as follows:

$$i_t = i_{t-1} - c(i_{t-1}) + \Delta i_{t-1}$$

The pack-size purchased in $t - 1$ is denoted by Δi_{t-1} . The other variables were defined previously.

²⁸Note that I do not model a Markov-process for prices and price expectations are not influenced by past prices at all. This is necessary as I cannot distinguish whether a product was included in the consideration set and not purchased or if it was not included in the first place. Therefore I do not know when the consumers information set was "updated" with new prices and when not. This makes the simplification of basing price expectations purely on the identity of the store visited necessary.

²⁹This is mainly done in order to reduce the computational burden, in principal a Markov-process could be estimated for the search cost.

4.3 Choice Probabilities and Likelihood Function

The probability that a consumer considers buying in time period t (the consumer index i is omitted) can now be determined using the standard logit formula.

$$d_{c,t} = \frac{\exp(E_{p_{k,t}}[EV_i(S = 1, i_t, p_{k,t}) | k_t] - s_t)}{\exp(u_{0,t} + \beta E_{k_{t+1}, s_{t+1}}[EV(S = 0, i_{t+1}, k_{t+1}, s_{t+1})]) + \exp(E_{p_{k,t}}[EV_i(S = 1, i_t, p_{k,t}) | k_t] - s_t)}$$

The probability that a consumer buys product $\tilde{j} \in J$ conditional on considering a purchase can be calculated in a similar way. J denotes the set of all available pack-sizes including the outside option of not purchasing anything.

$$d_{\tilde{j}|c,t} = \frac{\exp(u_{\tilde{j},t} + \beta E_{k_{t+1}, s_{t+1}}[EV(S = 0, i_{t+1}, k_{t+1}, s_{t+1})])}{\sum_{j \in J} \exp(u_{j,t} + \beta E_{k_{t+1}, s_{t+1}}[EV(S = 0, i_{t+1}, k_{t+1}, s_{t+1})])}$$

As the consideration set itself is not observable, it is only reflected in the purchase probabilities. In terms of observables we have the following probabilities:

$$P_{\tilde{j},t} = d_{\tilde{j}|c,t} * d_{c,t}$$

$$P_{no-purchase,t} = d_{nc} + (d_{0|c,t} * d_{c,t})$$

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

$$L = \prod_t \prod_j P_{j,t}^{y_{j,t}}$$

Or in logarithmic form:

$$LL = \sum_t \sum_j (\ln P_{j,t}) y_{j,t}$$

$y_{j,t}$ with $j \in \{no - purchase, purchase - \tilde{j} \in J\}$ is a variable that takes the value one for the decision actually taken in a particular period and zero otherwise. Therefore this maximization tries to make the probabilities of the observed choices as close as possible to one. Within the estimation routine I solve the value function using a fixed point algorithm.

5 Estimation

In order to take the model to the data, a specific form for the utility from consumption and the storage cost has to be chosen. For simplicity I am using a linear storage cost and define consumption to be determined by $c(i_t) = \tau * \log(i_t + 1)$, where τ is a parameter to be estimated. Because the storage technology and the utility from consumption together determine the average time between purchases, these parameters are only identified up to a scaling parameter.³⁰ Therefore I define $v(c(i_t)) = c(i_t)$. Also the search cost s_t has to be parameterized. I choose the following functional form:

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

Where x_t is a vector of trip characteristics, which were already used in the reduced-form regressions. Specifically I use the relative expenditure in the store as proxy for the time spend in the store. I also include a variable capturing whether the product is on promotion³¹ and the number of products bought in the same category. The latter captures whether the consumer is in the relevant part of the store. I define two variables for a wider and a more narrow definition of product category.³² \tilde{s} and β are parameters to be estimated. The functional form makes sure that for no values of x the search cost is negative, as this would not make any economic sense. This non-negativity is given as long as $\tilde{s} \geq 0$. For $\tilde{s} = 0$ there is no search cost. A simple t-test on this coefficient will therefore allow me to test for the relevance of search costs. The functional form also makes the incorporation of expectations regarding future search computationally easy as the search cost varies on the compact set $\tilde{s} * [0, 1]$.³³

5.1 Expectations

I have to compute the consumers' expected utility from purchasing at the relevant store in the current time period. Furthermore households form expectations regarding the identity of the store visited next period and the future search costs they have to incur.

For the expectations regarding $EV_i(\Omega_{S=1,t})$ in the current time period, i.e. the expected utility from searching for laundry detergent, it is assumed that the consumer knows the price distribution for each pack-size of the product at each store³⁴ The expected utility is calculated based on these store-specific price distributions, which are simply inferred from all the weekly prices over the whole sample period. In order to compute the expectation regarding the identity of the store that is visited next period, the sample frequencies of the visits to different supermarkets are taken, i.e. the number

³⁰See the section on identification for more explanation on this.

³¹The variable described at the end of section 2.1 is used here.

³²For more detail on these variables see the reduced-form regressions in section 3.

³³This makes it easy to define an appropriate grid. I will elaborate more on this in the following section.

³⁴As discussed earlier, I define five different categories of stores: the "big four" supermarket chains and a residual category for trips to other supermarkets or convenience stores that sell detergent.

of trips to a particular supermarket is divided by the total number of shopping trips over the whole sample period. Note that this includes the probability of not visiting any store in the next period. A similar approach is taken for the search cost. A grid for different possible values of the search cost is constructed. This is made particularly easy by the functional form chosen, as the search cost has to be an element of the interval $\tilde{s} * [0, 1]$. Then probabilities are assigned to each grid point using information from sample frequencies.

Note that a particularly simple form of non-time-varying expectations is used here. A first or higher order Markov-process on the evolution of prices is frequently used in the literature, but unfortunately not possible here as it is unknown when new price information is obtained (see earlier discussion).³⁵ It would be possible though, to use a Markov-process for the store-visit probabilities and the search costs. As this would increase the state space³⁶ and therefore the computational burden it is not done at this point.

5.2 Heterogeneity

In order to reduce the computational burden of the maximization routine I will not be able to calculate individual specific value functions for each one of the 257 household in my sample. In order to capture some heterogeneity I do consider different types of households though. Specifically I allow for different price coefficients α for above and below median income households, different τ , i.e. different per-period consumption and different storage costs for households of different size. As consumption is taken as exogenous here and in order to keep the computation simple I simply use average consumption over the whole sample period in order to form groups of high and low consumption households³⁷. For these two groups different values of τ are estimated. A different storage cost parameter is estimated for households above and below a size of four. Finally each household's value function in the next period will depend on the probability of visiting a particular store as the price distributions vary across stores. More importantly the probability of not visiting any store is relevant as it makes waiting more costly.³⁸ I therefore split households into two groups, depending on whether the probability of not visiting any supermarket is above or below the median value.

Together this yields a total of $2^4 = 16$ types of households. For each one of these groups a different value function has to be calculated. In terms of coefficients I obtain two values for α based on income,

³⁵It is also questionable whether price information from the same or a different store should be treated symmetrically, as it is usually done in the literature. Presumably a promotion in a particular store in time period t will imply different things for the price in the same store in $t + 1$ then in a different store in $t + 1$.

³⁶The value of the search cost and the identity of the store visited in the current time period would have to be included in the state space.

³⁷Average consumption is calculated simply by dividing the aggregate amount of detergent purchased by the duration of the household being in the sample. Household are assigned to one of the two groups depending on whether they have above or below median average consumption.

³⁸Both the identity of the stores visited and the frequency of shopping trips varies a lot across households.

two values for τ based on average consumption and two cost parameters for differently sized households. The heterogeneity in terms of store choice enters the value function through the expectations operator regarding k_{t+1} and is not captured by any specific coefficient.³⁹ It has an impact on the incentives for inter-temporal substitution though, as it specifies the duration until the next purchase. Despite the fact that store-visit probabilities are one dimension along which I group the households, I can even apply a different store-visit probability distribution to each of the 16 types. As no additional coefficients are estimated and type-specific value function have to be calculated anyway, this comes at no cost. The store choice probabilities for each of the types is based on their respective means.⁴⁰

5.3 Initial Inventory

I also try to deal with the problem of an unknown initial inventory. As a logarithmic form is used for the unobserved consumption, some convergence will happen over time⁴¹ and the impact of the initial inventory will fade. I start with the first observed purchase for each household and (arbitrarily) 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 weeks, the observed behavior is used in order to form the likelihood function. This helps to mitigate the initial inventory problem.

5.4 Results

Table 3 presents the results for the main regression using all the dimensions of heterogeneity presented above and covariates that move the search cost.

Lets first turn to the estimated coefficients other then the search cost terms. The coefficients all have the expected sign and are significant at conventional levels. The heterogeneity enters in the expected way. I find that richer households are less price sensitive as they receive less disutility from paying a higher price. Households with more than 4 members have lower storage costs. The utility from consumption is higher for high consumption households.⁴²

As for the variable influencing the search costs, again the coefficients have the expected sign and are precisely estimated. Search costs are lower when the relative expenditure is high. This makes intuitive

³⁹Note that the expectation regarding $EV_i(\Omega_{S=1,t})$ based on the sample frequencies of prices for each store and therefore do not vary across households. Also, the expectations regarding s_{t+1} are not type-specific, as the search cost is not allowed to vary across types for reasons of identification.

⁴⁰The distribution does not vary dramatically within the group of households with a low (high) probability of not visiting any store. I do use the additional information though, because this causes no increase in the computational burden.

⁴¹Households with a higher inventory will consume more then those with a low inventory. Over time the path of inventory will be less influenced by the initial inventory.

⁴²This result is unsurprising as I used the average consumption rates to group the household into high and low consumption types. It helps to capture heterogeneity in the consumption patterns more appropriately, which leads to more precise estimates on the other parameters. As consumption is assumed to be exogenous, this a priori grouping of high and low consumption household is consistent with the model.

sense as less time is spent in the store, which increases the hassle of searching for a particular product. If other products in the same product category are purchased, this also lowers the search costs. This might reflect that households group the purchases of similar kind of products together or simply that the consumer already spends time in the correct aisle, which lowers the search cost. Also in terms of magnitude the results seem reasonable as the number of products in the narrower product category lowers the search cost by a larger amount. Finally a promotion on one or several of the pack-sizes lowers the overall search costs as the households gain some information on the price prior to the store visit.

In terms of magnitude, the search cost varies on the interval $[0, 4.7583]$. A shopping trip where each of the trip characteristics takes on its respective average value in the sample has a search cost of 0.6294 associated with it. In order to get a feeling for the importance of the search cost, it can be compared with the disutility from purchasing detergent. A 1.92 kg pack causes a disutility of 2 at the moment of purchase, a 936 g pack is associated with a disutility of 1.12. These are the two most popular pack-sizes of laundry detergent.⁴³ The comparison with the disutility from paying a certain price for the product can be used to convert (dis-)utils into monetary values. When doing this I obtain a search cost of 1.5 pounds.

To illustrate the impact of the different trip characteristics on the search cost I calculate the search cost when all trip characteristics take on their mean values. I then increase one of the characteristics by one standard deviation while keeping the others constant. The result from this thought experiment are reported in table 4. It can be seen that the number of products purchased in the same narrow product category is the most important factor that shifts the search cost. The effect of promotions and the relevance of primary shopping trips might be understated by these numbers though. Most of the time, there is no promotion, and there is therefore a large mass of observation with a value of zero for this variable. The relative expenditure on a shopping trip is distributed bi-modally with a lot of trips with very little expenditure and many trips with above average expenditure. A change from a very small shopping trip (the lower mode of the distribution) to the second mode of expenditure involves a change in relative expenditure larger than one standard deviation for instance. In both cases using a standard deviation might therefore not capture the effect of the variables very well. I hence report results from two alternative experiments: a change in relative expenditure from one mode of the distribution to the other and a change in the promotion variable from zero to the average value conditional on there being a promotion.⁴⁴ It can be seen that the effect of relative expenditure (small vs "main" shopping-trip) is potentially quite large, whereas promotions have a relatively smaller impact on search costs.

I also run the estimation without the consideration stage in order to compare the results (see table

⁴³The disutility is calculated based on the average price and the estimated price coefficient.

⁴⁴i.e. conditional on the promotion variable to be strictly larger than zero.

3). Qualitatively the coefficients do not differ from the complete model. Of particular interest is the price coefficient. I find that it is larger relative to the case with search cost. This is due to the fact that a high price sensitivity is needed in order to explain why consumers buy large pack-sizes in non-promotion periods. The quantity discount is relatively small and even a small storage cost would always make the consumer prefer small packs. As even a small pack lasts for a much longer time than until the next shopping trip the fear of running out detergent unexpectedly also cannot explain this. In the presence of search costs one reason for buying large packs is to avoid future search costs. This is absent in the perfect information model. Therefore, when search costs are controlled for, the price sensitivity needed to explain to remaining variation in pack-size choice and temporal purchase patterns drops. This is closely related to the discussion on identification in the next section.

In order to check how sensitive the results are to the functional form assumptions, I vary the specification in several ways. I try to include a quadratic storage cost term for small and large size households. Both coefficients are positive but not significantly different from zero. I also excluded the first 20 weeks, instead of only 10 from the estimation. This also had little impact on the results.

With regards to the search costs I experiment with different functional forms and again find no effect on the qualitative results of the model.⁴⁵ Finally I estimate the model with a different search cost size parameter \tilde{s} for low and high income households. The results show that higher income households have a larger search cost, possibly reflecting a higher opportunity cost of time. As I am not sure about the strength of identification from a theoretical point of view (see discussion in section 6) I do not use this distinction in my main estimation.

6 Identification

The parameters to be estimated are the price coefficient α , the parameterized functions $v(c(\cdot))$ and $T(\cdot)$ and the search cost s as a function of trip characteristics. The consumers' *average* inventory holding will determine the parameters of the consumption utility and inventory costs.

As can be seen in Figure 1, consumers would like to always keep an "optimal" inventory, denoted by the dashed line in the graph.⁴⁶ This is not possible all the time because of the discrete set of pack-size choices. The consumers will therefore try to remain as close as possible to the optimal point, but will circulate around it over time. The pattern of their purchase will reveal the optimal inventory, which will identify $v(c(\cdot))$ and $T(\cdot)$ up to a scaling parameter.

There are two types of variation in purchase behavior left that will help me to identify both the

⁴⁵Specifically I use $\tilde{s} * \exp(x'_t \beta)$ and $\exp(\alpha + x'_t \beta)$. Note that I do need to specify a functional form that prevents the search cost from becoming negative for *any* possible value of the parameters.

⁴⁶For the purpose of illustration the graph shows a convex cost function. The same intuition is still true if the cost function is linear as long as the utility is concave. This is the case for the specification used here.

(average) search cost⁴⁷ and the price coefficient. For the same average inventory holding consumers differ in the pack-size they purchase and in how often they deviate from their average purchase pattern in reaction to a promotion. Bigger pack-sizes will lead to larger deviations from the optimal inventory on average. Consumers reacting to promotion causes both earlier and larger purchases in promotion periods relative to the average behavior. A very price sensitive consumer will buy larger packs on average as they are cheaper and he will bring his purchase forward when he encounters a promotion. A consumer with a high search cost will also want to buy large packs as he than has to search less often. At the same time a high search cost will narrow the window within which the consumer is aware of the price. The probability of bringing the purchase forward in a promotion period is therefore lower as the consumer is likely to be unaware of the price reduction. Together the average pack-size and the purchase accelerations allow to distinguish between search costs and price sensitivity. Consider for example a consumer buying large quantities of detergent each time he purchases but never reacts to promotion periods. This behavior is consistent with high search costs but cannot possibly be explained with price sensitivity alone.

Also, despite the fact that both low search costs and high price sensitivity lead to purchase acceleration in promotion periods, the impact of the two does differ somewhat. If we knew the correct underlying model of storage costs and the actual inventory holding, then the fact that a consumer misses out on a promotion but purchases shortly afterwards can only be explained by the existence of a consideration set formation. The presence of storage cost will lead to the reservation price rising slowly over time. The consideration set though leads to a discrete jump: in one time period the consumer does not even look at the price in the next he includes detergent in his consideration set. If we observe an upwards jump in the price (the end of a promotion period) and a purchase after the jump, this is only consistent with the existence of a consideration set. Unfortunately the distinction between the gradual effect of the storage technology on purchase behavior and the discrete one of the consideration stage is more difficult to make with imperfectly observed inventory. It has to be noted though, that the identification relies only on correct information regarding the *change* of inventory over time. An incorrectly inferred level would not be problematic. As the change in inventory is estimated from average consumption data this should be relatively reliable.

Finally the identification of the trip characteristics that enter the consideration stage comes from a kind of "exclusion restriction". As argued before there is no theoretical reason why the trip characteristics would enter into the optimization of a perfectly informed consumer (Bronnenberg and Vanhonacker (1996)). They can therefore not be part of the product characteristics which influence the choice after the consumer has searched (the second stage in this model) and have to be included in the considera-

⁴⁷Roughly speaking the average search cost is captured by the coefficient \tilde{s} in this model.

tion stage. If the consumer knows the *actual* prices of all products at every store he visits, whatever else he is doing on the shopping trip should not have an impact on his choice of a particular product. If search is costly though, the consumer will be less likely to consider a purchase on trips where his search costs are higher. This will lower his purchase probability on these trips. This would also lead to a higher likelihood of buying detergent on particular shopping trips, something that cannot be explained without a consideration stage. The impact of trip characteristics, that leads to variation in the search cost can hence be identified. They have an impact on the purchase probability only through the consideration set formation but do not effect utility from the purchase directly.

Although the consideration set itself is not observed it is standard in the literature to infer the decision in the consideration stage from observed purchases. Some early contributions have used direct survey data on the consideration decision (Roberts and Lattin (1991)). Nierop, Paap, Bronnenberg, Franses, and Wedel (2005) show though, that a direct approach using only purchase data yields similar estimation results as one which uses additional survey data on the considered options. This is reassuring as a direct approach is also employed here.

7 Price elasticities

Because of the search stage and the dynamic nature of the model, price elasticities cannot be computed analytically from the price coefficient. Instead I will simulate the change in market shares for different kinds of price changes by taking draws from the distribution of error terms and aggregating the choices of all simulated consumers into market shares. In order to look at the impact of both temporary and permanent price changes I compute the market shares for several time periods.

Apart from their type consumers also differ by the inventory they hold and the store they visit in a particular time period. Both of these things are going to have an impact on their choices. I therefore have to take the distribution of inventories and the probability of visiting a particular store into account when simulating the consumers' reaction. The discrete probability distribution of store-visits can easily be computed and was actually already used when the expectations regarding the store visit in the next period were formed. The distribution is computed from the sample frequencies of store visits and is therefore independent of any parameters of the estimation. For each simulated consumer and time period I randomly draw which store was actually visited from the type-specific distribution of store visits. Note that this also includes the option of not visiting any store. I also have to simulate values for the various extreme value error terms embedded in the model. There are a total of 15 error terms,⁴⁸ which have to be drawn for each time periods. In order to calculate the inventory distribution I start

⁴⁸There are 12 error terms for each of the different pack-sizes plus 1 for the outside option and 2 error terms from the consideration stage.

with assigning a particular gridpoint value of the inventory vector to each consumer. I (arbitrarily) give equal weight to each gridpoint, so I end up with a uniform distribution of the inventory gridpoints in my sample of simulated consumers. I then simulate the consumers' behavior over 40 time periods assuming that the price for each product is equal to the respective sample average in all time periods. I 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 40 simulated time periods.⁴⁹ I then use this "steady state" inventory distribution as the initial inventory for the simulation of price changes.

I can now simulate the choices consumers take in each of the time periods. This yields market shares for each of the different pack-size as well as the proportion of consumers that search but do not buy and those that do not even search. Note, that I am also able to compute store-specific market-shares. This has not been done before and can shed some light of the impact of store-specific promotions on competing stores.

I use a total of 10000 simulated consumers for each one of the 16 types. Total market shares are calculated by weighting the market share of each type-group of consumers with the frequencies of the respective type in the sample. I first run the simulation with prices equal to the averages of the respective price distributions for 10 time periods. As one would expect I see very little change in the market shares over time. I find that only about 5 percent of all consumers do buy detergent in a given time period (a half week), 37 percent do not go shopping, 45 do not search and 13 search but do not buy. The relative importance of search (or no search) is therefore quite high. This is in line with findings that the consideration stage explains a large part of the variation in purchase incidents such as Hauser (1978). Also, Dickson and Sawyer (1990) and Hoyer (1984) find that many consumers do not pay much attention to the price of a product they buy and spend very little time making their selection in the store.⁵⁰ This anecdotal evidence lends further support to the large relative importance of the consideration stage.

I then compute the reaction to both temporary and permanent price changes. I do use price *reductions* in my simulation as temporary promotion periods are common in the real world. For both temporary and permanent changes I look at product-specific, store-specific and product-store-specific price changes. I use the same set of random draws for each change and also for the initial path without a change. Any difference in market shares is therefore caused by the price change. The results regarding price elasticities and the dynamic adjustment are presented in table 5 and figure 2.

⁴⁹It takes a low consumption type consumer less than 20 time periods in order to run down his inventory completely, even if he started with the maximum possible inventory (the largest gridpoint value). Therefore it makes sense that the impact of the initial inventory (here the uniform distribution) should have faded completely after this time span.

⁵⁰They observe shoppers whilst they do their shopping in the supermarket and also interview them after they made their purchases.

I look at price changes for one particular pack size in all stores, for all sizes in one store and for one pack size in one store. I then plot the dynamic adjustment path for a 50 percent price decrease in all cases⁵¹ and calculate short term price elasticity for temporary and permanent changes and the long term price elasticities for permanent changes.⁵² For each case I look at the own price elasticity, the cross price elasticity with other products and/or stores, the cross elasticity with respect to the outside option and with respect to search. Remember that I do allow consumers to purchase 1 or 2 units of each pack-size. I do treat these as different goods in the estimation. As price changes will apply to one or several units of a certain pack size, I do aggregate all units of the same pack size in order to compute the market share.⁵³

Across all the different price changes considered I do observe almost no post-promotion dip, despite a very large contemporaneous reaction. This is driven by the small proportion of consumers that are actually aware of the price change. 82 percent do not go shopping at all or do not search for the product. There is thus a strong reaction from those that are aware of the change and bring their purchase forward, but on aggregate this does not cause a significant decrease in sales in later periods. This is consistent with lack of a post-promotion dip in aggregate data (see Hendel and Nevo (2003), Heerde, Leeflang, and Wittnik (2000) and Neslin and Stone (1996)).

When changing the price of one particular pack size (936g in this example) temporarily across all stores, there is a large contemporaneous reaction. The effect is particularly large for purchases of 2 packs of the same size, which accords with the incentives for stockpiling. Consumers mainly substitute away from other pack-sizes and to a lesser extent from the outside option. A permanent price change leads to a similar contemporaneous reaction with markets shares staying at the new level afterwards. Note that with lower expected price the incentive to search has become stronger. Although there is no stockpiling effect here which should lead to a weaker reaction, the increase in consumers searching makes up for this. The magnitude of the contemporaneous reaction for the permanent and temporary case is very similar.

For a temporary price change in only one store (Tesco in this case) the market share reaction is of similar magnitude as in the previous case. The difference is that consumers substitute away from the outside option. By construction they cannot substitute away from other stores contemporaneously. Only the products in the store they are visiting are available to them. An effect on other stores can only happen in later time periods because of now increased inventory holdings. As described above the post-promotion effects are very small. This is the case here as well. The effect of a permanent

⁵¹Note that the graphs show only 5 time periods as the various market shares settle to a new equilibrium very quickly.

⁵²By construction a temporary price change has no long term effect, therefore long term elasticities cannot be calculated for this case.

⁵³In order to do this I treat the outside option as one unit. The total market size is not fixed in this case, as even with a fixed number of consumers the total number of units (of inside and outside goods) can vary.

change is again similar to the reaction to a temporary change. Note that in this case of a store-specific promotion the market share peaks immediately and comes down somewhat afterwards. As consumers adjust to a more frequent purchases and higher consumption in the future this leads to an overshooting of the aggregate market share. Note also, that contrary to the temporary reduction we now see a large contemporaneous substitution away from other stores. This happens as consumers that visit stores other than Tesco now have a higher incentive to wait for a trip to Tesco. This reduces the share of other stores to the benefit of Tesco.

Finally I consider a change in a particular product in a particular store (a 936g pack at Tesco). The reactions are similar to the ones described above. The temporary increase in market share is now caused by both substitution away from other Tesco pack sizes and by substitution away from the outside option. A permanent change has a similar impact as in the previous two cases with a less pronounced contemporaneous peak. The effect on search behavior is smaller as the price reduction applies to fewer products in this case.

8 Conclusion

A structural model that incorporates search and stockpiling was presented. The framework can be used for the analysis of any storable grocery item. Modeling search will be particularly relevant for products where the inter-purchase duration is relatively long. When applying the model to purchase data for laundry detergent I find that search costs are statistically and economically significant. They have a big impact on purchase behavior, as around 70 percent of shopping trips without a purchase can be attributed to the consumer not searching for the product. I also find that the probability of considering to buy detergent is a function of trip characteristics. Overall expenditure in the store, the number of other product purchased in the same category and promotions all significantly lower the search cost, with the number of products in the same category having the biggest impact.

Furthermore I am able to analyze substitution patterns between products and stores in more detail than previous research. The simulation of price elasticities accords with the well-known absence of a post-promotion dip. Temporary price reduction have little impact on competing stores, but a large impact on the product / store for which the price was lowered. Permanent reductions also lead to substantially higher sales, with market shares adjusting almost immediately. Supermarkets suffer significantly from permanent price reductions in competing stores.

In future research it would be interesting to incorporate category level consideration (as in this paper) and brand level consideration in one unifying framework. This would allow the researcher to look at substitution across brands as well as pack-sizes and stores.

References

- ACKERBERG, D. A. (2003): “Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination,” *Management Science*, 44(3), 1007–1040.
- ALLENBY, G. M., AND J. L. GINTER (1995): “The Effects of In-Store Displays and Feature Advertising on Consideration Sets,” *International Journal of Research in Marketing*, 12, 67–80.
- 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.
- BEN-AKIVA, M., AND B. BOCCARA (1995): “Discrete choice models with latent choice sets,” *International Journal of Research in Marketing*, 12, 9–24.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63(4), 841–890.
- 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 (2006): “The Price Consideration Model of Brand Choice,” unpublished manuscript.
- 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.
- ERDEM, T., AND M. P. KEANE (1996): “Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets,” *Marketing Science*, 15(1), 1–20.
- 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.
- HAUSER, J. R. (1978): “Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information-Theoretic Approach,” *Operations Research*, 26(3), 406–421.
- HEERDE, H. J. V., P. S. H. LEEFLANG, AND D. R. WITTEK (2000): “The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data,” *Journal of Marketing Research*, 37(3), 383–395.
- HENDEL, I., AND A. NEVO (2003): “The Post-Promotion Dip Puzzle: What do the Data Have to Say?,” *Quantitative Marketing and Economics*, 1(4), 409–424.
- (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.
- HORTASU, 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.
- HOYER, W. D. (1984): “An Examination of Consumer Decision Making for a Common Repeat Purchase Product,” *The Journal of Consumer Research*, 11(3), 822–829.
- LACH, S. (2006): “Existence And Persistence Of Price Dispersion: An Empirical Analysis,” *The Review of Economics and Statistics*, 84(3), 433–444.
- 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.
- MITRA, A. (1995): “Advertising and the Stability of Consideration Sets over Multiple Purchase Occasions,” *International Journal of Research in Marketing*, 12, 81–94.
- NESLIN, S. A., AND L. G. S. STONE (1996): “Consumer inventory sensitivity and the postpromotion dip,” *Marketing Letters*, 7(1), 77–94.
- NIEROP, E. V., R. PAAP, B. BRONNENBERG, P. H. FRANSES, AND M. WEDEL (2005): “Retrieving Unobserved Consideration Sets from Household Panel Data,” Econometric Institute Report.

- 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.
- SCHIRALDI, P. (2009): “Automobile Replacement: A Dynamic Structural Approach,” *Unpublished Manuscript, London School of Economics*.
- 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.
- URBANY, J. E., P. R. DICKSON, AND A. G. SAWYER (2000): “Insights Into Cross- and Within-Store Price Search: Retailer Estimates Vs. Consumer Self-Reports,” *Journal of Retailing*, 76(2), 243–258.
- VARIAN, H. R. (1980): “A Model of Sales,” *American Economic Review*, 70(4), 651–659.
- WALTERS, R. G. (1991): “Assessing the Impact of Retail Price Promotions on Product Substitution, Complementary Purchase, and Interstore Sales Displacement,” *Journal of Marketing*, 55(2), 17–28.
- WALTERS, R. G., AND H. J. RINNE (1986): “An Empirical Investigation into the Impact of Price Promotions on Retail Store Performance,” *Journal of Retail*, 62(3), 237–266.

Supermarket Chain	Frequency of Trips (%)	Frequency of Purchase (%)	Price (Fairy Non-Bio Tabs 936 Gram Pack)			
			Average	Std. Dev.	Min	Max
Asda	19.88	25.83	2.6	0.189	2.42	2.88
Morrisons	11.56	13.08	2.64	0.213	1.99	2.99
Sainsburys	12.51	11.35	2.75	0.198	2.49	2.99
Tesco	37.51	39.54	2.62	0.203	2.42	3.02
Others	18.54	10.20	2.77	0.261	1.99	3.47

Table 1: **Descriptive Statistics on Shopping Behavior and Product Pricing.** The table shows the frequency of shopping trips and purchases incidences for the four main supermarket plus a residual category of other shops. It also shows the price of Fairy Non-Bio Tabs at different supermarket chains. There is variation in both the average price and also the depth and occurrence of promotions (as measured by the minimum price and the standard deviation). The price for the most popular pack size of 936kg is used here in order to illustrate the general pricing patterns. The descriptive statistics look very similar for the other pack sizes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable	Purchase Dummy	Purchase Dummy	Purchase Dummy	Purchase Dummy	Purchase Dummy	Purchase Dummy	Purchase Dummy	Purchase Dummy	Purchase Dummy
Sample	Full Sample	Full Sample	Full Sample	Full Sample	Full Sample (excl. first Purchase)	Full Sample (excl. first Purchase)	Full Sample (excl. first Purchase)	Full Sample (excl. first Purchase)	Weekend Trips only (excl. first Purchase)
Household Dummies	No	No	No	No	No	No	Yes	No	No
Price	-0.419*** (0.0468)	-0.442*** (0.0465)	-0.218*** (0.0616)	-0.461*** (0.0473)	-0.399*** (0.0502)	-0.222*** (0.0667)	-0.252*** (0.0708)	-0.261*** (0.0887)	-0.431*** (0.161)
Relative Expenditure		0.889*** (0.0294)				0.177*** (0.0430)	0.395*** (0.0589)	0.183*** (0.0554)	0.159 (0.112)
Promotion Dummy			0.496*** (0.0986)			0.565*** (0.110)	0.611*** (0.115)	0.620*** (0.144)	0.583** (0.266)
Other Products in the same Category (1)				0.221*** (0.0125)		0.212*** (0.0131)	0.234*** (0.0149)	0.213*** (0.0170)	0.202*** (0.0339)
Other Products in the same Category (2)				0.0470*** (0.00197)		0.0453*** (0.00261)	0.0325*** (0.00385)	0.0449*** (0.00321)	0.0449*** (0.00623)
Time since last Purchase					0.0395*** (0.00267)	0.0534*** (0.00306)	0.120*** (0.00479)	0.0497*** (0.00368)	0.0409*** (0.00642)
Volume previously purchased					-0.448*** (0.0383)	-0.455*** (0.0407)	-0.682*** (0.0591)	-0.481*** (0.0548)	-0.452*** (0.106)
Distance to the Supermarket								-0.855 (0.728)	-0.929 (1.531)
Observations	95210	95210	95210	95210	93875	93875	93875	57685	16344

Table 2: **Static Discrete Choice Estimation (Reduced-Form)**. The table shows results from multi-nominal logit regressions. *** denotes significance at the 1 percent level, ** at the 5 percent level and * at the 10 percent level. An observation is a shopping trip of a particular household, time periods without a shopping trip are dropped. The dependent variable is a dummy equal to one for one of the available options in each time period. The household can choose to purchase one of 12 different pack-sizes (see text for details) or not purchase anything. The utility of the option is normalized to zero. Some control variables are included in the regressions, but are not reported: The frequency of trips to a particular supermarket, the overall frequency of shopping trips (both measured at the household level), household size, a set of supermarket dummies. See text for more details on this. All reported covariates are discussed extensively in the text. The full sample of households and time periods is used in regressions (1)-(4), starting with the first purchase incidence for each household. In the remaining regressions lagged purchase characteristics are included, eliminating one observation per household. When household dummies are included in column (7), they enter all the options except the outside option and effectively capture the household specific average purchase probability. Some of the non-reported controls have to be dropped as they do not vary within the household. The results are to be interpreted as reduced-form as they show the importance of dynamic aspect by including proxies into a static framework. None of the variables (especially not the dynamic ones) have a structural explanation in this framework.

		with consideration stage	without consideration stage
alpha	below median income	0.4164*** (0.0219)	0.7416*** (0.0236)
	above median income	0.4032*** (0.0239)	0.7391*** (0.0248)
tau	low average consumption	0.3733*** (0.0013)	0.2867*** (0.0625)
	high average consumption	0.7099*** (0.0350)	0.6068*** (0.0835)
storage costs	household size less than four	0.2400*** (0.0219)	0.1871*** (0.0355)
	household size four or larger	0.2010*** (0.0229)	0.1719*** (0.0367)
search costs	s	9.5165*** (0.1390)	
	promotion	-0.1529*** (0.0455)	
	relative expenditure	-0.5593*** (0.0770)	
	# of products purchased in the same (narrow) product category	-1.0373*** (0.0698)	
	# of products purchased in the same (wider) product category	-0.0322*** (0.0049)	
observations	households	257	257
	weeks	27666	27666
	shopping trips	17802	17802

Table 3: **Estimation Results from the Dynamic Model.** The table shows results from estimating the full dynamic model with and without a consideration stage. *** denotes significance at the 1 percent level, ** at the 5 percent level and * at the 10 percent level. The full sample of households and time periods is used in the estimation. Time periods without a shopping trip do not provide any direct information on the model's parameters as no choice is taken (store choice and therefore also the choice not to go is assumed to be exogenous). It is still necessary to include them in order to model the inventory transition appropriately. The parametrization of the model is discussed in the text as well as the different types of heterogeneity. For α , τ and the storage cost different parameter values are estimated for different subsamples of the population. s denotes the overall scale of the search cost, whereas the four other covariates shift the search cost.

	baseline	change	difference
shopping trip with average characteristics	0.6294		
promotion		0.5712	0.0582
relative expenditure		0.4162	0.2132
# of products purchased in the same (narrow) product category		0.0864	0.5430
# of products purchased in the same (wider) product category		0.4192	0.2102
lower mode of relative expenditure	0.9734		
higher mode of relative expenditure		0.6132	0.3602
no promotion	0.6571		
average promotion conditional on any promotion		0.5316	0.1255

Table 4: **The Effect of Trip Characteristics on Search Costs.** The table shows the effect of changes in trip characteristics on search costs. In order to make the changes comparable, the effect of one standard deviation changes is shown. The first row reports the search cost for the baseline case of a shopping trip where all characteristics take on their average values. The following four rows report the new search cost and the difference to the baseline for a change in each characteristic. The second panel uses the search cost evaluated at the lower mode of the expenditure distribution as the baseline and compares it to a shift to the higher mode. The last panel uses no promotion as the baseline case and compares this to the case where the promotion variable is evaluated at its conditional average (conditional on there being any promotion, i.e. the variable being unequal to zero). In both of the lower panels all the other characteristics are evaluated at the average values as before.

	Own (total units)	1 Pack	2 Packs	Other Packs	Outside Option	Search
936g Pack in all Store						
Temporary price change (immediate)	3.204	1.543	10.622	-0.711	-0.433	0.000
Permanent price change (immediate)	2.994	1.645	39.038	-0.775	-0.384	-0.054
Permanent price change (long term)	2.774	1.561	36.938	-0.764	-0.362	-0.052
All Sizes in Tesco	Own (total units)	1 Pack	2 Packs	Other Stores	Outside Option	Search
Temporary price change (immediate)	2.826	0.570	3.796	-0.091	-0.379	0.000
Permanent price change (immediate)	3.865	0.548	4.514	-0.458	-0.443	-0.085
Permanent price change (long term)	3.146	0.536	3.878	-0.606	-0.309	-0.080
936g Pack in Tesco	Own (total units)	1 Pack	2 Packs	Other Packs	Other Stores	Outside Option
Temporary price change (immediate)	3.307	1.406	10.836	-0.671	0.000	-0.183
Permanent price change (immediate)	3.642	1.649	11.664	-0.080	-0.619	-0.063
Permanent price change (long term)	3.224	1.732	9.819	-0.072	-0.563	-0.065

Table 5: **Price Elasticities.** The table reports price elasticities for three types of changes: for one particular size across all stores, for all sizes at one store and for one size at one particular store. 936g and Tesco are the pack-size and supermarket used to illustrate the pattern. Price elasticities are reported for temporary changes and permanent changes in the time period of the change (short term elasticity) as well as after 5 time periods for the case of a permanent change (long term elasticity). Reactions of one and two unit purchases of the product on promotion are reported. The effect on the outside option is also reported in all cases. The reaction on other pack-sizes (at the same store) and on the pack-sizes on promotion in other stores are reported (where appropriate). Changes in the search behavior are reported in the final column.

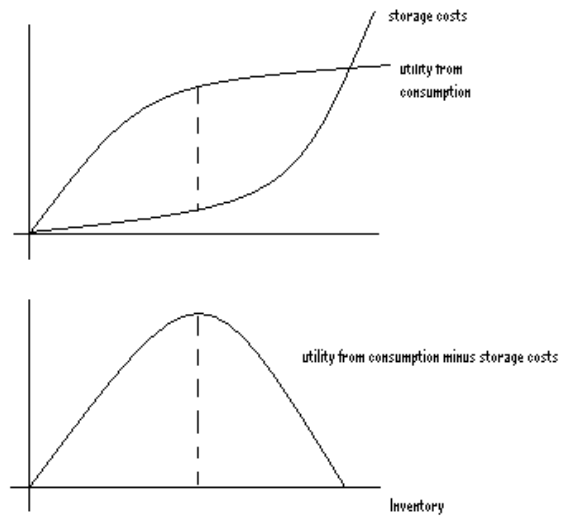
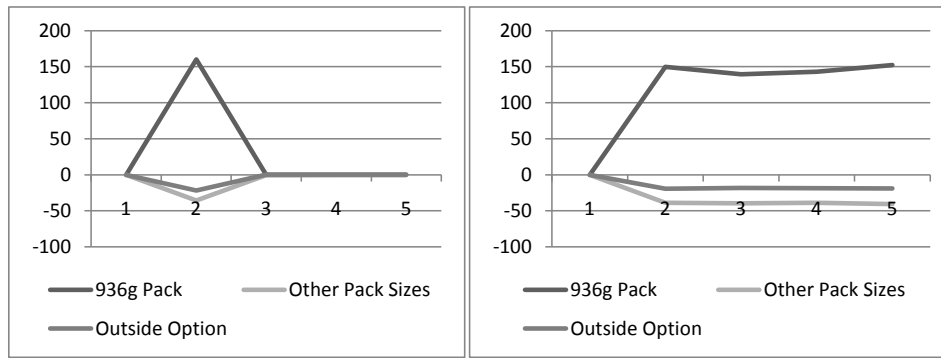


Figure 1: **Utility from Consumption and Storage Costs.** The graph illustrates the effect on utility of consumption and storage costs. Utility is linear in consumption and consumption is convex in inventory. Storage costs are assumed to be concave. The overall effect on utility is illustrated in the lower panel. The dashed line denotes the optimal inventory, i.e. the level of inventory that maximizes the net utility from consumption and storage.

Price change for one product in all stores: temporary and permanent



Price change for all products in one stores (Tesco): temporary and permanent



Price change for one product in one stores (Tesco): temporary and permanent

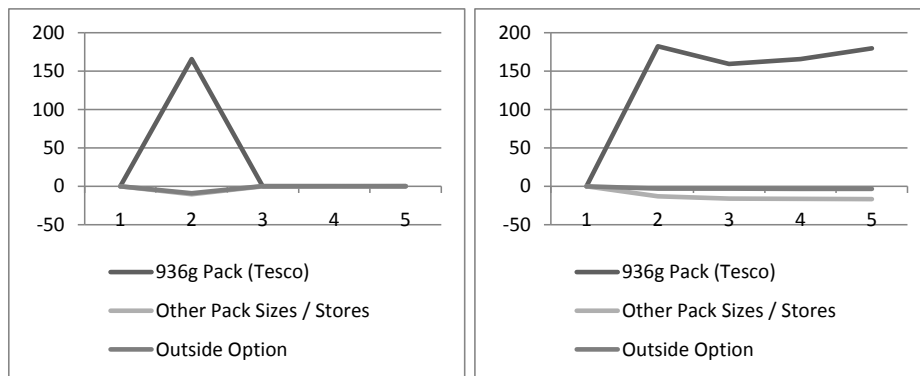


Figure 2: Dynamic Adjustment to Price Changes. The graphs show reaction to temporary (left column) and permanent (right column) price changes. The three rows of graphs show the dynamic adjustment for one particular size across all stores, for all sizes at one store and for one size at one particular store. 936g and Tesco are the pack-size and supermarket used to illustrate the pattern. All graphs show reactions to a 50 percent price reduction for the respective type of promotion. The units on the x-axis are time periods, with the change in price occurring in period 2. In case of a temporary change prices are back to the original level in time period 3. On the y-axis the reaction in terms of purchases is measured in percentages.