PRELIMINARY AND INCOMPLETE

Sales and Consumer Inventory[1]

Igal Hendel
University of Wisconsin, Madison and NBER

and

Aviv Nevo
University of California, Berkeley and NBER

March, 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. The sales vary in the discounted amount, the time from previous sale, and in their duration. Quantities sold tend to vary around a modal quantity when the price is at the modal, or regular, level. Somewhat unsurprising is the fact that during a sale the quantity sold increases significantly. 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 discounted amount as well as other promotional activities fixed.

---

[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 and University of Wisconsin 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.

There are several reasons, beyond pure academic curiosity, why we should care about why sales exist and what are their effects. As the example in the previous paragraph suggests, 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. Therefore, this 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).

Sales are often present in markets for fashion goods or markets where demand is seasonal, or follows some predictable pattern. We instead focus on retail goods, groceries, with highly frequent sales, 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 we described in the first paragraph, 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.

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. The consumer has two decisions to make in each period: how much to buy and how much to consume. Both these decisions are made to maximize the present expected value of future utility flows. Optimal behavior

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

is a function of the current price, the current inventory and a stochastic shock. In any period quantity purchased which is not consumed is stored as inventory. The quantity consumed can exceed the quantity purchased that period, but cannot exceed purchases plus current inventories. The stochastic shock, which is consumer-specific, allows for the possibility that the consumer might want to consume more in some periods due to reasons unobserved to the researcher and potentially unforeseeable to the consumer.

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. The consumer-specific costs of holding inventory include depreciation of the product (which vary by product), physical storage limitations and the costs of carrying the products home.

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 and how much was paid as well as whether on not a coupon was used. In addition we know when the households visited a supermarket but decided not to purchase a 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 for any given household. Finally, 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 the effect we want we have to properly control for the effect of other promotional activity. Furthermore, as predicted, this effect is larger during a sale period 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 either duration from previous purchase or duration to next purchase. The difference is smaller within a household compared to across 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.

Finally, we find evidence that the demand during a sale shifts out beyond what is explained by the change in price. Furthermore, sales change the slope of the demand function. Both of these are consistent with our predictions.

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 have offered 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. He shows that the optimal equilibrium strategy for the firms is to randomize between two prices: a high price which is the reservation price of the uninformed and a low price which will attract also the informed consumers. 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, 1991) develops a model in which suppliers periodically discount a durable good to clear the market of low valuation consumers. The model he develops assumes two types of consumers – high valuation and low valuation. Sobel also assumes that consumers arrive in the market over time and that low valuation consumers are willing to postpone their purchases. He demonstrates that sellers will periodically find it optimal to lower prices to clear out low valuation consumers.

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 and therefore model also the brand choice decision. Our ultimate interest lies in the supply side and therefore we focus on an industry in which we believe the brand choice is of secondary importance.

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 closer to their previous purchase. We also examine some of the same quantities, however, we control for 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

[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).

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. 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 exactly which product each household bought, where it was bought and how much was paid, as well as whether or not a coupon was used. In addition, we know when a household visited a supermarket and how much they spent overall.

2.2 *The Industry*

The data includes purchases of 24 different product categories. For now we 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 is 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 percent 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%), 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

[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.

of the non-linearity in prices is driven by the high prices of the smallest container size (32 oz.). The table also suggests that there are differences across stores in both the average price and the quantity discounts.

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 a change in the mix of brands. For this purpose Table 3 presents the same figures but focuses on one brand, Tide, the best selling product. 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 changes in the non-sale price and changes due to sales. 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 was indeed a mixture of non-linear non-sale prices and more sales for the larger brands. 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 implications 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 test indirectly.

3.1 The Basic Setup

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

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

where c_{it} is the quantity consumed of the good in question, ν_{it} is a shock to utility that changes the current marginal utility from consumption, θ_i is a vector of consumer-specific taste parameters and m_{it}

[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.

is utility from the outside good. The stochastic shock, ν_{it}, captures the possibility that the consumer might want to consume more in some periods due to reasons unobserved to the researcher. For simplicity we assume the shock to utility is additive in consumption, $u(c_{it}, \nu_{it}; \theta_i) = u(c_{it} + \nu_{it}; \theta_i)$. Low realizations of ν_{it} increase the household's need, increasing demand and making it more inelastic. In the application we make ν_{it} a parametric function of the household characteristics, to capture heterogeneity across households.

The product is offered in J different varieties, or brands. The consumer faces random and potentially non-linear prices. Let $p_{jt}(x)$ be the total price associated with purchasing x units of brand j.

The good is storable. Therefore, 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. Quantity not consumed is stored as inventory that can be consumed in the future. Consumption and purchase (and therefore inventory accumulation) decisions are made to maximize the (discounted) expected stream of future benefits.

For now we assume the consumer visits one store each week, i.e., we do not model the consumer's decision of where shop. Below 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

$$V(I(0)) = \max_{\{c_t, x_t, d_{jt}\}} \sum_{t=0}^{\infty} \delta^t E\left[u(c_t, \nu_t; \theta) - C(i_t) + \sum_j d_{jt}\left(\alpha p_{jt}(x_t) + A_{jxt}\right) | I(0)\right]$$

$$s.t. \quad 0 \leq i_t, \quad 0 \leq c_t, \quad 0 \leq x_t \qquad \textbf{(1)}$$

$$i_t = i_{t-1} + x_t - c_t$$

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. 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. Notice, 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, current prices, the current shock to utility from consumption, ν_t, and the vector of random shocks.[7]

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,[8] 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 next section.[9] Finally, the random shocks, ε_{jxt}, are assumed to be independently and identically distributed according to a type I extreme value distribution.

Define the value function as the discounted lifetime expected utility. This value function depends on all the relevant information known at time t. The state variables include the current

---

[7] In principle, the information set also includes the advertising variables, a_{jt}, (as well as 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.

[8] 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.

[9] 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.

inventory, current prices and current utility shocks.

3.2 Model with inter-visit consumption

The model presented in the previous section assumes 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 $W(I(t))$.[10] The optimal behavior can be characterized by the following Bellman equations

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

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

It is easy to allow the probability of a visit in the next period, q, to depend on consumer

[10] 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.

characteristics and to let it vary between visit and non-visit periods.

3.3 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 return to the basic model presented in Section 3.1 and 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 $p_t > p'_t$.

The solution of the consumer's problem is characterized by the following first order conditions

$$u'(c_t) = \eta_t - \lambda_t \qquad \textbf{(2)}$$

$$p_t = -\lambda_t + \psi_t \qquad \textbf{(3)}$$

$$-C'(i_t) + \lambda_t + \mu_t = \delta E(\lambda_{t+1}|I(t)) \qquad \textbf{(4)}$$

where μ_t, η_t, ψ_t, and λ_t are the Lagrange multipliers of the constraints defined in equation (1).

We first derive predictions about a given household behavior. (All proofs will be added in the Appendix).

Proposition 1: At any *t* with $x_t > 0$ the end of period inventory i_{t+1} equals $S(p)$ or 0, where $S(p)$ is a decreasing function *of* p_t, independent of i_t and ν_tt.

Proposition 2: $x(i_t, p_t, \nu_t)$ declines in the three arguments.

Corollary 1: There is a price $p^r < p^m$, where p^m is the modal, non-sale, price, such that at any price $p \geq p^r$ consumers only buy for current consumption.

Proposition 3: If only discrete quantities can be purchased or prices are non-linear in quantities then the target inventory S becomes a function of i_t and ν_t.

We next derive some cross household predictions.

Assumption 2: Storage costs are such that: $C_h(0) = 0,\ \forall h,$ and $C_h(i)$ increases in h for all i.

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

Proposition 5: $S(p)$ declines in the frequency the household visits the store.

The above claims have implications for patterns we should observe in the household and aggregate data. The implications of the model for households are:

1. Sales cause consumers to be more price sensitive. Since at the modal, non-sale, price a consumer only purchases for current consumption, a small price decrease causes the consumer to react by consuming more. During a sale, however, a consumer reacts to price changes not only by consuming more but also by accumulating higher inventories.
2. Sales affect the timing of purchases.
3. Non-sale purchases should have a higher probability that the previous purchase was not on a sale.

The implication for aggregate demand are:

4. Aggregate demand during both sale and non-sale periods will be higher the longer the duration from the previous sale. The marginal effect of duration on sales purchases is likely to decline in duration.
5. The effect of duration should be higher for demand during sales than demand during non-sale. In the latter, consumers only demand to cover consumption while on sale they hoard inventories.

4. Econometrics

Using the data described in Section 2.1 we test the model in two different 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[11] in the following way. 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 value for the initial inventory. We experiment with drawing this value from 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 conditional 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

[11] 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 random variable (Erdem et al, 2000).

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 to estimate the rest of the parameters. Rather than having the 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

$$Pr(d_{jt}=1,x_t|i_t,p_t)=\int \frac{\exp\left\{\alpha p_{jt}(x_t)+A^1_{jxt}+\max_{c_t}\left\{u(c_t+\nu_t)+\delta EV(I_t;d_{jt}=1,x_t,c_t)\right\}\right\}}{\sum_{j,x}\exp\left\{\alpha p_{jt}(x_t)+A^1_{jxt}+\max_{c_t}\left\{u(c_t+\nu_t)+\delta EV(I_t;d_{jt}=1,x_t,c_t)\right\}\right\}}\,dF(\nu_t) \quad \textbf{(5)}$$

where $A^1_{jxt}=A_{jxt}-\varepsilon_{jxt}$ and EV($\cdot$) is the expected future value given today's state variables and today's decisions. Note that the summation in the denominator of equation (5) 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 preferences 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 <u>conditional</u> on the size (quantity) purchased. In other words, we estimate a logit or random coefficients logit model, 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 (5)

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

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^1_{jxt} can be estimated by maximizing this probability. This amounts to estimating a (random coefficients) 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

$$\omega_{xt} = \log\left\{\sum\nolimits_k \exp\left(\alpha p_{kt}(x_t) + A^1_{kt}\right)\right\}.$$

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 category, 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.[12] 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

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

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

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 separate the probabilities of brands choices as oppose to size choices. Second, it helps reduce the computational burden of the dynamic problem. 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

---

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

probabilities, into the nested algorithm discussed above to compute the likelihood of purchasing a size (quantity). More precisely, using the definition of the inclusive values, and equation (5), 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 by indirect evidence constructed from both the aggregate and household-level data. Using the aggregate data we show that, as predicted, the duration from previous sale increases the quantity purchased, once we control for promotional activity. Next we turn our attention to the household data and use it to (1) characterize the differences in characteristics between households that buy more often on sale and those that do not; (2) characterize the difference in purchases between sale and non-sale purchases, both across households and for a given household; and (3) examine the decision of how much to buy conditional on a purchase.

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 some counterfactual experiments.

5.1 *Indirect Tests of the Theory*

One of the predictions of the model presented in Section 3 is that the demand during a sale should be higher as the duration from the previous sale increases. 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 periods of sale and not, the effect of duration from previous sale is negative.[13] This is driven by the following. The effect of sales we model in this paper implies that the quantity sold should increase with duration (i.e., as consumers run out of the quantity they piled up during the sale). However, since sales are correlated with other promotional activities, if we do not control for these activities their effect might confound the measurement of the effect of sales. Therefore, we might see demand increase even after a sale is over, if the sale was accompanied by some promotional activity. For example, before we include duration from previous feature, which is one of the promotional activities, the coefficient on duration from sale captures both

[13] Recall, that 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, when discussing the marginal effect of duration we limit the discussion to the linear coefficient.

effects. Indeed, once we include the duration from previous feature, in the regressions presented in Table 5, 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 highly insignificant.

Restricting the focus to only the sample of sales, the effect of duration from previous sale is positive even before we control for duration from previous feature. However, 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 is negative and becomes positive only once we control for duration from previous feature. The relative magnitude of the coefficients in the different samples is consistent with the theoretical predictions.

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 use these data in several ways. First, we try to distinguish between those households that tend to buy on sale and those that do not. We do this by regressing the fraction of times the household bought on sale during the observed period on various household characteristics.[14] The results are presented in Table 6. In column (i) we explore the explanatory power of demographic variables. The results suggest that households without a male will tend to buy more on sale, as will 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 statically significant, or at times 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 finding (Blattetberg and Neslin, 1990).

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

While the frequency a household buys on sale does not seem to be strongly correlated with standard household demographics it is correlated with two other household characteristics. Households that live in market I 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). This market has smaller homes with less rooms and bedrooms, relative to the other market. This is consist with our model that predicts that lower storage costs are correlated with purchasing more frequently on sale (and assuming a large home implies lower storage costs). We do not know anything about the house in which the households in the sample live so we cannot verify this effect directly using our household sample. However, we know the number of dogs owned. The results in columns (iii) suggest that owning a dog is positively, and significantly, correlated with purchasing on sale, even once we control for other household characteristics. At the same 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).

In columns (v) through (viii) we explore the correlation between frequency of purchasing on sale and other shopping decisions. 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.[15] 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,

[15]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.

to the median, with a concentration of 0.82, will decrease frequency of buying on sale by about 8 percentage points.

In column (vi) we show 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 increases by roughly 1.5 percentage points. The mean duration between purchases 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 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).

Next, we examine characteristics of purchase and compare them between sale and non-sale purchases. The results presented in Table 7 suggest that when purchasing during a sale households will buy more units and a larger size. 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). The last two rows of Table 7 suggest that the duration from previous purchase and to next purchase are higher if the current purchase was during a sale. The effects within a household are much smaller than the effects across, or between, households. Finally, we find that the probability your previous purchase was not on sale, given that your current purchase was not on sale is higher, as implied by the theory.

Table 8 presents the results of regressing the quantity purchased by a household on the price paid and various promotional activities. The results are for the sample with strictly positive purchases. The regression includes household-specific dummy variables (as well as dummy variables for each store and for each, broadly-defined, product). We also examined random effects models. The results were essentially identical, and therefore not reported. The effect of price is negative: the lower the price the more consumers will buy. The average elasticity implied by the results in the column (i) is roughly -0.8 (with a median of roughly -0.3).[16]

One of the predictions of our theory is that the inventory a household holds should impact the purchasing decision. We do not observe inventory therefore we constructed a variable that proxies for it in the following manner. 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.[17] 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. In reality since we include a household-specific dummy variable these corrections do not matter (as long as the inventory variable

[16]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.

[17]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.

enters the regression linearly).

The results, presented in column (ii) are consistent with our theory the more inventory you hold the less you 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. Indeed the effect of inventory during a sale is higher, as predicted. Both these effects are statistically significant, yet their magnitude seems low. This can be driven by several factors. First, we measure inventory in a very crude way, which is likely to suffer from measurement error. Assuming classical measurement error than the coefficient is biased towards zero.[18] This will be treated below, at least partly, by using the consumption decisions implied by the structural model to construct a slightly less arbitrary measure of inventory. Second, we are underestimating the effect of inventories because we are focusing on the decision of how much to buy conditional on a purchase. Indeed, conditional on being in store the probability of purchasing laundry detergent decreases by about 0.65 percentage points for every additional 16 ounces of inventory. The probability of purchase conditional on being in a store is roughly 9 percent.

The third reason why the effect of inventory on quantity purchased and probability of purchase might seem low is because the likely effect is highly non-linear. Our model predicts that the effect of inventory should be zero unless it falls below a ceratin threshold, which is a function of the other state variables. There are two difficulties in translating this into a linear effect. First, even if we observe, and could control, for the other state variables, the marginal effect of inventories should be zero unless the inventories are below the threshold. This is a non-linear effect that does not easily translate to the regression we propose. Second, since we do not observe one of the state

[18] 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 she shocks will yield classical measurement error.

variables we cannot control for it and we are therefore effecting our ability to measure the impact of inventories.

To properly control for this problem we estimate the structural model below. However, we also tried a few more arbitrary (but much simpler) measures. We tried to proxy for the threshold by computing the mean inventory for each household during periods of purchase. We also looked at the median and various other percentiles of the distribution of inventories. The results suggest that if a household's inventories are below its average it is likely to significantly increase its purchases. We performed the same analysis for the effect of inventory on probability of purchase, conditional on being in a store. 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 percent.

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. The two exceptions are the highly 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 will buy more during a sale in order to store the product in inventory.

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 theoretical prediction. We note, however, that there is at least one alternative interpretation. The price in the regression is the price paid for the product purchased. We saw, in Table 7, that on average purchases made during sales tend to be of products with higher non-sale price. Therefore, it is possible that we are picking up the fact that the price changes are not measured correctly. In other words, the actual price change for a product on sale is larger than we are measuring, because it is the difference between the sale price and non-sale price for the product bought, and therefore the reaction seems to be larger. We partly control for this effect by including product and store fixed effects in the regression.

5.2 *Structural Estimates*

TBA

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

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.

Additional proofs TBA.

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.

Erdem, T. and M. Keane (1996), "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science,* 15(1),1-20.

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.

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.

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

Size / **Store**	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

Store \ Size		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)
(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)
(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.