

The Dynamics of Markups and Inventories in Retailing Firms

VICTOR AGUIRREGABIRIA
University of Western Ontario

First version received February 1996; Final version accepted March 1998 (Eds.)

This paper is concerned with the interaction between price and inventory decisions in retailing firms and its implications for the dynamics of markups and the existence of sales promotions. We consider a model where a monopolistically competitive retailer decides price and inventories, and assumes lump-sum costs when placing orders or changing nominal prices. In this model, the existence of stockout probabilities and fixed ordering costs generate a cyclical price behaviour characterized by long periods without nominal price changes and short periods with very low prices (*i.e.* sales promotions). We estimate this model using a unique longitudinal dataset with information about retail and wholesale prices, inventories, orders, and sales for several brands in a supermarket chain. Based on the estimated model we perform several counterfactual experiments that show the important role that inventories and fixed ordering costs play in the dynamics of retail prices and the frequency of sales promotions in this dataset.

1. INTRODUCTION

Recent studies have presented empirical evidence supporting the existence of significant price dispersion and staggering in price changes across individual stores (Lach and Tsiddon (1992, 1996), Tommasi (1993)). The most common explanation for cross-sectional price dispersion and staggering in price changes builds on the existence of nominal price adjustment costs (*e.g.* menu costs) and not perfect correlation among the demand shocks at different firms. Sheshinski and Weiss (1977, 1983) show that, under certain conditions, an (S, s) rule is the optimal pricing rule for a monopolistically competitive firm who faces lump-sum costs of adjusting nominal prices.¹ If firms' demand shocks are not perfectly correlated, nominal price changes will not be synchronized across firms and, consequently, there will be cross-sectional dispersion in nominal prices.

Although there is increasing evidence for the existence of significant infrequency in nominal price changes (Cecchetti (1986), Lach and Tsiddon (1992, 1996), Kashyap (1995), Slade (1994, 1996)), there is also important evidence against the hypothesis of retailing firms following simple (S, s) pricing rules. The main evidence against the simple (S, s) model seems to be the relatively high frequency of sales promotions in retail stores. Using monthly price information from retail stores in Israel between 1978 and 1982, Lach and Tsiddon (1992) report that 5% of the observed nominal price changes were price reductions. Since annual inflation rate during that period was between 60% and 130%, this evidence seems puzzling. Slade (1994, 1996) shows that the prices of several brands of crackers in the supermarkets of a north american town tend to alternate between duration

1. Under an (S, s) pricing rule real price moves between an upper bound, S , and a lower bound, s . The nominal price is not changed when real price is above the lower threshold, but a nominal price increase occurs when real price reaches that threshold, and the new real price becomes S .

spells with a low nominal price and duration spells with a high price. Using also supermarket price datasets, Warren and Barsky (1995) and Pesendorfer (1996) report that *regular* retail prices (*i.e.* prices not under promotion) stay constant for relatively long periods (*i.e.* several months) but, during these periods, sales promotions are very frequent.

What are the main factors explaining sales promotions in retailing firms? How is it possible to explain the coexistence of long periods without nominal price changes and very short periods where a price reduction is followed by a price increase? There are different explanations for the existence and timing of sales promotions. In Varian (1980), promotions are viewed as a mechanism to discriminate among customers with different information or search costs. According to that model, sales must be costly to detect and thus they should be randomly distributed over time and not correlated with variables that are common knowledge for most customers (*e.g.* previous prices). Warren and Barsky (1995) show that an important proportion of sales promotions occur during weekends and holidays. They present empirical support for the hypothesis that these sales promotions are the result of exogenous shocks in consumers shopping intensity. These shocks make retailers' demands more elastic and, therefore, the optimal markup becomes smaller. Slade (1994) and Pesendorfer (1996) present strong evidence against the hypothesis that sales are randomly distributed over time. They show that sales promotions are state dependent (they depend on the history of previous prices) and duration dependent (they depend on the time duration since last promotion). In Slade, demand depends on a stock of goodwill that accumulates (erodes) when the firm charges low (high) prices. The existence of menu costs, together with the effect of the stock of goodwill on demand, can explain the coexistence of price rigidity and sales promotions. Pesendorfer shows that an important number of stylized facts associated with sales promotions are consistent with the hypothesis of intertemporal price discrimination of heterogeneous consumers (Sobel (1984)). This hypothesis is based on the existence of heterogeneous consumers and infrequent purchases of durable goods. In these models a sales promotion becomes optimal when a large enough number of low-willingness consumers (*i.e.* "shoppers") accumulates in the market.

In all these previous models inventory decisions do not play any role in the explanation for markdowns or sales promotions. However, there are important sources of interaction between price and inventory decisions that could contribute to explain markdowns. In this paper we consider this interaction as an additional element to the understanding of sales promotions in retailing firms. In particular, we show that the existence of stockout probabilities (that create substitutability between price and inventories in the profit function) and fixed ordering costs (that generate (S, s) inventory behaviour) can explain the puzzling coexistence of large periods without nominal price changes and short periods with very low prices. Furthermore, this hypothesis can also explain the state dependence and duration dependence of sales promotions. Our model has strong empirical predictions on the joint dynamics of prices, sales, and inventories. We test these predictions and estimate our structural model using a unique panel dataset that contains monthly information on sales, stocks, retail prices, wholesale prices, and orders to suppliers for 534 brands sold by a supermarket chain between January 1990 and May 1992.

Section 2 presents the theoretical model. In this model a monopolistically competitive retailer decides price and orders taking into account the existence of lump-sum costs of placing orders and changing nominal prices. The model combines the classical (S, s) inventory model (see Scarf (1959), Blinder (1981), among others) and the classical (S, s) pricing model (see Sheshinski and Weiss (1977, 1983)) to allow for joint price and inventory decisions with lump-sum ordering costs and menu costs. We characterize the form of the

optimal decision rule and analyse some of its empirical predictions. In particular, if fixed ordering costs are large relative to menu costs, sizable price reductions tend to occur when orders are placed, and the low price is maintained for a short period of time. After this period, infrequent price increases will occur until the next positive order. We show also that lump-sum ordering costs can generate cross-sectional price dispersion even when menu costs are negligible or zero.

The main emphasis of this paper is empirical. Using our supermarket dataset we obtain three sets of empirical evidence about our hypothesis. First, we present a descriptive analysis of the dynamic behaviour of prices and inventories and statistical tests of the predictions of our model in the context of reduced form estimations of the optimal decision rules. In these reduced form estimations we control for several potential sources of spurious state dependence (*i.e.* unobserved brand-heterogeneity and autocorrelation in demand shocks). Second, based on our theoretical model we estimate a discrete choice dynamic structural model. The structural parameters are estimated using a two-stage procedure similar to the sequential methods in Manski (1991, 1993), Hotz and Miller (1993), Hotz *et al.* (1994), and Ahn (1995)². We present specification tests and goodness-of-fit tests of the model, and analyse the implications of our parameter estimates. Finally, based on the estimated model, we obtain simulations of state and decision variables and perform several counterfactual experiments in order to evaluate the effect of different parameters on the dynamics of prices and markups.

Our main empirical results are the following. The preliminary analysis shows a very significant and robust effect of inventories at the beginning of the month on current price. The form of this state dependence is the one predicted by the theoretical model. We also obtain that markups follow a cyclical behaviour that matches the cyclical behaviour of inventories. Prices tend to be reduced when orders are placed, and they tend to increase during the period until the next positive order. In the estimation of the structural model we obtain a very significant and sizable estimate of the lump-sum ordering costs parameter, and a significant but much smaller estimate of menu costs parameters. Finally, our counterfactual experiments show that fixed ordering costs play a very important role in the dynamics of retail prices in this dataset. In particular, if there were not fixed ordering costs, the monthly frequency of sales promotions would drop from 50.8% to 26.8%, and the within-brand standard deviation of markups from 7.3% to 3.8%.

The rest of the paper is organized as follows. Section 2 introduces the theoretical model and characterizes the form of the optimal decision rule. Section 3 describes the dataset as well as the characteristics of the supermarket chain and the market where it operates. In Section 4 we present a descriptive analysis of the dynamic behaviour of inventories and prices, and reduced form estimations of the decision rules. Section 5 describes the econometric approach to estimate the structural model and presents estimates of the structural parameters. Section 6 presents several counterfactual experiments based on the estimated model. We summarize our results in Section 7.

2. A MODEL OF PRICE AND INVENTORY DECISIONS WITH LUMP-SUM COSTS

In this section we present a discrete time dynamic programming model where a retailer operating in a monopolistically competitive market decides retail price and orders to suppliers. The retailer takes these decisions at the beginning of every period (*e.g.* week) before

2. See also Heckman (1976), who used a synthetic cohorts method in the context of a dynamic structural model.

a demand shock is known. Although this demand shock becomes observable during the period, the retailer cannot change her price or place new orders until the beginning of next period. Therefore, an excess demand (*i.e. stockout*) may occur during the period. In addition, there are lump-sum costs associated with the decisions of placing orders and changing nominal prices. The model combines the classical (S, s) inventory model (see Scarf (1959), Blinder (1981), among others) and pricing models with menu costs (see Sheshinski and Weiss (1977, 1983, 1992)).

2.1. Decision problem

Consider a risk neutral retailing firm who sells an homogeneous good. At every period t the firm decides retail price and orders to wholesalers to maximize the expected and discounted stream of current and future real profits. The expected one-period real profit is equal to the real value of expected sales, minus ordering costs, storage costs, and price adjustment cost.

$$\pi_t = \exp \{p_t\} y_t^e - \exp \{c_t\} q_t - \alpha s_t - \eta^q I(q_t > 0) - \eta^p I(P_t \neq P_{t-1}), \quad (1)$$

where p_t and P_t are the logarithms of real and nominal retail prices at period t , respectively; y_t^e represents expected sales, in physical units; c_t is the logarithm of real wholesale price; q_t represents orders during period t ; s_t is the stock at the beginning of period t ; α is the unit storage cost; η^q and η^p are fixed ordering costs and fixed price adjustment costs, respectively, measured in real monetary units; and $I(\cdot)$ is the indicator function.

Fixed ordering costs result from transportation costs that are not proportional to the number of items in an order, and from other lump-sum costs associated with the classification and organization of new deliveries. Nominal price adjustment costs are the combination of decision costs and costs of making new lists of prices, new labels, etc. We assume that these costs do not depend on the magnitude of the nominal adjustment and, therefore, they are lump-sum costs or *menu costs*.

Sales are equal to the minimum of inventories and demand

$$y_t = \min \{y_t^*; s_t + q_t\}, \quad (2)$$

where y_t and y_t^* are sales and demand at period t , respectively. The firm is small relative to the size of the market where it operates. This implies that our firm takes into account the prices of its competitors when deciding its own price, but it does not consider the effects of its decisions on the behaviour of other firms because these effects are negligible. The demand for this firm depends on its own real price, on the average real price of the competitors, and on a consumers' demand shock. We consider an isoelastic demand function

$$y_t^* = \exp \{a_t + \gamma_1 p_t + \gamma_2 \tilde{p}_t\}, \quad (3)$$

where \tilde{p}_t is the logarithm of the average price at other firms, and a_t is a consumers' demand shock. We can rewrite the demand function in the following way

$$y_t^* = \exp \{a_t + \omega_t + \gamma_1 m_t + (\gamma_1 + \gamma_2) c_t\}, \quad (4)$$

where $m_t \equiv p_t - c_t$ is the markup in our firm, and ω_t is equal to $\gamma_2(\tilde{p}_t - c_t)$. Assuming that the wholesale price is the same for all the retailing firms in the market, $\tilde{p}_t - c_t$ represents the average markup at the rest of the firms.

Assumption 1. (Monopolistic competition in the product market). The average markup of the competitors does not depend on our firm's current or previous decisions. Therefore, ω_t follows an exogenous stochastic process. We assume that ω_t follows a stationary first order Markov process with transition density function $f_\omega(\omega'; \omega) \equiv pdf(\omega_{t+1} = \omega' | \omega_t = \omega)$.

Assumption 2. (Price-taker in the wholesale market). The firm is a price-taker in its relationship with suppliers. Therefore, real wholesale price does not depend on current or previous firm's decisions, i.e. it follows an exogenous stochastic process. We assume that c_t follows a stationary first order Markov process with transition density function $f_c(c'; c) \equiv pdf(c_{t+1} = c' | c_t = c)$.

Assumption 3. The consumers' demand shock a_t is independent of $(s_t, c_t, \omega_t, P_{t-1})$, and it is independently and identically distributed over time, with a probability density function $f_a(\cdot)$ that is continuous and twice differentiable.

Assumption 4. Aggregate inflation rate is constant over time. Therefore, $P_t = p_t + \rho t$ and $C_t = c_t + \rho t$, where ρ is the aggregate inflation rate, and C_t is the logarithm of nominal wholesale price at period t .

At the beginning of each period the retailer observes the level of inventories, wholesale price, nominal retail price at previous period, and average markup at other firms. Therefore, under Assumptions 1 to 4, the set of state variables at period t is $\{\omega_t, c_t, s_t, P_{t-1}\}$. Hereinafter, we will use as state variable $b_t \equiv P_{t-1} - C_t$, instead of the nominal retail price P_{t-1} . The variable b_t represents the markup at period t if the firm did not make any nominal adjustment in its retail price during this period. Using this notation the indicator for the existence of a nominal adjustment at period t becomes $I(m_t \neq b_t)$, and the vector of state variables is $k_t \equiv (\omega_t, c_t, s_t, b_t)'$. Therefore, the firm's decision problem is

$$\max_{\{m_t; q_t \geq 0\}} E\{\sum_{j=0}^{\infty} \beta^j \pi(k_{t+j}, m_{t+j}, q_{t+j}; \theta_\pi) | k_t, m_t, q_t\}, \quad (5)$$

where $\beta \in (0, 1)$ is the manager's discount factor,³ and $\theta_\pi \equiv (\gamma_1, \gamma_2, \alpha, \eta^q, \eta^p)'$. The one-period profit function $\pi(\cdot)$ can be written as follows

$$\begin{aligned} \pi(k_t, m_t, q_t; \theta_\pi) = & \exp\{c_t + m_t\} y^e(\omega_t, c_t, m_t, s_t + q_t) - \exp\{c_t\} q_t - \alpha s_t \\ & - \eta^q I(q_t > 0) - \eta^p I(m_t \neq b_t). \end{aligned} \quad (6)$$

The function $y^e(\cdot)$ represents expected sales, and under Assumption 3

$$\begin{aligned} y^e(\omega_t, m_t, c_t, s_t + q_t) = & \int_0^{\infty} \min\{y_t^*; s_t + q_t\} f_a(da_t) \\ = & y^{*e}(\omega_t, m_t, c_t) H\left(\frac{s_t + q_t}{y^{*e}(\omega_t, m_t, c_t)}\right), \end{aligned} \quad (7)$$

where $y^{*e}(\omega_t, m_t, c_t) = \exp\{\omega_t + \gamma_1 m_t + (\gamma_1 + \gamma_2) c_t\}$ is the expected demand; and, for any real value x , $H(x) \equiv E(\min\{e^a; x\} | x)$. Under Assumption 3, $y^e(\cdot)$ is increasing in $s + q$,

3. Since the one-period profit function represents real profits, the discount factor β does not include the inflation rate.

decreasing in m , strictly concave with respect to $s + q$ and m , and markup and orders are substitutes in the sales equation, *i.e.*, $\partial^2 y^e / \partial m \partial q = \gamma_1 f_a(\ln[s + q] - \ln[y^{*e}]) < 0$.⁴

To complete the model we should specify the transition rules for the endogenous state variables. The transition for the stock is

$$s_{t+1} = s_t + q_t - y_t = \max \{0; s_t + q_t - y_t^*\}. \quad (8)$$

And using the definition of b_t

$$b_{t+1} = P_t - C_{t+1} = m_t - \rho - (c_{t+1} - c_t). \quad (9)$$

Since time horizon is infinite, there is additive time-separability in preferences, state variables have first order Markov transition probabilities, and $\beta \in (0, 1)$, Blackwell's theorem applies and both the value function and the optimal decision rule are time invariant mappings on the space of state variables. The Bellman's equation of this problem is

$$V(k) = \max_{\{m; q \geq 0\}} \pi(k, m, q) + \beta \int V(\omega', c', m - \rho - [c' - c], s') f_a(da') f_c(dc'; c) f_\omega(d\omega'; \omega), \quad (10)$$

where we have omitted time sub-indexes and used ' to denote variables that are unknown to the manager when taking her decision. Equation (10) shows which are the sources of uncertainty in this decision problem. There is uncertainty about a component of current demand, a' , about future wholesale prices, c' , and about future markups at other firms, ω' . Finally, the decision problem can be characterized by the set of primitives

$$\theta \equiv \{\theta_\pi, \theta_{fc}, \theta_{f\omega}, \theta_{fa}, \beta, \rho\}, \quad (11)$$

where θ_{fc} , $\theta_{f\omega}$, and θ_{fa} are the parameters characterizing the transitional densities of c , ω , and a , respectively.

2.2. Characterization of the optimal decision rules

In this subsection we obtain the form of the optimal decision rules for markup and orders. For the sake of notational simplicity, we consider a version of the previous model where exogenous state variables, *i.e.* c_t and ω_t , are constant over time. However, if $f_c(\cdot)$ and $f_\omega(\cdot)$ are continuous and differentiable, it is straightforward to extend our results to the general case.

We can rewrite the Bellman's equation using the following expression

$$V(b, s) = \max_{\{m; z \geq s\}} \{Q(m, z) - \eta^q I(z > s) - \eta^p I(m \neq b)\} - (\alpha - 1)s, \quad (12)$$

where

$$Q(m, z) = \exp \{m\} y^e(m, z) - z + \beta \int V(m - \rho, \max \{0; z - \exp[a + \gamma_1 m]\}) f_a(da), \quad (13)$$

4. For $s + q/y^{*e}$ larger than a certain value \bar{x} , orders and markup are also substitutes in the profit function (where \bar{x} is implicitly defined by the equation

$$\frac{f_a(\ln \bar{x})}{1 - F_a(\ln \bar{x})} = \frac{-1}{\gamma_1},$$

and $F_a(\cdot)$ is the *cdf* of a).

and $z \equiv s + q$. Equations (12) and (13) define implicitly the value function $V(\cdot)$ as the fixed point of a contraction mapping that maps continuous functions into continuous functions. This property guarantees the existence and continuity of the value function (see Stokey and Lucas (1989), pages 49–55). However, the discontinuity of the one-period profit function implies that the value function is not concave and thus it is not possible to use the usual theorems (e.g. Stokey and Lucas (1989), Theorem 4.8) to prove that the optimal decision rule is a function and not a correspondence. To prove this property, as well as to characterize the optimal decision rule, we follow an approach similar to Scarf (1959) that exploits the properties of K -concave functions (see Appendix 1). We start considering the problem without menu costs.

Lemma 1. *Model without Menu Costs.* Under Assumptions 1 to 4, $\beta \in (0, 1)$, and $\eta^p = 0$, the model defined by equations (12) and (13) has a unique optimal decision rule with the following form:

$$q^*(s) = \begin{cases} z^* - s & \text{if } s < z^L \\ 0 & \text{if } s \geq z^L, \end{cases} \quad (14)$$

and

$$m^*(s) = \begin{cases} m^* & \text{if } s < z^L \\ \tilde{m}(s) & \text{if } s \geq z^L, \end{cases} \quad (15)$$

where m^* and z^* are optimal markup and inventories, respectively, when there are not lump-sum ordering costs and irreversibility restrictions (i.e. $z \geq s$); z^L is a constant lower than z^* ; and $\tilde{m}(\cdot)$ is a continuous and decreasing real function that represents the optimal markup for a given value of inventories.

Proof. See Appendix 1. ||

Lemma 1 shows that the endogeneity of markup does not affect the optimality of an (S, s) inventory rule. When c_t and ω_t are not constant, z^* , m^* , z^L and $\tilde{m}(\cdot)$ depend on c_t and ω_t , and we have an (S_t, s_t) decision rule. This simple model has interesting implications on the dynamic behaviour of markup. Figure 1 presents the path of markup and inventories when ω_t and c_t are constant. The markup increases between two orders (when the stock level decreases) and it drops down when new orders are placed (because $\tilde{m}(z^L) > \tilde{m}(z^*) \equiv m^*$). This markup behaviour results from two characteristics of the model: (1) the existence of a positive probability of stockout (and the negative effect of this probability on the price elasticity of sales), that creates substitutability between markup and inventories in the profit function; and (2) the existence of lump-sum ordering costs, that generates (S, s) inventory behaviour. When the level of inventories decreases between two orders, the probability of stockout increases and expected sales become more inelastic with respect to markup. Thus, the optimal markup increases between two orders, and decreases when the elasticity of sales goes up as the result of positive orders. Figure 1 shows also that the largest markup increases occur just after a positive order, and the increments tend to be smaller when we approach to the next positive order. The reason for this behaviour is that the cyclical path of markups generates a cyclical behaviour in sales. The largest sales and, consequently, the largest stock reductions and markup increases, occur just after a positive order.

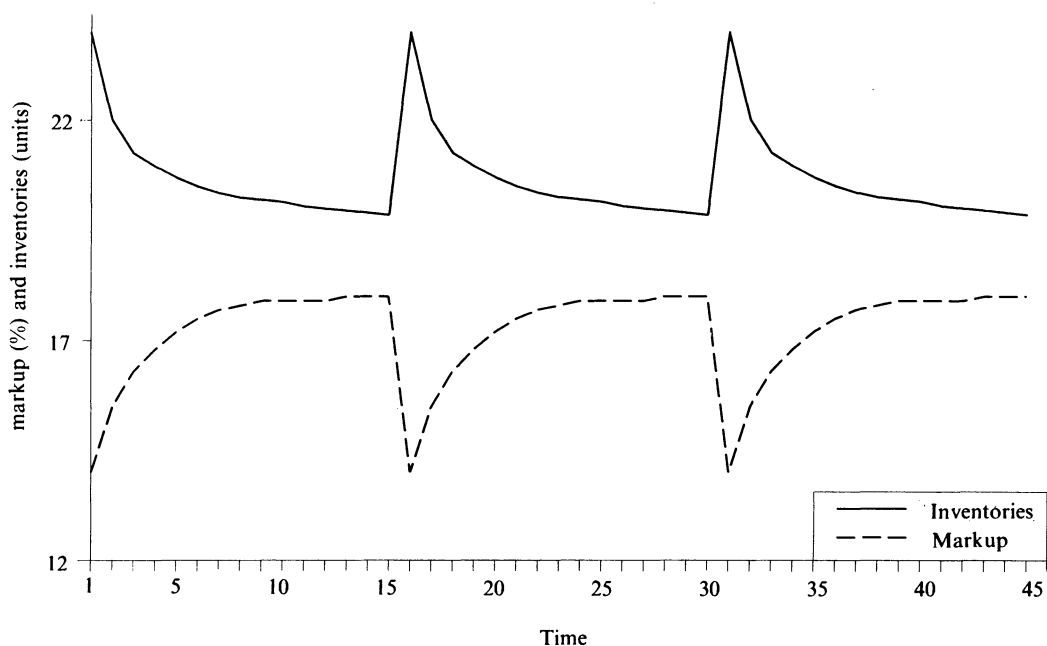


FIGURE 1
Time series of markup and inventories (without menu costs)

In the context of a monopolistically competitive market, the markup behaviour in our model can generate cross-sectional markup dispersion. The magnitude of this markup dispersion will depend on: (1) the magnitude of lump-sum ordering costs (*i.e.* $z^* - z^L$); (2) the sensitivity of the price elasticity of sales to changes in the probability of stockout (*i.e.* $\partial \tilde{m}(s)/\partial s$); and (3) the degree of correlation between the demand shocks at individual firms. The interesting result here is that there may be price (markup) cross-sectional dispersion when there are not menu costs.

However, this version of the model cannot explain infrequency in nominal price changes. In Lemma 2 we consider the model with both menu costs and lump-sum ordering costs.

Lemma 2. Model with Menu Costs. *Under Assumptions 1 to 4, and $\beta \in (0, 1)$, the model defined by equations (12) and (13) has a unique optimal decision rule with the following form:*

$$q^*(b, s) = \begin{cases} z^* - s & \text{if } (b, s) \in T_{11} \\ \tilde{z}(b) - s & \text{if } (b, s) \in T_{01} \\ 0 & \text{if } (b, s) \in T_{00} \cup T_{10}, \end{cases} \quad (16)$$

and

$$m^*(b, s) = \begin{cases} m^* & \text{if } (b, s) \in T_{11} \\ \tilde{m}(s) & \text{if } (b, s) \in T_{10} \\ b & \text{if } (b, s) \in T_{00} \cup T_{01}, \end{cases} \quad (17)$$

where m^* and z^* are the optimal markup and inventories, respectively, when there are not lump-sum costs and irreversibility restrictions (*i.e.* $z \geq s$); $\tilde{m}(\cdot)$ is a continuous and decreasing

real function that represents the optimal markup for a given value of inventories; $\bar{z}(\cdot)$ is a continuous and decreasing real function that represents the optimal inventories for a given value of the markup; and T_{00} , T_{10} , T_{01} , and T_{11} are four sets defining a partition of \mathbb{R}^2 . These sets are associated with the four discrete alternatives: “no price change and zero orders” (T_{00}); “no price change and positive orders” (T_{01}); “price change and zero orders” (T_{10}); and “price change and positive orders” (T_{11}).

Proof. See Appendix 1. ||

This optimal decision rule has clear similarities with the one obtained by Sheshinski and Weiss (1992) in the context of a menu costs pricing model with two goods. The decision rule defines a partition of the space (b, s) . Figure 2 presents this partition and the functions $\bar{m}(s)$ and $\bar{z}(b)$. The “inaction” region, T_{00} , consists of the set of points (b, s) for which the stock is relatively large and the vertical distance between markup at the beginning of the period and $\bar{m}(s)$ is not too large. The vertical distance between $\bar{m}(s)$ and the boundary of T_{00} depends on the magnitude of η^p , and the horizontal distance between $\bar{z}(b)$ and the boundary of T_{00} depends on η^q . The other three regions have also intuitive interpretations. In Figure 2 we have considered the case with ordering costs larger than menu costs. Figure 3 and points *A* to *F* in Figure 2 represent the dynamics of markup and inventories when the inflation rate is zero. Again, the markup increases between two orders and it drops down when orders are placed. But now there is also infrequency in nominal price adjustments.

When the inflation rate is positive the lines connecting points *A* and *B*, *C* and *D*, and *E* and *F* are not horizontal (i.e. *B* is below *A*, *D* below *C*, and *F* below *E*). In that case, it is possible that a positive order does not imply a nominal price reduction (i.e. point *F* might be below point *A*). The existence or not of nominal price reductions associated with positive orders depends on: (1) the magnitude of the inflation rate; (2) the relative magnitude of lump-sum ordering costs and menu costs; and (3) the slope of $\bar{m}(s)$. However, for any inflation rate, if ordering costs are large enough (i.e. the boundary of T_{00} is enough to the left) or if menu costs are small enough, there may be sizable nominal price

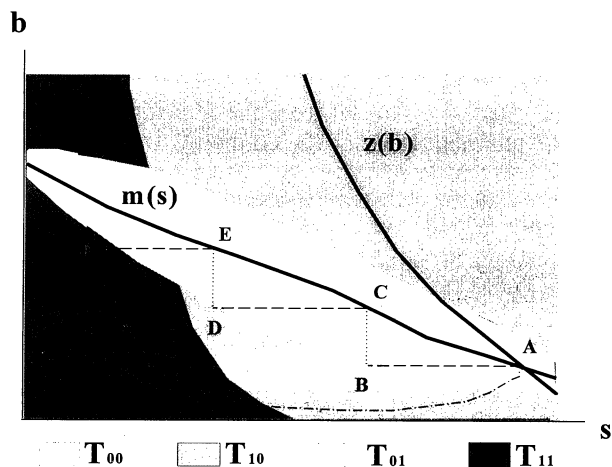


FIGURE 2
Optimal decision rule (model with menu costs)

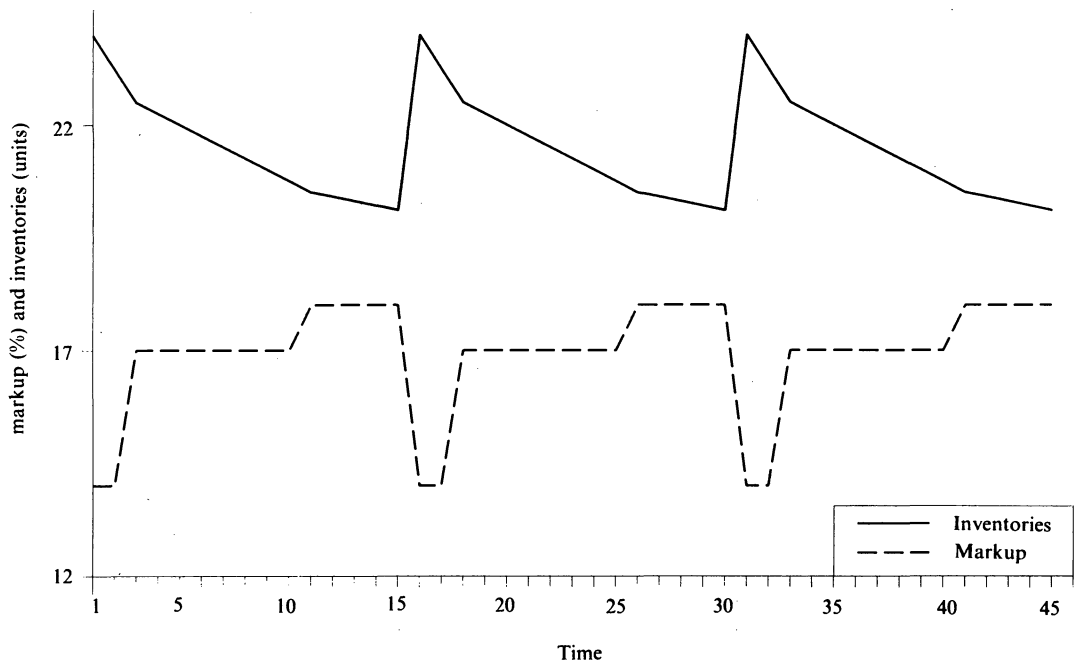


FIGURE 3

Time series of markup and inventories (with menu costs)

reductions when orders are placed. Notice that the larger the slope of $\tilde{m}(s)$ (in absolute value), the larger this price reduction.⁵

This model provides a potential explanation for the coexistence of long periods without nominal price adjustments (or infrequent price increases) and short periods with very low prices and markups. It can also explain the existence of positive duration dependence in sales promotions (see Pesendorfer (1996)). The model has strong empirical predictions on the joint dynamics of prices, inventories, orders and sales. The rest of this paper represents empirical evidence regarding this model.

3. THE FIRM AND THE DATASET

Our empirical analysis is based on a dataset from the central store of a supermarket chain. This section describes the characteristics of the company and the dataset.

3.1. *The supermarket chain*

Our company is a supermarket chain that operates in the Basque region of Spain. Between January 1990 and May 1992 (our sample period) the supermarket industry in this region has been characterized by the existence of a leader who captures more than 40% of the market. Our firm's market-share has been between 3% and 4%, the sixth in the market. During this period there have been no significant changes in the market shares of the top-ten supermarket chains in the region. The Spanish monthly inflation rate during this period was between 0.5% and 0.6% (Retail Price Index).

5. The slope of $\tilde{m}(s)$ depends on the sensitivity of the price elasticity of sales with respect to the stockout probability.

The company has a central store (headquarters of the firm) and over 60 outlets. Most of the company's decision making is centralized. The central store orders new deliveries from suppliers, stores the goods, sends to the outlets the orders that they make, and decides upon retail prices. The manager of an individual outlet sends her weekly orders to the central store. Since 1989 most of the communication between outlets and central store has taken place through a computer network. At the beginning of each week the central store sends to every outlet a list with the available brands and their retail prices (*i.e.* prices are decided weekly). Outlets place their orders to the central store through their terminals and, if the brands are available, orders are delivered to the outlets within twenty-four hours. The storage capacity in most of the outlets is small, and outlets usually place orders every week (*i.e.* one order per week in most of the cases).

In this paper we study the decision problem of the central store, that is, its decision about retail prices and orders to suppliers. When taking these decisions, managers at the central store do not have to know the demand functions for every brand at every specific outlet, but only the total demand of every brand at all the outlets in the supermarket chain. Indeed, we have seen how this is the information that these managers use to predict future sales and to make decisions about prices and inventories. Therefore, we abstract from the fact that the company sells through the outlets and we assume that its decision problem is equivalent to the problem of a retailer who sells directly to consumers.⁶

Every year (around June) the firm negotiates with each supplier an agreement that determines the conditions that will be applied during the next twelve months. These conditions include wholesale price, discounts, trade promotions, possibility of returns, and form of payment. Although the final contract specifies a wholesale price, the supplier maintains the right to change it during the year if there are variations in manufacturing costs or in competition in the wholesale market. According to the purchasing managers of the company, wholesale prices are generally the same for all the supermarket chains in the region and, in most of the cases, payment is made 60 or 90 days after delivery. The toughest part of the negotiation with suppliers is about the number of weeks during the year that the brand will be under promotion, and about the percentage of the cost of sales promotions that will be paid by the wholesaler (*e.g.* cost of posters, mailing, price labels).⁷

In our model the firm sells an individual good. However, the company in our dataset sells thousands of brands. Price and inventory decisions of different brands may not be separable. Although the synchronization/staggering of price and inventory decisions among different brands in a multiproduct firm is an interesting issue, we have not addressed this problem in this paper. Therefore, our econometric approach will be based on the assumption of separability between the decisions of the different brands. We consider the multiproduct case a topic for future research.

3.2. *The dataset*

The firm has maintained a database since January 1990. This database contains monthly information on prices, sales, orders to suppliers and inventories, for every brand sold by

6. We are assuming that the price and inventory decisions in the central store have very small effects on outlets' storage costs and on the costs of delivering items from central store to the outlets. Under these assumptions the relevant profit function that the central store tries to maximize includes sales, storage costs at the central store, menu costs, and costs of orders to suppliers, but not outlets' storage and ordering costs.

7. In our specification of the structural model we take into account that, from the point of view of the retailer, the menu costs associated with sales promotions can be smaller than the menu costs of changing the regular price.

the company.⁸ We use this information from January 1990 to May 1992 (29 months). During this period 8742 brands were sold by the supermarket chain. The raw dataset contains the following information for every brand and month: name and description, total sales of the brand in all the company's outlets, sales under promotion, orders from the central store to suppliers, stock in the central store at the beginning of the month, wholesale price, retail price under sales promotion, and *regular* retail price (*i.e.* not under sales promotion). All prices are measured in pesetas per item, and quantities are measured in number of items.

3.2.1. *The working sample.* Our working sample is a balanced panel of 534 brands. A brand has been included in our working sample if: (1) the central store keeps inventories of the brand; and (2) the brand has been sold at every month between January 1990 and May 1992. We discuss here the reasons why we use these two filters.

There are some brands for which the central store does not keep inventories. Some of them are very perishable goods which are delivered daily from wholesalers to outlets (*e.g.* fresh vegetables, fish, some types of bread, *etc.*) In other cases they are brands from manufacturers with efficient distribution networks that allow them to deliver their brands to individual outlets. From the point of view of the company's central store, there is not any inventory problem associated with these brands. Since we are interested in the relationship between price and inventory decisions we only consider those brands for which the central store keeps inventories. For the whole set of 8742 brands, there are 4966 brands for which the central store maintains inventories.

Some brands have been sold by the supermarket chain for the whole sample period (29 months), but most of them have been sold for shorter periods (even just one month). Therefore, infrequency of positive orders for certain brands is not associated with an inventory decision, but it is the result of the company's decision that the brand will not be sold in the future. If we considered these brands in our working sample we would introduce a spurious upward bias in the frequency of zero orders and thus an upward bias in our estimate of lump-sum ordering costs. In order to avoid this problem we concentrate our analysis on the set of brands that were sold by the firm at every month between January 1990 and May 1992. These are 534 brands of the 4966 brands for which the firm maintains inventories, which account for 15,486 brand-month observations.

The firm uses a two-level criterion to classify brands in its dataset. Individual brands are classified in groups that we call *products*, and these products are classified in seven large groups. Table 1 presents this classification as well as the number of brands of each product and of each large group in our working sample.

3.2.2. *Price data.* The two price variables in our dataset, regular retail price and price under promotion, are monthly averages (*i.e.* nominal value of sales divided by sales in physical units). In principle, the time average nature of this price information could lead to identification problems in our empirical analysis. However, the high nominal rigidity in regular retail prices implies that these monthly averages contain, indeed, very much information about actual nominal prices and nominal price changes.

We use three retail price variables in our empirical analysis: (1) retail price during the month, P , that we use to obtain the markup m ; (2) retail price at the beginning of the month, $P^{[b]}$, that we use to obtain the state variable b ; and (3) a categorical variable that indicates the existence of "no change", "increase" or "reduction" in nominal retail price

8. A brand is an individual commodity with specific label, size, taste, colour, etc.

TABLE 1

Classification of brands in products and groups of products

Names of product and groups	Number of brands	Names of products and groups	Number of brands	Names of products and groups	Number of brands
1. Health, beauty	48	5. Soft drinks	29	6. Food (cont.)	
101. Eau cologne	2	501. Fruit juices	10	626. Pets food	15
102. Deodorant	2	502. Water	3	627. Salt	3
103. Soap, Shampoo	10	503. Carb. drinks 2L	5	628. Mayonnaise	9
104. Hair spray	6	504. Carb. dks small	11	629. Vinegar	2
105. Beauty creams	9	6. Food	277	630. Eggs	3
106. Toothpaste	3	601. Sugar	4	631. Butter, margar.	6
107. Shaving creams	3	602. Coffee	12	632. Cheese	11
108. Shaving blades	8	603. Instant coffee	10	633. Milk	6
109. Bath sponges	5	604. Malt	6	634. Powder Milk	5
2. Cleaning	69	605. Tea	2	635. Sausages	3
201. Detergent	12	606. Olive oil	7	636. Foie-gras	2
202. Bleach	12	607. Sunflower oil	4	7. Alcohol, drinks	89
203. Softeners	7	608. Canned fish	11	701. Cognac	10
204. Dishwashers	5	609. Dry cod	3	702. Whisky	4
205. Cleaning sprays	15	610. Canned veget.	8	703. Other spirits	12
206. Floor polish	7	611. Marmalade	9	704. Sherry	5
207. Scourers	4	612. Beans, Lentils	9	705. Champagne	6
208. Dischcloths	2	613. Rice	4	706. Vermouth	4
209. Brooms	1	614. Olives	6	707. Sidra	3
210. Insecticides	3	615. Pickles	4	708. Beer	12
211. Kitchen paper	1	616. Biscuits	23	709. wine (regular)	15
3. Pharmacy	9	617. Cakes	5	710. Wine (quality)	18
301. Bandages	4	618. Candies	11		
302. Sanit. napkins	2	619. Chocolate	30		
303. Diapers	3	620. Toasted bread	6		
Others, no food	13	621. Flour	2		
401. Alumin. paper	6	622. Pasta	14		
402. Plastic objects	2	623. Instant soups	11		
403. Batteries	4	624. Instant rice	5		
404. Stick gum	1	625. Dried fruits	6		

during the month, $I^{\Delta P}$. Below we explain how we have used monthly average prices to construct these variables, and we make explicit certain assumptions.

Markups, $m_{it} = \ln(P_{it}) - \ln(C_{it})$, have been obtained using the average retail price during the month for P_{it} (i and t are the subindexes for brand and month, respectively). To construct $P_{it}^{[b]}$ we assume that sales promotions occur at the middle of the month. Therefore, at the beginning of the month the retail price is always equal to the regular retail price. Regular retail prices stay constant during relatively long periods. For 41.5% of the observations the monthly change in average regular price is zero, *i.e.* $\Delta ARP_{it} = ARP_{it} - ARP_{i,t-1} = 0$, where ARP_{it} is the monthly average regular price. It is clear that if $\Delta ARP_{it} = 0$ the actual regular price has been constant during months $t-1$ and t and, therefore, we have that: $P_{i,t-1}^{[b]} = P_{it}^{[b]} = P_{i,t+1}^{[b]} = ARP_{it}$. Using this information we identify the actual regular price at the beginning of the month for 67.7% of the observations in our sample (for 42% of the brands we observe the actual regular price for more than 21 of the 28 months). For the remaining 32.3% of the observations, we consider $P_{it}^{[b]} = ARP_{i,t-1}$.

To construct the categorical variable $I^{\Delta P}$ we use both ARP and the indicator for the existence of sales promotions during a month. We start constructing a categorical variable for the sign of the change in regular price, *i.e.* $I_{it}^{\Delta P \text{ Reg}} \in \{-1, 0, 1\}$. It is clear that, without

any assumption, $\Delta ARP_{it} = 0$ implies that $I_{it}^{APReg} = I_{i,t-1}^{APReg} = 0$. This accounts for 57.7% of the observations in the sample. According to company's managers, regular retail prices are rarely changed more than once during a month. This is consistent with the evidence we have mentioned above (e.g. more than 30% of the regular price quotations last more than two months). Therefore, we assume that regular retail prices change, at most, once per month. Exploiting this assumption we can identify I_{it}^{APReg} for an additional 26.0% of the observations in the sample. For instance, if $\Delta ARP_{it} = 0$ we know that $I_{i,t-2}^{APReg} = \text{sign}(\Delta ARP_{i,t-1})$, and $I_{i,t+1}^{APReg} = \text{sign}(\Delta ARP_{i,t+1})$. Figure 4 presents an example where we can identify I^{APReg} at every month except one (i.e. month number 4). For the remaining 16.3% of observations for which I^{APReg} is not identified we consider that $I_{it}^{APReg} = \text{sign}(\Delta ARP_{it})$.

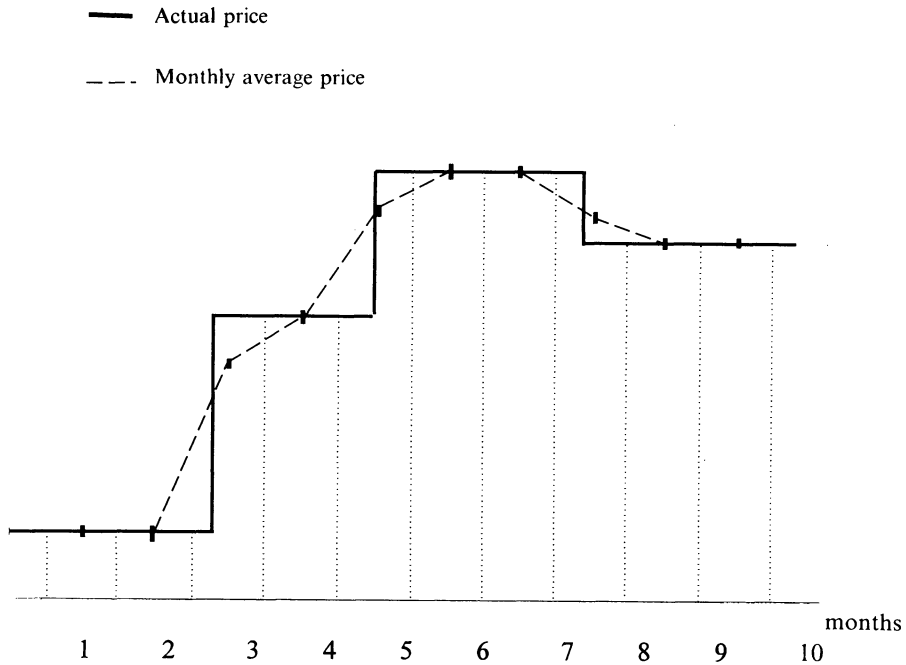


FIGURE 4

Actual price and monthly average price

We combine I^{APReg} and the indicator of sales promotions during a month, I_{it}^{SP} , to obtain a categorical variable that represents the behaviour of the nominal retail price within a month $I_{it}^{\Delta P} \in \{-1, 0, 1\}$. We consider that: $I_{it}^{\Delta P} = 1$ (price increase) if there is an increase in the regular retail price (i.e. $I_{it}^{APReg} = 1$); $I_{it}^{\Delta P} = 0$ if there is no change in regular retail price and no sales promotions (i.e. $I_{it}^{APReg} = 0$ and $I_{it}^{SP} = 0$); and $I_{it}^{\Delta P} = -1$ (price reduction) if there is a reduction in the regular retail price or there is a sales promotion but without a regular price increase (i.e. $I_{it}^{APReg} = -1$ or $\{I_{it}^{APReg} \neq 1 \text{ and } I_{it}^{SP} = 1\}$).⁹

4. PRELIMINARY EVIDENCE

Figure 5 presents the time series for average regular retail price and average retail price (left column) and orders and inventories (right column) for three brands in our sample.

9. For 6.1% of the observations both a sales promotion and an increase in regular price occur during the same month. The criterion described before implies that we have considered $I^{\Delta P} = 1$ for these observations.

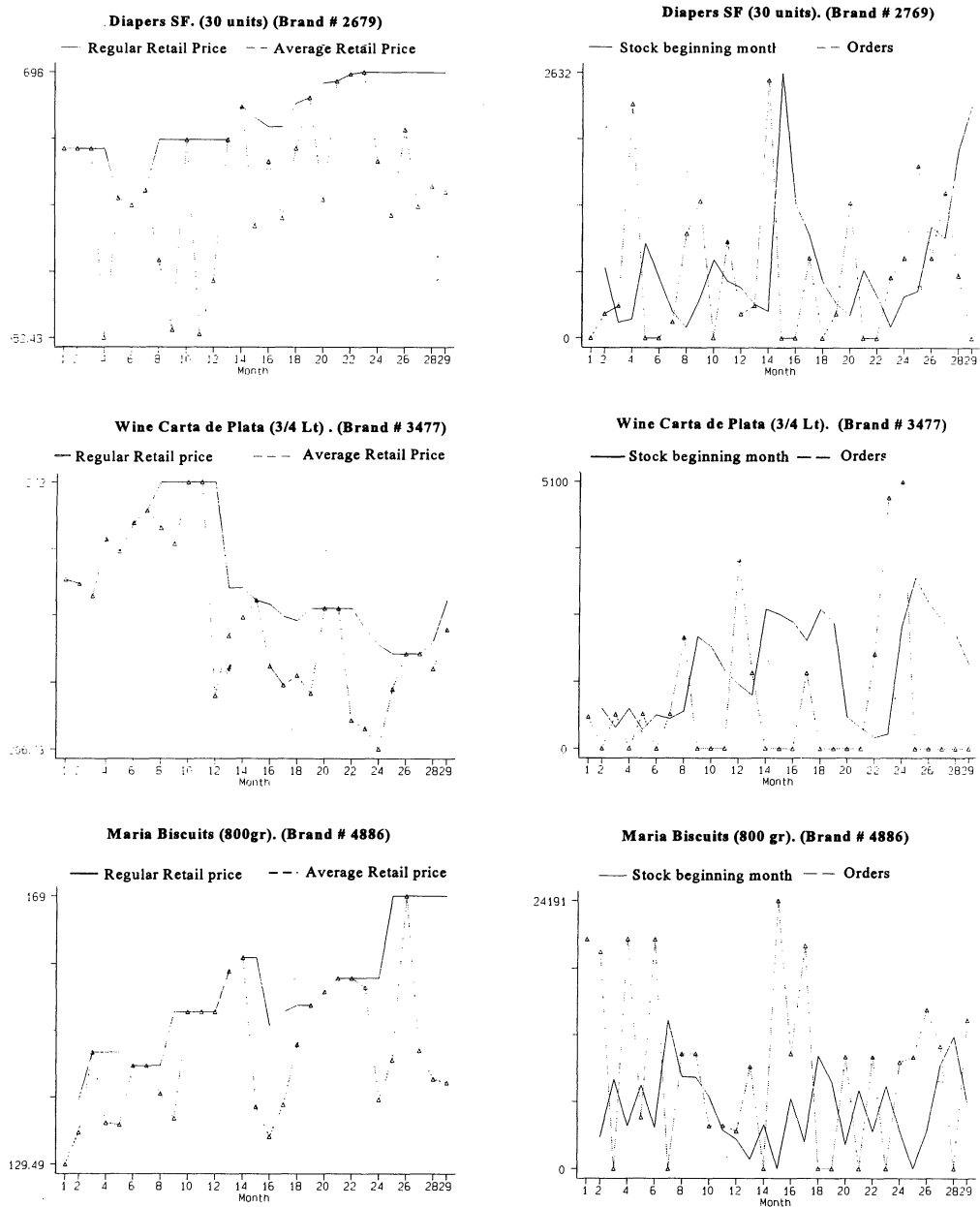


FIGURE 5
Time series for prices, inventories, and orders

This figure anticipates several results that will be discussed below using the whole sample. First, changes in regular retail prices occur very infrequently. Most nominal price changes are associated with sales promotions. Second, orders are infrequently placed and their volume tends to be larger than sales during one month (or even several months). Third, the largest reductions (and lowest levels) in average retail price tend to be associated with

large orders (*e.g.* months 4, 8, 9, 11, 20 and 25 for brand 2769; months 12, 22, 23, and 24 for brand 3477; and months 4, 15, 16, 17 and 24 for brand 4886). Of course, this is not a strong evidence in favour of the model presented in Section 2. The synchronization of positive orders and low prices could be explained by low prices leading to large sales, large stock reductions and, consequently, positive orders. In other words, there is a simultaneity problem and for that reason our empirical analysis will not place emphasis on the correlation between positive orders and price reductions. Instead, we will analyse how the state variables (*i.e.* predetermined variables) affect price and inventory decisions, and whether these effects are consistent with the predictions of our model.

TABLE 2

Descriptive statistics: Distribution of brand-specific averages 534 brands. January 1990–May 1992 (29 months)

		Mean	Pctile. 10	Pctile. 25	Median	Pctile. 75	Pctile. 90
Orders	(1) Frequency $q > 0$ (%)	76.60	48.28	65.52	79.31	89.66	96.55
	(2) Duration (in months)	1.38	1.04	1.08	1.22	1.44	2.00
Markups	(3) Regular retail price (%)	20.50	9.46	15.69	19.95	24.77	30.31
	(4) Price under promotion (%)	5.10	-0.31	1.82	4.36	7.65	11.43
	(5) Retail price (%)	16.68	5.70	12.12	16.65	21.76	27.14
Regular Retail Price	(6) Freq. price changes (%)	46.60	17.86	28.57	46.43	60.71	78.57
	(7) Duration (in months)	2.18	1.18	1.44	1.80	2.56	3.60
Wholesale Price	(8) Freq. price changes (%)	34.63	10.71	21.43	34.14	42.86	50.00
	(9) Duration (in months)	3.46	1.82	2.10	2.86	3.80	6.00
Sales Promotions	(10) Freq. prom. > 0 (%)	42.85	6.90	17.24	37.93	68.97	86.21
	(11) Ratio prom./total sales (%)	21.90	1.59	7.43	18.36	33.57	47.40
Retail Price	(12) Freq. price changes (%)	68.95	35.71	53.57	71.43	85.71	96.43
	(13) Duration (in months)	1.58	1.04	1.12	1.35	1.71	2.27

^a All the statistics in this table have been obtained using the 534 brand-specific means of the variables. Therefore, these statistics refer to the between-brands distribution of the variables.

^b Row 2 presents the average duration between two positive orders.

^c Rows 7, 9, and 13 present the average durations between two nominal price changes.

^d In order to avoid right-censoring bias due to incomplete duration spells, the average durations have been obtained without including those spells starting during the last 5 months.

^e The ratio in row 11 is equal to "sales under promotions (in phys. units)/total sales (in phys. units)".

Table 2 presents descriptive statistics for the decision variables. Rows 1 and 2 show the distributions of the ordering frequency and the duration between two orders, respectively. On average orders are placed every 1.38 months (5.5 weeks). Rows 3 to 5 refer to markups. Markups under promotions tend to be very small, and there is a sizable differential between these markups and the ones associated with regular retail price. In rows 6 and 7 we present the frequency of changes in regular retail prices and the average duration of a regular price quotation. Regular prices present important stickiness. For 50% of the brands the average duration of a price quotation is larger than 1.80 months (7.5 weeks). However, the stickiness of regular retail prices is significantly smaller than the one of wholesale prices (rows 8 and 9). Rows 10 and 11 present the frequency of sales promotions and the ratio between sales under promotion and total sales (in physical units). Sales promotions are very frequent and they account for 22% of total sales.¹⁰ The high frequency of sales promotions implies that the nominal rigidity of retail prices (rows 12 and 13) is significantly smaller than the one of regular retail prices.

10. Notice that the proportion of sales under promotion is lower than half the proportion of months with a promotion. This is consistent with the fact that, according to company's managers, a brand is maintained under promotion for no more than seven consecutive days and usually no more than once per month.

Our econometric approach to estimate the structural model requires the stationarity of the observable state variables, *i.e.* stock and markup at the beginning of the month, and real wholesale price. We have obtained Dickey–Fuller tests of stationarity for each of these variables and for every individual brand. These tests show that non-stationarity seems not to be a problem in this dataset. Stationarity is very clearly accepted (not rejected) for markups, inventories and orders of all the 534 brands in the sample. However, for 61 brands we have found evidence of non-stationarity in real wholesale prices. For these brands we have considered two alternative measures of real wholesale prices: the actual real wholesale price, and the deviation of real wholesale price with respect to its brand-specific time trend. All the estimations in this paper have been performed using both measures of real wholesale prices, obtaining negligible and statistically not significant differences between them.

TABLE 3

Fixed-Effects Probit Models for Discrete Choices (Standard errors in parentheses)

	Orders $q > 0$		Regular retail price				Sales promotions $I^{SP} = 1$	
	(1)	(2)	$I^{\Delta P \text{ Reg}} = -1$		$I^{\Delta P \text{ Reg}} = +1$		(7)	(8)
$\ln(s)$	-0.911 (0.020)	-0.927 (0.020)	0.026 (0.009)	0.018 (0.009)	-0.044 (0.008)	-0.044 (0.008)	0.109 (0.009)	0.088 (0.009)
b	-0.059 (0.243)	0.028 (0.244)	10.585 (0.275)	10.725 (0.276)	-11.728 (0.269)	-11.723 (0.270)	2.309 (0.213)	2.839 (0.219)
c	-0.183 (0.228)	-0.059 (0.232)	0.668 (0.206)	0.840 (0.208)	-0.942 (0.199)	-0.934 (0.201)	0.390 (0.206)	0.377 (0.212)
$\ln y(t-1)$		0.084 (0.024)		0.117 (0.022)		0.006 (0.021)		0.384 (0.021)
Log-lik.	-4,993	-4,985	-6,463	-6,448	-7,336	-7,336	-6,668	-6,498
Pseudo R^2 ^a	0.366	0.367	0.232	0.234	0.229	0.229	0.303	0.321
Number obs. ^b	14,056	14,056	14,868	14,868	14,952	14,952	13,888	13,888

^a $Pseudo - R^2 = 1 - RSS/TSS$, where $RSS = \sum_{i=1}^N \sum_{t=1}^T (D_{it} - \hat{D}_{it})^2$; $TSS = \sum_{i=1}^N \sum_{t=1}^T (D_{it} - \bar{D})^2$; D_{it} is the dummy dependent variable; and \hat{D}_{it} is the predicted index, between 0 and 1.

^b In these estimations we have not used those brands without enough variability in the dependent variable, *i.e.* with histories of the dependent variable that contain only 1 zero or only 1 one. This explains the different number of observations for each of the estimations.

Table 3 presents reduced form estimations of the discrete decisions for orders and prices. The explanatory variables in these models are the logarithms of the observable state variables; $\ln(s)$, c , and b . All the estimations include brand-specific fixed-effects to control for spurious state dependence due to unobserved individual heterogeneity. We have also included monthly dummies to control for “aggregate” shocks (*i.e.* that have the same effect on the decisions of all the brands in this company). Since the number of observations per brand is relatively large (*i.e.* 28), we expect the well-known bias of the fixed-effects estimator in dynamic panel data models to be small. For each binary choice we have estimated two probit models, depending on whether we include or not the logarithm of sales at previous month, $\ln(y_{i,t-1})$, as explanatory variable. This variable has been included to control for potential spurious state dependence due to forms of autocorrelation in unobservables not captured by brand fixed-effects (*e.g.* autocorrelation in unobservable demand shocks a_t and ω_t), as well as to test whether this autocorrelation is significant.

Columns 1 and 2 are consistent with an (S_t, s_t) inventory rule where changes in the thresholds are mainly associated with changes in expected demand. The stock at the beginning of the month is clearly the variable with the largest and most significant effect. Real

wholesale price has the expected negative effect on the probability of positive orders, but the effect is not significant. The effect of the markup at the beginning of the month is also not significant. The positive and significant effect of previous month sales (column 2) is consistent with positive autocorrelation in demand shocks. Nevertheless, including $\ln(y_{t-1})$ does not produce significant changes in the rest of the estimates, and the improvement in the goodness-of-fit of the model is negligible.

Columns 3 to 6 present the estimation of probit models for nominal adjustments in regular retail price. The markup at the beginning of the month is the state variable with larger explanatory power and it has the expected effect, *i.e.* the larger the markup the lower (higher) the probability of a positive (negative) price adjustment. The stock level has a significant effect on the probability of both types of nominal price adjustments. This effect is consistent with the predictions of our model. When the stock is large the firm tends to reduce the nominal retail price, but when the stock starts to decrease price increases become more likely. The signs of the estimates in the probits for sales promotions (columns 7 and 8) are the same as the ones in the probits for reduction in regular retail price. However, inventories have a much stronger effect on the probability of promotions.¹¹ Therefore, it seems that the type of retail price dynamics that result from inventory cycles is mainly associated with sales promotions. When we include $\ln(y_{t-1})$ the estimate of the coefficient associated with inventories decreases, both in the probit for promotions and in the probit for regular price reductions. However, the change in the parameter estimates is small and the qualitative results remain unchanged.

Given the markup at the beginning of the month, real wholesale price has a positive effect on the probabilities of promotions and regular price reductions, and a negative effect on the probability of regular price increases. However, notice that taking into account that $b_{it} = p_{i,t-1} - c_{it}$, the estimates in Table 3 show that real wholesale price has a positive (negative) effect on the probability of price increases (reductions).

Table 4 presents evidence about the behaviour of markups between two orders. For each brand we define an "ordering spell" as the set of observations between two positive orders. Let $r(i)$ be the r -th ordering spell for brand i , and let $\tau_{r(i)}$ denote the duration (in months) of spell $r(i)$. The subindex j denotes the j -th month in a certain spell, where $j = 0$ is the initial month (when the initial order was placed) and $j = \tau_{r(i)}$ is the last month (when a new order was placed). Let $m_{r(i),j}$ be the markup at the j -th month in the r -th ordering spell of brand i , and define

$$\beta_{\tau,j} \equiv E[m_{r(i),j} - m_{r(i),0} | \tau_{r(i)} = \tau]. \quad (18)$$

That is, $\beta_{\tau,j}$ is the average difference between the markup j periods after an order and the markup when that order was placed, given that the duration spell is τ . According to our model, if lump-sum ordering costs are large enough, we expect that for any value of τ : (1) $\beta_{\tau,1} < \beta_{\tau,2} < \dots < \beta_{\tau,\tau-1}$, because the markup increases between two orders; and (2) $(\beta_{\tau,2} - \beta_{\tau,1}) > (\beta_{\tau,3} - \beta_{\tau,2}) > \dots > (\beta_{\tau,\tau-1} - \beta_{\tau,\tau-2})$, because the increments in markup tend to be smaller when we approach to the next order.

Table 4 presents OLS estimates of these average markup differentials. The predictions of our model are clearly supported by these estimates. One month after an order has been placed (and if there has not been a new order) the markup is around 3 percentage points

11. The ratio between the parameter estimate associated to $\ln(s)$ and the parameter estimate associated to b is 0.002 (s.e. = 0.004) in the probit for regular price reductions, and 0.052 (s.e. = 0.004) in the probit for promotions.

TABLE 4
Markup behaviour between two orders

	Estimate of $\beta_{\tau,j}$ (in % points)	s.e.
$\tau = 1$		
After 1 month ($j = 1$)	-0.150	0.100
$\tau = 2$		
After 1 month ($j = 1$)	3.209	0.268
After 2 months ($j = 2$)	0.781	0.268
$\tau = 3$		
After 1 month ($j = 1$)	2.437	0.497
After 2 months ($j = 2$)	3.304	0.497
After 3 months ($j = 3$)	0.313	0.497
$\tau = 4$		
After 1 month ($j = 1$)	3.604	0.782
After 2 months ($j = 2$)	4.576	0.782
After 3 months ($j = 3$)	4.467	0.782
After 4 months ($j = 4$)	1.561	0.782

larger than in the previous month. The increases at the following months are less important but also significant. For $\tau = 4$, the average differential after 3 months is 4.47 percentage points (s.e. = 0.78).

How important are these markup differentials in terms of the within-brand variability of markups? The average within-brand standard deviation of markups is 5.61 percentage points (the median is 6.49 percentage points). Therefore, the variability of markups between two orders represent a very important proportion of the within-brand variability in markups.

In this section we have presented preliminary evidence about some of the predictions of our model. We have estimated significant effects of inventories on price decisions. These effects are particularly important for the decision about sales promotions. We have also shown that markups tend to have a cyclical path that matches the (S, s) behaviour of inventories, decreasing when orders are placed and increasing between two orders. Furthermore, this source of markup variability could account for a significant proportion of the time variability of markups. Our next step is to estimate the structural model and use the estimated model to analyse the contribution of several parameters to the variability of prices and markups and to the frequency of sales promotions.

5. ESTIMATION OF THE STRUCTURAL MODEL

The decision problem can be characterized in terms of three sets of parameters: (1) the vector of demand elasticities γ ; (2) the parameters in the transition probabilities of the state variables, θ_j ; and (3) storage costs and lump-sum costs, *i.e.* α , η^q and η^p . The first two sets of parameters can be estimated from the sales equation and the transition rules, respectively. However, the identification of storage costs and lump-sum adjustment costs requires one to exploit the optimal decision rule.

The decision rule in equations (16) and (17) is the combination of marginal conditions of optimality and optimal discrete choices. In this paper we obtain estimates of the structural parameters which exploit moment conditions associated with the optimal discrete

choice, but not moment conditions associated with the marginal conditions of optimality (*i.e.* Euler equations).¹²

It is also important to notice that lump-sum adjustment costs do not appear explicitly in the Euler equations. The identification of these parameters requires one to exploit moment conditions associated with the optimal discrete choice.

5.1. Optimal discrete choice

Consider the characterization of the optimal decision rule in equations (16) and (17). We can distinguish four regimes which result from combining the two discrete alternatives for orders and the two alternatives for nominal price change. When menu costs are asymmetric the number of regimes or discrete alternatives is six, *i.e.* the combination of two alternatives for orders and three alternatives for price change. Let $d \in \{1, 2, \dots, 6\}$ be the index for this discrete choice. We can represent the optimal decision rule in equations (16) and (17) using the following compact expression

$$\{m^*(k; \theta), q^*(k; \theta)\} = \sum_{d=1}^6 I(d^*[k; \theta] = d) \{\tilde{m}^d(k; \theta), \tilde{q}^d(k; \theta)\}, \quad (19)$$

where $d^*[k; \theta]$ is the optimal discrete choice, that is implicitly defined by a partition of the state space, such as the T sets in expressions (16) and (17); and $\{\tilde{m}^d(k; \theta), \tilde{q}^d(k; \theta)\}$ is the optimal decision for markup and orders conditional on the vector of state variables being in the sub-space associated with discrete alternative d .

Let $V^d(k; \theta)$ be the value function conditional on the hypothetical choice of discrete alternative d . By definition, $V(k; \theta) = \max_{\{d\}} \{V^d(k; \theta)\}$, and therefore

$$V^d(k; \theta) = \pi(k, \tilde{m}^d[k; \theta], \tilde{q}^d[k; \theta]; \theta_\pi) + \beta E \left(\max_{j \in \{1, \dots, 6\}} \{V^j(k'; \theta)\} | k, d; \theta \right). \quad (20)$$

Using the previous definitions the optimal discrete choice can be represented using the following expression

$$d^*(k; \theta) = d \Leftrightarrow d = \arg \max_{j \in \{1, 2, \dots, 6\}} \{V^j(k; \theta)\}. \quad (21)$$

Let x be our vector of observable state variables, *i.e.* $x = (c, s, b)'$. We do not observe ω , that is, we do not have certain information about demand shocks that is available to the firm when taking its decision. Given equations (20) and (21) and an assumption about the probability distribution of ω conditional on x , the econometric model will be completely defined.

Let $E\pi^d(x; \theta)$ be the function that represents the expected value of the one-period profit conditional on x and on the hypothetical choice of discrete alternative d .

$$E\pi^d(x; \theta) \equiv E(\pi(k, \tilde{m}^d[k; \theta], \tilde{q}^d[k; \theta]; \theta_\pi) | x, d; \theta). \quad (22)$$

Using this definition we can write

$$\pi(k, \tilde{m}^d[k; \theta], \tilde{q}^d[k; \theta]; \theta_\pi) = E\pi^d(x; \theta_\pi) + \varepsilon^d, \quad (23)$$

12. Our estimation of the Euler equations (see Aguirregabiria (1995)) provided imprecise estimates of some of the structural parameters. There are two simple reasons for this result. First, since corner solutions (*i.e.* zero price change or zero orders) are very frequent in our dataset, the subsample of observations that can be used to estimate the Euler equations is relatively small. Second, most of the within-brand variability in the decision variables is captured by the discrete choice. When we estimate the Euler equations in differences (and using the subsample of interior solutions) the variability of the transformed variables entering in these equations is very small, what results into imprecise parameter estimates.

where, by construction, the unobservable ε^d is orthogonal to x . Given our specification of the one-period profit function we have that

$$E\pi^d(x; \theta) = \Pi^d(x)' \mu(\theta_\pi), \quad (24)$$

where $\Pi^d(x) = (E[\exp\{m+c\}y - \exp\{c\}q|x, d], -s, -I(m < b), -I(m > b), -I(q > 0))'$; $\mu(\theta_\pi) = (1, \alpha, \eta^{p(-)}, \eta^{p(+)}, \eta^q)'$, and $\eta^{p(-)}$ and $\eta^{p(+)}$ are the menu costs associated with price reductions and price increases, respectively.

The unobservables $\{\varepsilon^d\}$ represent the uncertainty of the researcher about the actual expected profit that is observable to the firm. By construction they are mean independent of x , and their joint distribution depends on the distribution of ω conditional on x . We consider the following assumption about the joint distribution of x and $\varepsilon = (\varepsilon^1, \dots, \varepsilon^6)'$.

Assumption 5. Conditional Independence Assumption (Rust (1987):

$$pdf(x', \varepsilon'|x, \varepsilon, d) = pdf(\varepsilon'|x')pdf(x'|x, d), \quad (25)$$

where x' and ε' are next period values of x and ε .

In the context of our model, the conditional independence assumption implies that: (1) the dependence of the demand shock ω_{t+1} with respect to ω_t is captured by s_{t+1} , b_{t+1} , and c_{t+1} ; and (2) conditional on the discrete choice and on (s_t, b_t, c_t) , next period wholesale price, inventories and markup do not depend on ω_t . These two assumptions do not result from Assumptions 1 to 4 in Section 2. In particular, part of the dependence of ω_{t+1} with respect to ω_t could not be captured by x_{t+1} . However, without Assumption 5 the estimation of this decision problem would be computationally very demanding due to the existence of autocorrelated unobservables. We have tested the conditional independence assumption using a test proposed by Rust (1994a). In the context of our model this test consists on including previous period sales in the conditional choice profit functions and testing whether the parameters associated with this variable are significant.

Assumption 5 has some useful implications on the form of the discrete choice model in equations (20) and (21). Let $EV^d(k; \theta)$ be the second component of $V^d(k; \theta)$ in equation (20), i.e. the conditional choice expectation of next period value function. Under Assumption 5, $EV^d(k; \theta)$ does not depend on the vector of unobservables ε

$$EV^d(k; \theta) \equiv E \left(\max_{j \in \{1, 2, \dots, 6\}} \{V^j(k'; \theta)\} | k, d; \theta \right) = EV^d(x; \theta). \quad (26)$$

Taking into account equations (22) to (26), the optimal discrete choice becomes

$$d^*(k; \theta) = d \Leftrightarrow d = \arg \max_{j \in \{1, 2, \dots, 6\}} [\Pi^j(x)' \mu(\theta_\pi) + \varepsilon^j + \beta EV^j(x; \theta)]. \quad (27)$$

Another implication of the conditional independence assumption is that the probability of the history of discrete choices, conditional on observable state variables, is equal to the product of the conditional choice probabilities at each period

$$\Pr(d_{t1}, \dots, d_{tT} | x_t; \theta) = \prod_{t=1}^T \prod_{d=1}^6 P^d(x_{it}; \theta)^{I(d_{it}=d)}, \quad (28)$$

where $P^d(x_{it}; \theta)$ is the conditional choice probability $\Pr(d^*(k_{it}; \theta) = d | x_{it}; \theta)$.

Given a parametric specification for $pdf(\varepsilon|x)$, the econometric model in equations (20) and (27) can be estimated using a nested solution-estimation algorithm.¹³ However, the dimension of the state space and the decision space in our model makes this approach computationally very demanding. Instead, we will use an estimation method that does not require an explicit solution of the model at each iteration in the search for the parameter estimates. This method exploits a representation of the conditional choice value functions in terms of the structural parameters and conditional expectations of future paths of state and decision variables. Under Assumption 5, these conditional expectations depend only on observable variables. This property can be exploited to estimate the structural parameters using a sequential procedure. In a first stage the conditional expectations that enter in the value functions are estimated. In a second stage these estimates are solved in the expression of the conditional choice value functions, and estimates of the structural parameters are obtained exploiting moment conditions from the discrete choice model in equation (27). Different versions of this method have been previously proposed and implemented by Manski (1991, 1993), Hotz and Miller (1993), Hotz *et al.* (1994), and Ahn (1995). Our approach is closely related to Hotz–Miller method.

Lemma 3. *Under Assumption 5 and the multiplicative separability between θ and x in $E\pi^d(x; \theta)$, the optimal discrete choice can be represented using the following expression*

$$d^*(k_{it}; \theta) = d \Leftrightarrow d = \arg \max_{j \in \{1, 2, \dots, 6\}} \{ \Pi^j(x_{it})' \mu(\theta_\pi) + W_{it}^j \lambda(\theta_\pi) + \varepsilon_{it}^j \}, \quad (29)$$

where $\lambda(\theta_\pi) = (\mu[\theta_\pi]', 1)'$, and

$$W_{it}^d \equiv W^d(x, P, \theta_f^*), \quad (30)$$

where $W^d(\cdot)$ is a known closed-form function; P is the set of conditional probabilities $\{P^d(x): d = 1, \dots, 6; x \in X\}$, where $P^d(x) \equiv \Pr(d^*(k; \theta^*) = d | x; \theta^*)$; and $\theta^* \equiv (\theta_\pi^*, \theta_f^*)$ is the “true” vector of structural parameters in the population.

Proof. See Appendix 2. ||

The vector W_{it}^d has a straightforward interpretation. In our model this vector has six components. Each of these components is associated with one of the six components of the one-period profit function. For instance, the first component of W^d represents the expected and discounted stream of future sales minus variable ordering costs conditional on current value of x and on the hypothetical choice of discrete alternative d . The fifth component of W^d is the expected and discounted stream of the number of times that a positive order will be placed in the future (conditional on x and d). The last component of W^d is the expected and discounted stream of future ε 's conditional on x and d .¹⁴

It is important to notice that, in expression (29), lump-sum costs parameters are not only associated with the constant terms of the choice specific intertemporal profits. Lump-sum costs are also associated with some components of W_{it}^d . Since these components present significant sample variation, this characteristic is crucial for the identification of these parameters.

13. This approach has been successfully applied in the estimation of different dynamic discrete choice structural models. For excellent surveys of these methods see Eckstein and Wolpin (1989) and Rust (1994a, 1994b).

14. Notice that the expected stream of future ε 's is not zero because the firm will tend to choose those alternatives with relatively large values of ε .

5.2. Sequential estimation method

Based on Lemma 3 we can obtain a root- n -consistent estimator of θ_π using the following sequential procedure.

Stage 1: Nonparametric kernel estimation of $P^d(x)$; estimation of $\Pi_1^d(x) = E(\exp \{m + c\}y - \exp \{c\}q | x, d)$; and estimation of the transition probabilities of the state variables.

Stage 2: Using the previous estimates, construct the values $\hat{W}_{it}^d = W^d(x_{it}, \hat{P}, \hat{\theta}_f)$, and estimate θ_π using a GMM that exploits the following moment conditions:

$$E(Z_{it}[I(d_{it} = d) - p^{(d)}(\Pi_{it}, \hat{W}_{it}; \theta_\pi)]) = 0 \quad \text{for } d = 1, 2, \dots, 5, \quad (31)$$

where Z_{it} is a vector of instrumental variables (i.e., current and previous values of x_{it}); and, given an extreme value distribution for ε :

$$p^{(d)}(\Pi_{it}, \hat{W}_{it}; \theta_\pi) = \frac{\exp \{ \Pi_{it}^{d'} \mu(\theta_\pi) + \hat{W}_{it}^{d'} \lambda(\theta_\pi) \}}{\sum_{j=1}^6 \exp \{ \Pi_{it}^{j'} \mu(\theta_\pi) + \hat{W}_{it}^{j'} \lambda(\theta_\pi) \}}. \quad (32)$$

Hotz and Miller (1993) prove the consistency and asymptotic normality of a general class of estimators that includes this one. They also obtain the expression of the asymptotic covariance matrix of this estimator, that accounts for its sequential nature (Hotz and Miller (1993), equation 5.11).

5.3. Estimation results

We describe in Appendix 3 the details about the different estimations in the first stage of our procedure. We use a Gaussian kernel method for the estimation of the conditional choice probabilities and the transition probabilities. Here we discuss briefly some aspects related to the estimation of the first component of the vectors $\Pi^d(x)$ for $d = 1, \dots, 6$.

$$\Pi_1^d(x_{it}) = \exp \{c_{it}\} E(\exp \{m_{it}\} y_{it} - q_{it} | x_{it}, d). \quad (33)$$

This term represents expected (from the point of the econometrician) current profits gross of storage and lump-sum costs. Obviously, the form of these conditional expectations has an important effect on our estimates of the structural parameters α , $\eta^{p(-)}$, $\eta^{p(+)}$, and η^q . The main econometric problem associated with the estimation of $\Pi_1^d(x_{it})$ is the existence of sample selection bias. Since both “ $\exp \{m_{it}\} y_{it} - q_{it}$ ” and the actual choice, d_{it} , depend on the unobservable shock ω_{it} , a simple estimation of $\Pi_1^d(x_{it})$ using the subsample of observations with $d_{it} = d$ provides biased estimates due to the existence of a sample selection problem. In other words, we want to estimate expectations conditional on hypothetical choices, which are different to the expectations conditional on actual choices. This econometric problem is equivalent to the one in a model of occupational choice, where one should estimate the expected wages of an individual under the different hypothetical occupations before estimating the rest of parameters of the model.

Our theoretical model implies certain exclusion restrictions that allow us to control for this selection bias. In particular, the optimal decision rules in equations (16) and (17) imply that the discrete choice depends on all the observable state variables, but the optimal continuous decisions do not depend on all the observable state variables.¹⁴ We exploit

14. Notice that the expected stream of future ε 's is not zero because the firm will tend to choose those alternatives with relatively large values of ε .

TABLE 5

*Structural parameters estimates GMM Estimation of discrete choice model 534 brands.
January 1990–May 1992 (29 months)*

	Symmetric menu costs		Asymmetric menu costs	
	(1) Estimate (s.e.)	(2) Estimate (s.e.)	(3) Estimate (s.e.)	(4) Estimate (s.e.)
α	33.44 (6.39)	21.06 (6.69)	36.99 (6.23)	24.09 (6.59)
η^q	319.25 (24.76)	311.73 (24.97)	282.26 (21.94)	278.11 (21.40)
η^p	68.55 (25.05)	72.62 (26.24)		
$\eta^p(-)$			49.72 (19.87)	51.46 (19.42)
$\eta^p(+)$			138.94 (18.05)	138.67 (18.71)
σ_e	0.145 (0.017)	0.110 (0.020)	0.105 (0.016)	0.097 (0.011)
$\ln y_{i,t-1} (d=1)$		0.303 (0.064)		0.261 (0.087)
$\ln y_{i,t-1} (d=3)$		-0.067 (0.027)		-0.055 (0.024)
$\ln y_{i,t-1} (d=4)$		0.398 (0.162)		0.242 (0.089)
$\ln y_{i,t-1} (d=5)$		0.024 (0.089)		0.052 (0.084)
$\ln y_{i,t-1} (d=6)$		-0.033 (0.047)		-0.021 (0.056)
D-H-W test (p -value)	2.54 (0.469)		3.13 (0.535)	
Hansen test (p -value)	64.13 (0.213)	53.87 (0.365)	54.86 (0.480)	49.09 (0.510)
Goodness of fit (p -value)	18,091 (0.196)	18,078 (0.201)	17,973 (0.388)	17,970 (0.394)
Observations	13,350	13,350	13,350	13,350

^a Instrumental variables: $\{c_{i,t-j}, \ln s_{i,t-j}, b_{i,t-j} : \text{for } j = 0, 1, 2, 3\}$

^b The number of moment conditions is 60, *i.e.* 12 instruments times 5 equations.

^c Goodness-of-fit test: $\hat{\chi}^2 = NT \sum_{m=1}^M \sum_{d=1}^6 [\hat{p}^d(x^m) - p^d(x^m, \hat{\theta})]^2 / \hat{p}^d(x^m)$, where $M = 3584$ (the number of cells in the discretized state space), $\hat{p}^d(x^m)$ and $p^d(x^m, \hat{\theta})$ are the kernel estimate and the structural estimate, respectively, of the conditional choice probabilities. Under the H_0 , $\hat{\chi}^2$ is a χ^2 with $5M-K$ degrees of freedom, where K is the dimension of θ .

these exclusion restrictions to estimate $\Pi_1^d(x_{it})$ controlling for selection bias in a nonparametric form. We also test these restrictions under a parametric specification (see Appendix 3, Section A.3.3).¹⁵

The rest of this section concentrates on the estimation of the structural parameters. Table 5 presents these estimates. The discount factor β has not been estimated, and has been fixed at 0.985. However, we have estimated the model with other values for β , obtaining that the estimates of the rest of parameters do not change significantly for values of β between 0.950 and 0.999. The aggregate inflation rate, p , has been also fixed at 0.55% (Spanish average monthly inflation rate during the period). All the parameters are measured in 1990 US dollars. We have considered four different specifications: with and without asymmetric menu costs, and including or not $\ln(y_{i,t-1})$ (to test for the existence of spurious effects due to autocorrelation in demand shocks). All the estimations are two-stage GMM based on a White's estimator of the covariance matrix that allows for conditional heteroscedasticity. We have considered the same set of instrumental variables in the four specifications: current and lagged state variables up to lag 3.

We present three specification tests. The first is a Durbin–Hausman–Wu (DHW) test of the null hypothesis “no change in parameter estimates when $\ln(y_{i,t-1})$ is included”. The estimates of the parameters associated with $\ln(y_{i,t-1})$ are always significantly different from zero. However, the changes in the estimates of the structural parameters when including this variable are negligible and not significant, except for the unit storage costs α . The DHW test presents strong evidence in favour of the hypothesis of no change in

15. For instance, when $q^*(b, s, c, \omega) > 0$ and $m^*(b, s, c, \omega) \neq b$, the model predicts that $q^*(b, s, c, \omega) + s$ and $m^*(b, s, c, \omega)$ do not depend on b and s .

the parameter estimates. The second specification test is a Sargan-Hansen test of overidentifying restrictions. For all the estimations the p -value of this test is larger than 0.2, but it increases when we include $\ln(y_{i,t-1})$ and, specially, when we allow for asymmetric menu costs. Our third specification test is a Chi-square goodness-of-fit test (see Gourieroux and Monfort (1995) Vol. 2, pages 109–114). Including $\ln(y_{i,t-1})$ leads to small improvements in the goodness-of-fit of the model, but the fit of the model increases very importantly when we allow for asymmetric menu costs. In general terms, the specification tests provide evidence in favour of the model with asymmetric menu costs. Furthermore, although the assumption of no autocorrelation in demand shocks is rejected, the effect of this misspecification on our parameter estimates appears to be negligible.

The parameter estimates are very precise in all the specifications. The least precise estimate is the one for unit storage costs, α . Since this parameter is associated with the ratio s_{it}/\bar{s}_i (where \bar{s}_i is the average stock level for brand i), our estimates imply that if this ratio goes from 50% to 150% (*i.e.* an increase of 100 percentage points) monthly storage costs associated with an individual brand would increase between \$11 and \$37 (*i.e.* 95% confidence interval using estimates in column 4). Lump-sum ordering costs are significant and quantitatively very important. In column 4 the 95% confidence interval for η^q is [\$235, \$321] that represents between 4% and 6% of average monthly sales of an individual brand in the sample. This estimate is very robust to the consideration of symmetric/asymmetric menu costs and to including or not $\ln(y_{i,t-1})$.

Columns 3 and 4 present a significant difference between menu costs associated with nominal price increases and those for price reductions. Menu costs associated with a price reduction are approximately one third of the menu costs of a price increase. Since our estimation of $\Pi_1^d(x_{it})$ in the first stage does not impose symmetry in the response of sales to price changes, it is unlikely that the difference in our estimates of $\eta^{p(+)}$ and $\eta^{p(-)}$ is spuriously capturing this type of misspecification in the sales equation. Although, up to a certain extent, menu cost parameters in our model are “black boxes”, the significant and sizable difference between $\eta^{p(+)}$ and $\eta^{p(-)}$ is a relevant empirical result. Based on our conversations with company’s managers we consider that the contractual relationship between retailers and wholesalers can be particularly important to explain the relatively small costs of sales promotions for retailers. The contract between a wholesaler and the supermarket chain specifies the proportion of certain costs associated with sales promotions that will be supported by the wholesaler. These costs include menu costs as well as marketing costs (*e.g.*, mailing). According to these managers, the percentage of the cost supported by the wholesaler is generally larger than 50%. Menu costs might be symmetric, but the part of menu costs supported by the retailer is asymmetric¹⁶

However, we will show in next section that asymmetric menu costs are not the only factor, and in fact not the most important, to explain sales promotions in our dataset. Fixed ordering costs and (S, s) inventory dynamics contribute to explain a very important proportion of these promotions.

In Table 6 we present some statistics related to our estimates of fixed ordering costs and menu costs, and compare them with those obtained in previous studies. Total fixed ordering costs represent 3.15% of the total value of sales in the supermarket chain, and 3.35% of total variable ordering costs. For each item sold by the chain, almost 4 cents are spent in fixed ordering costs. Before we discuss the statistics associated with menu costs,

16. In principle, an alternative approach would be to exploit our specification of the demand equation and estimate $\Pi_1^d(x_{it})$ jointly with the rest of parameters of the model. However, this approach is not feasible in our case. Even if we exploit a parametric specification of the demand, we still have to estimate the expected values of m_{it} and q_{it} conditional on x and d in order to obtain $\Pi_1^d(x_{it})$.

TABLE 6

Menu costs and fixed ordering costs

Fixed ordering costs				
	Fixed OC per order	Fixed OC/ sales	Fixed OC/ items sold	Fixed OC/ variable OC
	\$278.11	3.15%	\$0.0365	3.35%
Menu costs				
	MC per price change and per store	MC/sales	MC/items sold	
Our study	Total	0.70%	\$0.0081	
	$\Delta P > 0$	0.31%	\$0.0035	
	$\Delta P < 0$	0.39%	\$0.0046	
Levy <i>et al.</i>	(a)	0.70%	\$0.0119	
	(b)	0.72%	\$0.0123	
Slade	\$2.70			

^a Supermarket chains where price tags are placed on the shelves but not on each individual item.

^b Supermarket chains where a separate price tag is placed on each item (in addition to the shelf price tag).

it is important to take into account the following consideration. In the studies by Slade (1994) and Levy *et al.* (1997) menu costs are measured at the level of individual stores, but our estimates represent menu costs for the whole supermarket chain. In order to obtain an estimate of menu costs per price change and per store we have divided our estimate by the number of stores in our supermarket chain. *i.e.* 62. This measure relies on the assumption that there are not economies of scale in menu costs with respect to the number of stores in a supermarket chain. This assumption is strongly supported by the results in Levy *et al.* These authors show that almost 100% of the menu costs considered in their study (that exclude managers' decision costs) are at the level of individual stores. Notice that only our estimate of menu costs per price change and per store relies on this assumption. The rest of statistics in Table 6 do not depend on this assumption.

In general terms the statistics in Table 6 show that the magnitude of our estimates is very similar to the magnitude of menu costs estimated in previous studies. Our estimate of menu costs per price change (and per store) is closer to the one in Slade than to the estimates in Levy *et al.*¹⁷ This could be explained by the fact that both our estimates of menu costs and the one by Slade include managers' decision costs, but the measures in Levy *et al.* do not include this type of menu costs. Our estimate of the ratio of total menu costs over total value of sales is very close to the one reported by Levy *et al.* (0.70%). Our estimate of the ratio of menu costs over number of items sold is approximately 33% smaller than the one in Levy *et al.*

6. COUNTERFACTUAL EXPERIMENTS

In this section we use the previous estimates to analyse the contribution of several structural parameters to the frequency of sales promotions and to the variability of markups.

17. The contract specifies the minimum number of weeks that the supermarket chain will have a certain brand under promotion during the year, and the amount of money per sales promotion that the wholesaler will pay to the supermarket if the conditions of the contract have been fulfilled.

We have solved numerically the model and generated simulations under four different scenarios. We start describing several features which are common to all the scenarios.

First, real wholesale prices are constant over time. This assumption has been made to reduce the computational cost of solving the model (3 state variables instead of 4). The results of our experiments should be robust to this simplifying assumption. The discretization grids for the state variables (b, s) and for the decision variables (m, q) are the same as in the estimation of the model (see Appendix 3). We have discretized ω_t in five values $\{-2\sigma_w, -\sigma_w, 0, \sigma_w, 2\sigma_w\}$ where σ_w is the standard deviation of ω_t .

We assume that ω_t follows a stationary AR(1) process, $\omega_t = \phi_w \omega_{t-1} + \xi_t$, where ξ_t is $iid N(0, \sigma_\xi^2)$. We also assume that $a_t \sim iid N(0, \sigma_a^2)$. Unit storage costs have been fixed at 24.09 (*i.e.* point estimate in column 4 of Table 5), and aggregate monthly inflation rate at 0.55% (*i.e.*, average monthly inflation rate during the sample period). The estimations in Table 5 have been obtained using a semi-reduced-form estimation of expected sales. However, we cannot use these estimates of expected sales when performing our simulations because the experiments would be subject to the Lucas critique. For that reason we have obtained parametric estimates of the demand parameters $\{\gamma_1, \gamma_2, \sigma_a^2 + \sigma_\xi^2, \phi_w\}$ in Appendix 3, and used these estimates in the simulations presented below. These estimates are: $\hat{\gamma}_1 = -7.806$ (0.648), $\hat{\gamma}_2 = 2.519$ (0.414), $\hat{\phi}_w = 0.082$ (0.018), and $\sqrt{\hat{\sigma}_a^2 + \hat{\sigma}_\xi^2} = 0.447$. It is clear that we cannot identify separately σ_a and σ_ξ , and thus we should make an assumption on the decomposition of the variance of the error term in the sales equation in the variance of the shock that is observable to the firm when taking its decisions (σ_ξ^2) and the variance of the shock that is unobservable to the firm, σ_a^2 . the larger the ratio $r_a \equiv \sigma_a^2 / (\sigma_a^2 + \sigma_\xi^2)$, the larger firm's uncertainty about current demand and the larger stockout probabilities. Our simulations in scenarios 1 to 3 are based on $r_a = 0.10$, that implies $\sigma_a = 0.141$ and $\sigma_\xi = 0.306$.

Scenario 1 is our benchmark model. It corresponds to the model estimated in column 4 of Table 5, and $\sigma_a = 0.141$, $\sigma_\xi = 0.306$. In the other three scenarios we have considered three experiments: zero lump-sum ordering costs (scenario 2); symmetric menu costs (scenario 3); and constant markups of competitors, *i.e.* $\sigma_\xi = 0.00$, (scenario 4).

TABLE 7
Counterfactual experiments

	Scenario 1 $\eta_q = 278.1$ $\eta^{p(-)} = 51.5$ $\sigma_\xi = 0.306$	Scenario 2 $\eta_q = 0.0$ $\eta^{p(-)} = 51.5$ $\sigma_\xi = 0.306$	Scenario 3 $\eta_q = 278.1$ $\eta^{p(-)} = 138.7$ $\sigma_\xi = 0.306$	Scenario 4 $\eta_q = 278.1$ $\eta^{p(-)} = 51.5$ $\sigma_\xi = 0.0$
Frequency of positive orders (%)	70.1	100.0	64.2	69.7
Frequency of negative ΔP (%)	50.8	26.8	48.5	44.4
Frequency of positive ΔP (%)	29.6	10.9	19.4	14.2
Std. dev. of markup (%)	7.3	3.8	5.6	4.1

Table 7 presents several statistics obtained using 1000 replications of the variables. The comparison of scenarios 1 and 2 shows that if there were not fixed ordering costs the proportion of periods with nominal price reductions would drop down from 50.8% to 26.8%, and the standard deviation of markups would decrease in 3.5 percentage points. The effects of having symmetric menu costs (scenario 3) are much smaller than the ones of removing fixed ordering costs. Finally, when markups of competitors are constant, the reductions in the variability of markups and in the proportion of sales promotions are very significant, with a similar magnitude to those associated with removing fixed ordering costs.

According to these experiments almost 50% of the sales promotions and 50% of the variability of markups are associated with the dynamic of inventories in the presence of lump-sum ordering costs. Of course, it is not the only important factor, and exogenous demand shocks have also an important contribution to sales promotions and markups variability.

7. CONCLUSIONS

This paper presents empirical support for the hypothesis that (S, s) dynamics of inventories can explain an important proportion of sales promotions and of the time variability of markups in the supermarket chain. We use a unique panel dataset that contains monthly information on sales, stocks, retail prices, wholesale prices, and orders to suppliers for 534 brands sold by a supermarket chain between January 1990 and May 1992. Reduced form estimations of the optimal pricing rules show significant effects of inventories on price decisions, specially for the decision about sales promotions. We have also presented descriptive evidence on markups following a cyclical path that matches the (S, s) behaviour of inventories, decreasing when orders are placed and increasing between two orders. Finally, simulations based on the estimated structural model show that the existence of fixed ordering costs explain around half of the time variability of markups and half of the sales promotions in our dataset. It would be unwise to draw general conclusions from a single case study. However, the two characteristics of our model that explain the important role of inventories in price dynamics, *i.e.* stockout probabilities and lump-sum ordering costs, are not specific of our supermarket.

Since our dataset comes from only one supermarket chain, we are not able to measure how much of the cross-sectional variability of markups in a certain market can be explained by (S, s) inventory behaviour. We consider this issue an interesting topic for further research.

To the best of our knowledge our estimate of fixed ordering costs is the first in the literature. Our estimate of these costs is sizable and significantly larger than our estimate of menu costs. According to these estimates annual fixed ordering costs and annual menu costs account for 3.2% and 0.7%, respectively, of the annual sales in this supermarket chain. The magnitude of our estimates of menu costs is similar to the one obtained in previous studies, though they are closer to the ones obtained in studies that include managers' decision costs in the definition of menu costs (Slade (1994)). We find significant asymmetry in menu costs: menu costs associated with price reductions are approximately one third of those associated with price increases. The contractual relationship between retailers and wholesalers could explain the relatively small costs of sales promotions for retailers. In particular, a significant proportion of the menu costs of sales promotions in supermarkets is supported by wholesalers.

APPENDIX

1. Proofs of Lemmas 1 and 2.

The proofs of Lemmas 1 and 2 build on the properties of K -concave functions. We start presenting the definition and some properties of these functions (see Scarf (1959) and Bertsekas (1976)).

Definition. Let $f(x)$ be a real function on X , where $X \subseteq \mathbb{R}^N$, and let K be a positive integer. We say that $f(x)$ is K -concave if for any x_0 and x_1 in X , and any $\lambda \in (0, 1)$:

$$\lambda f(x_0) + (1 - \lambda)f(x_1) \leq (1 - \lambda)K + f(\lambda x_0 + [1 - \lambda]x_1).$$

By definition concavity is equivalent to 0-concavity. It is also clear that if $f(x)$ is K-concave it is also M-concave for any $M \geq K$. We will use the following properties of K-concave functions.

- (i) If $f(\cdot)$ is K-concave and increasing and $g(\cdot)$ is concave then: $h = g \circ f$ is K-concave.
- (ii) If $f(\cdot)$ is K-concave, $g(\cdot)$ is M-concave, and α_1, α_2 are two positive scalars then: $\alpha_1 f + \alpha_2 g$ is $[\alpha_1 K + \alpha_2 M]$ -concave.
- (iii) If $f(x_1, x_2)$ is K-concave, and $x_2^*[x_1] \equiv \arg \max_{x_2} \{f(x_1, x_2)\}$ then: $g(x_1) \equiv f(x_1, x_2^*[x_1])$ is K-concave.
- (iv) If $f(x)$ is strictly K-concave it has a unique global maximum, x^* .
- (v) If $f(x)$ is strictly K-concave, $X \subseteq \mathfrak{R}$, and x^* is the global maximum, then: the equation $f(x) = f(x^*) - K$ has only two solutions, x^L and x^H where $x^L < x^H$. Furthermore, $f(x) > f(x^*) - K$ if and only if $x \in (x^L, x^H)$.

(a) *Proof of Lemma 1.* We proceed in two stages. first, we prove that if $V(s)$ is strictly η^q -concave the optimal decision rules for markup and orders are the real functions in equations (14) and (15). Second, we prove that $V(s)$ is η^q -concave. We define the following mappings

$$\begin{aligned} \tilde{z}(m) &= \arg \sup_{z \in \mathfrak{R}} Q(m, z); & \tilde{m}(z) &= \arg \sup_{m \in \mathfrak{R}} Q(m, z); \\ (m^*, z^*) &= \arg \sup_{(m, z) \in \mathfrak{R}} Q(m, z). \end{aligned}$$

Given these definitions and the assumption $\eta^p = 0$, the optimal decision for inventories and markup can be written as

$$z^*(s) = \arg \sup_{z \geq s} \{Q(\tilde{m}[z], z) - \eta^q I(z > s)\}; \quad m^*(s) = \tilde{m}(z^*[s]).$$

(a.1) If $V(s)$ is strictly η^q -concave, properties (i) and (ii) guarantee that $Q(m, z)$ is strictly η^q -concave. Therefore, m^* and z^* are unique (i.e. property [iv]), $\tilde{m}(s)$ is a real function (i.e. property [iv]), and $Q(\tilde{m}[z], z)$ is strictly η^q -concave. Let z^L be the smaller of the two solutions to the equation $Q(\tilde{m}[z], z) = Q(m^*, z^*) - \eta^q$. Using property (v) and the definition for $z^*(s)$ above, it is simple to verify that $z^*(s) = I(s < z^L)z^* + I(s \geq z^L)s$. Therefore, $q^*(s) = I(s < z^L)(z^* - s)$, and $m^*(s) = I(s < z^L)m^* + I(s \geq z^L)\tilde{m}(s)$.

(a.2) It remains to prove that $V(s)$ is strictly η^q -concave. We exploit the fact that

$$\sup_{s \geq 0} \sup_{\{z \geq s; m \in \mathfrak{R}\}} |my^q(m, z) - z - \eta^q I(z > s)| \leq \text{const} \tan t < \infty$$

This property guarantees (see Puterman (1994), page 151) that for any value of s

$$V(s) = \lim_{T \rightarrow \infty} V_T(s)$$

where $V_T(s)$ is the value function for the finite-horizon problem with time-horizon equal to T . Our proof is inductive. For $T = 1$ we have that $Q_1(m, z) = my(m, z) - z$ is strictly concave. Therefore, using our result in (a.1), the optimal decision for this one-period problem has the form in equations (14) and (15), where now: $(m_1^*, z_1^*) = \arg \max_{(m, z) \in \mathfrak{R}^2} Q_1(z, m)$; $\tilde{m}_1(z) = \arg \max_{m \in \mathfrak{R}} Q_1(z, m)$, and z_1^L is the smaller of the two solutions to the equation $Q_1(\tilde{m}_1[z], z) = Q_1(m_1^*, z_1^*) - \eta^q$. The value function of this one period problem is:

$$V_1(s) = I(s < z_1^L) \{Q_1(m_1^*, z_1^*) - \eta^q\} + I(s \geq z_1^L) Q_1(\tilde{m}_1[s], z)$$

And it is simple to verify that this function is strictly η^q -concave. Assume that, for any $T \geq 1$, $V_T(s)$ is strictly η^q -concave. Again, by properties (i), (ii) and (iii), $Q_{T+1}(\tilde{m}_{T+1}[z], z)$ is also strictly η^q -concave, and the optimal decision has the form in equations (14) and (15). The value function of this finite-horizon problem is:

$$V_{T+1}(s) = I(s < z_{T+1}^L) \{Q_{T+1}(m_{T+1}^*, z_{T+1}^*) - \eta^q\} + I(s \geq z_{T+1}^L) Q_{T+1}(\tilde{m}_{T+1}[s], z)$$

that is also a strictly η^q -concave function. Therefore, $V(s) = \lim_{T \rightarrow \infty} V_T(s)$ is strictly η^q -concave.

(b) *Proof of Lemma 2.* We follow the same strategy that in the proof of Lemma 1. First, we prove that if $V(b, s)$ is strictly $(\eta^q + \eta^p)$ -concave the optimal decision rules for markup and orders are the real functions in equations (16) and (17). Second, we prove that $V(b, s)$ is strictly $(\eta^q + \eta^p)$ -concave. The values m^* and z^* , and the functions $\tilde{m}(z)$ and $\tilde{z}(m)$ have the same definitions that in the proof of Lemma 1.

Notice that, by definition: (1) conditional on changing price and placing orders the optimal decision is (m^*, z^*) ; (2) conditional on changing price but not placing orders the optimal decision is $(\tilde{m}(s), s)$; and (3)

conditional on placing orders but not changing price the optimal decision is $(b, \tilde{z}(b))$. Therefore, we can define implicitly the regions in \mathfrak{R}^2 associated with the four possible discrete choices using $Q(\cdot)$, m^* , z^* , $\tilde{m}(s)$, and $\tilde{z}(b)$. For instance, the region T_{11} (i.e. price change and positive orders) is

$$T_{11} = \{(b, s) : s < z^* \text{ \& } Q(m^*, z^*) - \eta^q - \eta^p > \max \{Q(b, s); Q(\tilde{m}(s), s) - \eta^p; Q(b, \tilde{z}(b)) - \eta^q\}\}.$$

The sets T_{00} , T_{10} , and T_{01} can be defined in a similar way.

(b.1) If $V(b, s)$ is strictly $(\eta^q + \eta^p)$ -concave, properties (i) and (ii) guarantee that $Q(m, z)$ is strictly $(\eta^q + \eta^p)$ -concave. Therefore, m^* and z^* are unique, $\tilde{m}(s)$ and $\tilde{z}(b)$ are real functions, and $Q(\tilde{m}[z], z)$ and $Q(m, \tilde{z}[m])$ are $(\eta^q + \eta^p)$ -concave. Using these results, property (v) and the definitions of the four T -sets, it is possible to verify that $z^*(b, s)$ and $m^*(b, s)$ are the real functions presented in equations (16) and (17) (see Aguirregabiria (1995) for details about this part of the proof).

(b.2) The proof of the $(\eta^q + \eta^p)$ -concavity of $V(b, s)$ is equivalent to the one in (a.2).

2. Proof of Lemma 3.

By Bellman's principle we can define implicitly the conditional choice value functions as follows

$$EV^d(x; \theta) = \int \left\{ \max_{j \in \{1, 2, \dots, 6\}} \Pi^j(x') \mu(\theta_\pi) + \varepsilon^{j'} + \beta EV^j(x'; \theta) \right\} p(x' | x, d; \theta_f),$$

where x' and $\varepsilon^{j'}$ represent next period values of x and ε^j , respectively. Under Assumption 5 this expression can be written as

$$EV^d(x; \theta) = \int \left\{ \sum_{j=1}^6 P^j(x'; \theta) (\Pi^j(x') \mu(\theta_\pi) + \varepsilon^{j'} + \beta EV^j(x'; \theta)) \right\} p(x' | x, d; \theta_f),$$

where $\{P^d(x'; \theta)\}$ are the conditional choice probabilities, and $\tilde{e}^d(x; \theta) \equiv E(\varepsilon^d | x, d^*[k, \theta] = d)$. Under general conditions (see Hotz and Miller, Proposition 1) the function $\tilde{e}^d(x; \theta)$ can be written in terms of the vector of conditional choice probabilities, $P(x; \theta) \equiv \{P^1(x; \theta), \dots, P^6(x; \theta)\}$. Therefore, we define $e^d(P[x; \theta]) \equiv \tilde{e}^d(x; \theta)$. For instance, if the unobservables ε^d are iid with a double exponential distribution, $e^d(P[x; \theta]) = \text{Euler cons.} - \ln(P^d[x; \theta])$.

Let $G(x; \theta)$ be the Social Surplus function (defined by McFadden [1974]). By definition

$$G(x; \theta) \equiv E \left[\max_{j \in \{1, 2, \dots, 6\}} \{\Pi^j(x) \mu(\theta_\pi) + \varepsilon^j + \beta EV^j(x; \theta)\} | x; \theta \right].$$

Taking into account the previous expressions, it is straightforward that

$$EV^d(x; \theta) = \int G[x'; \theta] p(x' | x, d; \theta_f),$$

and

$$G(x; \theta) = \sum_{j=1}^6 P^j(x; \theta) [\Pi^j(x') \mu(\theta_\pi) + e^d(P[x; \theta]) + \beta \int G(x'; \theta) p(x' | x, j; \theta_f)].$$

This expression defines $G(\cdot; \theta)$ as the fixed point of a contraction mapping.

Now, consider the numerical solution of this contraction mapping for a particular value of θ . Let $\{x^1, x^2, \dots, x^M\}$ be a discretization grid in the space of observable state variables. Using this discretization, we can write the previous contraction mapping in matrix form as follows

$$G(\theta) = \sum_{d=1}^6 P^d(\theta) * [\Pi^d \mu(\theta_\pi) + e^d(P[\theta]) + \beta F^d(\theta_f) G(\theta)],$$

where $G(\theta)$, $P^d(\theta)$ and $e^d(P[\theta])$ are $M \times 1$ vectors; “*” is the Hadamard product (element-by-element); Π^d is an $M \times 5$ matrix; and $F^d(\theta_f)$ is the $M \times M$ matrix of transition probabilities conditional on discrete choice d . It is simple to verify that the solution to this contraction mapping is

$$G(\theta) = (I_M - \beta F(\theta))^{-1} \left\{ \sum_{d=1}^6 P^d(\theta) * (\Pi^d \mu(\theta_\pi) + e^d(P[\theta])) \right\},$$

where $F(\theta) = \sum_{d=1}^6 P^d(\theta) * F^d(\theta_f)$ is the matrix of unconditional transition probabilities.

Let θ^* be the true vector of structural parameters in the population, and define $P^{d*} \equiv P^d(\theta^*)$, $F^{d*} \equiv F^d(\theta_f^*)$ and $F^* \equiv F(\theta^*)$. Therefore, for any θ close enough to θ^* we have that

$$G(\theta) \approx (I_M - \beta F^*)^{-1} \left\{ \sum_{d=1}^6 P^{d*} * (\Pi^d \mu(\theta_\pi) + e^d(P^*)) \right\},$$

and

$$EV^d(\theta) \approx W^d \lambda(\theta_\pi),$$

where $\lambda(\theta_\pi) = \{\mu(\theta_\pi), 1\}'$ and:

$$W^d = F^{d*} (I_M - \beta F^*)^{-1} (\sum_{d=1}^6 P^{d*} * \Pi^d, \sum_{d=1}^6 P^{d*} * e^d(P^*))$$

Since Π^d 's are known and we can obtain nonparametric consistent estimates of F^* , F^{d*} and P^* (without estimating the structural model), we can obtain consistent estimates of W^d .

3. First stage estimations

A.3.1. Discretization of the state variables. We should discretize the space of observable state variables to obtain estimates of W^d . Since the range of variation of c_{it} and $\ln(s_{it})$ is very different for the different brands, we have considered the variables b_{it} , \tilde{c}_{it} , and $\tilde{l}s_{it}$, where \tilde{c}_{it} and $\tilde{l}s_{it}$ are the logarithms of real wholesale price and stock, respectively, in deviations with respect to their brand-means. We have used uniform grids to discretize each of these variables. The following table presents the details of this discretization.

TABLE A.3.1.
Discretization of the state variables

	b	\tilde{c}	$\tilde{l}s$
Minimum value (percentile)	-0.03 (2)	-0.27 (2)	-5.50 (2)
Maximum value (percentile)	0.36 (98)	0.18 (98)	2.00 (98)
Step (cell size)	0.03	0.03	0.50
Number of cells	14	16	16
Correlation between first differences of actual and discretized variables	0.94	0.93	0.98

The total number of cells in the discretized space of observable state variables is 3584. It is important to notice the very large correlation coefficients between the first time differences of the discretized and the actual variables. This discretization is capturing most of the within-brand variability of these variables.

A.3.2. Nonparametric estimation of the conditional choice probabilities. The conditional choice probabilities that we use to compute W^d (i.e. vectors P^{d*}) have been estimated using Gaussian kernel estimators. We have used the actual variables, not the discretized ones, and estimates have been evaluated at each of the 3584 grid points in the discretization of the state variables. The matrix of bandwidths is $hS^{1/2}$, where S is the sample covariance matrix of the conditioning variables, $(b_{it}, \tilde{c}_{it}, \tilde{l}s_{it})$. The smoothing parameter, h , has been chosen using Silverman's rule of thumb.

The estimates are very precise at most of the grid points. The existence of thick tails in the empirical distribution of markups, and a very thick left tail in the distribution of the stock, make possible to estimate with relative accuracy the conditional choice probabilities at small values of the stock and small and large values of the markup. The larger standard errors are obtained for the estimates at those cells associated with large values of the stock.

A.3.3. Estimation of $\Pi_1^d(x) = \exp(c)E(\exp\{m\}y - q|x, d)$. We estimate separately $E(\exp\{m_{it}\}y_{it}|x_{it}, d)$ and $E(q_{it}|x_{it}, d)$. Here we illustrate our econometric approach for the case of $E(q_{it}|x_{it}, d)$. Notice that for those discrete alternatives with $q = 0$ this function is trivially equal to zero. According to our model

$$q_{it} = \tilde{q}^d(x_{it}, \omega_{it}) = E(q_{it}|x_{it}, d) + u_{it}^d \text{ if } d_{it} = d,$$

where u_{it}^d is, by construction, orthogonal to x_{it} . However, $\{u_{it}^d|d_{it} = d\}$ is not orthogonal to x_{it} due to the existence of selection bias (i.e. d_{it} depends also on ω_{it}). Therefore, a regression of q_{it} on x_{it} using the subsample of observations with $d_{it} = d$ will provide biased estimates of $E(q_{it}|x_{it}, d)$.

Our theoretical model implies certain restrictions on the functions $\tilde{q}^d(x_{it}, \omega_{it})$ (i.e. exclusion restrictions) that allow us to control for this selection bias in a nonparametric form. In particular, the optimal decision rule in equations (16) and (17) implies that: (a) if $\{q_{it} > 0 \text{ and } \Delta P_{it} \neq 0\}$ then $q_{it} + s_{it}$ does not depend on b_{it} and s_{it} ;

and (b) if $\{q_{it} > 0 \text{ and } \Delta P_{it} = 0\}$ then $q_{it} + s_{it}$ does not depend on s_{it} . The decision rule implies also exclusion restrictions that can be exploited in the estimation of $E(\exp\{m_{it}\}y_{it}|x_{it}, d)$.

Using these exclusion restrictions and a linear specification of $\ln(q_{it} + s_{it})$ in terms of x_{it} , we estimate $E(q_{it}|x_{it}, d)$ controlling for sample selection bias through the nonparametric method in Ahn and Powell (1993). We also estimate this equation *a la Heckman* in order to test the exclusion restrictions. The following table presents the estimation results for $E(q_{it}|x_{it}, d)$. Not very surprisingly, there is strong evidence for the existence of selection bias. Furthermore, the exclusion restrictions are not rejected in any of the cases.

TABLE A.3.3
Estimation of $E(q|x, d)$ Dependent variable: $\ln(q + s)$ (within-brand estimations)

	$q > 0 \text{ and } \Delta P < 0$			$q > 0 \text{ and } \Delta P = 0$			$q > 0 \text{ and } \Delta P > 0$		
	(1) No control	(2) Heckman	(3) Ahn-Powell	(4) No control	(5) Heckman	(6) Ahn-Powell	(7) No control	(8) Heckman	(9) Ahn-Powell
$\ln(s_{it})$	0.053	-1.767		0.031	0.135		0.039	-0.180	
(s.e.)	(0.005)	(1.330)		(0.004)	(0.154)		(0.006)	(0.289)	
b_{it}	0.466	-11.778		-0.217	0.733	-0.033	-0.431	-2.235	
(s.e.)	(0.115)	(10.213)		(0.169)	(0.307)	(0.167)	(0.139)	(1.947)	
c_{it}	-1.780	-3.578	-1.548	-0.890	0.282	-1.022	-1.555	-1.204	-0.998
(s.e.)	(0.111)	(0.877)	(0.096)	(0.107)	(0.333)	(0.106)	(0.130)	(0.131)	(0.129)
LM test ^a		6.70	364.48		22.26	153.28		28.60	56.80
(p-value)		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
LM test ^b		3.09			0.76			1.85	
(p-value)		(0.21)			(0.38)			(0.40)	
Number Observ.	6185	6185	6185	3428	3428	3428	1848	1848	1848

^a Langrange Multipliers test of the null hypothesis "Selection terms = 0".

^b Langrange Multipliers test of the exclusion restrictions.

A.3.4. Transition probabilities. Nominal wholesale prices (C_{it}) are very lumpy (the frequency of no changes is 66%). The specification of the stochastic process for real wholesale prices ($c_{it} = C_{it} - p_t$) should take into account this characteristic. We consider the following specification for the transitional probability

$$f_c(c_{i,t+1}; c_{it}) = \begin{cases} p_c^0(c_{it}) & \text{if } c_{i,t+1} = c_{it} - \rho \\ f_c^1(c_{i,t+1}; c_{it}) & \text{if otherwise,} \end{cases}$$

where $p_c^0(c_{it})$ is the conditional probability function for the event "no change in nominal wholesale price", and $f_c^1(c_{i,t+1}; c_{it})$ is a transitional density for the case in which we have nominal changes in wholesale price. We have estimated these two functions using Gaussian kernel estimators. Again, we use Silverman's rule of thumb to fix the smoothing parameter. The resulting estimates show that c_{it} has a very significant and strong monotonic effect on the probability $p_c^0(\cdot)$, that goes from zero to one over the discretized range of variation of c_{it} . The effect of c_{it} on the conditional density of $c_{i,t+1}$ is much smaller and almost not significant.

A.3.5. Parametric estimation of the demand. According to our specification, at those periods in which there are not stockouts

$$\ln y_{it} = \gamma_0 + \gamma_1 p_{it} + \gamma_2 c_{it} + \omega_{it} + a_{it}.$$

There are three econometric issues to take into account in the estimation of this equation: (1) the existence of brand fixed-effects; (2) endogeneity of retail prices; and (3) possible autocorrelation in ω_{it} . We consider two different specifications for the stochastic process of ω_{it} , iid and AR(1). For the iid case we estimate the equation in first differences using as instrumental variables the lags of prices, stock, and sales at $t-2$ and $t-3$. For the AR(1) case we estimate the equation

$$\Delta \ln y_{it} = \phi_\omega \Delta \ln y_{i,t-1} + \gamma_1 \Delta p_{it} + (-\phi_\omega \gamma_1) \Delta p_{i,t-1} + \gamma_2 \Delta c_{it} + (-\phi_\omega \gamma_2) \Delta c_{i,t-1} + \Delta u_{it},$$

where ϕ_ω is the autoregressive parameter in the AR(1) process of ω_{it} ; and $\Delta u_{it} = \Delta a_{it} + \Delta \xi_{it}$, where ξ_{it} is the iid shock in the AR(1) process. We estimate $\{\phi_\omega, \gamma_1, \gamma_2\}$ in two stages. First, we estimate by IV the parameters in

the previous equation without imposing the nonlinear restrictions among them. Second we impose these restrictions using a Minimum Distance estimator.

The following table contains our estimation results. The Sargan test rejects the overidentifying restrictions for the specification with iid ω_{it} . Furthermore the sign of γ_2 in that estimation is negative. In the specification with ω_{it} AR(1), the overidentifying restrictions are not rejected and the parameters have the expected signs. We use these estimates in our simulations of the model in Section 6.

TABLE A.3.6
Estimated demand parameters

Parameter	Within groups iid	Within groups AR(1)	IV in Fist dif. iid	IV in Fist dif. AR(1)
ϕ_ω		0.183 (0.080)		0.082 (0.018)
γ_1	-5.509 (0.079)	-6.889 (0.085)	-2.219 (0.220)	-7.806 (0.648)
γ_2	3.116 (0.092)	2.144 (0.094)	-1.109 (0.164)	2.519 (0.414)
<i>p</i> -value Sargan test			0.000	0.366

Acknowledgements. This paper is a revised version of chapter 3 of my Ph.D. dissertation. I would like to thank Namkee Ahn, Cesar Alonso, Tim Bresnahan, Emilio Cerda, Andreas Hornstein, Angelo Melino, Albert Marcet, Costas Meghir, Robert Miller, Alfonso Novales and John Rust, as well as two anonymous referees, for helpful comments. I am specially grateful to my thesis advisor Manuel Arellano for his help and support. I also want to acknowledge the firm Sebastian de la Fuente S.A. for providing the dataset used in this study.

REFERENCES

- AGUIRREGABIRIA, V. (1995), "Estimation of Sequential Decision Models with Limited Dependent Variables: An Application to a Model of Price and Inventory Decisions" (Ph.D Thesis, Universidad Complutense de Madrid).
- AHN, H. (1995), "Nonparametric Two Stage Estimation of Conditional Choice Probabilities in a Binary Choice Model under Uncertainty", *Journal of Econometrics*, **67**, 337–78.
- AHN, H. and POWELL, J. L. (1993), "Semiparametric estimation of censored selection models with a non-parametric selection mechanism", *Journal of Econometrics*, **58**, 3–30.
- BERTSEKAS, D. (1976), "Dynamic programming and stochastic control" (New York: Academic Press).
- BLINDER, A. (1981), "Retail Inventory Investment and Business Fluctuations", *Brookings Papers on Economic Activity*, **2**, 443–505.
- CECCHETTI, S. G. (1986), "The Frequency of Price Adjustment: A Study of Newstand Prices of Magazines", *Journal of Econometrics*, **31**, 255–274.
- ECKSTEIN, Z. and WOLPIN, K. I. (1989), "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey", *Journal of Human Resources*, **24**, 562–598.
- GOURIEROUX, C. and MONFORT, A. (1995), "Statistics and Econometric Models" (Cambridge: Cambridge University Press).
- HECKMAN, J. (1976), "A Life Cycle Model of Earnings, Learning and Consumption", *Journal of Political Economy*, **84**, 11–44.
- HOTZ, J. and MILLER, R. A. (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies*, **60**, 497–529.
- HOTZ, J., MILLER, R. A., SANDERS, S. and SMITH, J. (1994), "A Simulation Estimator for Dynamic Models of Discrete Choice", *Review of Economic Studies*, **61**, 265–289.
- KASHYAP, A. (1995), "Sticky Prices: New Evidence from Retail Catalogs", *Quarterly Journal of Economics*, **110**, 245–274.
- LACH, S. and TSIDDON, D. (1992), "The Behavior of Prices and Inflation: An Empirical Analysis of Disaggregated Price Data", *Journal of Political Economy*, **100**, 349–389.
- LACH, S. and TSIDDON, D. (1996) "Staggering and Synchronization in Price-Setting: Evidence from Multi-product Firms", *American Economic Review*, **86**, 1175–1196.
- LEVY, D., BERGEN, M., DUTTA, S. and VENABLE, R. (1997), "The Magnitude of Menu Costs: Direct Evidence From Large U.S. Supermarket Chains", *Quarterly Journal of Economics*, **113**, 791–825.
- MANSKI, C. (1991), "Nonparametric Estimation of Expectations in the Analysis of Discrete Choice Under Uncertainty", in W. Barnett, J. Powell, and G. Tauchen, (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, (Cambridge: Cambridge University Press).
- MANSKI, C. (1993), "Dynamic Choice in Social Settings: Learning from the Experience of others", *Journal of Econometrics*, **58**, 121–36.
- McFADDEN, D. (1974), "Conditional Logit Analysis and Qualitative Choice Behaviour", in P. Zarembka (ed.), *Frontiers in Econometrics* (New York: Academic Press).

- PESENDORFER, M. (1996), "Retail Prices. A Study of Price Behaviour in Supermarkets" (Mimeo, Yale University).
- PUTERMAN, M. (1994) *Markov Decision Processes* (New York: Wiley).
- RUST, J. (1987), "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher", *Econometrica*, **55**, 999–1033.
- RUST, J. (1994a) "Estimation of Dynamic Structural Models, Problems and Prospects: Discrete Decision Processes", in J. J. Laffont and C. Sims (eds.), *Advances in Econometrics: Sixth World Congress*, Vol. II, (Cambridge: Cambridge University Press).
- RUST, J. (1994b) "Structural Estimation of Markov Decision Process", in R. E. Engle and D. L. McFadden (eds.), *Handbook of Econometrics*, Vol. 4, (Amsterdam: North-Holland).
- SCARF, H. (1959), "The Optimality of (S, s) Policies in the Dynamic Inventory Problem", in K. Arrow, S. Karlin, and P. Suppes (eds.), *Mathematical Methods for the Social Sciences*, (Stanford: Stanford University Press), 196–202.
- SHESHINSKI, E. and WEISS, Y. (1977), "Inflation and Costs of Price Adjustment", *Review of Economic Studies*, **44**, 287–303.
- SHESHINSKI, E. and WEISS, Y. (1983), "Optimum Pricing Policy Under Stochastic Inflation", *Review of Economic Studies*, **50**, 513–529.
- SHESHINSKI, E. and WEISS, Y. (1992) "Staggered and Synchronized Price Policies under Inflation: The Multiproduct Monopoly Case", *Review of Economic Studies*, **59**, 331–359.
- SLADE, M. (1994), "Optimal Pricing with Costly Adjustment and Persistent Effects: Empirical Evidence" (Mimeo, University of British Columbia).
- SLADE, M. (1996), "Sticky Prices in a Dynamic Oligopoly: An Empirical Investigation of Fixed and Variable Adjustment Costs" (Mimeo, University of British Columbia).
- SOBEL, J. (1984), "The Timing of Sales", *Review of Economic Studies*, **51**, 353–368.
- STOKEY, N. and LUCAS, R. E. (1989) *Recursive Methods in Economic Dynamics* (Cambridge: Harvard University Press).
- TOMMASI, M. (1993), "Inflation and Relative prices: Evidence from Argentina", in E. Sheshinski and Y. Weiss (eds.), *Optimal Pricing, Inflation and Cost of Price Adjustment* (Cambridge: MIT Press).
- VARIAN, H. (1980), "A Model of Sales", *American Economic Review*, **70**, 651–659.
- WARREN, E. and BARSKY, R. (1995), "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays", *Quarterly Journal of Economics*, **110**, 321–352.