PRELIMINARY
COMMENTS WELCOME[*]

Sales and Consumer Inventory[1]

Igal Hendel
University of Wisconsin, Madison and NBER

and

Aviv Nevo
University of California, Berkeley and NBER

May 2001

1. Introduction

For most, if not all, non-durable consumer products prices tend to be at a modal level with occasional short-lived price reductions, namely, sales. Unsurprisingly, during a sale the quantity sold increases significantly. However, in those instances where the duration of the sale is longer than the standard duration, we find that the quantity sold decreases over time, holding the discount and other

[*]This version is a rough draft. We are in the process of revising the text and splitting the paper, as a result the current version might be somewhat confusing. We include a few comments in Section 3, which hopefully provide an idea of how the different parts fit in. Future versions will be made available thorough our web pages.

[1]We wish to thank David Bell for the data and seminar participants in the Berkeley-Stanford IOfest 2000, the spring 2001 NBER Productivity workshop, Berkeley, BYU, Chicago GSB, Northwestern, Wharton, University of Virginia, University of Wisconsin, Yale and NYU for comments and suggestions. The second author wishes to thank the Center for the Study of Industrial Organization at Northwestern University, for hospitality and support. Comments are very welcome and should be directed to either igal@ssc.wisc.edu or nevo@econ.berkeley.edu.

promotional activities fixed. While sales are often present in markets for fashion goods or markets where demand is seasonal, or follows some predictable pattern, our focus is on goods where seasonality and fashion play less of a role. In particular, our empirical analysis will focus on laundry detergents where we believe that fashion and seasonality play no role.[2] Motivated by the price and quantity patterns, observed in the data, we focus on intertemporal price discrimination. Specifically, we examine heterogeneity in consumers storage costs and its implications for the consumers' willingness to purchase products when prices are low, store them as inventory and consume them later when the prices are higher.

There are several reasons, beyond pure academic curiosity, why we should care about sales and their effects. First, when a good is storable, there is a distinction between the short run reaction to a price change during a sale and the long run reaction to the same price change. This has direct implications for demand estimation that relies on data with this pattern, and affects analysis based on these estimates, whether it is merger analysis or computation of welfare gains from introduction of new goods and services. Second, understanding the impact of sales allows us to study the issue of optimal sales. Finally, the way in which data with this pattern are used to construct price indexes depends on our interpretation of what is driving sales (Feenstra and Shapiro, 2001).

We consider a consumer's dynamic problem when she has an expected stream of future demands, is able to store a consumption good and faces uncertain future prices. In each period the consumer decides how much to buy, which brand to buy and how much to consume. These decisions are made to maximize the present expected value of future utility flows. Optimal behavior is a function of the current price, the current inventory and a stochastic shock. In any period quantity

[2] See, for example, Warner and Barsky (1995), Chevalier, et al. (2000) and MacDonald (2000), for papers studying the relation between seasonality and sales.

purchased that is not consumed is stored as inventory. The quantity consumed can exceed the quantity purchased that period, but cannot exceed purchases plus current inventories. In this model the consumer will purchase for two reasons: for current consumption and to build inventories. Consumers increase inventories when the difference between the current price and the expected future price is lower than the cost of holding inventory.

In order to test the model we use weekly store-level price and quantity scanner data on laundry detergents. This data set was collected using scanning devices in nine supermarkets, belonging to different chains, in two sub-markets of a large mid-west city. Besides the price charged and quantity sold we know other promotional activities that took place. In addition to these data we use a household-level data set. We follow the purchasing patterns of over 1,000 households over a period of 104 weeks. We know exactly which product was bought, where it was bought, how much was paid and whether a coupon was used. In addition we know when the households visited a supermarket but decided not to purchase laundry detergent.

We use these data in several ways. First, we test the implications for household and aggregate behavior derived from the model. In the process we provide evidence on the difference between sale and non-sale purchases, both across households and within a household over time. Second, in order to further test the theory we examine household behavior directly by structurally estimating the model. The major difficulty in estimating the model is that while purchases are observed, consumption decisions, and therefore inventory holdings, are not. As described in Section 4 we propose to use the structure of the model, to construct the unobserved household decisions.

Results of our preliminary analysis suggest the following. First, in line with the model's prediction, we find that, all else constant, a longer duration from previous sale, has a positive effect on the aggregate quantity purchased. In order to find this effect we have to properly control for the effect of other promotional activity. Furthermore, as predicted, this effect is larger during sale periods than during non-sale periods.

Second, standard household demographics are not powerful in explaining the difference across households in the fraction of purchases on sale. However, in support of our theory, we find that two indirect measures of storage costs are positively, and significantly, correlated with a household's tendency to buy on sale. We also find that the difference across households in the fraction of purchases on sale is positively correlated with number of different stores visited, number of brands bought, and frequency of going to store.

Third, when comparing both purchases for a given household over time, and across households, we find that purchases on sale tend to be of more units and larger sizes. We find a difference between sale and non-sale purchases in both duration from previous purchase and duration to next purchase. The difference is smaller within a household than the difference across, or between, households.

Fourth, assuming constant consumption over time we construct an inventory for each household. We find, as predicted by the model, that this variable is negatively correlated with the quantity purchased (conditional on a purchase) and with the probability of buying conditional on being in a store. Also, as predicted, the size of the effect varies depending if the purchase was on sale or not.

In summary, we find evidence that is consistent with our model. Some of the effects we find, while statistically significant are smaller than we expected. Our analysis suggests that this is driven by a combination of measurement error and a non-linear effect. Both of these will be handled, at least partly, in the structural model. Furthermore, the structural model will allow us to perform some counterfactual experiments, which will address the questions we used to motivate the analysis.

1.1 Literature Review

There are several theoretical papers that offer explanations for sales. Varian (1980) develops a model in which there are two types of consumers: those that have a low cost of search for information, and are informed about prices, and those that have high search costs, and are therefore uninformed. He assumes that uninformed consumers choose a store at random and buy if the price is below their reservation, while the informed consumers go to the store with the lowest price. In equilibrium, firms randomize over prices. Randomization is important in order to justify why the uninformed do not become informed after finding a single low price. Several papers have used similar models, but rather than informed and uninformed consumers they use switchers and non-switchers (for example, see Narasimhan, 1988, and Rao, 1991).

Sobel (1984) presents a model of a competitive industry selling a durable good. He assumes that two types of consumers –high valuation and low valuation– arrive in the market over time, and that low valuation consumers are willing to postpone their purchases if the price is above their valuation. He demonstrates that sellers will periodically, and randomly, find it optimal to lower prices to clear out low valuation consumers. Conlisk, Gerstner and Sobel (1984) and Sobel (1991) present a monopolistic version of this model. Equilibrium involves cyclical prices, although, in contrast to Sobel (1984) the cycle is deterministic.

Salop and Stiglitz (1982) present a competitive industry model with price dispersion. Firms randomize between two prices. Low prices generate higher sales, as consumers purchase to store an extra unit when they find a lower price. High prices generate higher per unit profits but lower sales, as buyers facing a high price only buy for current consumption. Hong, McAfee and Nayyar (2000) also present a storable good model that generates demand dynamics, namely, a (negative) link between current prices and future demand. As in Varian there are shoppers and non-shoppers, and as in Salop and Stiglitz consumers, the shoppers, can keep up to one unit of inventories. It is shown that there exist equilibria where firms use random pricing. Moreover, prices are negatively correlated over time. This is the only model that allows for storage in a dynamic set up. Price randomness, in this model, is driven by two assumptions. Consumers are assumed to chose a store based on the price of a single item and firms are informed about other firms' prices and hence sales. We analyze the complementary case where supermarkets cannot monitor the rival prices', due to the number of products sold, and consumers decision of where to shop is exogenous.

There are not many empirical studies of sales in the economics literature. Hosken et al. (2000) study the probability of a product being put on sale as a function of its attributes. Papers closer to our approach are Pesendorfer (forthcoming), which studies sales in the ketchup market using similar, but not identical, regressions to the indirect evidence we consider below, and Erdem et al (2000), who construct a structural model. Besides several modeling assumptions we differ from the latter paper in the focus. Erdem et al focus on the demand side while our ultimate interest lies in the supply side.

There are numerous studies in the marketing literature that examine the effects of sales, or more generally the effects of promotions (for example, see Blatteberg and Neslin, 1990, and references therein). Closest to our approach are the papers that examine the effect of sales on household stockpiling. Several papers[3] use household-level data to show that when purchasing during a promotion households tend to buy more units, larger sizes and in shorter duration to their previous purchase. We also examine some of the same quantities, however, we control for

[3]For example, see Ward and Davis (1978), Shoemaker (1979), Wilson, Newman and Hastak (1979), Neslin, Henderson and Quelch (1985), Gupta (1988), Chiang (1991), Grover and Srinivasan (1992) and Bell, Chiang and Padmanabhan (1999).

differences across households by using the panel structure of the data. Blattberg, Eppen and Lieberman (1981) are concerned with the relationship between retailer and household inventory policies. Their model is somewhat similar to ours and like the above mentioned work they present evidence similar to some of our indirect evidence.

Based on the results from the household-level data, there have been some attempts to find a dip in the (aggregate) quantity sold following a sale. The difficulty is finding this effect is noted in Blattberg, Briesch and Fox (1995). Neslin and Schneider Stone (1996) discuss eight possible arguments for why this might be the case. Van Heerde, Leeflang and Wittink (2000) empirically examine the importance of these arguments. Our results directly shed light on this “puzzle”.

Finally, several recent papers have studied price adjustment and its implications from various perspectives. These include Warner and Barsky (1995), Chevalier, et al. (2000) and MacDonald (2000), who study the seasonality of price adjustments. Feenstra and Shapiro (2001) study the implications of sales for computation of a price index.

2. The Data and the Industry

2.1 *Data*

The main data set used in this paper consists of price and quantity store scanner data and has two components, store and household-level data. The first was collected using scanning devices in nine supermarkets, belonging to different chains, in two separate sub-markets in a large mid-west city. Besides the price charged and (aggregate) quantity sold we know promotional activities that took place, for each detailed product (brand-size) in each store in each week. The second component of the data set is household-level data in which we observe the purchases of roughly 1,000 households over a period of 104 weeks. We know when a household visited a supermarket and how much they spent each visit. The data includes purchases in 24 different product categories for which we know exactly which product each household bought, where it was bought, how much was paid, and whether a coupon was used.

2.2 *The Industry*

We first focus on laundry detergents.[4] Laundry detergents come in two main forms: liquid and powder. Liquid detergents account for 70 percent of the quantity sold. Unlike many other consumer goods there are a limited number of products offered. Table 1 shows the market shares of the top selling detergents. The top eight (six) brands account for 75 percent of the liquid (powder) volume sold.

Most brand-size combinations have a regular price. In our sample 71 percent of the weeks the price is at the modal level, and above it only approximately 5 percent of the time. Defining as a sale as any price below the modal price, we find that in our sample 43 and 36 percent of the volume sold of liquid and powder detergent, respectively, was sold during a sale. There is some variation over time and across products in the percent sold on sale, as can be seen in Table 1. The median discount during a sale is 40 cents, the average is 67 cents, the 25 percentile is 20 cents and the 75 percentile is 90 cents. In percentage terms the median discount is 8 percent, the average is 12 percent, and the 25 and 75 percentiles are 4 and 16 percent, respectively.

Detergents come in many different sizes. However, about 97 percent of the volume of liquid detergent sold was sold in 5 different sizes:[5] 128 oz. (55%), 64 oz. (31%), 96 oz. (8%), 256 oz. (2%),

---

[4] In the future we would like to extend the analysis to other categories.

[5] Towards the end of our sample Ultra detergents were introduced. These detergents are more concentrated and therefore a 100 oz. bottle is equivalent to a 128 oz. bottle of regular detergent. For the purpose of the following numbers we aggregated 128 oz. regular with 100 oz. Ultra, and 68 oz. with 50 oz.

32 oz. (2%). Sizes of powder detergent are not quite as standardized. 56 different sizes are available, with the top 10 sizes accounting for approximately 70 percent of sales. Prices are non-linear in these sizes. Table 2 shows the price per 16 oz. unit for several container sizes. The figures are computed by averaging the, un-weighted and quantity-weighted, per unit price in each store over weeks and brands. The numbers suggest some per unit discount for the largest sizes. However, most of the non-linearity in prices is driven by the high prices of the smallest container size (32 oz.). The table also suggests that average price and quantity discounts differ across stores.

The figures in Table 2 are averaged across different brands and therefore might be slightly misleading since not all brands are offered in all sizes or at all stores. Therefore, some of the previous conclusions could be driven by differences in the mix of brands. For this purpose Table 3 presents the same figures but focuses on one brand, Tide. For this product there does seem to be, at least for some stores, a per unit discount even when comparing 64 oz. containers to 128 oz. containers. The difference in the share of quantity sold of each size seems to be highly correlated with the size of the discount.

The figures in both Tables 2 and 3 average across sale and non-sale periods. Defining a sale as any price below the modal price of a brand-size-store combination, on average 23 percent of the observations are sales. 9 percent of the observations have a price that is less than 90 percent of the modal price. Table 4 presents the modal, non-sale, price and sale frequency by store and size. The quantity discounts, which we observed in the previous two tables, are present also in the modal price although on a much smaller scale. Smaller sizes tend to have less sales with the smallest size having essentially no sales. Therefore, the quantity discounts we observe in the previous two tables were indeed a mixture of non-linear non-sale prices and more sales for the larger sizes. Finally, we can see from Table 4 that there are differences in the frequency of sales across stores.

Our data records two types of promotional activities: *feature* and *display*. The *feature* variable measures if the product was advertised by the retailer (e.g., in a retailer bulletin sent to consumers that week.) The *display* variable captures if the product was displayed differently than usual within the store that week.[6] The correlation between a sale, defined as a price below the modal, and being featured is 0.38. Conditional on being on sale, the probability of being featured is less than 20 percent. While conditional on being featured the probability of a sale is above 93 percent. The correlation with *display* is even lower at 0.23. However, this is driven by a large number of times that the product is displayed by not on sale. Conditional on a display, the probability of a sale is only 50 percent. If we define a sale as the price less than 90 percent of the modal price, both correlations increase slightly, to 0.56 and 0.33, respectively.

3. The Model

In this section we present a model of consumer inventory behavior. We want to use this model for two somewhat different purposes. We plan to structurally estimate the parameters of the model, and therefore the model has to be rich enough to deal with the complexity of the data. We also need the model in order to derive predictions which can be taken to the data without imposing as much structure. We start by discussing a richer model, which we will take to the data directly, and then we simplify the model somewhat in order to obtain some analytic predictions, which we can take to the data in order to test the model indirectly.

3.1 The Basic Setup

We consider a model in which a consumer obtains the following per period

[6] These variables both have several categories (for example, type of display: end, middle or front of aisle). For now we treat these variables as dummy variables.

$$u(c_{it}, \nu_{it}; \theta_i) + \alpha m_{it}$$

where c_{it} is the quantity consumed of the good in question, ν_{it} is a shock to utility, θ_i is a consumer-specific vector of taste parameters and m_{it} is the outside good consumption. The stochastic shock, ν_{it}, captures demand shocks unobserved to the researcher. For simplicity we assume the shock is additive in consumption, $u(c_{it}, \nu_{it}; \theta_i) = u(c_{it} + \nu_{it}; \theta_i)$. The shock affects the marginal utility from consumption. Low realizations of ν_{it} increase the household's need, increasing demand and making it more inelastic. We also assume $\frac{\partial u(c_{it} + \nu_{it}; \theta_i)}{\partial c} \geq \alpha p \ \forall p$ and $\forall \nu$, which is a sufficient condition for positive consumption every period. This assumption has no major impact on the predictions of the model, while it avoids having to deal with corner solutions.

The product is offered in *J* different varieties, or brands. Below we show how the consumer chooses between brands. Note that consumption is not indexed by brand, i.e., the utility depends only on the total consumption of detergents. This is key to deriving the estimation algorithm we propose below. The consumer faces random and potentially non-linear prices. Let $p_{jt}(x)$ be the total price associated with purchasing size *x* of brand *j*.

Given the prices of all brands and all sizes, the consumer at each period has to decide which brand to buy, where $d_{ijt} = 1$ denotes a choice of brand *j*, how much to buy, denoted by x_{it} (if $x_{it} > 0, \sum_j d_{ijt} = 1$), and how much to consume. Rather than letting consumption be endogenously determined we could let it be set at a fixed rate or let it be described by a random variable distributed (independently of prices). Both these alternative assumptions, which have been made by previous work, are nested within our framework. All the results below hold, indeed the proofs and estimation of the structural model are both simpler. We feel it is important to allow consumption to vary in response to prices since this is the main alternative explanation to why consumers buy more during sales, and we want to make sure that are results are not driven by assuming it away.

Since the good is storable, quantity not consumed is kept as inventory for future consumption. Facing different prices over time, consumers have to decide whether to purchase immediately or wait for a lower price in the future. For now we assume the consumer visits one store each week, i.e., we do not model the consumer's decision of when to shop. In Appendix A we discuss an extension that allows consumers to vary in the frequency of visiting stores. Facing different prices over time, consumers have to decide whether to purchase immediately or wait for a low price in the future. After dropping the subscript *i*, in order to simply notation, the consumer's problem can be represented as

$$
\begin{aligned}
V(I(0)) = \max_{\{c_t, x_t, d_{jt}\}} \sum_{t=0}^{\infty} \delta^t E\Big[u(c_t, \nu_t; \theta) - C(i_t) + \textstyle\sum_j d_{jt}\big(\alpha p_{jt}(x_t) + A_{jxt}\big) \big| I(0)\Big] \\
s.t. \quad 0 \le i_t, \quad 0 \le x_t \\
i_t = i_{t-1} + x_t - c_t
\end{aligned}
\tag{1}
$$

where α is the marginal utility from income, A_{jxt} captures the effects of brand differentiation and advertising, $I(t)$ denotes the information at time *t*, $\delta > 0$ is the discount factor, and $C(i_t)$ is the cost of storing inventory.[7] The effect of differentiation and advertising is $A_{jxt} = \xi_j + \beta a_{jxt} + \varepsilon_{jxt}$ where ξ_j, is a fixed taste of brand *j* that could be a function of brand characteristics and could vary by consumer, βa_{jxt}, the effect of advertising variables on the consumer choice, and ε_{jxt}, a random shock that impacts the consumer's choice. Note that the latter is size specific, namely, different sizes get different draws introducing randomness in the size choice.

The information set at time *t* consists of the current inventory, i_{t-1}, current prices, the current

[7]Notice we do not need to impose $c \geq 0$ since we assumed u' is such that there is always positive consumption.

shock to utility from consumption, ν_t, and the vector of random shocks.[8] Consumers face two sources of uncertainty: future utility shocks and unpredictable future prices. We assume that the consumer knows the current shock to utility from consumption, ν_t,[9] and that these shocks are independently distributed over time. Prices are (exogenously) set according to a first-order Markov process, which we describe in the Section 4.[10] Finally, the random shocks, ε_{jxt}, are assumed to be independently and identically distributed according to a type I extreme value distribution.

3.2 Deriving predictions

Before we estimate this model structurally we would like to indirectly test its relevance. Therefore, we derive predictions and their implications for patterns we can potentially observe in both aggregate and household-level data. To study the properties of the model we solve for optimal consumer behavior in equation (1) under two simplifying assumptions:

Assumption 0: Only one brand of the good is offered, it is not advertised (i.e., $A_t(x_t) = 0$), and it is sold in continuous amounts at linear prices.

Assumption 1: Assume $F(p_{t+1}|p_t)$ first order stochastically dominates $F(p_{t+1}|p'_t)$ for all $p_t > p'_t$.

Under these simplifying assumptions the consumer's behavior can be described as follows. In each period a consumer weights the costs of holding inventory against the (potential) benefits

[8] In principle, the information set also includes the advertising variables, a_{jt}, (and the taste for the brand if it varies over time). However, as we show below for the purpose of specifying the dynamic process we collapse all these variables into a "quality" adjusted price. Therefore, in order to simplify notation we include only price in the information set.

[9] It is quite reasonable to assume that at the time of purchase the current utility shock has still not been realized. This will generate an additional incentive to accumulate inventory – the cost of a stock out. Since this is not our focus, we ignore this effect for now, but it can easily be included in the application.

[10] In principle we can deal with the case where utility shocks are correlated over time. However, this significantly increases the computational burden since the expectation in equation (1) will also be taken conditional on ν_t (and potentially past shocks as well). In future extensions we would like to endogenize the price process.

from buying at the current price instead of future expected prices. She will buy for storage only if the current price and her inventory are sufficiently low. At high prices the consumer might purchase for immediate consumption, depending on her inventory and the realization of the random shock to utility. We now formally characterize this behavior.

The solution of the consumer's inventory problem is characterized by the following Lagrangian

$$\max_{\{c_t, x_t, i_t\}} E(\sum_{t=0}^{\infty} \delta^t \{u(c_t + \nu_t;\theta) - C(i_t) - \alpha p_t x_t + \lambda_t(i_{t-1} + x_t - c_t - i_t) + \psi_t x_t + \mu_t i_t\} \mid I(t)) \quad \textbf{(2)}$$

where μ_t, ψ_t, and λ_t are the Lagrange multipliers of the constraints in equation (1). From equation (2) we derive the first order conditions with respect to consumption,

$$u^{/}(c_t + v_t;\theta) = \lambda_t, \quad \textbf{(3)}$$

purchase,

$$\alpha\, p_t = \lambda_t + \psi_t, \quad \textbf{(4)}$$

and inventory,

$$C'(i_t) + \lambda_t = \delta E(\lambda_{t+1} | I(t)) + \mu_t. \quad \textbf{(5)}$$

Using these conditions we derive the basic predictions of the model. We show that consumers follow a (conditional) S-s type of behavior, where the target inventory level is a function of current price only, $S(p)$, and the trigger inventory level depends both on prices and the utility shock, $s(p,v)$.

Let $c^*(p_t,\nu_t)$ be the consumption level such $u^{/}(c^*(p_t,\nu_t) + \nu_t) = \alpha p_t$ and let $S(p)$ be the inventory level such $C'(S(p)) + \alpha\, p_t = \delta E(\lambda_{t+1} | I(t))$.

Proposition 1 In periods with purchases, $x_t > 0$, the target level of inventory, i_t, equals $S(p_t)$, a decreasing function of p_t, independent of the other state variables i_{t-1} and ν_t. Moreover, the inventory level that triggers a purchase is $s(p_t,\nu_t) = S(p_t) + c^*(p_t,\nu_t)$, which is decreasing in both arguments.

Proof of Proposition 1: If $x_t > 0$ then $\psi_t = 0$. If $i_t = 0$, there is nothing to show, simply $S(p_t) = 0$. In the complementary case, $i_t > 0$, we know $\mu_t = 0$. Using equation (4) and $\mu_t = \psi_t = 0$, equation (5) becomes: $C'(i_t) + \alpha p_t = \delta E(\lambda_{t+1} | I(t))$, which shows the end-of period inventory, i_t, is independent of the states variables i_{t-1} and ν_t. Moreover, since $F(p_{t+1}|p_t)$ increases in p_t, by equation (3) we get that the right hand side of the last equality declines in p_t. Hence, since $C^{//}>0$ the end of period inventory, i_t, declines in price.

To show that the inventory level that triggers a purchase is $S(p_t) + c^*(p_t,\nu_t)$, assume first that the consumer is willing to buy when she has an initial inventory $i_{t-1} > S(p_t) + c^*(p_t,\nu_t)$. In such a case, $i_t > S(p_t)$, which violates equation (5) since it would hold with equality for $i_t = S(p_t)$, but the left-hand side is bigger and the right-hand side smaller for $i_t > S(p_t)$. Now suppose the consumer does not want to purchase when $i_{t-1} < S(p_t) + c^*(p_t,\nu_t)$. Since $x_t = 0$ we know $\psi_t > 0$, which in turn, by equation (3), implies $c_t > c^*(p_t,\nu_t)$. Hence, $i_t < S(p_t)$, which implies equation (5) cannot hold. By definition, it holds for $S(p_t)$, but for $i_t < S(p_t)$ the left-hand side is lower than the right-hand side. We conclude that the inventory $i_{t-1} = S(p_t) + c^*(p_t,\nu_t)$ triggers purchases. ■

Remark: If only discrete quantities are available or prices are non-linear in quantities then the target inventory $S(\cdot)$ becomes a function of i_{t-1} and ν_t. As we see below this implies that we can not test this proposition directly with our data.

Inventories are an important dimension of household behavior, however since we do not observe inventories in our data, we present next predictions about purchases on which we have detailed data.

Proposition 2 $x(i_{t-1}, p_t, \nu_t)$ declines in the three arguments.

Proof of Proposition 2: There are two cases to consider. Case 1: $x_t > 0$ and $i_t = 0$. In this case purchases equal consumption minus initial inventories: $x(i_{t-1}, p_t, \nu_t) = c(i_{t-1}, p_t, \nu_t) - i_{t-1}$. Since $x_t > 0$ we can combine equations (3) and (4) to get $u^{/}(c_t + \nu_t; \theta) = \alpha p_t$, which implies that $c(i_{t-1}, p_t, \nu_t)$ declines in ν_t and p_t, and is independent of i_{t-1}. Thus, $x(i_{t-1}, p_t, \nu_t)$ declines in ν_t, p_t, and i_{t-1}.

Case 2: $x_t > 0$ and $i_t > 0$. From Proposition 1 we know $x(i_{t-1}, p_t, \nu_t) = S(p_t) + c(i_{t-1}, p_t, \nu_t) - i_{t-1}$. The result follows from Case 1 and Proposition 1, which showed $S(p_t)$ declines in p_t. ■

Corollary 1: There is a price $p^r < p^m$, where p^m is the highest (non-sale) price, such that at any price $p \geq p^r$ if consumers buy they do so for current consumption exclusively.

Proof of Corollary 1*:* At p^m if $x_t = 0$ there is nothing to show. If $x_t > 0$ we can combine equations (4) and (5) to get $C'(i_t) + \alpha p_t = \delta E(\lambda_{t+1} | I(t)) + \mu_t$. Moreover, from equation (4) we know $E(\lambda_{t+1} | I(t)) \leq \alpha E(p_{t+1} | I(t))$. The right hand side of the last inequality is strictly lower than αp^m (as long as prices lower than p^m, arise with positive probability). Hence, since $\delta \leq 1$ we know $C'(i_t) + \alpha p^m > \delta E(\lambda_{t+1} | I(t))$ for any $i_t \geq 0$. Therefore, equation (5) can hold with equality only if $\mu_t > 0$, i.e., when $i_t = 0$. Since the inequality is strict it holds also for some $p^r < p^m$. Concluding the proof that if any quantity is purchased, it is for consumption only, since no inventories will be left at the end of the period. ■

Proposition 3: Holding p_t and ν_t constant, if $i^{/}_{t-1} > i^{//}_{t-1}$ then $i_t(i^{/}_{t-1}, p_t, v_t) \geq i_t(i^{//}_{t-1}, p_t, v_t)$, namely the target level of inventory, i_t, is an increasing function of i_{t-1}.

Proof of Proposition 3: There are three cases to consider. Case 1: Both levels of inventory trigger purchase. By Proposition 1 the target level of inventory is *S*(*p*), which is independent of initial inventorty, and therefore the result holds. Case 2: $i_{t-1}^{//}$ triggers purchase but $i_{t-1}^{/}$ does not. By the second part of Proposition 1this implies that $i_{t-1}^{/} > S(p) + c^*(p_t, \nu_t)$. Since no purchase was made optimal consumption will be (weakly) less than $c^*(p_t, \nu_t)$. Therefore, $i_t(i_{t-1}^{/}, p_t, \nu_t) > S(p) = i_t(i_{t-1}^{//}, p_t, \nu_t)$.

Case 3: Neither inventory level triggers purchase. If the optimal consumption $c(i_{t-1}, p_t, \nu_t)$ is decreasing in i_{t-1} then since there is no purchase the result trivially holds. Consider the case where the optimal consumption is increasing in i_{t-1}. Suppose that i_t decreases in i_{t-1}. Plugging equation (3) into equation (5), we see that the left-hand side of equation (5) declines in i_{t-1}. However, the right hand side increases in i_{t-1}. Since we supposed that consumers with higher i_{t-1} have a lower i_t, as i_t decreases the consumer will have a higher expected future marginal utility from consumption. Moreover, if the non-negativity constraint binds, it adds another positive term in the right hand side. This leads to a contradiction, which implies i_t increases in i_{t-1}, during non-purchase periods. ■

Up to now we have focused on characterizing a single consumers optimal behavior. Next we generate cross household comparative statics. We start by studying how the optimal behavior changes with a change in storage cost. Assume storage costs are such that: $C'_h(i)$ increases in *h* for any inventory *i*.

Proposition 4: Consumers with higher storage costs (higher h), all else equal, hold lower inventories, purchase more frequently and in lower quantities.

Proof of Proposition 4: Follows directly from equation (5).

Next, we examine how the optimal inventory behavior changes with changes in the household consumption needs, captured by consumer characteristics θ_i. Assume that these characteristics are such that $\frac{\partial^2 u(c_{it}, \nu_{it}; \theta_i)}{\partial\theta\partial c} \geq 0$, namely, consumers with higher θ have a higher marginal utility from consumption.

Proposition 5: Consumers with marginal utility from consumption (higher θ) hold higher inventories, i.e. all else equal, $S(p_t)$ is higher.

Proof of Proposition 5: Suppose $S(p_t)$ decreases (weakly) in θ. Plugging equation (3) into equation (5), we see that the left-hand side of equation (5) declines in θ. However, the right hand side increases in θ, as consumers with higher θ (i) hold lower inventories, hence expect to consume (weakly) less in the next period, (ii) have a higher marginal utility from consumption and (iii) are more likely to have the constraint on inventory binding. This leads to a contradiction, which implies $S(p_t)$ increases in θ. ■

3.3 Testable Implications

In this section we state the model's predictions that can test with the data described in Section 2. We first present the implication of Propositions 1 and 2 on the impact of inventories on purchases. As we do not observe consumer inventories we use two different strategies. First, we will make several assumptions that will allow us to come up with a proxy for the unobserved inventory. Second, we resort to prediction on other aspects of consumer purchase behavior, which indirectly inform us about the stockpiling behavior. Therefore, before turning to the indirect predictions which match our data, we present those regarding inventories to be tested by proxies, which follow Propositions 1 and 2.

Implication I1: *Quantity purchased and the probability of purchasing decline in inventories*.

We now turn to predictions more relevant for our data, which does not include information on inventories. From Proposition 2 we know that during sales quantity purchased is bigger. Quantity purchased can increase simply because consumption increases when the price decreases (see equation (3)). However, our model predicts an additional effect: purchases are bigger during sales because stockpiling occurs only then (Corollary 1).[11] Since we do not know the size of the consumption effects, showing that quantity purchased increases during sales is a weak test of our theory.

The consequence of stockpiling is a higher end of period inventory, which , all else equal, implies a longer duration until the next time the consumer hits the threshold for purchase, *s*. Alternatively, if sales involved a higher purchase only due to consumption effects, duration would be unaffected by sales. This gives the following implication, which indirectly testifies to the

[11] Corollary 1 describes the stockpiling effect for every price less than p^r, which we have implicitly equated with a sale. In the empirical analysis we do not observe p^r and therefore we will experiment with several definitions of a sale.

presence of stockpiling.

Implication I2: *Duration until next purchase is longer during a sale.*

From Proposition 1 we know that the inventory that triggers a purchase, $s(\cdot)$, is lower at non-sale prices. Hence, according to our model, since during a non-sale purchase inventory is lower, on average the duration from previous purchase will be longer. Furthermore, if the previous purchase was on sale then, all else equal, their inventory would have been higher. Then by Proposition 3 the consumer's inventory would be higher today, relative to their inventory if the previous purchase was not during a sale. Therefore, conditional on purchasing on non-sale today, it is more likely that the previous purchase was not during a sale. This leads to the following implications.

Implication I3: *Duration from previous purchase is shorter at sale periods.*

Implication I4: *Non-sale purchases have a higher probability that the previous purchase was not on a sale, namely:* $Pr(NS_{t-1}|S_t) < Pr(NS_{t-1}|NS_t)$, *where S =sale purchase and NS=non-sale purchase.*

In the model we take the frequency with which a household observes a price as given, and equal across households. However, in our sample some households are exposed to larger numbers of price draws; either because they visit the supermarket more frequently, or because they visit more supermarkets per week, or because they perceive a larger number of products as close substitutes. Households observing more draws are able to purchase at sale prices more often, they may not need to store as much when they find a sale. Another dimension in which households are likely to differ is their ability to store. The impact of storage cost is given by Proposition 4. The following implication summarizes these dimensions of household heterogeneity.

Implication I5: *The proportion of purchases on sale increases with: (i) the number of price draws a consumer gets and (ii) the ability to store.*

From Proposition 2 and Corollary 1 we know that at the non-sale prices consumers only purchase for current consumption, hence price only affects consumption. During sales, however, consumers react to price reductions not only by consuming more but also by accumulating inventories.

Implication I6: *Consumers are more price sensitive during sales*.

We now turn to implications on aggregate demand. The aggregation of implication I2 over a population that visits the supermarket at different periods leads to implication I7, namely, that store level demand increases with duration since last sales. Moreover, since at non-sale periods consumers only demand for current consumption, while, on sale they hoard inventories (Corollary 1) we expect duration to have stronger effects during sales.

Implication I7: *Aggregate demand increases in the duration from the previous sale.*

Implication I7': *Duration effects are stronger during sale*.

4. Econometrics

Using the data described in Section 2.1 we test the model in two ways. First, we examine the theory indirectly by examining some of its implications. This step is described in detail in the next section. For reasons we motivate below, we go beyond testing the implications of the theory and structurally estimate the model. The structural estimation is based on the nested algorithm proposed by Rust (1987), but has to deal with issues unique to our problem. We start by providing a general overview of our estimation procedure and then discuss some of the more technical details.

4.1 An overview of the estimation

We base our estimation on the "nested algorithm" proposed by Rust (1987). The procedure is based on nesting the (numerical) solution of the consumer's dynamic programming problem within the parameter search of the estimation. The solution to the dynamic programming problem yields the consumer's deterministic decision rules, i.e., for any value of the state variables the consumer's optimal purchase and consumption . However, since we do not observe the random shocks, which are part of the state variables, from our perspective the decision rule is stochastic. Assuming a distribution for the unobserved shocks we derive a likelihood of observing the decisions of each consumer (conditional on prices and inventory). We nest this computation of the likelihood into a non-linear search procedure that finds the values of the parameters that maximize the likelihood of the observed sample.

We face two main hurdles in implementing the above algorithm. First, we do not observe inventory since both the initial inventory and consumption decisions are unknown. We deal with the unknown inventories using the model to derive the optimal consumption[12] in the following way.

[12] There are at least a couple of alternative ways to construct a consumption series. First, we could assume that weekly consumption is constant, for each household over time, and estimate it by the total purchase over the whole period divided by the total number of weeks. Alternatively, we could assume that consumption is an exogenously given

Assume for a second that the initial inventory is observed. Therefore, we can use the procedure described in the previous paragraph to obtain not only the likelihood of the observed purchases, but also the probability of different consumption levels, and therefore the likelihood of different inventory levels, at time $t = 1$. For each inventory level we can again use the procedure of the previous paragraph to obtain the likelihood of the observed purchase, but now we account for the distribution of the inventory level when computing the likelihood. We can continue this procedure to obtain the likelihood of observing the sequence of purchases for each household. In order to start this procedure we need a distribution for initial inventory. We experiment with using the ergodic distribution of inventory for each household, and with using part of the data in order to simulate the initial distribution of inventory.

The second, and more difficult, problem is the dimensionality of the state space. If there were only a few brand-size combinations offered at a small number of prices, then the above would be computationally feasible. In the data over time households buy several brand-size combinations, which are offered at many different prices. This makes the "standard" approach computationally infeasible. We therefore propose the following three-step procedure. The first step, consist of maximizing the likelihood of observed brand choice <u>conditional</u> on the size (quantity) bought in order to recover the marginal utility of income, α, and the parameters that measure the effect of advertising, β and ξ's. As we show below, we do not need to solve the dynamic programming problem in order to compute this probability. In the second step, using the estimates from the first stage, we compute the "inclusive values" for each size (quantity) and their transition probabilities from period to period. This allows us, in the final step, to apply the nested algorithm discussed above to the a simplified problem in order to estimate the rest of the parameters. Rather than having the

random variable (Erdem et al, 2000).

state space include prices of all available brand-size combinations, it includes only a single "price" for each size. For our data set this is a considerable reduction in the dimension of the state space. We use this simplified problem to define and maximize the likelihood of purchasing a size (quantity).

4.2 The three step procedure

For a given value of the parameters the probability of observing the purchase decision (which brand and what size) as a function of the observed state variables (prices) is

$$Pr(d_{jt}=1,x_t|p_t)=\sum_{i_t} Pr(d_{jt}=1,x_t|i_t,p_t)Pr(i_t).$$

Given the assumption that ε_{jxt} follows an i..i.d. extreme value distribution, the probability of the purchase decision conditional on prices and inventory is

$$Pr(d_{jt}=1,x_t|i_t,p_t)=\int \frac{\exp\left\{\alpha p_{jt}(x_t)+A^1_{jxt}+\max_{c_t}\left\{u(c_t+\nu_t)+\delta EV(I_t;d_{jt}=1,x_t,c_t)\right\}\right\}}{\sum_{j,x}\exp\left\{\alpha p_{jt}(x_t)+A^1_{jxt}+\max_{c_t}\left\{u(c_t+\nu_t)+\delta EV(I_t;d_{jt}=1,x_t,c_t)\right\}\right\}}\, dF(\nu_t) \quad \textbf{(6)}$$

where $A^1_{jxt}=A_{jxt}-\varepsilon_{jxt}$ and EV(·) is the expected future value given today's state variables and today's decisions. Note that the summation in the denominator of equation (6) is over all brands and all sizes. The probability of inventory $Pr(i_t)$ is computed, as described above, by starting with an initial distribution and updating it using observed purchases and optimal consumption computed from the model. This probability can be used to form a likelihood, but as was pointed above (and as we can see from this equation) it requires keeping as state variables all the prices of all brand-size combinations. We therefore propose an alternative three-step procedure.

In the first step, we estimate part of the preference parameters (the marginal utility of income, α, and the parameters that measure the effect of advertising, β and ξ's) using a static model of brand choice conditional on the size (quantity) purchased. In other words, we estimate a logit, restricting the choice set to options of the same size (quantity) actually bought in each period. The static estimation yields consistent, but potentially inefficient, estimates of these parameters.

We now want to justify the first step of our algorithm. Let $c_k^*(x_t, \nu_t)$ be the optimal consumption conditional on a realization of ν_t and purchase of size x_t of brand *k*.

Lemma 1: $c_j^*(x_t, \nu_t) = c_k^*(x_t, \nu_t)$.

Proof: (in the Appendix).

Conditional on the size purchased the optimal consumption is the same regardless of which brand is chosen.

Given this lemma and that in our model $EV(I_t; d_{jt}=1, x_t, c_t) = EV(I_t; x_t, c_t)$, namely the brand of the inventory does not affect future utility. All terms involving future expected utility in the brand choice cancel, thus, from equation (6)

$$Pr(d_{jt}=1 | x_t, i_t, p_t, \nu_t) = \frac{\exp\left\{\alpha p_{jt}(x_t) + A_{jxt}^1\right\}}{\sum_k \exp\left\{\alpha p_{kt}(x_t) + A_{kxt}^1\right\}} = Pr(d_{jt}=1 | x_t, p_t)$$

where the summation is over all brands available in size x_t at time *t*. In order to compute this probability we do not need to solve the dynamic programming problem, nor do we need to generate an inventory series. Therefore, the marginal utility of income, and the parameters that enter A_{jxt}^1 can be estimated by maximizing this probability. This amounts to estimating a brand choice logit using only the choices with the same size as the size actually purchased.

In the second step, using the estimates from the first stage, we compute the "inclusive values" for each size (quantity) and their transition probabilities from period to period. The inclusive value

$$\omega_{xt} = \log\left\{\sum_k \exp\left(\alpha p_{kt}(x_t) + A_{kt}^1\right)\right\}.$$

can be thought of as a "quality" adjusted price index for all brands in that size category. Note, that since the parameters might vary by observed or unobserved consumer characteristics these values will differ by consumer.

The usefulness of the inclusive value, is that it collapses the state space to a single index per size, therefore reducing the computational cost. For example, instead of keeping track of the prices of nine brands times five sizes (roughly the dimensions in our data), we only have to follow five quality adjusted prices. The main loss is that transition probabilities have to be defined in a somewhat limited fashion. Two price vectors that yield the same vector of inclusive values will have the same transition probabilities to next period state, while a more general model will allow these to be different.[13] In reality, however, we believe this is not a big loss since it is not practical to specify a much more general transition process. For the results presented below we use the inclusive values estimated from the first step to estimate the following transition

$$Pr(\omega_{1,t}, ..., \omega_{S,t} | \omega_{1,t-1}, ..., \omega_{S,t-1}) =$$

$$N(\gamma_{10} + \gamma_{11}\omega_{1,t-1} + ... + \gamma_{1,S}\omega_{S,t-1}, \sigma_1) \cdot ... \cdot N(\gamma_{S0} + \gamma_{S1}\omega_{1,t-1} + ... + \gamma_{SS}\omega_{S,t-1}, \sigma_S)$$

where S is the number of different sizes and $N(\cdot,\cdot)$ denotes the normal distribution.

The usefulness of the inclusive values is twofold. First, it helps us to separate the probability of observed choices into the probability of choosing a brand conditional on size and the probability of choosing a size. Second, it helps reduce the computational burden of the dynamic problem since we need to solve the dynamic problem only in order to compute the latter probability. Each individual, maximizing her expected value of utility stream, computes her expected value function with respect to the future evolution of the inclusive value of each for each size only, as oppose to having to record as state variables all the characteristics (price, advertizing and feature) of each size and brand.

In the third, and final, step we feed the inclusive values, and the estimated transition probabilities, into the nested algorithm discussed above to compute the likelihood of purchasing a

---

[13] There is also a loss of a efficiency in the estimates, mentioned in the first step.

size (quantity). More precisely, using the definition of the inclusive values, and equation (6), we can write

$$Pr(x_t | i_t, \omega_{xt}) = \int \frac{\exp\left(\omega_{xt} + \max_{c_t}\left\{u(c_t + \nu_t) + \delta EV(I_t; x_t, c_t)\right\}\right)}{\sum_x \exp\left(\omega_{xt} + \max_{c_t}\left\{u(c_t + \nu_t) + \delta EV(I_t; x_t, c_t)\right\}\right)} \; dF(\nu_t).$$

It is this probability that we use to construct a likelihood function in order to consistently estimate the remaining parameters of the model.

Our estimates for the parameters of the utility from consumption, the cost of holding inventory and discount factor are those that maximize this likelihood. The likelihood is a function of the expected value function, which despite the reduction in the number of state variables, is still computationally burdensome to solve. We use approximation and simulation methods (Keane and Wolpin, 1994; Rust 1996, 1997; Bertsekas and Tsitsiklis, 1996) and parallel processing to reduce the computation time. We also hope in the future, once we allow for more heterogeneity in the dynamic parameters, to use the methods proposed by Ackerberg (2000) to reduce the number of times needed to solve the dynamic programming problem.

5. Results

In this section we use the data previously described to present evidence on the relevance of the theory outlined in Section 3. We start with indirect evidence constructed from both the aggregate and household-level data. The evidence generally confirm the implications derived in Section 3. Motivated by the indirect evidence we impose more structure, which allows us to examine the relevance of our theory directly. Also, the direct estimation yields estimates of the parameters of the model, which allows us to perform counterfactual experiments.

5.1 *Indirect Tests of the Theory*

5.1.1 Aggregate data: the effect of duration from previous sales

According to implication I7 demand should increase with the duration from the previous sale (i.e., as consumers run out of the inventory stockpiled during the last sale). Moreover, the impact of duration is stronger during sales, (implication I7'). Table 5 presents the results of regressing the log of quantity sold, measured in ounces, as a function of price, current promotional activity and duration from past promotional activity. Different columns present the results for different samples.

Using the whole sample, i.e., both sales and non-sales periods, contrary to the model's predictions the effect of duration is negative.[14] However, this result is driven by the correlation between sales and other promotional activities, like feature. Without controlling for duration from previous feature, which is one of the promotional activities, the coefficient on duration from sale captures both effects. Indeed, once we include the duration from previous feature, in column 2, the coefficient on duration is positive and significant as expected. We also tried to include duration from last display in the regression, but the coefficient was insignificant.

Restricting attention to sales periods, the effect of duration from previous sale is positive even before we control for duration from previous feature. Once we control for duration from previous feature the coefficient increases in magnitude. For the non-sale sample, the effect of duration from previous sale becomes positive only once we control for duration from previous feature. Consistent with implication I7', the effect of duration is stronger during sales periods.

5.1.2 Household sales proneness

[14]Duration is measured in weeks/100. In all the columns, even though the coefficient on duration squared is significant, the implied marginal effect will be of the same sign as the linear term for the range of duration values mostly observed in the data. Therefore, we limit the discussion to the linear coefficient on duration.

In addition to the aggregate data used to produce the results in Table 5, we also have data on the purchases of roughly 1,000 households over a period of two years. We first use these data to distinguish between those households that tend to buy on sale and those that do not. We regress the fraction of times the household bought on sale during the observed period on various household characteristics.[15] Results, presented in Table 6, show that demographics have little explanatory power. In column (i) we see that households without a male tend to buy less on sale, as do households with a female working less than 35 hours a week. Households with higher per person income are less likely to buy on sale, and so are households with a female with post high school education. These effects are just barely statistically significant, and some not significant, at standard significance levels. Overall the observed demographics explain less than 3 percent of the variation, across households, in the fraction of purchases on sale. Both the direction and lack of significance of these results is consist with previous findings (Blattetberg and Neslin, 1990).

While the frequency a household buys on sale is not strongly correlated with standard household demographics it is correlated with two other household characteristics, relevant from the theory perspective. First, households that live in market 1 tend to buy less on sale. This is true even after controlling for many demographic variables including income, family size, work hours, age and race, as seen in column (ii). Market 1 has smaller homes with less rooms and bedrooms, relative to the other market. Under the assumption that home size proxies storage costs, this finding is consist with our model that predicts that lower storage costs are correlated with purchasing more frequently on sale (I5.ii). Second, though we know nothing about each households' house, we know the number of dogs they own. Columns (iii) shows that the having a dog is positively, and significantly, correlated with purchasing on sale, even after we control for other household characteristics. At the

---

[15]We also looked at the fraction of quantity purchased on sale. The results were essentially identical.

same time owning a cat is not. Assuming that dog owners have larger homes, while cat owners do not, this further supports our theory. Dog ownership is not just a proxy for the market since the effects persist once we also include a market dummy variable, as seen in column (iv).[16]

To test the first part of implication I5 we explore three proxies for the number of price draws a household observes. In columns (v) through (viii) we explore the correlation between frequency of purchasing on sale and the proxies: the number of stores, the frequency of visits to the stores and the number of different products the household purchased.

Households that bought in more than one store tend to buy more on sale: increasing the number of stores visited during the two year period by one, increases the frequency of purchasing on sale by 5 percentage points.[17] The percentage of households that buy in one, two or three stores is 22, 40 and 23, respectively. The relationship continues to hold if instead of number of stores visited we measure the concentration of expenditures across stores with an Herfidendal-like measure. Going from the 25 percentile household, with a concentration of 0.58, to the median, with a concentration of 0.82, will decrease frequency of buying on sale by about 8 percentage points.

Column (vi) shows that households that shop more frequently tend to buy more on sale. If the average duration between visits to a store increases by a day the frequency of purchasing on sale decreases by roughly 1.5 percentage points. The mean duration between visits is 6.2 days, the median is 5.7 and the 25 and 75 percentiles are 4.1 and 7.9, respectively.

Finally, the frequency of purchasing on sale is also correlated with the number of different

---

[16]Dogs might alternatively be a proxy for spare time, which may reflect a higher propensity to search. However, if dog was capturing propensity to search it would loose importance once we control for measures that proxy the propensity search (e.g., frequency of visits and number of stores). In fact the number of dogs is uncorrelated with those proxies, moreover, dogs' significance is not affected by controlling for search proxies (see column (viii)).

[17]The mean fraction of purchases on sale is 0.48, with a median of 0.5, 25 and 75 percentiles of 0.2 and 0.74, respectively.

brands a household purchased over the observed period. Each additional brand increases the probability of purchase by 2 percentage points. The percent of households that buy one through five brands is 17, 22, 21, 16 and 11, respectively. Since we want to distinguish between a household that buys the same brand almost always except for rare occasions, from the household that buys equal amount of two brands, we also constructed a Herfindel-like measure of the concentration of quantity purchased of different brands. The results suggest that moving from the 25 percentile (0.35) to the median (0.50) to the 75 percentile (0.82) of the brand concentration decreases the frequency of purchasing on sale by 3 and 9 percentage points, respectively. All these effects also hold once we control for the characteristics used in columns (i) - (iv). These findings regarding storage cost proxies and the frequency of shopping support the predictions in I5.

5.1.3 Sale vs. non-sale purchases

Next, we compare sale and non-sale purchases. The results presented in Table 7 suggest that when purchasing on sale households buy more units and larger sizes. This is true both when comparing between households (households that make a larger fraction of their purchases during sales tend to buy larger sizes) and within a household over time (when buying during a sale a household will tend to buy a larger size), as predicted by Proposition 2.

It is also shown in Table 7 that duration to next purchase is bigger for purchases on sale, while duration from previous purchase are shorter while on sale. These finding match the within household duration predictions implied by implications I2 and I3.

Notice that both implications I2 and I3 are within household implications. However, they have between households counterparts, namely, those households that consume more, higher θ, buy more on sale (Proposition 4). Indeed all the between duration effects are positive. This is quite natural, as households more prone to buy on sale buy bigger quantities, hence less frequently.

Finally, we find that the probability the previous purchase was not on sale, given that current purchase was not on sale is higher, as implied by prediction I4. The reasoning behind the prediction is that since non-sale purchases have a lower inventory threshold (namely, inventories have to be low for the buyer not to be willing to wait for a sale) a non-sale purchase informs us that inventories are low, which in turn means, other things equal, that the last purchase was not on sale. Notice the between household differences are a lot bigger. Suggesting a large cross-household heterogeneity in sales proneness, as those households buying today on sale, are a lot more likely to have purchased last time on sale as well.

5.1.4 Inventories, purchases and promotional activities

In Table 8 we present a set of regressions that study the link between quantity purchased by a household, conditional on a purchase, the price paid and promotional activities. The dependent variable in all the regressions is the quantity and the dependent variables include household-specific dummy variables (as well as dummy variables for each store and for each, broadly-defined, product).[18] The average price elasticity implied by the results in the column (i) is roughly -0.8 (with a median of roughly -0.3).[19]

One of the predictions of our theory is that the inventory a household holds should impact the quantity purchased (implication I1). We do not observe inventory therefore we construct a proxy,

---

[18] We also examined random effects models. The results were essentially identical, and therefore not reported.

[19] In a log-log equation to coefficient is roughly -1. Note, that none of these numbers should be interpreted as a demand elasticity. First, we restrict the sample to strictly positive purchases, i.e., we are examining the decision of how much to buy conditional on purchase. Second, the prices, as well as other variables, are for the product actually purchased and not a fixed product.

under the assumption of constant consumption over time. For each household we sum the total quantity purchased over the two year period. We divide this quantity by 104 weeks to get the average weekly consumption for each household.[20] Assuming the initial inventory for each household was zero, we use the consumption variable to construct the inventory for each household at the beginning of each week. This generated some negative inventories, which we can treat by adding a household specific initial inventory that assures that we do not get any negative inventories. Since we include a household-specific dummy variable these corrections do not matter (as long as the inventory variable enters the regression linearly).

The results, presented in column (ii) are consistent with implication I1: the higher the inventory a household holds the less they buy. The estimated coefficient suggests that each unit of (16 ounce) inventory reduces the quantity purchased by about 4.3 percent (or roughly two thirds of an ounce). In column (iii) we interact this variable with purchase on sale. The effect of inventory during a sale is higher than during non-sale periods. Although not predicted by the model, this difference seems reasonable, mainly because of the discreteness of units offered. During non-sale purchases, consumers are predicted to buy only for current consumption, namely, which in our data maps into a small container. Hence, as long as a single small container size is offered, inventories are likely to affect the probability of purchase, but not the quantity purchased.

The effect of inventories on quantity purchased is statistically different than zero, but the magnitude of the effect is quite low. From Proposition 1 we know that assuming continuous quantities, the model predicts a slope of minus one: conditional on purchasing the target is not a function of inventory and therefore every additional unit of inventory reduces the quantity purchased

[20]By regressing this measure on household demographics we can check that we get something reasonable. Indeed, our measure of consumption increases with family size, if there is a teenager in the family and if the female works more than 35 hours a week.

by one. There are several data and modeling reasons that could explain the difference between Proposition 1 and the estimated coefficient. First, we measure inventory in a very crude way, leading to measurement error. Classical measurement error biases the coefficient towards zero.[21] Second, as we pointed out in Section 3, the result of Proposition 1 does not hold once only discrete quantities are offered, as is the case in practice. This will make purchases less sensitive to inventories, in particular during non-sales, where no matter what the initial inventory is, consumers are predicted to buy only for consumption. Finally, as we saw in Table 7, a significant component of increased purchase during a sale is buying a larger size, and not more units. Normally not all sizes are offered on sale at the same time. Therefore, the result of Proposition 1 does not directly apply here. These problems are treated in the structural estimation by using the consumption decisions implied by the structural model and controlling for the actual choice set faced each period.

We performed the same analysis for the effect of inventory on the probability of purchase, conditional on being in a store (still implication I1). If the household's inventory was below its average it was almost twice as likely to buy. The overall probability of purchase is roughly 9 percent. If the inventory was above the average (for that household) it went down to 7 percent and if the inventory was below average it increased to over 13. The probability of purchasing laundry detergent decreases by about 0.65 percentage points for every additional 16 ounces of inventory.

The model does not incorporate other promotional activities than sale, but naturally they affect purchasing behavior. Columns (iv)- (ix) in Table 8 add the promotional variables to the regression. In columns (iv)-(vi) these variables are not interacted with price. The price coefficient is effected only slightly and for the most part the effects of the promotional variables are as expected.

---

[21]The model we presented in Section 3 predicts that consumption will respond to unobserved shocks. This implies that the assumption of constant consumption used to generate the inventory series will be right on average but will generate measurement error. The assumptions we made on the shocks will yield classical measurement error.

The two exceptions are the non-significant coefficient on feature, which is somewhat at odds with our finding from the aggregate data, and the negative effect of the interaction of sales and display. The first is driven by the high correlation between the feature variable and the interaction with sale. As we see at the bottom of the table, for this sample conditional on feature there is 0.89 probability of a sale. The latter becomes positive, as expected, once we interact the promotional variables with the price. Once we interact the promotional variables with price the effect of a sale is to shift out demand. This is consist with the theory presented in Section 3, which suggests that households buy more during a sale in order to store the product in inventory (Proposition1 and Corollary 1).

In columns (vii)- (ix) we allow the price sensitivity to vary with promotional activity. We find that sales tend to increase the price sensitivity, especially if they are combined with a feature or display promotion. Taken literally this implies that households tend to increase their purchase more if a price cut is during a sale, compared to a cut in the regular (non-sale) price. Once again this interpretation is consist with the model (implication I6).

5.2 *Structural Estimates*[22]

In order to further test our model and in order the derive implications we present in this section preliminary structural estimates. We estimates a restricted form of the model described in Section 3 using the algorithm described in Section 4. For now we did not allow for any heterogeneity across households and used a restrictive functional form for the storage costs. The results are based on based on a sample of 100 households. The parameters are of the expected sign. Since the results are preliminary and the parameters are of little direct interest we only present (some) implications.

---

[22]This section is incomplete and the results are preliminary.

Table 9 presents a comparison of the own-price elasticities computed from the dynamic model and those computed from various static models. All static models include all brand-size combinations, they differ in how price is measured. Model 1 used the price of each choice, Model 2 uses the price per unit, and Model 3 uses the price but also includes a size fixed effect. It is not surprising that the own-price elasticity increase (in absolute value) as we move from Model 1. Households buy larger sizes which are move expensive, without controlling for this the estimation yields less price sensitive demand. Model 2 controls for this in a "standard" way by using price per unit. Indeed the demand is more elastic. Model 3 controls for the differences across sizes by using a size fixed effect. If the EV term in the dynamic model was constant over time, this model would perfectly control for it. For this reason it is the most comparable with the dynamic model.

The mean own-price elasticity from the dynamic model is about 5% lower than Model 3 and a fair bit higher that the other 2 models. This might not seem like much, but when we look at the distribution of

the percent difference we see that the mean is somewhat misleading. The differences can be quite large, and the static model can either over or under estimate the true elasticity, because of the inventory. When inventory is high the static model will over estimate the elasticity, and visa versa when inventory is low.

Overall, we find this results encouraging. Even with a very restricted version of our model the results seem to suggest that the dynamic are economically important. We believe that as we estimate more complex versions of the model, which allow for more heterogeneity, the economic significance will only increase.

6. Preliminary Conclusions and Extensions

In this paper we propose a model of consumer inventory holding. We use the model to derive several implications, which we take to the data. Our data consists of an aggregate detailed scanner data and a household-level data set. Using these data sets we find several pieces of evidence consistent with our model. (1) Aggregate demand increases as a function of duration from previous sale, and this effect differs between sale and non-sale periods. (2) Fraction of purchases on sale are higher in one market (the market that on average has larger houses) and if there is a dog in the house. Both of these could potentially be correlated with lower storage costs. (3) When buying on sale households tend to buy more units, larger sizes and increase the duration to next purchase. (4) Sales seem to shift demand and change the price sensitivity. (5) Inventory constructed under the assumption of fixed consumption over time, is negatively correlated with quantity purchased and the decision to buy conditional on being in a store.

The main negative result is that the effect of inventory while statistically significant seems small. We discussed several reasons that could be driving this result including measurement error in the construction of the inventory variable and non-linear effects. Both of these will be handled, at least partly, in the structural model by relying on the model described in Section 3 to predict the non-linear effects and to construct an inventory variable assuming optimal behavior by the consumers. Furthermore, the structural model will allow us to better interpret the estimates, as well as perform some counterfactual experiments. The latter will allow us to return to some of the questions we used to motivate the analysis.

We are currently exploring extensions along several dimensions. First, we are extending our theoretical analysis to include the supply side. This, jointly with the structural estimates, will allow us to examine questions like what proportion of the variation in sales can be explained by our estimates, and given our estimates what are the optimal patterns of sales. Second, the analysis in this paper focuses on one product category, laundry detergents. We choose this category because we thought it justified some of the assumptions we had to make to focus the analysis on consumer inventory. However, our theory has predictions across categories, which we can test using the additional categories our data set contains.

Appendix A

The model presented in the Section 3.1assumes that consumers visit the store every period. In the data we observe variation in the time between store visits. This variation impacts the previous model in two ways. First, consumption should vary with the duration between visits. In principle this could be handled by allowing the distribution of the consumption shock, ν_t, to depend on the duration for previous visit. This approach does not account for the effect of duration to next visit on the expected distribution of prices in the next visit. Therefore, we propose the following model.

In periods when the consumer visits the store his behavior is described by the above model. In each period there is a probability, q, that he will visit the store next period. If he does not visit the store he only chooses consumption and does so as to maximize the current utility, minus inventory cost, plus future gains, subject to the same constraints as before. As before let the value function in periods of store visits be $V(I(t))$, and the value function during non-visit periods be $W(I(t))$.[23]

The optimal behavior can be characterized by the following Bellman equations

$$V(I(t)) = -C(i_t) + \max_{\{c_t, x_t, d_{jt}\}} u(c_t + \nu_t) + \alpha p_{jt}(x_t) + A_{jt}(x_t) + \delta E\left[qV(I(t+1)) + (1-q)W(I(t+1)) | I(t) \right]$$

$$W(I(t)) = -C(i_t) + \max_{\{c_t\}} u(c_t + \nu_t) + \alpha p_{jt}(x_t) + \delta E\left[qV(I(t+1)) + (1-q)W(I(t+1)) | I(t) \right].$$

It is easy to allow the probability of a visit in the next period, q, to depend on consumer characteristics and to let it vary between visit and non-visit periods.

---

[23] We abuse notation here since the information in each period is different. In store visit periods it includes the random shocks, ε_{jxt}, while in non-visit periods it does not. More importantly, during store visits the information set includes actual prices, while during non-store visits prices might not be observed by the consumer and therefore expected, or imputed, prices enter the information set.

Appendix B

Proof of Lemma 1: Suppose there exists *j* and *k* such that $c_j^* = c_j^*(x_t, \nu_t) \neq c_k^*(x_t, \nu_t) = c_k^*$. Then

$$\alpha p_{jt}(x_t) + A_{jxt} + u(c_j^* + \nu_t) + \delta EV(I_t; d_{jt} = 1, x_t, c_j^*) >$$

$$\alpha p_{jt}(x_t) + A_{jxt} + u(c_k^* + \nu_t) + \delta EV(I_t; d_{jt} = 1, x_t, c_k^*)$$

and therefore

$$u(c_j^* + \nu_t) - u(c_k^* + \nu_t) > \delta EV(I_t; d_{jt} = 1, x_t, c_k^*) - \delta EV(I_t; d_{jt} = 1, x_t, c_j^*).$$

Similarly, from the definition of $c_k^*(x_t, \nu_t)$

$$u(c_j^* + \nu_t) - u(c_k^* + \nu_t) < \delta EV(I_t; d_{jt} = 1, x_t, c_k^*) - \delta EV(I_t; d_{jt} = 1, x_t, c_j^*),$$

which is a contradiction.

References

Ackerberg, D. (2000), "Importance Sampling and the Simulated Method of Moments," Boston University, mimeo.

Bell D., J. Chiang and V. Padmanabhan (1999), "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18, 504-26.

Bertsekas, D. and J. Tsitsiklis (1996), *Neuro-Dynamic Programming,* Athena Scientific.

Blattberg R., G. Eppen and J. Lieberman (1981), "A Theoretical and Empirical Evaluation of Price Deals in Consumer Nondurables," *Journal of Marketing*, 45, 116-129.

Blattberg, R. and S. Neslin (1990), *Sales Promotions,* Prentice Hall.

Blattberg R., R. Briesch and E. Fox (1995), "How Promotions Work," *Marketing Science*, 14, G122-132.

Chevalier, J., A. Kashyap and P. Rossi (2000), "Why Don't Prices Rise During Peak Demand Periods? Evidence from Scanner Data," NBER Working Paper no. 7981.

Chiang J. (1991), "A Simultaneous Approach to the Whether, What and How Much to Buy Questions, *Marketing Science*, 10, 297-315.

Conlisk, J., Gerstner, E. and Sobel, J., "Cyclic Pricing by a Durable Goods Monopolist," *Quarterly Journal of Economics*; 99(3), August 1984, pages 489-505.

Erdem, T., M. Keane and S. Imai (2000), "Consumer Price and Promotion Expectations: Capturing Consumer Brand and Quantity Choice Dynamics under Price Uncertainty," University of California at Berkeley, mimeo.

Feenstra, R. and M. Shapiro (2001), "High Frequency Substitution and the Measurement of Price Indexes," NBER Working Paper no. 8176.

van Heerde H., Leeflang P. and D. Wittink (2000), "The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data," *Journal of Marketing Science*, 37, 383-395.

Hong, P., P. McAfee and A. Nayyar (2000), "Equilibrium Price Dispersion with Consumer Inventories," working paper.

Hosken, D., D. Matsa, and D. Reiffen (2000) "How do Retailers Adjust Prices: Evidence from Store-Level Data," working paper.

Grover R. and V. Srinivasan (1992), "Evaluating the Multiple Effects of Retail Promotions on Brand Loyal and Brand Switching Segments, *Journal of Marketing Research*, 29, 76-89.

Gupta S. (1988), "Impact of Sales Promotions on When, What, and How much to Buy," *Journal of Marketing Research*, 24, 342-55.

Keane, M. and K. Wolpin (1994), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *Review of Economics and Statistics*, 76(4), 648-72.

MacDonald, J. (2000), "Demand, Information, and Competition: Why Do Food Prices Fall At Seasonal Demand Peaks?," *Journal of Industrial Economics*, 48 (1), 27-45.

Narasimhan, C. (1988), "Competitive Promotional Strategies," *Journal of Business,* 61 (4), 427-49.

Neslin S., C. Henderson and J. Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science,* 4(2), 147-165.

Neslin and Schneider Stone (1996), "Consumer Inventory Sensitivity and the Postpromotion Dip," *Marketing Letters*, 7, 77-94.

Pesendorfer, M. (forthcoming), "Retail Sales. A Study of Pricing Behavior in Supermarkets," forthcoming *Journal of Business*.

Rao, R. (1991), "Pricing and Promotions in Asymmetric Duopolies," *Marketing Science*,10(2), 131-44.

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

Rust, J. (1996), "Numerical Dynamic Programming in Economics," in H. Amman, D. Kendrick, and J. Rust (eds.), *Handbook of Computational Economics*, Volume 1, 619-729.

Rust, J. (1997), "Using Randomization to Break the Curse of Dimensionality," *Econometrica*, 65, 487-516.

Salop, S. and Stiglitz,-J., "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical," *American Economic Review*; 72(5), December 1982, pages 1121-30

Sobel, J. (1984), "The Timing of Sales," *Review of Economic Studies,* 51, 353-368.

Sobel, J. (1991), "Durable Goods Monopoly with Entry of New Consumers," *Econometrica,* 59(5), 1455-85.

Shoemaker R. (1979), "An Analysis of Consumer Reactions to Product Promotions," in *Educators' Conference Proceedings,* Chicago: American Marketing Association, 244-248.

Ward, R. and J. Davis (1978), "A Pooled Cross-Section Time Series Model of Coupon Promotions," *American Journal of Agricultural Economics*, 60, 393-401.

Wilson D., L. Newman and M. Hastak (1979), "On the Validity of Research Methods in Consumer Dealing Activity: An Analysis of Timing Issues," in *Educators' Conference Proceedings*, Chicago: American Marketing Association, 41-46.

Warner, E. and R. Barsky (1995), "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays,"*Quarterly Journal of Economics*;110(2), 321-52.

Table 1
Brand Volume Segment Shares and Fraction Sold on Sale

Brand	Firm	7/91-12/91		1/92-6/92		7/92-12/92		1/93-6/93	
		Share in Seg	% on Sale	Share in Seg	% on Sale	Share in Seg	% on Sale	Share in Seg	% on Sale
Liquid:			34.4		37.8		40.0		56.9
Tide	P & G*	20.7	19.6	22.2	40.6	22.8	34.8	19.5	47.1
All	Unilever	11.3	45.7	11.2	39.8	17.8	52.4	18.5	65.0
Purex	Dial	5.9	84.2	10.0	73.6	8.2	62.9	11.4	73.3
Wisk	Unilever	12.3	47.2	12.5	53.4	11.2	63.1	9.8	69.1
Solo	P & G*	12.8	18.2	10.9	10.5	11.4	11.7	5.6	2.1
Cheer	P & G*	5.1	14.3	4.8	45.2	4.1	36.2	4.3	42.6
A & H*	C & D*	5.8	36.1	4.5	20.9	4.2	30.2	3.6	52.3
Surf	Unilever	5.4	56.7	4.1	36.2	3.8	60.5	2.8	73.9
Other	–	20.6	29.7	19.9	23.4	16.5	25.2	24.5	59.8
Powder			31.2		33.7		36.1		43.7
Tide	P & G*	37.5	26.3	42.0	35.3	40.1	37.5	39.2	39.8
Cheer	P & G*	11.0	39.1	8.6	39.0	9.5	37.1	13.2	59.9
A & H*	C & D*	18.9	29.9	13.7	17.2	13.7	10.6	12.0	13.7
All	Unilever	3.6	24.8	5.4	24.8	5.4	69.5	6.0	89.6
Surf	Unilever	3.2	39.8	4.2	30.3	4.2	53.5	4.6	71.1
Purex	Dial	1.2	37.4	0.7	40.9	0.7	17.0	0.4	34.4
Other	–	24.7	35.5	26.3	40.2	26.4	37.8	24.6	39.5

* A & H = Arm & Hammer; P & G = Procter and Gamble; C & D = Church and Dwight.
Columns labeled *Share in Seg* are segment market share of volume sold in the nine store in our sample and columns labeled *% on Sale* are the percent of the volume, for that brand in that quarter, sold on sale. The category *Other* includes all other brands, including those produced by some of the manufacturers listed.

Table 2
Non-linear Pricing by Store

Size **Store**	32 oz.			64 oz.			96 oz.			128 oz.			256 oz.		
Mrkt I	$/16 oz.		share (%)	disc (%)		share (%)	disc (%)		share (%)	disc (%)		share (%)	disc (%)		share (%)
	uw	qw		uw	qw		uw	qw		uw	qw		uw	qw	
1	1.21	1.20	1.8	29.8	36.1	21.1	28.0	29.6	8.7	33.7	41.1	59.5	27.1	33.0	2.3
2	1.46	1.51	1.1	43.6	46.3	23.0	42.0	45.9	7.3	44.9	57.9	54.1	44.1	47.1	2.8
3	1.82	1.63	2.3	49.1	49.6	44.5	43.8	41.8	6.5	52.8	51.2	35.8	–	–	–
4	1.57	1.62	3.2	38.0	41.7	41.6	35.9	36.9	6.2	39.6	49.6	39.9	–	–	–
5	1.62	1.62	2.8	40.0	42.1	43.2	39.0	38.7	7.9	43.2	49.1	36.5	–	–	–
Mrkt II															
1	1.86	1.55	1.4	48.3	48.6	26.7	49.2	57.5	10.1	53.3	66.1	58.8	–	–	–
2	1.51	1.38	2.6	44.2	42.8	50.2	42.2	38.0	15.6	43.0	40.1	29.6	36.8	30.9	1.2
3	1.63	1.57	1.2	48.8	50.4	38.5	44.2	45.0	7.9	52.7	53.6	41.8	–	–	–
4	1.60	1.64	1.0	46.0	49.0	29.7	44.0	47.2	8.2	47.9	54.3	41.5	39.2	40.6	1.6

Data from all brands of liquid detergent. The column labeled *$/16 oz.* presents the average per unit, un-weighted(*uw*) and quantity-weighted(*qw*), price of a container size in a store. The average is taken over weeks and across different brands. The column labeled *disc* presents the percentage discount in, un-weighted(*uw*) and quantity-weighted(*qw*), price per 16 oz. unit, relative to the, un-weighted(*uw*) and quantity-weighted(*qw*), price of a 32 oz. packet, respectively. The column labeled *share* presents the share of quantity sold in each store as a total of total quantity of liquid detergent sold in that store.

Table 3
Non-linear Pricing for TIDE by Store

Store / **Size**	32 oz.			64 oz.			96 oz.			128 oz.			256 oz.	
	$/16 oz.		share (%)	disc (%)		share (%)	disc (%)		share (%)	disc (%)		share (%)	disc (%)	
Market I	uw	w		uw	w		uw	w		uw	w		uw	w
1	1.20	1.20	0.6	12.6	14.0	7.1	11.7	11.7	4.9	24.7	35.0	72.6	17.4	17.4
2	1.20	1.20	1.0	13.2	14.4	13.7	17.5	17.5	6.3	19.0	27.6	40.5	24.8	26.2
3	1.33	1.33	4.2	12.0	16.2	26.8	12.3	19.7	15.2	14.3	23.2	25.9	–	–
4	1.35	1.35	2.6	14.1	15.4	26.6	13.8	14.0	13.6	14.8	17.7	33.9	–	–
Market II														
1	1.35	1.35	3.2	15.8	21.5	32.4	13.1	12.9	19.6	17.1	27.7	25.8	–	–
2	1.25	1.25	6.2	17.2	17.5	34.5	16.6	16.4	22.8	16.9	17.2	27.9	23.2	23.3
3	1.25	1.25	1.7	16.3	18.9	24.0	17.1	21.1	18.3	19.7	22.8	27.2	–	–
4	1.25	1.25	0.6	14.1	23.4	19.6	16.8	18.5	6.5	17.0	24.9	35.3	21.7	21.6

The column labeled *$/16 oz.* presents the average per unit, un-weighted(*uw*) and quantity-weighted(*qw*), price of a container size in a store. The average is taken over weeks and across different brands. The column labeled *disc* presents the percentage discount in, un-weighted(*uw*) and quantity-weighted(*qw*), price per 16 oz. unit, relative to the, un-weighted(*uw*) and quantity-weighted(*qw*), price of a 32 oz. packet, respectively. The column labeled *share* presents the share of quantity sold in each store as a total of total quantity of liquid detergent sold in that store.

Table 4
Non-Sale Prices, Frequency of Sale and Quantity Sold, by Store and Size

Size / Store		32 oz.				64 oz.					96 oz.					128 oz.				
	price	sale (%)		big sale (%)		disc (%)	sale (%)		big sale (%)		disc (%)	sale (%)		big sale (%)		disc (%)	sale (%)		big sale (%)	
Market I		freq	q	freq	q		freq	q	freq	q		freq	q	freq	q		freq	q	freq	q
1	0.95	0.0	0.0	0.0	0.0	18.0	16.1	38.9	9.8	27.4	13.1	8.6	10.8	0.0	0.0	25.8	15.6	14.3	7.4	11.2
2	1.35	15.4	6.7	15.4	6.7	40.4	29.9	36.1	15.6	16.5	39.6	24.3	29.3	1.1	5.6	44.2	27.8	62.3	9.5	49.6
3	1.28	3.0	2.7	3.0	2.7	32.5	19.6	42.3	13.9	37.4	17.7	21.4	52.0	12.6	43.8	25.4	39.9	71.1	25.9	62.2
4	1.69	8.9	8.3	8.9	8.3	43.8	9.6	19.7	4.3	13.6	40.2	4.0	6.8	1.7	6.0	47.5	17.5	36.2	11.2	31.1
5	1.68	10.8	8.7	10.8	8.7	43.7	10.1	19.9	3.8	13.2	41.9	3.5	8.3	0.8	4.4	46.1	22.8	38.6	16.4	32.8
Market II																				
1	1.54	10.5	4.4	0.8	0.1	43.0	15.0	44.4	7.2	39.7	60.5	14.1	11.7	0.7	2.0	52.5	23.7	81.6	12.0	77.4
2	1.28	0.0	0.0	0.0	0.0	34.0	37.2	52.9	19.7	34.2	34.4	24.6	28.1	0.0	0.0	34.2	22.4	32.5	10.9	18.9
3	1.56	9.5	10.4	2.0	2.8	51.1	20.5	35.9	7.6	22.0	42.5	18.6	39.6	6.8	23.7	48.1	36.6	64.5	15.9	42.7
4	0.99	0.4	0.5	0.4	0.5	13.2	25.6	44.3	12.4	29.9	7.8	26.9	42.4	13.8	24.5	13.7	31.5	63.6	11.8	46.8

The column labeled *price* presents the modal price per 16 oz. for a 32 oz. container in each store. Columns labeled *disc.* present the discount in the per unit modal price for each size. Columns labeled *sale* and *big sale* present the frequency (*freq*) of the price being below its modal value (by size and store) and the frequency of it being at less than 90 percent of the modal price, respectively, and quantity sold (*q*) at those instances.

Table 5
Demand as a Function of Duration from Previous Promotional Activity

Variable	full sample		sale=1 sample		sale=0 sample	
log(price)	-2.79	-2.81	-2.76	-2.73	-2.44	-2.35
	(0.07)	(0.07)	(0.12)	(0.12)	(0.11)	(0.16)
duration from previous sale	-0.48	1.00	1.10	2.70	-0.83	0.75
	(0.19)	(0.26)	(0.41)	(0.50)	(0.21)	(0.31)
$(\text{duration from previous sale})^2$	0.32	-1.82	-2.92	-5.08	0.86	-1.43
	(0.44)	(0.55)	(0.96)	(1.13)	(0.49)	(0.64)
feature	0.49	0.49	0.50	0.52	0.77	0.66
	(0.03)	(0.03)	(0.04)	(0.04)	(0.15)	(0.16)
display	0.99	0.97	0.92	0.90	1.04	1.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
duration from previous feature	–	-2.06	–	-2.55	–	-1.95
		(0.24)		(0.43)		(0.29)
$(\text{duration from previous feature})^2$	–	2.78	–	3.05	–	2.66
		(0.44)		(1.05)		(0.52)
N =	10,684	10,178	3,225	3,047	7,459	7,131

The dependent variable in all regressions is the natural logarithm of quantity purchased (measured in ounces). Each observation is a brand-size combination in a particular store. Duration from previous sale (feature) is measured as number of weeks, divided by 100, from previous sale (feature) for that brand in that store for any size. All regressions include brand-size and store dummy variables.

Table 6
Correlation Between Households Fraction of Purchases on Sale and Household Characteristics

Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
constant	0.46 (0.03)	0.53 (0.03)	0.44 (0.03)	0.52 (0.03)	0.36 (0.03)	0.56 (0.03)	0.41 (0.03)	0.43 (0.05)
male head of household	0.07 (0.02)	0.04 (0.02)	0.06 (0.02)	0.04 (0.02)				0.03 (0.02)
female works <35 hrs/week	0.06 (0.03)	0.05 (0.03)	0.05 (0.03)	0.04 (0.03)				0.04 (0.03)
female works >35 hrs/week	-0.01 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.02)				-0.01 (0.02)
income per person	-0.009 (.009)	0.002 (.009)	-0.005 (.009)	0.005 (.009)				0.006 (0.009)
female post HS education	-0.03 (0.02)	-0.02 (0.02)	-0.03 (0.02)	-0.02 (0.02)				-0.02 (0.02)
Latino	-0.12 (0.05)	-0.05 (0.05)	-0.12 (0.05)	-0.05 (0.05)				-0.04 (0.05)
market I		-0.14 (0.02)		-0.14 (0.02)				-0.13 (0.02)
dog dummy variable			0.08 (0.02)	0.06 (0.02)				0.07 (0.02)
cat dummy variable			-0.02 (0.03)	-0.01 (0.02)				-0.003 (0.027)
# of stores					0.05 (0.01)			0.03 (0.01)
avg days b/ shopping						-0.014 (0.004)		-0.009 (0.004)
# of brands							0.021 (0.006)	0.021 (0.006)
R-squared	0.026	0.067	0.037	0.075	0.023	0.015	0.012	0.103

The dependent variable is the fraction of purchases made during a sale. Each household is an observation.

Table 7
Differences in Purchasing Patterns Between Sale and Non-Sale Purchases

	Average during non-sale purchases	Difference during sale purchases			Average during non-big-sale purchases	Difference during big-sale purchases		
		Total	Within households	Between households		Total	Within households	Between households
Units purchased	1.04 (0.01)	0.07 (0.01)	0.05 (0.01)	0.10 (0.02)	1.04 (0.01)	0.08 (0.01)	0.07 (0.01)	0.09 (0.02)
Size (16 oz.)	4.54 (0.03)	0.77 (005)	0.50 (0.04)	1.20 (0.20)	4.61 (0.03)	0.88 (0.05)	0.61 (0.05)	1.10 (0.20)
Quantity (16 oz.)	4.73 (0.04)	1.21 (0.06)	0.81 (0.60)	1.97 (0.26)	4.82 (0.04)	1.43 (0.07)	1.01 (0.07)	1.77 (0.27)
Duration from previous purchase (days)	44.26 (0.70)	5.97 (1.07)	-1.62 (0.98)	25.91 (8.32)	44.68 (0.64)	7.12 (1.17)	-2.56 (1.08)	29.61 (8.30)
Duration to next purchase (days)	43.94 (0.71)	7.50 (1.10)	1.19 (0.99)	30.46 (8.64)	43.97 (0.64)	10.66 (1.20)	3.04 (1.10)	33.15 (8.70)
Duration to next purchase, conditional on it being non-sale (days)	41.94 (0.80)	10.99 (1.50)	3.11 (1.23)	28.00 (7.96)	42.20 (0.75)	14.86 (1.70)	5.11 (1.43)	25.72 (8.00)
Previous purchase not on sale	0.69 (0.01)	-0.28 (0.01)	-0.06 (0.01)	-0.74 (0.02)	0.65 (0.01)	-0.27 (0.01)	-0.03 (0.01)	-0.66 (0.02)

Based on all purchases of liquid and powder detergents by households observed in our sample. A sale is defined as a price below the modal price, of a UPC in a store over the observed period. A big sale is defined as a price 10 percent below the modal price. The column labeled *Within households* controls for an household fixed effect, while the column labeled *Between households* is the regression of household means. Standard errors are provided in parentheses.

Table 8
Quantity Purchased by Household as a Function of Price and Promotional Activities

Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
price	-2.64	-2.57	-2.58	-2.32	-2.26	-2.26	-1.85	-1.80	-1.79
	(0.10)	(0.10)	(0.10)	(0.11)	(0.11)	(0.11)	(0.12)	(0.12)	(0.12)
price*sale							-0.91	-0.89	-0.91
							(0.23)	(0.23)	(0.23)
price* feature							0.14	0.22	0.21
							(0.73)	(0.72)	(0.72)
price* display							0.19	0.17	0.16
							(0.33)	(0.32)	(0.32)
price*sale *feature							-2.06	-2.19	-2.16
							(0.85)	(0.84)	(0.84)
price*sale *display							-1.34	-1.31	-1.30
							(0.55)	(0.54)	(0.54)
sale				0.28	0.31	0.40	1.24	1.25	1.35
				(0.12)	(0.12)	(0.12)	(0.25)	(0.25)	(0.25)
feature				-0.04	-0.07	-0.06	0.06	-0.02	-0.007
				(0.23)	(0.23)	(0.23)	(0.53)	(0.52)	(0.52)
display				0.49	0.51	0.52	0.37	0.40	0.41
				(0.13)	(0.13)	(0.13)	(0.29)	(0.29)	(0.29)
sale * feature				0.89	0.90	0.88	1.94	2.04	2.00
				(0.27)	(0.27)	(0.27)	(0.63)	(0.63)	(0.63)
sale * display				-0.22	-0.25	-0.27	0.64	0.59	0.56
				(0.19)	(0.19)	(0.19)	(0.47)	(0.47)	(0.46)
inventory		-0.043	-0.037		-0.043	-0.034		-0.043	-0.034
		(0.003)	(0.004)		(0.003)	(0.004)		(0.003)	(0.004)
inventory *sale			-0.015			-0.021			-0.021
			(0.005)			(0.005)			(0.005)

Pr(sale | feature) = 0.89; Pr(feature | sale) = 0.63;
Pr(sale | display) = 0.68; Pr(display | sale) = 0.54

The dependent variable in all regressions is the quantity purchased (measured in 16 oz units.) The regressions have 8012 observations, where an observation is a purchase of a strictly positive quantity of detergent by a household. All regressions also include household-specific dummy variables, 8 (broadly defined) product-specific dummy variables and store dummy variables. Prices ($/16 oz) and promotional variables are for the product purchased.

Table 9
A Comparison of Elasticities Computed from the Structural Model and from Static Models

	Model 1	Model 2	Model 3
Average Own-Price Elasticity (dynamic model == -2.15)	-0.46	-1.84	-2.25
Percent difference (percent):[a]			
Average	-78.8	-7.3	3.7
Median	-77.8	3.8	9.1
5 percentile	-84.6	-55.4	-26.3
95 percentile	-76.7	20.4	11.8

All static models are conditional logit models, model 1includes the price of a brand, model 2 includes price per ounce, model 3 includes price and a size dummy variable. The elasticities are evaluated for purchases at the purchase price.

[a]Computed as (static model elasticity - dynamic model elasticity)/dynamic model elasticity.