Sales promotions in supermarkets: Estimating their effects on profits and consumer welfare

Victor Aguirregabiria
Boston University

November 2002
Preliminary and incomplete version

Abstract

Some countries have regulations that impose restrictions on the use of sales promotions in retail markets. The main motivation of these policies has been the protection of small retailers. This paper studies empirically the welfare implications of this type of policies. We present a model of dynamic price competition among retailers who sell several varieties of a differentiated storable good. In this model, firms use sales promotions as a mechanism to discriminate intertemporally among heterogeneous consumers. The model is estimated using scanner data from the food retailers of a US town. The estimated model is used to compute counterfactual equilibria under different restrictions on the use of sales promotions. We compare consumer surplus and the profitability of small and large retailers under the factual and the counterfactual equilibria.

KEYWORDS: Sales promotions; Intertemporal price discrimination; Market power; Estimation of dynamic demand models; Estimation of game theoretic models.

1 Introduction

While retailers in Britain and US enjoyed of great flexibility in the use of sales promotions, most countries in continental Europe have restrictions on the form, the frequency and the magnitude of these temporary price discounts. In Germany, for instance, sales are limited to twelve working days every six months, and cannot start on the last Mondays in January or in July. In Belgium, non seasonal sales are prohibited, and in France, Italy and Spain a prior authorization is required. In most of these countries the price discount cannot exceed 20% of the regular price. The main motivation of these restrictions has been the protection of the centers of towns and villages, and thereby the livings of small shopkeepers. In fact, the regulation of sales promotions can be seen as part of a set of legislative constraints that tries to protect small retailers and that includes also restrictions on hours of operation and location. Not surprisingly, small retailers have been the main supporters of the restrictive regulations. They argue that sales promotions are a form of predatory conduct from large supermarkets, and that consequently they have negative effects on competition and welfare.[1] Of course, large supermarkets are opposed to this argument and they consider that consumers benefit very significantly from these temporary price discounts. The regulation of sales promotions in retail markets has recently become an important policy issue in Europe, and the European Commission has started a consultation process aimed to homogenize the regulation on sales promotions in the European Union.

This paper studies the welfare implications of policies that restrict the use of sales promotions. The main objective of this paper is to study empirically the contribution of sales promotions to consumers' welfare and to the profitability of small and large retailers in a specific food retail market. The study uses scanner from supermarkets and stores in Sioux Falls, South Dakota, between 1985 and 1987.

Economists have studied different reasons why retailers use sales promotions. The one that has received greater attention in the literature is *intertemporal price discrimination (IPD)*.[2] Retailers use temporary price discounts to discriminate among consumers who are heterogeneous in their valuations of the good (Conslik et al, 1984, and Sobel, 1984) or in their information about prices (Varian, 1980), or both (Pesendorfer, 2001). The model in this paper builds on these previous studies on sales promotions and IPD, and it extends them to the context of a differentiated product market of a storable good. Similar dynamic demand models have been considered in Erdem, Imai and Keane (2002), Hendel and Nevo (2002), and Keane (2002). In this paper we present the complete equilibrium model and define the Markov Perfect equilibrium.

The typical sales promotion in our data occurs once every six weeks, the price discount is around 20%, and the increase in weekly sales is 500% (see Figure 1). Of course, part of the large increase in sales associated with these price discounts is the re-

[1] For instance, it has been argued that the flexibility of the law on sales promotions in Britain is an important factor to explain why profit margins in UK supermarkets are between two and three times larger than in continental Europe (see Burt and Spark, 1997, and Seth and Randall, 1999, pp. 214-217).

[2] Other reasons why retailers use sales promotions are: to avoid waste; to respond to shocks in consumers' shopping intensity (Warren and Barsky, 1995, and MacDonald, 2000); to encourage consumer loyalty (Slade, 1998); or to manage inventories efficiently (Aguirregabiria, 1999).

sult of consumers' substitution over brands, over stores and over time. Any approach that ignores or underestimates these substitution effects will provide an upward biased estimate of the effect of sales promotions on consumers welfare. Therefore, the consideration of a model where the good is differentiated (i.e., consumers can substitute over brands and stores) and storable (i.e., consumers hold inventories and can substitute over time) is crucial for our analysis. One of the main econometric issues in this paper is the identification of these substitution effects.

Another objective of this paper is to study how the practice of sales promotions affect firms' profits and market structure. In particular, the main interest is to test whether small retailers are affected negatively by this practice. To study empirically this issue, we use the estimated model to compute a counterfactual equilibrium under the assumption that stores keep a constant price. Then, we compare firms' profits and market structure in the factual and in the counterfactual equilibrium.

The rest of this preliminary and incomplete version of the paper is organized as follows. Section 2 presents the model. The econometric issues and the estimation procedure are explained in section 3. Section 4 describes the market and the data. Finally, section 5 presents the estimation results.

2 Model

Consider a retail market with S stores, each one selling B brands of a storable good. Time is discrete and indexed by t, and we index stores by s and brands by b. Stores are price-takers in their relationship with wholesalers, but they have market power in the retail market. Every week, stores decide prices simultaneously. Consumers hold inventories of the good, consume from these inventories and decide, every week, how many units to purchase, which brand to purchase and at which store.

2.1 Consumer behavior

Let $U_t(k_t, c_t, b_t, s_t)$ be a consumer utility function, where k_t is consumption of the storable good, c_t is consumption of the rest of the goods (i.e., composite good), and b_t and s_t represent brand and store choices, respectively. We omit the consumer subindex in this section. For a given week, all the units of the good that a consumer purchases belong to the same store and brand.[3] I adopt a *multi-stage budgeting* ap-

[3] This is the case for practically all (99.8%) households' purchases of ketchup and canned tuna in the data used in this paper.

proach to model consumers' quantity, store and brand choices.[4] Under the following assumptions on preferences, the conditions for multi-stage budgeting (Gorman, 1971) apply to this dynamic model.

Assumption 1: The one-period utility is additively separable:

$$U_t(k_t, c_t, b_t, s_t) = u_t^K(k_t) + u_t^C(c_t) + u_t^B(b_t) + u_t^S(s_t) \tag{1}$$

where $u_t^K(.)$, $u_t^C(.)$, $u_t^B(.)$, and $u_t^S(.)$ are functions.

Notice that this assumption implies that utility from consumption of the storable good does not depend on the brand or the store where it was purchased. Brand and store choices affect utility at the moment of purchase but not at subsequent periods. This assumption simplifies substantially the model. Also, we ignore brand and store switching costs.[5]

Assumption 2: Consumer preferences for brands and stores have the following form. For $b \in \{1, 2, ..., B\}$ and any $s \in \{1, 2, ..., S\}$,

$$u_t^B(b) = \bar{\xi}(b) + \xi_t(b) \qquad ; \qquad u_t^S(s) = \bar{\eta}(s) + \eta_t(s) \tag{2}$$

where $\{\bar{\xi}(b) : b = 1, 2, ..., B\}$ and $\{\bar{\eta}(s) : s = 1, 2, .., S\}$ represent the consumer's time-invariant or average taste for the different brands and stores; and $\{\xi_t(b) : b = 1, 2, ..., B\}$ and $\{\eta_t(s) : s = 1, 2, ..., S\}$ are idiosyncratic shocks in tastes, which are independently distributed over consumers, time, brands and stores with extreme value distribution.

Assumption 3: Consumers spent all their disposable income. Therefore $c_t = y - p_t^{b_t s_t} q_t$, where y is the consumer weekly income, p_t^{bs} is the price of brand b at store s, and $q_t \in \{0, 1, 2, ...Q\}$ is the number of units of the storable good that the consumer purchases at week t. The specification of the utility from consumption of the composite good is:

$$u_t^C(c_t) = \alpha_1^C \; c_t^{\alpha_2^C} \tag{3}$$

where α_1^C and α_2^C are parameters.

[4] See Hausman, Leonard and Zona (1994), and Hausman, Leonard and McFadden (1995) for this approach in static models, and Hendel and Nevo (2001) for a dynamic demand model similar to the one in this paper.

[5] In our model, all the persistence in consumers' brand and store choices is explained by time-invariant heterogeneity in tastes.

Assumption 4: We do not model explicitly the optimal consumption for the storable good. Instead, we assume that consumption of the storable good is proportional to the inventories of the good: $k_t = \lambda(i_t + q_t)$, where $\lambda \in (0, 1)$ is a parameter, and i_t is the level of inventories at the beginning of the period. Utility from consumption of the storable good is:

$$u_t^K(k_t) = \alpha_1^K \; k_t^{\alpha_2^K} + \varepsilon_t(q_t) \tag{4}$$

where α_1^K and α_2^K are parameters, and $\{\varepsilon_t(0), \varepsilon_t(1), ..., \varepsilon_t(Q)\}$ are idiosyncratic shocks which are independently and identically distributed over consumers, over time and over q with extreme value distribution.

Under these assumptions, the indirect utility of buying q units of brand b at store s is:

$$U_t = \alpha_1^K \left[\lambda \left(i_t + q\right)\right]^{\alpha_2^K} + \alpha_1^C \left(y - p_t^{bs} q\right)^{\alpha_2^C} + \varepsilon_t(q) + \bar{\xi}(b) + \xi_t(b) + \bar{\eta}(s) + \eta_t(s) \tag{5}$$

Finally, to complete the demand part of the model we should incorporate an assumption about consumers' beliefs of future prices. We assume that consumers have rational expectations and know the true transition probability of prices. Let p_t be the vector of prices for all brands and stores at period t. For the moment we only consider that $\{p_t\}$ has discrete support and follows a Markov process with transition probability $F_p(p_{t+1}|p_t)$.

Under these assumptions, a consumer's decisions can be represented in terms of three sets of choice probabilities. First, the brand choice probabilities are:

$$P_B(b \mid q_t, s_t, p_t) \equiv \Pr(b_t = b \mid q_t, s_t, p_t) = \frac{\exp\left\{\alpha_1^C \left(y - p_t^{bs_t} q_t\right)^{\alpha_2^C} + \bar{\xi}(b)\right\}}{\sum_{j=1}^{B} \exp\left\{\alpha_1^C \left(y - p_t^{js_t} q_t\right)^{\alpha_2^C} + \bar{\xi}(j)\right\}} \tag{6}$$

Second, store choice probabilities:

$$P_S(s \mid q_t, p_t) \equiv \Pr(s_t = b \mid q_t, p_t) = \frac{\exp\left\{m(s, q_t, p_t) + \bar{\eta}(s)\right\}}{\sum_{j=1}^{S} \exp\left\{\; m(j, q_t, p_t) + \bar{\eta}(j)\right\}} \tag{7}$$

where $m(s, q_t, p_t) \equiv (1/\sigma_\eta) \ln \left[\sum_{b=1}^{B} \exp\left\{\alpha_1^C \left(y - p_t^{bs} q_t\right)^{\alpha_2^C} + \bar{\xi}(b)\right\}\right]$. And third, the quantity choice probabilities:

$$P_Q(q \mid i_t, p_t) = \Pr(q_t = q \mid i_t, p_t) = \frac{\exp\left\{v(q, i_t, p_t) \;/\; \sigma_\varepsilon\right\}}{\sum_{j=0}^{Q} \exp\left\{v(j, i_t, p_t) \;/\; \sigma_\varepsilon\right\}} \tag{8}$$

where $\{v(q, i_t, p_t)\}$ are the conditional choice value functions, i.e., the expected present value of current and future utility if the consumer buys q units of the good when the current state is (i_t, p_t). These value functions are implicitly defined by the Bellman equations:

$$v(q, i_t, p_t) \equiv \alpha_1^K \left[\lambda\left(i_t + q\right)\right]^{\alpha_2^K} + M(q, p_t) + \beta\sigma_\varepsilon \sum_{p_{t+1}} F_p(p_{t+1}|p_t) \;\ln\left[\sum_{j=0}^{Q} \exp\left\{v(j, i_t, p_{t+1})\;/\sigma_\varepsilon\right\}\right] \tag{9}$$

with $M(q, p_t) \equiv \sigma_\eta \ln\left[\sum_{s=1}^{S} \exp\left\{m(s, q, p_t) + \bar{\eta}(s)\right\}\right]$.

In this model, brand and store choice probabilities are used to aggregate prices of different stores and brands. That is, these probabilities are used to obtain a single price index for each household, $M(., p_t)$, that captures the way in which the household perceives different brands and stores as substitutes of each other.

The aggregate demand of brand b at store s and period t is:

$$D_t^{bs}(p_t) = \sum_{q=1}^{Q} \left[\sum_i P_Q(q|i, p_t)\; G_t(i)\right]\; P_S(s|q, p_t)\; P_B(b|q, s, p_t) \tag{10}$$

where $G_t(i)$ is the distribution of consumer inventories at period t. Consumers, or households, are heterogeneous in their tastes. In particular, there are N groups of consumers according to the values of the parameters $\{\bar{\xi}(b)\}$, $\{\bar{\eta}(s)\}$, α_1^K and λ. Therefore, we have a different distribution of inventories and a different demand for each consumer type. The market demand $D_t^{bs}(p_t)$ is just the result of aggregating the demands from each consumer type.

2.2 Store pricing decisions

Let $p_t^s = \{p_t^{bs} : b = 1, 2, ..., B\}$ be the vector of prices at store s, and let p_t^{-s} represent prices at all the stores except store s. The set of feasible prices is discrete: $p_t^s \in \{p^1, p^2, ..., p^J\}$. Current profits have four components: revenue, purchasing costs, price adjustment costs (i.e., the so called *menu costs*), and a profitability shock that is private information of the store. That is, profits of store s at period t are:

$$R_t^s(p_t, p_{t-1}^s, \omega_t^s) = \sum_{b=1}^{B}(p_t^{bs} - w^b)\; D_t^{bs}(p_t) - AC\left(\sum_{b=1}^{B} 1(p_t^{bs} \neq p_{t-1}^{bs})\right) + \omega_t^s(p_t^s) \tag{11}$$

where w^b is the unit cost of brand b; $AC(.)$ represents price adjustment costs, which depend on the number of price changes; $1(.)$ is the indicator function; and $\omega_t^s(p_t^s)$

is a component of profitability that is private information of store s and may depend on prices at this store. We assume that these private information shocks are independently and identically distributed over time and over stores.

Every week, stores decide prices simultaneously. When taking this decision, stores know the demand functions $D_t^{bs}(.)$, previous prices at every store, p_{t-1}, and their own shocks $\{\omega_t^s\}$, but they have uncertainty about the shocks ω at other stores. Therefore, they have uncertainty about current prices at other stores.

The game has a Markov structure, and we assume that firms play Markov strategies. That is, if $\{p_{t-1}, \omega_t^s\} = \{p_{t+j-1}, \omega_{t+j}^s\}$ then the prices that store s chooses at periods t and $t+j$ are the same. Let $\pi = \{\pi_s(p_{t-1}, \omega_t^s)\}$ be a set of strategy functions or decision rules. Associated with a set of strategy functions π we can define a set of *conditional choice probabilities* $F_\pi = \{F_\pi^s(p_t^s|p_{t-1})\}$ such that,

$$F_\pi^s(p^j|p_{t-1}) \equiv \Pr\left[\pi_s(p_{t-1}, \omega_t^s) = p^j \mid p_{t-1}\right] = \int I\left\{\pi_s(p_{t-1}, \omega_t^s) = p^j\right\} \; g(\omega_t^s) \; d\omega_t^s \tag{12}$$

where $I\{.\}$ is the indicator function. The probabilities $\{F_\pi^s(p^j|p_{t-1}) : j = 1, 2, ..., J\}$ represent the expected pricing behavior of store s from the point of view of the rest of the firms when firm i follows its strategy in π.

Following Milgrom and Weber (1985) we can represent a Markov Perfect equilibrium (MPE) in probability space. Let π^* be a set of MPE pricing strategies, and let F^* be the set of conditional choice probabilities associated with these strategies. Then, F^* solves a fixed point mapping in probability space: $F^* = \Psi(F^*)$. See Aguirregabiria and Mira (2002b). In this reformulation of the equilibrium, price transition probabilities represent stores' strategies. An equilibrium is a fixed point of a suitably modified best response function $\Psi()$.

3 Estimation of structural parameters

Our data set is a long panel with information on consumers' purchasing histories and demographic characteristics, and stores' prices and sales for all brands in several product lines. It also reports total expenditure of individual consumers at every store visit. The parameters of the model are: (1) tastes for brands and stores, $\{\bar{\xi}(b)\}$ and $\{\bar{\eta}(s)\}$; (2) parameters in the utility functions, α_1^C, α_2^C, α_1^K, α_2^K; (3) consumption rate for the storable good, λ; (4) dispersion parameters σ_η and σ_ε; (5) marginal costs, $\{w^b\}$; and (6) menu costs, $AC(.)$. We estimate these parameters following a sequential procedure.

In this section we explain our sequential procedure for the estimation of these parameters. But before we describe our measure of consumers' inventories.

Let $\tau_{gt} \in \{1, 2, ...\}$ be the variable "time since last purchase of good g", and let $\tilde{n}_{gt}$ represent the number of units the consumer bought in his last purchase (i.e., $\tilde{n}_{gt} = n_{g,t-\tau_{gt}}$). The transition rule of inventories implies that:

$$\ln(k_{gt}) = \tau_{gt} \ \ln(1 - \lambda_g) + \ln(\tilde{n}_{gt})$$

Based on this expression we can use τ_{gt} and $\tilde{n}_{gt}$ as state variables, and treat λ_g as an unknown parameter to be estimated. Notice that we do not observe initial values for $\tilde{n}_{gt}$ and τ_{gt}. However, given that these variables are observable we can use their sample means (for each type consumer type) as initial conditions. Also notice that τ_{gt} and $\tilde{n}_{gt}$ are discrete variables, and we do not have to make any discretization of the state space to estimate consumers' dynamic programming problem.

3.1 Estimation of demand parameters

(A). Brand choices: Given prices, a consumer most preferred brand does not depend on the store, or on whether he has high or low shopping costs. It is also independent of the quantity chosen, and of whether this quantity is positive or zero. The most preferred brand only depends of the values of $\{\bar{u}_b + \varepsilon_{bt} - p_{bt}\}$ for different brands. Therefore, we can estimate preferences $\{\bar{u}_b\}$ using the histories of brand choices for all households in the sample. The model is a static multinomial logit. We specify preference parameters $\{\bar{u}_b\}$ as linear functions of a vector of household characteristics x_i (i.e., household income, family size, education and working status of the head(s) of the household). More specifically, for household i, $\bar{u}_{bi} = x_i' \gamma_b$ where γ_b is a vector of parameters. We allow the dispersion parameter σ_ε to depend on household income. We also allow for aggregate shocks in brand preferences (e.g., promotions and advertisement from manufacturers) by including time dummies for each brand.

(C). Store choices: Using information on choice of store for the subsample of observations with HSC we can estimate (up to scale) the values $\{W_i^s\}$ using the probabilities in (**??**). Although we have a long panel (i.e., 93 weeks for each household), there are some households which have HSC for only very few weeks. Therefore, we specify the values W_i^s as parametric functions of household characteristics.

For the subsample of observations with LSC we use the probabilities in (**??**) to estimate preference parameters $\{\bar{u}^s\}$. First, we use our estimates of brand preferences to obtain estimates for $m_i^s(p_t^s) \equiv \ln\left[\sum_b \exp\{(\bar{u}_{bi} - p_{bt}^s)/\sigma_\varepsilon\}\right]$. Then, use solve these estimates in the probabilities $P^{Ls}(p_t)$ and estimate the parameters $\{\bar{u}^s\}$. Again, we specify $\{\bar{u}^s\}$ as linear functions of a vector of household characteristics: $\bar{u}_i^s = x_i'\gamma^s$ where γ^s is a vector of parameters.

(D). Quantity choices: The function $c(.)$, the consumption rate λ, and the discount factor are estimated from the quantity choice probabilities. These probabilities result from a discrete choice dynamic programming model with observable state variables τ_{gt} and $\tilde{n}_{gt}$. This model has a similar structure to the one studied by Rust (1994), with the particular feature that it is and ordered discrete choice model. In principle, the NFXP algorithm in Rust (1994) or the NPL algorithm in Aguirregabiria and Mira (2001) can be used to obtain maximum likelihood estimates. The space of the observable state variable $(\tau_{gt}, \tilde{n}_{gt})$ is relatively small (i.e., approx. 200 cells), but given that we allow for heterogeneity in c and λ the computational cost of NFXP is significantly larger than NPL, and therefore we use the later algorithm.

Our specification of the consumption rate is, $\lambda_i = \exp\{\lambda_0 + \lambda_1 hsize_i\}$, where λ_0 and λ_1 are parameters and $members_i$ is the number of members in household i. The specification of function $c(.)$ is: $c(k_{it}; x_i) = (x_i'\delta)\ln(k_{it})$, where δ is a vector of parameters.

3.2 Estimation of supply parameters

There are two main econometric and computational issues in the estimation of the supply parameters: (1) the dimension of the state space in a store decision problem is very large; and (2) the model has multiple equilibria, what creates an *indeterminacy problem* and makes maximum likelihood estimation unfeasible. To deal with the first problem we use the randomization techniques proposed by Rust (1997). For the second problem, we use a sequential estimation procedure proposed in Aguirregabiria and Mira (2002b) to deal with multiple equilibria in the estimation of dynamic discrete choice games. The main assumption behind this method is that there are not *sunspots* associated with the multiple equilibria. That is, for given structural parameters, players (or nature) always select the same equilibrium and they do not jump between different equilibria. We describe below in more detail our econometric approach.

(A) First stage: Expected demands: To obtain expected demands we should estimate the distribution of consumers' inventories conditional on the state variables (p_{t-1}, x_t). We run regressions for $\ln(\tau_{it})$ and for $\ln(\tilde{n}_{it})$ on these state variables and a vector household characteristics. Then we combine our estimates of demands $d_b^s(p_t, k_t|i)$ with conditional distribution of τ_{it} and $\tilde{n}_{it}$ that results from the previous regressions to obtain the expected demands $D_b^s(p_t, p_{t-1}, x_t)$.

(B) Second stage: reduced form estimation of price transition probabilities: Given that private information shocks $\{\omega_t^s\}$ are *iid* over stores, it is clear that $\Pr(p_t|p_{t-1}, x_t) = \prod_{s=1}^S \Pr(p_t^s|p_{t-1}, x_t)$, and therefore we can estimate price transition probabilities separately for each store. Furthermore, our parametric assumption on the distribution of $\{\omega_t^s\}$ provides useful information that can be exploited in the reduced form estimation of these probabilities. In particular, if $\{\omega_t^s(a)\}$ are *iid* extreme value we have that $\Pr(p_t^s = p_a|p_{t-1}, m_t) = \exp(h_a^s(p_{t-1}, x_t) \left[\sum_{j=1}^A \exp(h_j^s(p_{t-1}, x_t)\right]$. We use polynomials in (p_{t-1}, x_t) for the functions $h_a^s(.)$. We use $\hat{P}^s(p_t^s|p_{t-1}, x_t)$ to denote these estimated reduced form probabilities.

(C) Third stage: estimation of structural parameters: Let $\hat{y}_{bt}^s(a)$ represent (estimated) expected sales of brand b at store s when this store chooses the vector of prices p_a and the other stores behave according to $\hat{P}^{-s}$. That is,

$$\hat{y}_{bt}^s(a) \equiv \sum_{p_{-s}} \hat{P}^{-s}(p^{-s}|p_{t-1}, x_t)\ D_b^s(p_a, p^{-s}, p_{t-1}, x_t)$$

Define also, $\hat{y}_{0t}^s(a) \equiv \sum_b p_{ba} \hat{y}_{bt}^s(a)$, which is the estimated expected revenue for store s at period t if it chooses prices p_a. For the sake of presentation, we first consider the myopic case, i.e., $\beta = 0$. Under the assumption that $\{\omega_t^s(a)\}$ are *iid* extreme value distributed, the probability that store s chooses prices p_a at period t is:

$$P^s(p_a|p_{t-1}, x_t) = \frac{\exp\left\{\dfrac{1}{\sigma_\omega}\left[\hat{y}_{0t}^s(a) - \sum_b c_b\ \hat{y}_{bt}^s(a) - \eta^s\ i_t^s(a)\right]\right\}}{\sum_{j=1}^A \exp\left\{\dfrac{1}{\sigma_\omega}\left[\hat{y}_{0t}^s(j) - \sum_b c_b\ \hat{y}_{bt}^s(j) - \eta^s\ i_t^s(j)\right]\right\}}$$

where η^s is a menu cost parameter, and $i_t^s(a) = \sum_b 1(p_{ba}^s \neq p_{b,t-1}^s)$. It is clear that the parameters $\{c_b\}$ and $\{\eta^s\}$ are theoretically identified. Given estimates of structural parameters we can obtain more informed estimates of the variables $y_{bt}^s(a)$ and $y_{0t}^s(a)$ and apply again the same procedure.

When stores are forward looking, i.e., $\beta > 0$, the expression for price transition probabilities is more complicated:

$$P^s(p_a|p_{t-1}, x_t) = \frac{\exp\left\{\frac{1}{\sigma_\omega}\left[\hat{y}^s_{0t}(a) - \sum_b c_b \hat{y}^s_{bt}(a) - \eta^s i^s_t(a) + \beta\left(\hat{Y}^s_{0t}(a) - \sum_b c_b \hat{Y}^s_{bt}(a) - \eta^s \hat{I}^s_t(a)\right)\right]\right\}}{\sum_{j=1}^{A}\exp\left\{\frac{1}{\sigma_\omega}\left[\hat{y}^s_{0t}(j) - \sum_b c_b \hat{y}^s_{bt}(j) - \eta^s i^s_t(j) + \beta\left(\hat{Y}^s_{0t}(j) - \sum_b c_b \hat{Y}^s_{bt}(j) - \eta^s \hat{I}^s_t(j)\right)\right]\right\}}$$

where $\{\hat{Y}^s_{0t}(a)\}$, $\{\hat{Y}^a_{bt}(a)\}$ and $\{\hat{I}^s_t(a)\}$ are functions of price transition probabilities and market demands. These functions are more complicated than their static counterparts, but they can be obtained solving a system of linear equations with dimension the one of the space of observable state variables (p_t, x_t). Given these variables, the model is just a multinomial logit with nonlinear restrictions among the parameters, and the estimation is very standard.

The main computational cost of this procedure appears in the computation of the values $\hat{y}^s_{bt}(a)$, and specially in the computation of $\{\hat{Y}^s_{0t}(a)\}$, $\{\hat{Y}^a_{bt}(a)\}$ and $\{\hat{I}^s_t(a)\}$. It is in these computations where we exploit the randomization techniques in Rust (1997).

4 Market, data and preliminary evidence

The data in this paper comes from the ERIM database collected by A. C. Nielsen between 1985 and 1988 in Sioux Falls, South Dakota.[6] In 1985, Sioux Falls had a population of 81,340 people. More than 2,500 households were issued magnetic ID cards to be presented at the checkout counter when shopping at participating stores. Purchases were scanned and sent along with the panelist ID number to a central computer. Table 2 presents descriptive statistics of household demographics and shopping behavior.

The number of participating stores was 19. According to A.C. Nielsen, these stores represent 80% of the grocery and drug retail sales in Sioux Falls. Thirteen stores belong to four regional supermarket chains, and six are independent stores. Table 3 presents market shares for the ten products in our data set. Chains A and B are the leaders in terms of market shares, and they account for approximately two

[6] Nielsen selected this location because it was representative of the US population in terms of household income, size and age.

thirds of the market in each of the product lines. Average market shares are very similar for the different product lines.

Figures 1 and 2 present the time series of prices and sales for the leading brands of ketchup and canned tuna in three supermarkets chains. Several features of these series are shared by most brands and products in our data. While supermarkets A and B tend to have sales promotions quite frequently, chain D rarely has sales promotions. Although sales promotions are common practice by many retailers, not all stores follow this pricing strategy, or at least not with the same intensity. For instance, supermarkets that advertise themselves as "everyday low price" keep prices constant for longer periods than their competitors, and they rarely apply sales promotions. In our data set, we find this type of heterogeneity between supermarkets' pricing strategies: some supermarket chains practice sales promotions very intensively while others tend to keep constant prices. In particular, firms with larger market shares are the ones who practice sales promotions, and they tend to charge significantly higher *regular prices* than those firms that charge constant prices. Interestingly, once we obtained average prices weighted by sales, these price differentials disappear.

5 Estimation results

We estimated a preliminary version of the model. Results will be presented in the seminar.

REFERENCES:

Aguirregabiria, V. (1999): "The dynamics of markups and inventories in retailing firms," *The Review of Economic Studies*, 66, 275-308.

Aguirregabiria, V., and P. Mira (2002a): "Swapping the nested fixed point algorithm: A class of estimators for Markov decision models", *Econometrica*, forthcoming.

Aguirregabiria, V. and P. Mira (2002b): "Simulated Pseudo Maximum Likelihood Estimation of Dynamic Discrete Games," manuscript, Department of Economics, Boston University.

Anderson, S., A. de Palma, and J.-F. Thisse (1992): "Discrete choice theory of product differentiation", Cambridge, MA. MIT Press.

Berry, S. (1994): "Estimating discrete choice models of product differentiation", *RAND Journal of Economics*, 25, 242-262.

Berry, S., J. Levinshon and A. Pakes (1996): "Automobile prices in market equilibrium", *Econometrica*, 64, 841-890.

Burt, S., and L. Sparks (1997): "Performance in food retailing: A cross national consideration and comparison of retail margins," *British Journal of Management*, 8, 133-150.

Chevalier, J. (1995): "Capital structure and product-market competition: Empirical evidence from the supermarket industry", *American Economic Review*, 85(2), 415-435.

Chesher, A., and J. Santos-Silva (2002): "Taste variation in discrete choice models", *Review of Economic Studies*, 69, 147-168.

Coase, R. (1972): "Durability and monopoly", *Journal of Labor and Economics*, 15, 143-149.

Conslik, J., E. Gertsner and J. Sobel (1984): "Cyclic pricing by a durable good monopolist", *Quarterly Journal of Economics*, 99, 489-505.

Cotterill, R. (1986): "Market power in the retail food industry: Evidence from Vermont", *Review of Economics and Statistics*, 68(3), 379-386.

Cotterill, R, and L. Haller (1992): "Barrier and queue effects: A study of leading US supermarket chain entry patterns", *Journal of Industrial Economics*, 40(4), 427-440.

Ericson, R., and A. Pakes (1995): "Markov-Perfect industry dynamics: A framework for empirical work", *Review of Economic Studies,* 62, 53-82.

Erdem, T., S. Imai and M. Keane (2002): “Consumer Price and Promotion Expectations: Capturing Consumer Brand and Quantity Choice Dynamics under Price Uncertainty,” manuscript, University of California Berkeley.

Hausman, J., G. Leonard, and D. McFadden (1995): “A utility-consistent, combined discrete choice and count data model: Assesing recreational use losses due to natural resource damage,” *Journal of Public Economics*, 56, 1-30.

Hausman, J., G. Leonard, and J. Zona (1994): “Competitive Analysis with Differentiated Products,” *Annales de Economie et de Statistique*, 34, 159-80.

Hendel, I. and A. Nevo (2002a): “Sales and Consumer Stokpiling,” manuscript, University of California at Berkeley.

Hendel, I. and A. Nevo (2002b): “Measuring the Implications of Sales and Consumer Stockpiling Behavior,” manuscript, University of California at Berkeley.

Holton, R. (1957): ”Price discrmination at retail: the supermarkets case”, *Journal of Industrial Economics*, 6(1), 13-32.

Keane, M. (2002): “Brand and Quantity Choice Dynamics under Price Uncertainty,” manuscript. Yale University.

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

Pakes, A. and P. McGuire (1994): ”Computing Markov-Perfect Nash equilibria: Numerical implications of a dynamic differentiated product model, *Rand Journal of Economics*, 555-589.

Pesendorfer, M. (2000): ”Retail Prices: A Study of Price Behavior in Supermarkets”, *Journal of Business*, Forthcoming.

Rust, J. (1994): ”Estimation of dynamic structural models, problems and prospects: discrete decision processes”, in C. Sims (ed.) *Advances in Econometrics. Sixth World Congress*, Cambridge University Press.

Seth, A., and G. Randall (1999): “The Grocers: The Rise and Rise of the Supermarket Chains,” Kogan Page Limited. London.

Slade, M. (1998): ”Optimal Pricing with Costly Adjustment: Evidence from Retail Grocery Stores”, *Review of Economic Studies*, 65, 87-108.

Smith, H. (2001): ”Supermarket choice and supermarket competition in market equilibrium”, mimeo University of Oxford.

Sobel, J. (1984): "The timing of sales", *Review of Economic Studies*, 51, 353-368.

Sobel, J. (1991): "Durable good monopoly with entry of new consumers", *Econometrica*, 59, 1455-1485.

Stokey, N. (1981): "Rational expectations and durable goods pricing", *Bell Journal of Economics*, 12, 112-128.

Tirole, J. (1988): "The theory of industrial organization", MIT Press, Cambridge, Massachusetts.

Varian (1980): "A Model of Sales", *American Economic Review*, 70, 651-659.

Warner, E., and R. Barsky (1995): "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays," *The Quarterly Journal of Economics*, 110, 321-352.

Table 1
The food supermarket industry in the US economy*

Employment				*Revenue*			
Rank	NAICS	Industry	Million workers	Rank	NAICS	Industry	Sales Billion $
1	4521	Dept. stores	2.5	1	5241	Insurance	996
2	**44511**	**Food supers.**	**2.5**	2	336	Transp. equip.	572
3	5221	Banks	2.0	3	4411	Auto dealers	554
4	336	Transp. equip.	1.8	4	5221	Banks	533
5	334	Computers	1.7	5	334	Computers	438
6	5241	Insurance	1.6	6	311	Food manuf.	424
7	311	Food manuf.	1.5	7	325	Chemicals	420
8	333	Machinery	1.4	**8**	**44511**	**Food supers.**	**351**
9	4411	Auto. dealers	1.1	9	333	Machinery	270
10	325	Chemicals	0.9	10	4521	Dept. stores	220

*Source: 1997 US Economic Census

Table 2
Household demographics and shopping behavior
Sioux Falls (South Dakota): Year 1985

Demographics		**Shopping behavior**	
Variable	*Avg (std)*	*Variable*	*Avg (std)*
Family members	3.34 (1.36)	Expenditure (in $ per week)	58.6 (59.1)
Annual income (in $)	28,998 (15.477)	Expenditure per capita (in $ per week)	19.3 (20.5)
Female and male heads	86.5 %	Number of store visits per week	2.8 (1.9)
Only female head	12.2 %		
Only male head	1.3 %		
Male working	83.8 %		
Female working	71.2 %		
Male high school graduate	88.9 %		
Female high school graduate	91.1 %		
Male college graduate	28.6 %		
Female college graduate	23.3 %		

Table 3
Store Market shares (in %)
Sioux Falls. From 1985 (week 35) to 1987 (week 23)

	Chain A (3 stores)	*Chain B* (6 stores)	*Chain C* (2 stores)	*Chain D* (2 stores)
Ketchup	35.8	27.9	8.6	19.4
Sugar	35.4	28.6	6.7	19.8
Tuna	35.9	30.6	10.5	13.9
Brownies	33.1	36.7	9.0	14.7
Dry detergent	36.3	30.8	11.5	15.2
Margarine	35.8	31.6	7.7	18.2
Peanut butter	36.2	30.3	8.3	17.2
Soups	33.7	32.0	10.5	16.4
Tissue paper	33.7	30.1	11.0	15.7
Yoghurt	39.5	27.6	14.4	10.1

Market shares are based on sales in physical units.

Table 4
Chain market shares for different brands of ketchup and tuna
Sioux Falls. Period 1985-1987

	Ketchup			
	Chain A	*Chain B*	*Chain C*	*Chain D*
Heintz (32 oz)	34.7	26.7	7.8	23.8
Hunt's (32 oz)	41.4	29.0	12.2	10.1
Del Monte (32 oz)	46.9	32.1	1.9	14.7
	Tuna			
	Chain A	*Chain B*	*Chain C*	*Chain D*
Star-Kist in water (6.5 oz)	34.3	33.0	13.9	10.6
Chicken of sea in water (6.5 oz)	42.0	30.6	6.5	14.3
Star-Kist in oil (6.5 oz)	34.1	34.3	12.2	12.5
Chicken of sea in oil (6.5 oz)	41.9	29.8	5.0	16.9

Table 5
Relative average prices (Chain A = 100)
Sioux Falls. Period 1985-1987

		Chain B	*Chain C*	*Chain D*
Ketchup	Not weighted	101.5	109.5	101.5
	Weighted	98.8	104.4	100.0
Sugar	Not weighted	98.2	96.7	72.3
	Weighted	98.0	96.8	73.5
Tuna	Not weighted	107.3	103.6	96.0
	Weighted	103.3	95.6	100.2
Brownies	Not weighted	97.1	92.5	92.3
	Weighted	98.2	91.4	96.0
Dry detergent	Not weighted	99.0	97.4	97.6
	Weighted	98.4	99.1	98.7
Margarine	Not weighted	96.3	98.3	91.4
	Weighted	93.6	90.4	96.6
Peanut butter	Not weighted	103.3	99.8	97.4
	Weighted	100.0	100.0	96.9
Soups	Not weighted	100.8	94.4	92.5
	Weighted	98.9	92.7	93.2
Tissue paper	Not weighted	102.8	99.0	93.1
	Weighted	99.6	93.5	95.1
Yoghurt	Not weighted	91.5	94.2	96.2
	Weighted	90.4	93.4	97.0
10 products	Not weighted	**99.8**	**98.5**	**93.0**
	Weighted	**97.4**	**95.7**	**95.1**

(1) Not weighted relative price between chain s and chain A is:
$\sum_b (\bar{p}_{sb}/\bar{p}_{Ab})w_b$, where $\bar{p}_{sb}$ is the average price (not weighted by sales) of brand b at chain s, and w_b is the market share of brand b.

(2) Weighted relative price between chain s and chain A is:
$\sum_b (\hat{p}_{sb}/\hat{p}_{Ab})w_b$, where $\hat{p}_{sb}$ the average price (weighted by sales) of brand b at chain s.

Table 6
Conditional Logit for Brand Choice
Product: Ketchup. Market: Sioux Falls (1985-1987)

	Without taste heterogeneity			
	Linear Utility		*Nonlinear utility*	
Parameter	*Estimate*	*Std. error*	*Estimate*	*Std. error*
Brand 1 ($\bar{\xi}^1$)	2.0730	(0.0438)	2.0459	(0.0441)
Brand 2 ($\bar{\xi}^2$)	1.1738	(0.0467)	1.1508	(0.0470)
Brand 3 ($\bar{\xi}^3$)	0.9818	(0.0483)	0.9503	(0.0486)
Brand 4 ($\bar{\xi}^4$)	1.1191	(0.0472)	1.1294	(0.0473)
Brand 5 ($\bar{\xi}^5$)	0.1621	(0.0573)	0.1872	(0.0575)
Brand 6 ($\bar{\xi}^6$)	0.4397	(0.0522)	0.4097	(0.0526)
α_{C1}	1.8293	(0.1483)	6.7728	(1.2430)
α_{C2}	1.0	-	0.8658	(0.2528)
α_{C3}	0.0	-	0.1374	(0.0541)
# obs.	13,496		13,496	
Log likel.	-22,657.4		-22,629.8	
Like. Ratio Index	0.0034		0.0047	

	With taste heterogeneity			
	Linear Utility		*Nonlinear Utility*	
Parameter	*Estimate*	*Std. error*	*Estimate*	*Std. error*
Brand 1 ($\bar{\xi}^1$)	1.9465	(0.0479)	1.9185	(0.0484)
Brand 2 ($\bar{\xi}^2$)	0.8782	(0.0516)	0.8552	(0.0520)
Brand 3 ($\bar{\xi}^3$)	0.7263	(0.0526)	0.6918	(0.0534)
Brand 4 ($\bar{\xi}^4$)	0.9591	(0.0518)	0.9664	(0.0519)
Brand 5 ($\bar{\xi}^5$)	0.0671	(0.0624)	0.0953	(0.0629)
Brand 6 ($\bar{\xi}^6$)	0.2265	(0.0572)	0.1930	(0.0580)
α_{C1}	2.6231	(0.1745)	2.7068	(1.4024)
α_{C2}	1.0	-	0.9244	(0.3052)
α_{C3}	0.0	-	0.5516	(0.3491)
Cov. matrix $\{\bar{\xi}_i^b\}$	$\Omega_\xi = \begin{bmatrix} 0.42 & & & & & \\ -0.06 & 0.29 & & & & \\ -0.01 & 0.08 & 0.23 & & & \\ -0.01 & 0.04 & 0.05 & 0.23 & & \\ 0.02 & 0.04 & 0.04 & 0.04 & 0.11 & \\ 0.02 & 0.07 & 0.09 & 0.04 & 0.04 & 0.16 \end{bmatrix}$			
Score test: $H_0 : \Omega_\xi = 0$	$p - value = 0.0000$			
# obs.	13,496		13,496	
Log likel.	-15,934.0		-15,920.6	
Like. Ratio Index	0.2991		0.2998	

Table 7
Conditional Logit for Store Choice
Product: Ketchup. Market: Sioux Falls (1985-1987)

	Without taste heterogeneity		*With taste heterogeneity*	
Parameter	*Estimate*	*Std. error*	*Estimate*	*Std. error*
Store 1 ($\bar{\eta}^1$)	1.8359	(0.0363)	1.6570	(0.0434)
Store 2 ($\bar{\eta}^2$)	1.6570	(0.0369)	1.4000	(0.0454)
Store 3 ($\bar{\eta}^3$)	0.6433	(0.0435)	0.2303	(0.0536)
Store 4 ($\bar{\eta}^4$)	0.0012	(0.0477)	-0.0658	(0.0540)
σ_ξ	0.5359	(0.1846)	0.6323	(0.1354)
Cov. matrix $\{\bar{\eta}_i^s\}$			$\Omega_\eta = \begin{bmatrix} 0.68 & & & \\ -0.20 & 0.74 & & \\ 0.00 & 0.05 & 0.38 & \\ 0.08 & 0.07 & 0.09 & 0.16 \end{bmatrix}$	
Score test: $H_0 : \Omega_\eta = 0$			$p - value = 0.0000$	
# obs.	13,496		13,496	
Log likel.	-18,103.9		-9,484.8	
Like. Ratio Index	0.0002		0.4762	

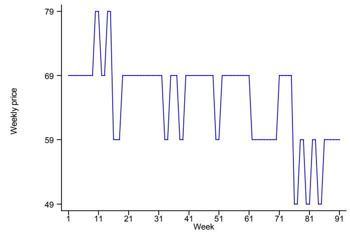
FIGURE 1

Time series of prices and sales for the two leading brands of Tuna

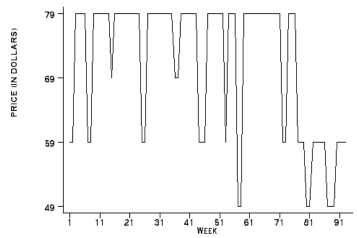
Brand: Star-Kist. Water

Prices

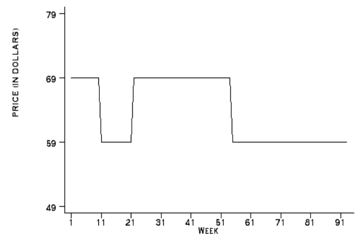
Chain A



Chain B

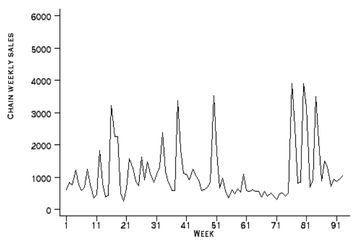


Chain D

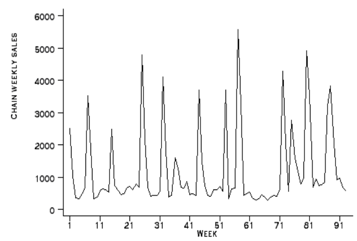


Sales

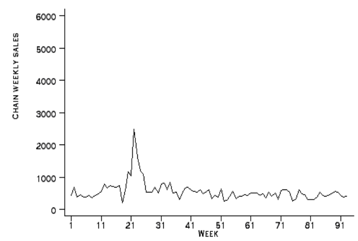
Chain A



Chain B



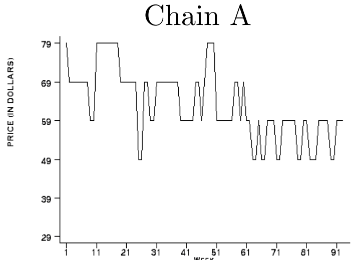
Chain D



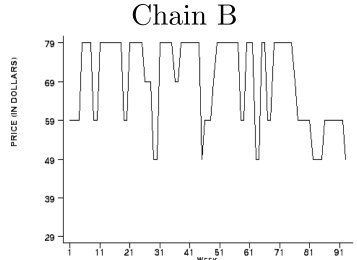
Brand: Chicken of Sea. Water

Prices

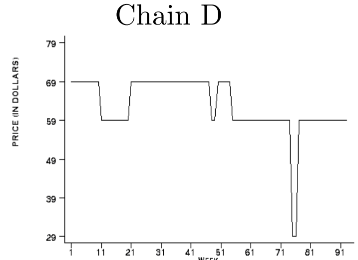
Chain A



Chain B

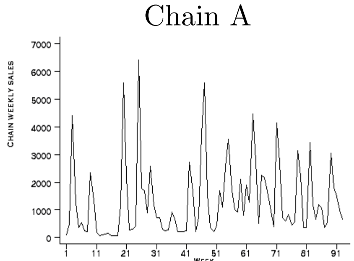


Chain D

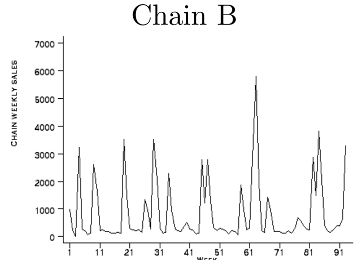


Sales

Chain A



Chain B



Chain D

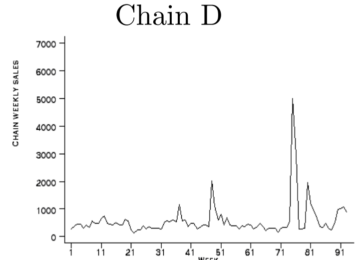


FIGURE 2

Time series of prices and sales for the leading brand of Ketchup

Brand: Heinz (32 ounzes)