Dynamic Pricing in a Distribution Channel in the Presence of Switching Costs

Koray Cosguner †

Assistant Professor of Marketing, Georgia State University

Tat Chan

Professor of Marketing, Washington University in St. Louis

P.B. (Seethu) Seetharaman

W. Patrick McGinnis Professor of Marketing, Washington University in St. Louis

Forthcoming at Management Science

Permalink: http://dx.doi.org/10.1287/mnsc.2016.2649

† Corresponding Author: Koray Cosguner, email: kcosguner@gsu.edu; Ph: 404-413-7697. This article is based on the first author's doctoral dissertation, written at Olin Business School under the supervision of the second and third authors. The authors thank seminar participants at Olin, Georgia State, Yale, Rice, Alberta, Johns Hopkins, Emory, Carnegie Mellon, Koc, Sabanci, HEC, and Ozyegin for their comments and feedback on earlier versions of this article.

Dynamic Pricing in a Distribution Channel in the Presence of Switching Costs

Abstract

We advance the literature on dynamic oligopoly pricing models in the presence of switching costs by additionally modeling the strategic pricing role of the retailer within the distribution channel. In doing this, we study the relative dynamic pricing implications of how current retail and wholesale prices for a brand must optimally take into account, respectively, past versus future demand for the brand. Using scanner data from the cola market, we find that while the retailer exploits the benefit of inertial demand by appropriately increasing the retail profit margin, the cost of investing is borne entirely by the manufacturers. We use simulation studies to show how the retailer will lose its ability to leverage the benefits of inertial demand as consumers become more price sensitive. We also show that when inertia of the more price sensitive customer segment increases, the aggregate welfare of consumers, retailer and manufacturers may all increase.

Keywords: Dynamic Pricing, Switching Costs, Distribution Channel, Structural Econometric Models.

Introduction

Dynamic pricing incentives among firms arise in product markets characterized by switching costs that are due to *inertia* in consumers' brand choices over time. Inertia refers to the phenomenon of consumers often repeat-purchasing the same brand on successive purchase occasions, leading to the market-level demand for a brand being positively correlated over time. A pricing implication of such inertia in demand, for example, is that reducing the price of Coke in the current week will increase the demand for Coke not only in the current week but also in the subsequent weeks when the price reduction on Coke has been retracted. Thus, Coke faces a trade-off between charging a low price to attract new customers and locking them for future purchases ("invest"), and charging a higher price to extract higher profits from its already locked-in customers ("harvest"). The optimal pricing strategy for Coke then depends on (1) the actual extent of inertia in consumers' brand choices in the cola market, as well as (2) the pricing strategies of competing cola manufacturers (such as Pepsi). Furthermore, since Coke sells its products through common, independent retailers, the pricing strategy of Coke also depends on (3) the pricing strategies of retailers in determining *retail* prices of Coke (and Pepsi) for end consumers, given that Coke (or Pepsi) sets *wholesale* prices for the retailer. Since each independent retailer also faces similar trade-offs between charging a low retail price on Coke (or Pepsi) to attract customers to that brand and locking them in to that brand, versus charging a higher retail price on Coke (or Pepsi) to extract higher retail profits from that brand's already locked-in customers, this creates interesting dynamic interactions between the pricing strategies of manufacturer firms and retailers.

Previous studies of pricing in the distribution channel are based on myopic pricing policies (see, for example, McGuire and Staelin 1983, Iyer 1998, Sudhir 2001, Desai, Koenigsberg and Purohit 2004, Dukes, Gal-Or and Srinivasan 2006 etc.). Although a few studies indeed account for the dynamic effects of consumer switching costs on price competition among manufacturers (see, for example, Wernerfelt 1991, Beggs and Klemperer 1992 etc.), they ignore the strategic role of the retailer. Such an omission is not innocuous because a retailer with a category management objective, unlike a manufacturer with a brand management objective, may have little to no incentive to lock in customers to a particular brand by lowering its current retail price. Knowing that the retailer may not cooperate in passing through lowered wholesale prices on their brands to end consumers, therefore, may increase the harvesting (i.e., increasing wholesale prices) over investing (i.e., decreasing wholesale prices) incentives for manufacturers. Our objective in this research is to model the harvesting versus investing pricing incentives of not only competing manufacturers, but also the retailer, in the presence of inertial demand.

We use a full equilibrium approach that simultaneously accounts for the dynamic incentives of competing manufacturers in setting wholesale prices, as well as the dynamic incentives of the common retailer through whom they sell in determining retail prices, for the different brands in the category. For this purpose, we propose a structural econometric model of dynamic pricing decisions of manufacturers and a common retailer in the presence of inertia in consumers' brand choices. To recover parameters from the model, our first step is estimating a consumer-level brand choice model that includes the effects of inertia. We use a control function approach (Petrin and Train 2010, the idea of which is similar to the approach in Villas-Boas and Winer 1999) to solve the price endogeneity problem. Our second step is estimating the structural econometric pricing model that accounts for the dynamic interactions both among manufacturers as well as between each manufacturer and the retailer. We use a technique similar to the Euler equation approach (see, for example, Hall 1979 and Hansen and Singleton 1982) while take account of the competitive interactions. This approach was first proposed by Berry and Pakes (2000). We differ from their method by using a polynomial approximation for the optimal pricing strategy of every firm and, based on the approximation, forward simulating future profits to obtain the optimality conditions for the price decisions. We further impose a penalty function in the model estimation to ensure that the polynomial price functions closely approximate the optimal price decisions predicted by the model.

We apply the model to study the cola category in a local market over a period of two years. We first use a *scanner panel data* on the household brand choice to estimate a two-segment brand choice model. Results show that the cola category is characterized by significant inertia in demand. We then use a *retail tracking data* on store-level prices of cola brands of the same local market to estimate the pricing model. Based on the estimation results, we use counterfactuals to decompose the harvesting and investing incentives of each channel member. We find that, while manufacturers have both harvesting and investing incentives, the retailer always only harvests. In other words, the retailer effectively free rides on the manufacturers' efforts by taking a lion's share of the additional profits that accrue to the channel from the existence of inertial demand. This result highlights an interesting vertical conflict in the distribution channel, between manufacturers and the retailer, in the presence of inertial demand. We also find that the less preferred brand is pressured to invest for the future by cutting its current wholesale price. Its profit margin, therefore, is reduced when compared to a market without inertia, illustrating that asymmetric brand preferences can lead to significant differences among the brands in harvesting versus investing incentives.

Further simulation studies show that the retailer significantly loses its ability to leverage the benefits of inertial demand as consumers become more price sensitive. On the other hand, each manufacturer's harvesting and investing incentives both decrease with an increase in consumer price sensitivity, yielding a smaller decrease in wholesale prices. In another simulation study, we show that when inertia of the light user / more price sensitive customer segment increases, retail prices of both Coke and Pepsi decrease. This is because the retailer is willing to sacrifice profit margin in this case for the benefit of expanding market demand. This ends up increasing the aggregate welfare of all 3 parties – consumers, retailer and manufacturers – in the distribution channel. This finding departs from the usual finding in this literature that an increase in switching costs hurts consumers by increasing retail prices (Viard 2007, Dube et al. 2008). This can have useful managerial implications such as, for example, enabling the prediction of how retail prices and profits of each channel member would change when manufacturers change customer brand inertia and price sensitivity using marketing activities.

Literature Review

There are two seminal game-theoretic studies, namely Beggs and Klemperer (1992) and Wernerfelt (1991), that centrally motivate the importance of our econometric research from the standpoint of pricing strategies of brand manufacturers. We discuss these two studies first, and later studies on the same issue next.

Beggs and Klemperer (1992) derive the normative pricing implications of inertial demand in a differentiated duopoly using an infinite period game-theoretic framework, which generalizes the two-period game-theoretic framework of Klemperer (1987a, b), where new customers arrive and a fraction of old consumers leave in each period. Furthermore, in each period, old (locked-in) customers and new (uncommitted) customers are intermingled and the two firms cannot discriminate between these groups of customers. The authors study whether the firms' temptation to exploit their current customer bases would lead to higher prices (*harvesting* incentive), or whether the firms' desire to attract new customers would lead to lower prices (*investing* incentive), than in the case of no inertia. The authors show that under a wide range of parametric assumptions, both firms – each with an installed base of existing customers – face demand functions that are relatively price inelastic compared to their counterparts in an otherwise identical mature market without inertia. This decreased price elasticity reduces the price rivalry among the firms, leading to higher prices and profits for both firms. The authors show that inertial demand could lead to vigorous price competition in the early growth stages of a market, as competing firms aggressively try to build market shares for their brands. When the modeling framework allows for rational (i.e., "forward-looking") consumers, the prices of the two firms are shown to become less competitive because consumers who realize that firms with higher market shares will charge higher prices in the future are less price elastic than naïve consumers. The authors find that prices rise as (1) firms discount the future more, (2) consumers discount the future less, (3) turnover of consumers decreases, and (4) the rate of growth of the market decreases.

In contrast to the discrete-time, game-theoretic framework adopted by Beggs and Klemperer (1992), Wernerfelt (1991) adopts a continuous-time, game-theoretic framework to study price competition between firms in inertial markets. Consistent with the findings in Beggs and Klemperer (1992), Wernerfelt (1991) also derives higher equilibrium prices for firms, as well as a positive effect of the extent of firms' future discounting behavior on equilibrium prices, in inertial markets. This shows that the equilibrium pricing results are robust to whether the game-theoretic pricing models are solved in discrete or continuous time. Using banking and telecommunications data, respectively, Kim, Kliger and Vale (2003) and Viard

(2007) find that the harvesting incentive dominates the investing incentive in dynamic pricing in the presence of switching costs.

Unlike Beggs and Klemperer (1992) and Wernerfelt (1991), who show that the harvesting incentive outweighs the investing incentive for manufacturers under a wide range of parametric assumptions, Chintagunta and Rao (1996) and Dube, Hitsch and Rossi (2009) find the opposite to be the case under some parametric assumptions that are based on actual demand estimates. They show that myopic pricing strategies of firms that fail to recognize the long-run impact of their current prices lead to prices that are higher than those implied by dynamic pricing strategies. Doganoglu (2010) obtains the same result when the degree of inertia in demand in his model is assumed to be sufficiently low. Villas-Boas (2004) also derives the same result for the case where inertial demand endogenously arises out of consumers learning about how well different brands fit their preferences, and when the distribution of consumer valuations for each product is negatively skewed. In sum, with inertial demand, manufacturer prices could increase or decrease depending on whether harvesting or investing incentives dominate, respectively (see Villas-Boas 2015 for a state-of-the-art review of this literature).

All of the models of dynamic pricing discussed above ignore the strategic role of the retailer in the distribution channel. In other words, the pricing implications of inertial demand are derived for manufacturer pricing, tacitly assuming that manufacturers sell directly to end consumers. However, in reality, retail prices for end consumers are set by retailers while manufacturers only set wholesale prices for retailers. The wholesale pricing implications of switching costs for competing manufacturers may be very different from those derived for direct selling manufacturers in the dynamic pricing literature. Explicitly modeling the strategic pricing role of the retailer is important in this regard. For example, the retailer could serve as a "strategic buffer" in terms of cushioning wholesale price competition among manufacturers (Sudhir 2001). Two recent studies look at the consequences of inertial demand for retailers' pricing decisions, namely, Che, Sudhir and Seetharaman (2007) and Dube, Hitsch, Rossi and Vitorino (2009). However, Dube et al. (2009) derive optimal retailer prices, ignoring the role of manufacturers and, therefore, treating the retailer's costs as exogenously specified. Che et al. (2007), on the other hand, simultaneously account for the strategic role of competing manufacturers in setting wholesale prices, while deriving optimal retail prices in the distribution channel. Furthermore, Che et al. (2009) take an econometric, as opposed to a purely game-theoretic, approach in explaining retail pricing decisions of retailers. In this sense, the Che et al. (2007) study is closely related to this research. However, given the computational challenges associated with the estimation of a structural econometric model of pricing in the distribution channel (as will be explained in the next paragraph), Che et al. (2007) formulate their pricing model for a finite number of decision periods only (as opposed to infinite periods). Our study relaxes this restrictive assumption and derives the appropriate pricing model for the distribution channel under the general case of infinite period decision-making of manufacturers (as in the game-theoretic literature, see, for example, Beggs and Klemperer (1992) and Wernerfelt (1991)), as well as the retailer. We find from the empirical application that the optimal pricing decisions of the retailer and manufacturers, when they consider long-term profits, are significantly different from the case when they only consider one or two periods ahead, as in the model of Che et al. (2007). This long-term profit approach helps us better understand the tension between the harvesting and investing incentives of manufacturers and the retailer in driving the observed retail prices of brands in our data. More generally, our approach can be used by manufacturers and retailers in order to correctly assess the long-run consequences of alternative pricing strategies of brands, something that cannot be accomplished using the Che et al. (2007) approach.

Estimating econometric models of dynamic pricing in the distribution channel in the presence of inertial demand presents a significant estimation challenge, arising when some firm actions (such as wholesale prices, as in our case) are unobserved by the researcher. In fact, as we will discuss in detail later, this difficulty renders recently developed econometric techniques in the econometrics literature – Pakes and McGuire (2001), Bajari, Benkard and Levin (2007) etc. – inapplicable to our context. We use a new estimation strategy to recover the parameters when some of the agents' actions are unobserved. We apply our structural econometric model of dynamic pricing in the distribution channel to the cola market.

The rest of the paper is organized as follows. In the next section, we present our econometric model of inertial demand, as well as the associated estimation procedure. In the third section, we present our structural econometric model of dynamic pricing in the distribution channel in the presence of inertial demand, as well as the associated estimation procedure. Section 4 presents the estimation results from applying our proposed structural econometric models of inertial demand and dynamic distribution channel pricing on scanner panel data from the cola market. In Section 5, we discuss the policy implications of our estimation results based on a counterfactual simulation. Section 6 concludes with caveats and directions for future research.

Econometric Model of Demand with Switching Costs

To develop an econometric model of brand choice with the no-purchase option for scanner panel data in the cola category, we recognize that the typical household h ($h = 1, 2, \ldots, H$), which is observed over $t = 1, 2, \ldots, T_h$ shopping trips, either buys or does not buy one of J cola brands. On any given shopping trip, we observe an outcome variable y_{ht} that takes the value j ($j = 0, 1, 2, \ldots, J$). When $y_{ht} = 0$ it means that the household does not purchase in the cola category during shopping trip t. Further, during each shopping trip of a household, we observe the price (P_{jt}), display (D_{jt}), and feature (F_{jt}) covariates that the household faces, regardless of whether the household purchases in the cola category. Our econometric approach models the multinomial outcome y_{ht} as explained next.

Let U_{hjt} denote the (indirect) utility of household h for brand j at shopping trip t. We assume that we can express this utility as a function of brand-specific covariates, (P_{jt}, D_{jt}, F_{jt}), as well as the household's lagged brand choice outcome, which represents the brand that was most recently purchased by the household, also referred to as the household's *state dependence*, S_{ht}, as follows.

$$U_{hjt} = \mu_{hj} + \beta_{1h} * P_{jt} + \beta_{2h} * D_{jt} + \beta_{3h} * F_{jt} + \gamma_h * Summer_t + \lambda_h * I[S_{ht} = j] + \xi_{jt} + \varepsilon_{hjt}, \qquad (1)$$

where μ_{hj}, $j = 1, 2, \ldots, J$, is the household's brand preference, $Summer_t$ is an indicator variable that takes the value 1 if week t belongs to the summer season, and 0 otherwise, $I[A]$ is the indicator function that takes the value of 1 when event A occurs and the value of 0 otherwise. For model parameters, β_{1h}, β_{2h}, β_{3h} represent the household sensitivities to marketing mix, γ_h the seasonality effect, and λ_h the household-specific inertia parameter.[i] This utility function suggests that the switching cost is equal to $-\lambda_h/\beta_{1h}$. We assume that a demand shock, ξ_{jt}, which is common to all households, may exist and correlate with price P_{jt}. The idiosyncratic error ε_{hjt} is independent from P_{jt} and are distributed *iid* across households, brands, and periods.

Let U_{h0t} denote the (indirect) utility of household h for the no-purchase option (also called "outside good") *0* at shopping trip t. We assume that we can express this utility as follows.

$$U_{h0t} = \varepsilon_{h0t}, \tag{2}$$

where the random error ε_{h0t} is distributed *iid* across households, brands, and periods.

The multinomial outcome y_{ht} is determined by the principle of maximum utility. We observe $y_{ht} = j$ when the utility of the j^{th} option to the household exceeds that of the remaining options. This yields the following probabilistic model for brand choice,

$$\Pr\left(y_{ht} = j\right) = \Pr\left(U_{hjt} = \max\left\{U_{h1t},\ldots,U_{hJt},U_{h0t}\right\}\right). \tag{3}$$

This inertial demand model, which has been used, for example, by Seetharaman, Ainslie and Chintagunta (1999), Shum (2004), Dube, Hitsch and Rossi (2010) etc., captures inertia as a *first-order* behavioral phenomenon, i.e., only the household's most (and not the second-most, third-most etc.) recent brand choice (i.e., $I[S_{ht} = j]$) influences its current brand choice probabilities. Past research in packaged goods categories has demonstrated that higher-order lagged brand choices capture little additional explanatory variance beyond the most recent lagged choice outcome, in terms of explaining current brand choices of consumers (see, for example, Kahn, Kalwani and Morrison 1986, Seetharaman 2003 etc.).

Following the latent class approach of Heckman and Singer (1984), we assume that households belong to M segments. This simplifies our empirical objective to estimate the utility parameters for each of M segments (rather than H households), as well as the associated segment sizes. This is done by maximizing the following sample log-likelihood function.[ii]

$$\ln L = \sum_{h=1}^{H} \ln\left(\sum_{m=1}^{M} \pi_m * \left[\prod_{t=1}^{T_h} \prod_{j=0}^{J} \left\{\Pr^m\left(y_{ht} = j\right)\right\}^{Y_{hjt}}\right]\right), \tag{4}$$

where $\pi_m \in [0, 1]$ stands for the size of segment m, and $\Pr^m(y_{ht} = j)$ is the conditional choice probability of household h buying brand j at shopping trip t, given that the household belongs to segment m. The probability is derived from equation (3), where the household-specific parameters in the utility function in equation (1) are replaced by segment-specific parameters. Since households usually undertake shopping trips at weekly intervals, we will interchangeably use t, for expositional purposes, to refer to shopping trip or week.

Structural Econometric Model of Dynamic Distribution Channel Pricing in the Presence of Switching Costs

To develop a structural econometric model of distribution channel pricing in the cola category, we recognize that each manufacturer j sets a wholesale price W_{jt} for a retailer, while the retailer then sets a retail price P_{jt}, for their brand, during each of $t = 1, 2, \ldots, T$ weeks in the data.[iii] The retailer is a monopolist in a local market. During each week, we observe the retail price as an outcome variable but not the wholesale price. Our econometric approach models the continuous outcome P_{jt} as explained next. We do this in two steps. We first derive a predictive model of aggregate brand market share, which is an aggregation of individual-level brand demand, as derived in the previous section. We then embed this predictive model of aggregate brand market share within a dynamic pricing game within a distribution channel involving competing manufacturers and a common retailer. This game assumes that manufacturers engage in Bertrand price competition with each other while setting their wholesale prices, while the retailer plays the role of a Stackelberg follower while setting retail prices for the manufacturers' brands (taking their wholesale prices as given). As in past empirical work on pricing games, we assume that display and feature activities of brands are exogenous to their pricing decisions. Usually display and feature decisions tend to be jointly coordinated across categories. For these reasons, our exogeneity assumption may not be unreasonable.

Predictive Model of Aggregate Brand Market Share

Let S_{jt}^{m} denote a *state variable* that represents the (segment-specific) *installed base* for brand *j* during week *t*. This installed base variable represents the number of consumers in segment *m*, as of week *t*, whose most recent brand choice is brand *j*. Further, let $S_t^m = (S_{1t}^m, S_{2t}^m, ..., S_{Jt}^m)$ represent the vector of installed base variables across all *J* brands during week *t*. The following equation, called the *state equation*, captures the evolution of S_{jt}^m from week *t* to week *t+1*.

$$S_{j,t+1}^{m} = \sum_{k \neq j}^{J} S_{kt}^{m} * \Pr_t^m(k \to j) + S_{jt}^{m} * \left(1 - \sum_{k \neq j}^{J} \Pr_t^m(j \to k)\right), \tag{5}$$

where $\Pr_t^m(k \to j)$ stands for the *switching probability*, for a consumer in segment *m*, of switching from brand *k* to brand *j*, during week *t*, and is given by

$$\Pr_t^m(k \to j) = \Pr^m\left(\mathrm{y}_{ht} = \mathrm{j} \mid I[S_{ht} = k] = 1\right). \tag{6}$$

where $I[S_{ht} = k] = 1$ represents that the household's previous purchased brand is *k* (and $I[S_{ht} = k'] = 0$ for any other brand *k'*), and Pr^m is the brand choice probability in equation (3). When $\lambda > 0$, the above equation implies that $\Pr_t^m(j \to j) > \Pr_t^m(k \to j)$. That is, for a given consumer, the probability of choosing a particular brand now is greater if the same brand had been chosen the previous time than if another brand had been chosen the previous time, indicating the inertia in brand choice. Equation (5) represents how the installed base of brand *j* changes from week *t* to week *t+1*. This happens in two ways: one, customers currently in the installed bases of the other brands (S_{kt}^{m}) switch to the installed base of brand *j* by buying brand *j* in week *t*, which happens with probability $\Pr_t^m(k \to j)$, as shown in equation (6); two, customers currently in the installed base of brand *j* (S_{jt}^{m}) continue being in the installed base of brand *j*, by either repeat-purchasing brand *j*, or choosing the no-purchase option, in week *t*, with the collective probability of the two events being $1 - \sum_{k \neq j}^{J} \Pr_t^m(j \to k)$.

Given the state equation (5), aggregate market share for brand *j* in week *t*, MS_{jt}, is given by

$$MS_{jt} = \sum_{m=1}^{M} \pi_m * MS_{jt}^{m}, \tag{7}$$

where MS_{jt}^{m} stands for segment-level market share for brand *j* in week *t* in segment *m*, and is given by

$$MS_{jt}^{m} = \sum_{k=1}^{J} S_{kt}^{m} * \Pr_t^m (k \to j). \tag{8}$$

This completes our discussion of the predictive model of aggregate brand market share. In summary, aggregate market share for brand *j* in week *t* is predicted using equation (7), which, in turn requires equation (8) as an input, which, in turn, requires equations (5) and (6) as inputs. The unknown parameters in these equations – which include all parameters in equation (6), as well as the parameter π_m in equation (7) -- are estimated using household-level scanner panel data, as explained in the previous section.[iv]

Markov-Perfect Equilibrium of the Dynamic Pricing Game

Let C_{jt} denote the marginal cost of the manufacturer for brand *j* during week *t*. It is specified as

$$C_{jt} = \alpha_{0j} + c_{jt} \cdot \alpha + v_{jt}, \tag{9}$$

where α_{0j} stands for average marginal cost of the manufacturer over time, c_{jt} a vector of cost variables, and v_{jt} a time-varying cost shock. We assume that v_{jt} is *iid* across all *j* and *t*. Let $v_t = (v_{1t}, v_{2t}, \ldots, v_{Jt})'$. We assume that v_t is public information for all manufacturers and the retailer (but not to the researcher). However, the value of future v_s, $s>t$, is still unknown when manufacturers and retailer are making pricing decisions in week *t*.

We assume a discrete-time, infinite-horizon framework (with $t = 1, 2, \ldots, \infty$), with manufacturers making simultaneous wholesale pricing decisions in each period (week), and playing a repeated Bertrand game with discounting. Let $Z_t = \left(D_{jt}, F_{jt}, \xi_{jt}, c_{jt}, v_{jt}, S_{jt}^{m}, j = 1, \ldots, J, m = 1, \ldots, M; Summer_t\right)$ be a vector of state variables that will impact manufacturers' and the retailer's demand, profits and pricing decisions. Given wholesale price W_{jt} and retail prices $P_t = (P_{1t}, P_{2t}, \ldots, P_{Jt})'$, the manufacturer's single-period profit in period *t* is given by

$$\Pi^{j}(W_t, P_t, Z_t) = (W_{jt} - C_{jt}) * MS_{jt}. \tag{10}$$

where MS_{jt} is the aggregate brand market share defined in equation (7). We assume that the retailer's marginal cost of selling each unit is equal to the wholesale price W_{jt}.[v] The retailer chooses retail prices in each period, taking wholesale prices as given. The retailer's single-period profit in period t is given by

$$\Pi^R(W_t, P_t, Z_t) = \sum_{k=1}^{J} (P_{kt} - W_{kt}) * MS_{kt}. \tag{11}$$

On account of inertial demand, W_{jt} and P_{jt} will not only influence the current market shares of brands, MS_{jt}, but also change the installed bases of all brands in all consumer segments, S_{jt}^{m}, which, in turn will affect the future stream of profits of, as well as future strategic interactions among, the manufacturers and the retailer.

All channel members are assumed to have full information about the current state variables Z_t before making their pricing decisions. The retailer is assumed to choose $(P_{1t},\ldots, P_{Jt})$ to maximize the expected discounted sum of single-period category profits,

$$E\left[\sum_{k=t}^{\infty} \rho^{k-t} * \Pi^R(P_k, W_k, Z_k) \mid Z_t\right], \tag{12}$$

while the manufacturer is assumed to choose W_{jt} to maximize the expected discounted sum of single-period brand profits.

$$E\left[\sum_{k=t}^{\infty} \rho^{k-t} * \Pi^j(P_k, W_k, Z_k) \mid Z_t\right], \tag{13}$$

where the expectation is taken over all the other channel members' current actions, all future values of observed and unobserved states, and all future actions of all channel members. We assume that the manufacturers and the retailer have a common discount factor $\rho < 1$.

We focus our attention on Pure-Strategy Markov-Perfect Equilibria (MPE), noting that there could be multiple such equilibria. In our case, a Markov strategy for a channel member describes their pricing strategy for week t – wholesale or retail, depending on whether manufacturer or retailer -- as a function of current Z_t. Formally, the retailer's strategy can be written as $\Omega_R:\ Z \rightarrow P \in R^J$, where $P = (P_1, P_2\ldots, P_J)$'

is the vector of retail prices, while each manufacturer's strategy can be written as $\Omega_j : \mathrm{Z} \to W_j \in R$, where W_j is the wholesale price charged by manufacturer *j*. A Markov profile $\Omega = \Omega_R \times \prod_{j=1}^{J} \Omega_j$, which is defined as $\Omega : \mathrm{Z} \to (\mathrm{P,W})$, is an MPE if there is no channel member *i* (retailer or manufacturer) who prefers an alternative strategy Ω_i' over Ω_i, when all other channel members are choosing their strategies according to Ω_{-i}. This can be formally written as follows.

$$V_i(\mathrm{Z} \mid \Omega_i, \Omega_{-i}) \geq V_i(\mathrm{Z} \mid \Omega_i', \Omega_{-i}), \forall i, \forall Z, \forall \Omega_i', \tag{14}$$

where V_i is the value function of the channel member that is defined below.

Given that the behavior is a Markov profile, for each manufacturer *j*, the discounted sum of profits can be written in the form of the following Bellman equation.

$$V_j(\mathrm{Z}) = \sup_{W_j} \left\{ (W_j - C_j) * MS_j + \rho * \mathrm{E} V_j(\mathrm{Z}' \mid \mathrm{Z}, P) \right\}, \tag{15}$$

Similarly, the retailer's discounted sum of profits can be written as the form of the following Bellman equation.

$$V_R(\mathrm{Z}) = \sup_{P_1, P_2, \dots, P_J} \left\{ \sum_{k=1}^{J} (P_k - W_k) * MS_k + \rho * \mathrm{E} V_R(\mathrm{Z}' \mid \mathrm{Z}, P) \right\}. \tag{16}$$

Note that the payoff relevant states for the retailer are identical to those for each manufacturer. While the retailer's pricing decisions are influenced by manufacturers' wholesale prices, since the manufacturers' wholesale prices are functions of *Z*, the value function of the retailer is also a function of *Z*.

Model Estimation

We use a two-step model estimation procedure. In the first step, we estimate the parameters in the utility function in equation (1) by maximizing the log-likelihood function in equation (4). While doing this, we correct for the potential endogeneity of the price variable using the control function approach (see Petrin and Train 2010). This involves running a first-stage linear regression of price P_{jt} versus instruments, which include the demand covariates D_{jt}, F_{jt}, and $Summer_t$ and the cost covariates c_{jt}. Let $\hat{\xi}_{jt}$ be the fitted price residual for each week. We then use a linear function, $\varphi \cdot \hat{\xi}_{jt}$, to approximate the demand shock ξ_{jt}. The assumption of the control function approach is that, conditional on $\hat{\xi}_{jt}$, the error term $\hat{\varepsilon}_{hjt} = \xi_{jt} - \varphi \cdot \hat{\xi}_{jt} + \varepsilon_{hjt}$ (see equation (1)) is independent of P_{jt}. This approach is similar to that in Villas-Boas and Winer (1999) and has been used, for example, by Pancras and Sudhir (2007) and Ma, Seetharaman and Narasimhan (2012). Based on these demand estimation results, we can calculate how the installed base S_{jt} evolves from week t to week $t+1$, based on equations (5) and (6).

The second step is to estimate the cost parameters ($\alpha_{0j}, j = 1, 2, \ldots, J, \alpha$) in equation (9) from the pricing equation. The biggest challenge in this step is that wholesale prices are unobserved, but as they will determine the retailer's retail pricing policy, if the retailer's continuation value in equation (16) is known, one can invert the wholesale prices by using the retailer's optimality conditions for setting retail prices. Similarly, if each manufacturer's continuation value in equation (15) is known, one can invert unobserved manufacturer costs by using the manufacturer's optimality conditions for setting wholesale prices. Thereafter, putting the retailer's and manufacturers' value functions together, and by jointly exploiting the optimality conditions of the retailer and the manufacturer, one can infer the marginal costs of the manufacturers using observed retail prices. This estimation strategy is adopted in Sudhir (2001) and Villas-Boas (2007) for static games in the retail channel. Since we are dealing with a dynamic game, the value functions of the retailer and the manufacturers do not have a closed form.

There are two estimation methods that have been previously developed for multi-agent problems, such as ours, and have become well-established in the literature. First, the Pakes and McGuire (1994) algorithm, which uses the nested fixed point (NFXP) algorithm in a multi-agent context, allows us to calculate optimal dynamic policies conditional on given parameters. However, to use this algorithm we would need to search for the fixed points which represent the optimal policies for different combinations of state variables in a numerical search routine, and repeat the procedure until we get a close match between the computed and observed prices. That is, conditional on model parameters, we must compute the MPE for all channel members. The algorithm requires us to iterate over σ_j many times until the best response function is satisfied. In our case, the fixed point search must account for both horizontal (among manufacturers) as well as vertical (between manufacturers and the retailer) interactions, which increases the computational burden. Even more strikingly, with large dimensions of the state space, such as in our case (with two manufacturers and a retailer, as well as two consumer segments, for each of which we need to track both observed and unobserved states), the curse of dimensionality problem becomes severe. For this reason, adapting the NFXP algorithm to estimate our dynamic game presents a non-trivial challenge.

The computational burden associated with the NFXP algorithm can be mitigated using the second method, the two-step approach of Hotz and Miller (1993), which has been extended for multi-agent problems by Bajari, Benkard and Levin (2007). Under this approach, the policy functions of agents are *estimated* for various points in the state space using the observed data. With the estimated policy functions and state transitions, one can calculate the values of the agents, for a given set of structural parameters, using the idea of forward simulation, which was first proposed by Hotz, Miller, Sanders and Smith (1994). However, to obtain reliable estimates from the first step requires a large number of observations with sufficient variation in state variables in the data, which makes it difficult to apply in our application. More importantly, this method cannot be directly applied to our model where unobserved state variables and policies (i.e., wholesale prices of manufacturers are unobserved in the data) exist.

We develop a new estimation method that can handle both the existence of unobserved policies of agents as well as a high-dimensional state space, with a lower computational burden. We describe our estimation method next.

Proposed Estimation Method for the Dynamic Pricing Game

Under the MPE assumption, optimal actions of the manufacturers are functions of payoff relevant states. Similar to Hotz and Miller (1993), we approximate the optimal policies of manufacturers and the retailer using a parametric polynomial function of observed and unobserved states. Then, conditional on any trial set of parameters of the policy functions, we forward simulate the value functions. Our estimation strategy searches the parameters of policy functions and structural parameters simultaneously through one numerical search routine, by minimizing a criterion function based on moment conditions, together with a penalty function if the optimality conditions are not satisfied. At true structural parameters, it is required that our parametric policy functions are consistent with the policies from the data, and the parametric policy functions satisfy the first-order conditions. This strategy differs from Bajari, Benkard and Levin (2007) as policy functions are not estimated in the first step; instead, the parametric policy functions for unobserved wholesale prices are identified from the constraint in our criterion function that the functions have to be consistent with the necessary optimality conditions (we will discuss this below). Therefore, our approach is similar to the MPEC approach proposed by Su and Judd (2012).

The objective of the estimation is to estimate the parameters (α_{0j}, $j = 1, 2, \ldots, J$, α). We parameterize the retailer's retail pricing policies for brands as flexible functions of state variables Z as shown below.

$$P_j = \hat{P}_j(Z \mid \theta_j^R), \tag{17}$$

where $\hat{P}_j(\cdot)$ denotes the parametric approximation of the optimal retail policies, and θ_j^R is a vector of parameters. We also parameterize manufacturers' wholesale pricing policies as flexible functions of state variables as shown below.

$$W_j = \hat{W}_j(Z \mid \theta_j^M), \tag{18}$$

where $\hat{W}_j(\cdot)$ denotes the parametric approximation of the optimal wholesale pricing policies of manufacturer *j*, and θ_j^M is a vector of parameters.

Given the policy functions in equations (17) and (18), as well as the structural parameters, the expected continuation values, $\mathrm{E}V_j(Z'|Z,P)$ and $\mathrm{E}V_R(Z'|Z,P)$, which are represented by the second terms on the right-hand side of equations (15) and (16) respectively, can be computed using forward simulation (see Online Appendix 1 for details).

Similar to the Euler equation estimation approach used in Berry and Pakes (2000), we take the derivative of the value function of the retailer in equation (16) with respect to retail price P_j in order to construct the *J* first-order conditions for the retailer, as shown below.

$$F_j = \frac{\partial V_R(Z)}{\partial P_j} = MS_j + \sum_{k=1}^{J}(P_k - W_k) * \frac{\partial MS_k}{\partial P_j} + \rho * \frac{\partial \mathrm{E}V_R(Z'|Z,P)}{\partial P_j} = 0. \tag{19}$$

The derivative $\partial \mathrm{E}V_R(Z'|Z,P)/\partial P_j$ can be obtained using the chain rule, as shown below.

$$\frac{\partial \mathrm{E}V_R(Z'|Z,P)}{\partial P_j} = \frac{\partial \mathrm{E}V_R(Z'|Z,P)}{\partial S'(Z,P)} * \frac{\partial S'(Z,P)}{\partial P_j}. \tag{20}$$

Rearranging terms, we can write the first-order condition for retail price such that it expresses the retailer's optimal retail price for a brand as a function of *Z,* as shown below.

$$P_j^*(Z) = W_j - \left\{ MS_j + \sum_{k \neq j}(P_k - W_k) * \frac{\partial MS_k}{\partial P_j} + \rho * \frac{\partial \mathrm{E}V_R(Z'|Z,P)}{\partial P_j} \right\} * \left[\frac{\partial MS_j}{\partial P_j} \right]^{-1}. \tag{21}$$

If the parametric policy function in equation (17) is optimal, for any given set of state variables the computed retail prices should match the retail prices from equation (21), after allowing for approximation error due to the parametric policy function, as shown below.

$$P_j^*(Z) \cong \hat{P}_j(Z|\theta_j^R) \tag{22}$$

We take the derivative of the value function of the manufacturer in equation (15) with respect to wholesale price W_j in order to construct the first-order conditions for the manufacturer, as shown below.

$$\frac{\partial V_j(\mathrm{Z})}{\partial W_j} = MS_j + (W_j - C_j) * \left\{ \sum_{k=1}^{J} \frac{\partial MS_j}{\partial P_k} * \frac{\partial P_k}{\partial W_j} \right\} + \rho * \sum_{k=1}^{J} \frac{\partial \mathrm{E} V_j(\mathrm{Z}' \mid \mathrm{Z}, P)}{\partial P_k} * \frac{\partial P_k}{\partial W_j} = 0. \tag{23}$$

The derivative $\partial \mathrm{E} V_j(\mathrm{Z}' \mid \mathrm{Z}, P) / \partial P_k$ can be obtained using the chain rule, as shown below.

$$\frac{\partial \mathrm{E} V_j(\mathrm{Z}' \mid \mathrm{Z}, P)}{\partial P_k} = \frac{\partial \mathrm{E} V_j(\mathrm{Z}' \mid \mathrm{Z}, P)}{\partial S'(\mathrm{Z}, P)} * \frac{\partial S'(\mathrm{Z}, P)}{\partial P_k}. \tag{24}$$

Rearranging terms, we can write the first-order condition for wholesale price such that it expresses the manufacturer's optimal wholesale price for a brand as a function of *Z*, as shown below.

$$W_j^*(\mathrm{Z}) = C_j - \left[MS_j + \rho * \sum_{k=1}^{J} \frac{\partial \mathrm{E} V_j(\mathrm{Z}' \mid \mathrm{Z}, P)}{\partial P_k} * \frac{\partial P_k}{\partial W_j} \right] \left\{ \sum_{k=1}^{J} \frac{\partial MS_j}{\partial P_k} * \frac{\partial P_k}{\partial W_j} \right\}^{-1}. \tag{25}$$

The above equation involves the retailer's retail pricing responses to manufacturers' wholesale price changes, i.e., $\partial P / \partial W$. We take the derivative of the retailer's first-order condition, F_j (see equation 19), with respect to all of the *J* retail prices ($dP_1, \ldots, dP_J$) and with respect to a single wholesale price W_j, with variation dW_j. The following equation system can be derived.

$$\begin{aligned}
&F_{j1} dP_1 / dW_1 + F_{j2} dP_2 / dW_1 + \cdots + F_{jJ} dP_J / dW_1 = \partial MS_1 / \partial P_j \\
&F_{j1} dP_1 / dW_2 + F_{j2} dP_2 / dW_2 + \cdots + F_{jJ} dP_J / dW_2 = \partial MS_2 / \partial P_j \\
&\vdots \\
&F_{j1} dP_1 / dW_J + F_{j2} dP_2 / dW_J + \cdots + F_{jJ} dP_J / dW_J = \partial MS_J / \partial P_j
\end{aligned}, \tag{26}$$

where $F_{ij} = \partial^2 V_R(\mathrm{Z}) / \partial P_i \partial P_j$. Therefore, we can represent the above *JxJ* total derivatives in matrix form as follows

$$\frac{dP}{dW} = F^{-1} A, \tag{27}$$

where the $[j, i]^{\text{th}}$ elements of the above matrices are as shown below.

$$\frac{dP}{dW}[j,i]=\frac{\partial P_j}{\partial W_i},\ F[j,i]=F_{ji},\ A[j,i]=\frac{\partial MS_i}{\partial P_j}\text{ for } i,j=1,2,...,J. \tag{28}$$

Substituting from equations (27) and (28) in to equation (25), we obtain the manufacturer's optimal wholesale pricing policy function. If the parametric policy function in equation (18) is optimal, for any given set of state variables, the computed wholesale prices should match the wholesale prices from equation (25), after allowing for approximation error from the parametric policy function, as shown below.

$$W_j^*(\mathrm{Z})\cong \hat{W}_j(\mathrm{Z}\,|\,\theta_j^M). \tag{29}$$

Combine equations (21) and (25), we obtain

$$\begin{aligned}\nu_j = P_j - C_j &+ \left\{MS_j+\sum_{k\neq j}(P_k-W_k)*\frac{\partial MS_k}{\partial P_j}+\rho*\frac{\partial \mathrm{E}V_R(\mathrm{Z}'\,|\,\mathrm{Z},P)}{\partial P_j}\right\}*\left[\frac{\partial MS_j}{\partial P_j}\right]^{-1}\\ &+\left[MS_j+\rho*\sum_{k=1}^{J}\frac{\partial \mathrm{E}V_j(\mathrm{Z}'\,|\,\mathrm{Z},P)}{\partial P_k}*\frac{\partial P_k}{\partial W_j}\right]\left\{\sum_{k=1}^{J}\frac{\partial MS_j}{\partial P_k}*\frac{\partial P_k}{\partial W_j}\right\}^{-1}.\end{aligned} \tag{30}$$

Let $X_t=(D_{jt},F_{jt},c_{jt},S_{jt}^m,j=1,...,J,m=1,...,J,Summer_t)$, as a subset of Z_t, be the vector of observed (to the researcher) state variables. To recover the structural parameters of interest, we construct the following two moment conditions.

$$E[\nu_j\,|\,\mathrm{X}]=0,\ E[\nu_j^{\ 2}\,|\,\mathrm{X}]-\sigma_j^2=0\,, \tag{31}$$

where σ_j^2 is the variance of the cost shock ν_j that we also estimate in the model. A key point of departure of our estimation approach from the GMM estimator that is typically used in the literature lies in equations (22) and (29). Given a set of state variables Z_q, where $q=1,\ldots,K_Z$, and K_Z is the dimensionality of the set, our estimates are obtained by minimizing not only a criterion function that is based on the moment conditions in equation (31), but also the following two "penalty" functions.

$$\sum_{q=1}^{K_Z}[P_j^*(\mathrm{Z}_q)-\hat{P}_j(\mathrm{Z}_q\,|\,\theta_j^R)]^2,$$

$$\sum_{q=1}^{K_Z}[W_j^*(Z_q)-\hat{W}_j(Z_q\,|\,\theta_j^M)]^2, \tag{32}$$

At the true policy functions and true values of model parameters, the errors associated with the moment conditions in equation (31), as well as the approximation errors in equation (32), will be minimized. Note that we treat the differences between optimal prices and the corresponding parameterized policy functions as approximation errors that should be minimized in the estimation routine. This is different from the approach in Su and Judd (2012), since we do not estimate $P_j^*(Z_q)$ and $W_j^*(Z_q)$ as constrained parameters. Instead, our targets are parametric functions $\hat{P}_j(Z_q\,|\,\theta_j^R)$ and $\hat{W}_j(Z_q\,|\,\theta_j^M)$ that approximate the optimal prices. This strategy enables us to only estimate parameters θ_j^R and θ_j^M, which will vastly reduce the number of model parameters.

The asymptotic distribution of our estimator is difficult to derive (as in BBL). Furthermore, we have to account for the estimation error in the estimated demand function. Therefore, we use bootstrap to calculate the standard errors of the estimated supply side parameters (see Online Appendix 2 for the details).

Model Identification

Regarding model identification, the moment conditions in equation (31) indicate that the estimation of the cost parameters $\Theta^C = (\alpha_{0j},\ j = 1,\ 2,\ \ldots,\ J,\ \alpha)$ relies on the identification assumption that $X_t = (D_{jt}, F_{jt}, c_{jt}, S^m_{jt}, j = 1,\ldots,J, m = 1,\ldots,J, Summer_t)$ is independent from the cost shock ν_{jt}. In particular, we assume that the cost shock does not affect S^m_{jt} directly. Instead, it has transitory effects on manufacturers' payoffs, as well as the retailer's payoff, by affecting their pricing decisions and thus will impact future S^m_{jt}, for $s>t$. In other words, as in Rust (1987), we assume that the observed states X_t and the unobserved states (i.e., ν_{jt}), are conditionally independent. Model identification also requires that there is sufficient variation in X_t in the data. In the empirical application in the next section, our data shows intertemporal fluctuations in the cost variables (c_t) and seasonality ($Summer_t$); there are also large variations across brands and time periods in demand shifters including D_{jt}, F_{jt}, and S^m_{jt}.

We conduct a series of Monte Carlo simulations to study how well our proposed estimation approach can recover the model parameters under a wide range of assumed structural parameters, i.e., high versus low average cost, high versus low variance of cost shocks, using a sample size similar to the empirical application that we will describe in the next section. Under each tested case in our simulations, we find that the estimated cost parameters are very close to their true (assumed) values. The results are reported in Online Appendix 3. This Monte Carlo simulation exercise gives us confidence regarding the efficiency of our proposed estimator.

The goal of our paper is to disentangle the harvesting and investing incentives of manufacturers and the retailer. The identification of the incentives relies on the assumptions of the dynamic pricing game that manufacturers are Stackelberg leaders and the retailer is a Stackelberg follower, and that manufacturers engage in repeated Bertrand price competition. These assumptions let us construct the optimality condition for the retailer in equation (19). This condition is different from that which corresponds to myopic profit maximization on account of the last term, i.e., $\rho * [\partial \mathrm{E} V_R(Z' \mid Z, P) / \partial P_j]$. When inertia exists, the derivative is negative, i.e., higher current price will reduce the future installed customer base $S' \in Z'$ thus lowering the expected continuation value of the next period $\mathrm{E}V_R(Z'|Z,P)$. Given the demand and cost parameters, one can infer the harvesting and investing incentives of the retailer with the following procedure: First, the difference between the computed profit margin $(P_j - W_j)$, as the solution of the equation system (for all J equations) in the full dynamic model, and the computed margin when the retailer maximizes the myopic profit (i.e. fix ρ to zero), will represent the investing incentive of the retailer. Next, with ρ fixed to zero, the effect of inertia on demand is captured by the second term, $\sum_{k=1}^{J}(P_k - W_k)*[\partial MS_k / \partial P_j]$, in equation (19). Under inertial demand, the partial derivative $\partial MS_j / \partial P_j$ will be less negative than when inertia is not present (i.e., λ_h in equation (1) is zero), and the retailer will set a higher retail price P_j. The difference between the computed profit margin $(P_j - W_j)$, when the retailer maximizes the myopic profit (i.e. fix ρ but not λ_h to zero), and the computed margin when inertia does not exist (i.e. fix both ρ and λ_h to zero), will represent the harvesting incentive of the retailer.

We can also disentangle the harvesting and investing incentives of manufacturers based on the optimality condition for manufacturer j in equation (23). The last term $\rho * \sum_{k=1}^{J}[\partial \mathrm{E}V_j(Z'|Z,P)/\partial P_k]$ $*[\partial P_k / \partial W_j]$ captures the influence of the current wholesale price, W_j, on retail prices, and thus change the next period's state $S' \in Z'$ and the expected continuation value, $\mathrm{E}V_j(Z'|Z,P)$, of the next period. By similar rationale as above, the difference between the computed profit margin $(W_j - C_j)$, as the solution of the above equation system (for all J equations) in the full dynamic model, and the computed margin when manufacturer j maximizes the myopic profit (i.e. fix ρ to zero), represents the investing incentive of the manufacturer. Next, with ρ fixed to zero, the difference between the computed profit margin $(W_j - C_j)$, when the manufacturer maximizes the myopic profit (i.e. fix ρ but not λ_h to zero), and the computed margin when inertia does not exist (i.e. fix both ρ and λ_h to zero), will represent the harvesting incentive of the manufacturer. To summarize, given the demand and cost parameters, we can quantify the harvesting and investing incentives of manufacturers and the retailer based on the assumptions of the game structure.

Our proposed estimation approach allows the researcher to invert strategies that are unobserved in the data from the optimality conditions. Thus, unlike BBL we can allow for unobserved actions in our model. Furthermore, since we estimate the parameterized policy functions from the optimality conditions, given arbitrary state variables, it does not require large number of observations in data. Compared with the nested fixed-point approach in Pakes and McGuire (1994), the researcher does not need to iterate the value or policy functions to search for “fixed points” as equilibrium prices in an inner algorithm. We instead simultaneously estimate the structural model parameters and the parameterized policy functions. We find that this strategy vastly reduces the computational time in model estimation.

Since our methodology relies on first-order conditions from the Bellman equation, it is designed specifically for problems with continuous policies such as pricing, advertising, R&D investment etc. For problems involving both continuous and discrete (e.g., entry and exit) policies, one can use a hybrid algorithm that uses inequality constraints for discrete actions, together with the first order conditions for continuous actions. To further decrease the computational burden, the numerical search routine should start with a good set of initial values. For example, one can start with the parameters from the static counterpart of the dynamic game. Since the parameters from the static game may be fairly close to the dynamic counterpart, it reduces the convergence time of the numerical search routine significantly. Another issue concerns how to flexibly model the pricing policy functions. Employing a high-order polynomial approximation may lead to too many estimable parameters, especially when the dimensionality of the state space is large. Therefore, we start with a low-order (e.g., linear) polynomial, and then gradually increase the order of the polynomial until the optimal and the parametric policy functions closely match each other.[vi]

Empirical Results

We use scanner panel data from Information Resources Incorporated's (IRI) scanner-panel database on cola purchases of 370 households making 31062 shopping trips at a supermarket store in a suburban market of a large U.S. city. The dataset covers a two-year period from June 1991 to June 1993. The supermarket is a local monopolist in the sense of not having other supermarkets nearby and, therefore, drawing a loyal core group of shoppers to the same store for their grocery shopping. Table 1 presents some descriptive statistics on weekly marketing variables and market shares of four major cola brands in the data. The 370 households are observed to purchase cola during 5784 (18.62%) of their shopping trips. In terms of average prices, we see that Coke, Pepsi and Royal Crown occupy a high price-tier, while the Private Label occupies a low price-tier, at the store.[vii] In terms of display and feature promotions, we see that Pepsi is displayed and featured more frequently than the other brands by the retailer. In terms of average weekly market shares, Pepsi is observed to be the dominant cola brand (with an average market share of 0.4569), while the Private Label is the smallest brand (with an average market share of 0.0671).

[Insert Table 1 and Figure 1 here]

Figure 1 shows the shares of Coke, relative to that of Pepsi, in terms of quantity sales and installed customer base, over the sample period. Due to the demand and cost shocks in the market, there is a large fluctuation in Coke's share of sales across weeks. Changes in Coke's share of installed customer base are smaller, as the last brand choice of households who do not make purchases is taken into account; still, we observe a significant variation that is necessary to estimate the state dependence in the demand function.

Estimation Results for the Demand Model with Switching Costs

We use the control function approach to control for the price endogeneity. Retail prices of every brand are regressed against the covariates in the utility function, including brand intercepts, displays, features, and the indicator of summer, as well as the covariates in the cost function, including the price of sugar and petroleum.[viii] Assuming that the demand shocks are independent across geographical markets, we also include the retail prices in another store in a suburban area that has different demographics from the area where the focal store is located. We then use $\varphi \cdot \hat{\xi}_{jt}$, where $\hat{\xi}_{jt}$ is the price residual, as a control function that enters the utility function for brand j. We assume that the error term $\hat{\varepsilon}_{hjt} = \xi_{jt} - \varphi \cdot \hat{\xi}_{jt} + \varepsilon_{hjt}$ is independent of P_{jt} and, together with the error term ε_{h0t} for the no-purchase option, are distributed *iid* Gumbel with location 0 and scale 1. Therefore, the choice probability function in equation (3) becomes the well-known multinomial Logit share function.

The first two columns of Table 2 present the estimates of our proposed demand model under the 2-support heterogeneity specification.[ix] Although the brand preference for Pepsi is higher than for Coke for the two segments, the difference is not statistically significant, suggesting that the main reason that Pepsi enjoys a higher market share is because of the lower retail price and more frequent displays and features (see Table 1). The estimated price coefficient is negative, and the estimated display and feature coefficients are positive, as expected, for both segments. Between the two segments, segment 2 consists of 71.5 % of the households. It is found to be more sensitive to price, display, and feature than segment 1. The smaller values of brand intercepts (except for Private Label) imply that households in this segment purchase less cola than segment 1. The seasonality effect of summer is negative in both segments, which is an unexpected result. We check from data and find that the average weekly sales of the cola category during summer are indeed lower than that in other seasons.[x]

The estimated inertia coefficients are positive for both segments. The estimated parameters translate to switching costs -- which can be interpreted as the price premium that a brand can charge to a consumer who bought that same brand last time, relative to a consumer who bought another brand last time – of \$0.28 and \$0.12 in segments 1 and 2, respectively. These are substantively significant, given the average prices of cola brands (see Table 1). The switching costs have a major impact on consumers' brand choice and firms' revenue. For the more inertial segment 1, the purchase probabilities of Pepsi and Coke, assuming that their previous purchases were the same brand, are about three times higher than the purchase probabilities when their previous purchases were a different brand. Segment 2 has a lower switching cost; still, the purchase probabilities when the previous purchase was the same brand are twice as high as the purchase probabilities when the previous purchase was a different brand. These give manufacturers strong investing as well as harvesting incentives. The economic significance of the switching costs motivates us to model the price competition between Coke and Pepsi as a dynamic pricing game. As very experienced and knowledgeable manufacturers in the industry for over a century, it is reasonable to believe that both Coke and Pepsi have taken into account the consumer inertia when deciding wholesale prices.

[Insert Table 2 here]

To check the robustness of our specification of the inertia, we estimate an alternative model assuming that the state dependence will only last for one period. That is, the inertia will reduce to zero in the utility function if a household does not purchase any cola brand in the current week. The last two columns in Table 2 reports the results. Estimated model parameters are qualitatively the same as the proposed demand model. Differences in the parameters between the two segments are also similar. However, this alternative model has a lower likelihood function value than the proposed model, and it can be rejected based on the BIC.

Estimation Results for the Dynamic Pricing Model in the Presence of Switching Costs

Using the results from the proposed inertial demand model, we estimate the dynamic pricing model in the distribution channel as the next step. Given any set of trial parameters, we use forward simulation to simulate the expected value functions for Coke and Pepsi and the retailer. Conditional on the parametric policy functions of players, the installed customer base of either Coke or Pepsi evolves according to the state equation (5). Future displays and features are simulated from the empirical distributions in the data. Future demand shocks $\hat{\xi}$'s are also simulated from the empirical distribution which is estimated from the price regression we discussed above. Finally, we draw future cost shocks (assumed to be *iid* over time) from the distribution *normal* $(0, \sigma_j^2)$, where σ_j^2 is the variance of the cost shocks for brand j.[xi]

The second column of Table 3 reports the estimated marginal cost parameters for Coke and Pepsi of the proposed model ("Dynamic Model (with Retailer)"). The average marginal cost of Pepsi is estimated to be 15.5% lower than Coke, offering a partial explanation why its average retail price in Table 1 is lower than Coke's. Given the average retail prices, the estimated costs translate to estimated channel profit margins of $0.367 and $0.374 for Coke and Pepsi, respectively. The estimated costs are in the ball-park of published estimates of marginal costs in this industry during that period (see, for example, Yoffie 1994), and lend face validity to our estimates. The estimates for the cost variables petroleum and sugar are insignificant and small in magnitude, suggesting that, while it is well known that both Coke and Pepsi not only have high capital costs but also spend a lot on advertising, the contribution of cost variables to marginal costs is relatively much lower.

For comparison, the third column of Table 3 ("Myopic Model (with Retailer)") reports the estimated marginal costs yielded by a structural econometric model assuming that all players are myopic. Because the benefit of installed customer base for future profits is not considered, the manufacturers and the retailer in the model do not have investing incentive. The estimated average marginal costs for Pepsi and Coke are lower than the estimates from the proposed dynamic model. This is because, when players in the game only have harvesting incentive, the predicted profit margins from the myopic model have to be higher than the dynamic model, and the estimated marginal costs will be adjusted accordingly.[xii] We also estimate another dynamic pricing model (see the fifth column of Table 3) that ignores the existence of the retailer and assumes, instead, that manufacturers directly set retail prices (as in Chintagunta and Rao 1996, Dube, Hitsch and Rossi 2009). The estimated average marginal costs yielded by this model are much higher than their counterparts under the full-channel dynamic model. This is because, since the double marginalization problem is absent, the estimated costs must be higher in order to compensate for the absence of the retailer mark-up so that the model can explain the observed (high) retail prices. This finding is reminiscent of the result in Che, Sudhir and Seetharaman (2007) that an empirical test of whether a collusive or a competitive model of dynamic pricing among manufacturers better explains observed prices may lead one to spuriously pick the collusive model if one ignores the strategic pricing role of the retailer. This happens in their case because the estimated oligopolistic conduct parameter falls under the regime of manufacturer collusion in order to compensate for the absence of the retailer mark-up so that the model can explain the observed (high) retail prices of brands.

[Insert Table 3 here]

In order to understand the substantive implications, we use the estimated structural parameters of the proposed model (from the second column in Table 3), assuming that it is the first week of summer, and compute the resulting equilibrium wholesale and retail prices for Coke and Pepsi.[xiii] We compare the associated retail margins and manufacturer margins to those yielded by the manufacturers-only model (from the fifth column in Table 3). Under the full-channel dynamic pricing model, the equilibrium channel profit margins of Coke and Pepsi are \$0.397 (=\$0.229 for retailer + \$0.168 for manufacturer) and \$0.427 (=\$0.230 for retailer + \$0.197 for manufacturer), respectively, while under the manufacturers-only dynamic pricing model, the corresponding profit margins are \$0.146 and \$0.163, respectively. These findings imply that the profit margins of channel members are under-estimated if the strategic pricing role of the retailer in the distribution channel is ignored. When evaluating the welfare impact of a public policy (e.g., increasing sales tax) or possible anti-trust pricing effects of impending mergers, it is important for policy makers to understand both the total channel profit as well as how it is distributed between manufacturers and retailers. Ignoring the strategic role of retailers can lead to substantive distortions in predicted policy consequences.

Policy Implications

We conduct a series of counterfactual studies, based on the estimation results, in this section to examine the policy implications from our full channel dynamic pricing model. Our objective is to first investigate the harvesting and investing incentives of each of the channel members under the empirical context in our data. This helps us investigate the potential conflicts of interests in the distribution channel when inertia exists in the demand. We then investigate how the harvesting versus investing incentives change under different market environments, reflected in manipulated changes in consumer price sensitivity and inertia. We also study how retail and wholesale prices, as well as the profits of channel members, change as market conditions change.

To conduct counterfactuals, we compute, given a state variable Z (see equation (5)), the optimal wholesale prices for the manufacturers and retail prices for the retailer. For the simplicity of analysis we assume there are no demand shocks ξ's nor cost shocks ν's. We also assume that it is the first week of summer. We use a modified NFXP algorithm proposed in Pakes and McGuire (1994). Computational details are provided in Online Appendix 4.

Harvesting versus Investing

We quantify and, therefore, decompose the harvesting and investing incentives for the manufacturers and the retailer in this section. To achieve this objective, we have to compare different scenarios. First, we assume that one or all firms become *myopic* thus the discount factor is zero, i.e., the last term of either equation (19) or equation (23) is set to zero, and let all firms re-optimize their pricing strategy. Myopic players have no investing incentive. Comparing the profit margin, when the market is at equilibrium, under this scenario with that from the proposed dynamic model will help to quantify the investing incentive. Next, we assume a *static* scenario in which the inertia (λ_{SD} in Table 2) of the two customer segments is fixed to zero. The derivative $\partial MS_k / \partial P_j$ in equation (19) or equation (23) has a larger magnitude than when inertia is present. Static players have neither investing nor harvesting incentive. Comparing the profit margin under this scenario with that from the myopic scenario will help us to separate the harvesting from investing incentives.

We assume the current period is the first week of summer, and compute the market equilibrium of the week under different scenarios. Results are reported in Table 4. We use two decomposition methods. First, we assume that all firms change to myopic or static players simultaneously. Results are reported at the upper panel. We then assume either Coke, Pepsi, or the retailer become myopic or static, while the other firms remain fully dynamic. Results of this method are reported at the lower panel. Results from either method are qualitatively similar. As far as the retailer's profit margins are concerned, the harvesting incentive significantly increases its margins for Coke and Pepsi, while the investing incentive slightly increases the margins further, the net effect being that the retailer's profit margins are *higher* than those in the absence of inertia. In other words, the retailer does not have the incentive to reduce current margin for the investing purpose for any brand. This is because increasing the installed customer base for one brand implies reducing the installed customer base for another brand that the retailer also carries. The retailer therefore is unwilling to sacrifice the current margin to help the future sales for any brand. The harvesting incentive contributes the lion's share of the increase in retail prices, yielding equilibrium retail profit margins for both brands that are higher than those in the absence of inertia. The investing effort comes entirely from the manufacturers, which ends up decreasing the profit margin of Coke by 22-23 % and Pepsi by about 17 %, respectively. This finding highlights a new type of vertical conflict that arises in the distribution channel, between manufacturers and the retailer, in the presence of inertial demand. This conflict arises on account of the retailer's ability to free ride, by charging higher retail prices in the future, on manufacturers' current investments in the form of reduced wholesale prices. The asymmetry in harvesting and investing incentives are also found between manufacturers. Pepsi, as a more preferred brand[xiv], has a stronger harvesting incentive. Coke, on the other hand, has less ability to harvest and is pressured to invest in building consumer installed base for future. The net effect is that the manufacturer profit margin for Coke is reduced much more than for Pepsi using either decomposition method.

To conclude, we observe that, while manufacturers have both significant harvesting and investing incentives that move in opposite directions, the retailer always only harvests[xv], illustrating a vertical conflict between manufacturers and the retailer, in the presence of inertial demand, that has not been documented in the literature. Furthermore, brands with a low consumer preference will have stronger investing incentive and smaller harvesting incentive. This highlights how brand heterogeneity can contribute to significant differences in profit margin among manufacturers.

Finally, we compare the decomposition results with a finite-period dynamic model as in Che, Sudhir and Seetharaman (2007), in which we assume the manufacturers and the retailer only look forward for one to five periods. Results from these models also suggest that there is little investing incentive for the retailer. The investing incentive, however, is under-estimated from 14% for Coke and 10% for Pepsi, when players look forward for five periods, to 23% for Coke and 17% for Pepsi, when players only look forward for the next period. This result demonstrates the importance of modeling the full dynamic price competition, instead of finite periods, in order to recover the true investing and harvesting incentives of manufacturers and the retailer.

Effects of Increasing Price Sensitivity

Suppose that as a result of changes in the socio-economic environment, households in our data become more price sensitive. We study in this section how such a change will impact the harvesting and investing incentives of channel members and, consequently, the equilibrium prices. Based on the estimation results, we vary the price sensitivity of the two customer segments from -40% to +40% of their estimated magnitudes, and compute the steady-state equilibrium prices, market shares, and profits for Coke and Pepsi. First, the harvesting incentive of the retailer and manufacturers declines as households become more price-sensitive. However, while manufacturers' investing incentive also declines, the retailer's investing incentive remains unchanged at its negligible value. The net impact is that as price sensitivity increases, the difference between the retailer and the manufacturers, in terms of how sensitive their profit margins are to accounting for dynamic incentives in pricing, narrows (see Figure 2a).

Unsurprisingly, we find that both wholesale and retail prices decrease as consumer price sensitivity increases (see Figure 2b). However, retail prices are observed to decrease at a faster rate than wholesale prices. This is because the retailer's harvesting incentive decreases, while the retailer's investing incentive continues to be negligible, with an increase in consumer price sensitivity. Consequently, the retailer significantly loses its ability to leverage the benefits of inertial demand as consumers become more price sensitive. On the other hand, each manufacturer's harvesting and investing incentives both decrease with an increase in consumer price sensitivity, with the resulting price changes in opposing directions roughly cancelling each other, thus yielding a smaller decrease in wholesale prices.

[Insert Figure 2 here]

Effects of Increasing Inertia

With the technology development, manufacturers have the ability to identify to which customer segment each household belongs, based on past purchases, and then follow up with targeted marketing activities to increase their inertia. One way of increasing inertia is to offer a coupon each time households purchase Coke or Pepsi, for the use of next purchases. This may increase "top of mind" recall among the installed bases of each brand toward the brand they purchased last time and, therefore, make them repeat purchase the favored brands with greater likelihood.[xvi] We compute the equilibrium outcomes for Coke and Pepsi at various values of the inertia parameter for one segment at a time.

Figures 3a and 3b illustrate the changes in profit margins and prices, respectively, for the retailer and manufacturers when the level of the inertia of customer segment 1 (heavy user / less price sensitive) increases from 0 to 2 (estimated inertia is 1.39). We find that the harvesting incentive of each channel member increases, and while the investing incentive of each manufacturer also increases, that of the retailer remains almost unchanged at its negligible value, as the inertia of customer segment 1 increases. Consequently, retail prices rise but wholesale prices remain relatively stable, implying that the retailer gains disproportionately more than the manufacturers from increasing levels of inertia.

Figures 3c and 3d illustrate the changes in profit margins and prices, respectively, for the retailer and manufacturers when the level of the inertia of customer segment 2 (light user / more price sensitive) increases from 0 to 2 (estimated inertia is 0.82). Unlike the findings in Figures 3a and 3b, we see here that when inertia of customer segment 2 increases, retail prices of both Coke and Pepsi decrease. This is because the retailer is willing to sacrifice profit margin in this case for the benefit of expanding the demand from the price sensitive segment 2, which is also a much larger segment. The demand expansion effect will dominate the benefit of exploiting the increased inertia; thus, the retailer's profit will also increase. For manufacturers, as the inertia of customer segment 2 increases, the profit margin increases for Pepsi but decreases for Coke. The total profit increases for both manufacturers since market demand has expanded because of the lower retail prices.

As far as consumer welfare is concerned, the implications are interestingly different across the two cases. In Figure 3b, it is observed that as the inertia of customer segment 1 increases, retail prices for both brands increase, which decreases consumer welfare. However, in Figure 3d, it is observed that as the inertia of customer segment 2 increases, retail prices of both brands decrease, which increases consumer welfare. In other words, increasing inertia in customer segment 2 increases the aggregate welfare of all 3 parties – consumers, retailer and manufacturers – in the distribution channel. Our finding that consumer welfare can increase, under some conditions, as switching costs increase, is new to the literature on switching costs. They depart from the usual finding in this literature that an increase in switching costs hurts consumers by increasing retail prices (Viard 2007, Dube et al. 2008).

[Insert Figure 3 here]

Conclusions

In this study, we propose and estimate a structural pricing model for the full distribution channel when consumers have switching costs. For this purpose, we study the cola market, which is characterized by significant inertia in consumers' brand choices over time. We estimate a consumer-level brand choice model, which includes the effects of inertia, using *scanner panel data* on cola brand choices of consumers in a local market over a period of two years. We then estimate a structural econometric pricing model, that accounts for the pricing interactions, both among manufacturers, as well as between each manufacturer and the retailer, using *retail tracking data* on store-level prices of cola brands from the same local market over the same period of two years.

We find that the cola category is characterized by significant inertia in demand. Our study reveals that while the retailer exploits the benefit of inertial demand by appropriately increasing the retail profit margin, the cost of investing is borne entirely by the manufacturers. The retailer effectively free rides on the manufacturers' efforts by taking a lion's share of the additional profits that accrue to the channel from the existence of inertial demand. This result highlights an interesting vertical conflict in the distribution channel, between manufacturers and the retailer, in the presence of inertial demand. Further, we also find that the less preferred brand does not have the same luxury of harvesting that is enjoyed by the more preferred brand. The less preferred brand is, therefore, pressured to invest for the future by cutting its current wholesale price. This decreases its profit margin when compared to a market without inertia. This finding illustrates how asymmetric brand preferences can lead to significant differences among the brands in harvesting versus investing pricing incentives.

Based on the estimation results, we use a simulation study to show that the retailer significantly loses its ability to leverage the benefits of inertial demand as consumers become more price sensitive. On the other hand, each manufacturer's harvesting and investing incentives both decrease with an increase in consumer price sensitivity, yielding a smaller decrease in wholesale prices. In another simulation study, we find that when inertia of the more price sensitive customer segment increases, it can increase the aggregate welfare of all 3 parties – consumers, retailer and manufacturers – in the distribution channel. These findings are new to the literature on consumer switching costs.

Some caveats are in order. First, we acknowledge that our estimates of marginal costs may be biased if consumers are, in fact, forward-looking in their decision-making. For example, Villas-Boas (2015) shows that if consumers foresee the future pricing consequences of their inertia, they become less price sensitive, which raises the equilibrium prices of firms. Second, our model does not capture the consumer stockpiling behavior, which has implications for dynamic pricing. In the cola category, however, stockpiling is not pervasive as revealed in our data. Therefore, ignoring the effects of consumer stockpiling may not be a critical omission in our case. That said, while extending our model to product categories where consumer stockpiling is, in fact, significant, explicitly modeling stockpiling behavior, as well as its implications for dynamic pricing, would be necessary. Last, but not least, modeling non-price competition including display and feature decisions jointly with the pricing decisions would be an important, but challenging, area of future research.

References

Bajari, P., Benkard, C. L., Levin, J. (2007). Estimating Dynamic Models of Imperfect Competition, Econometrica, 75, 5, 1331-1370.

Beggs, A., Klemperer, P. (1992). Multi-Period Competition with Switching Costs, Econometrica, 57, 3, 651-666.

Berry, S., Levinsohn, J., Pakes, A. (1995). Automobile Prices in Market Equilibrium, Econometrica, 63, 4, 841-890.

Berry, S., & Pakes, A. (2000). Estimation from the optimality conditions for dynamic controls. Manuscript, Yale University, 841-890.

Che, H., Sudhir, K., Seetharaman, P. B. (2007). Bounded Rationality in Pricing Under State Dependent Demand: Do Firms Look Ahead? How Far Ahead?, Journal of Marketing Research, 44, 3, 434-449.

Chintagunta, P. K., Rao, V. R. (1996). Pricing Strategies in a Dynamic Duopoly: A Differential Game Model, Management Science, 42, 11, 1501-1514.

Desai, P. S., Koenigsberg, O., Purohit, D. (2004). Strategic Decentralization and Channel Coordination, Quantitative Marketing and Economics, 2, 1, 5-22.

Doganoglu, T. (2010). Switching Costs, Experience Goods and Dynamic Price Competition, Quantitative Marketing and Economics, 8, 1, 167-205.

Dube, J. P., Hitsch, G. J., Rossi, P. E. (2009). Do Switching Costs Make Markets Less Competitive?, Journal of Marketing Research, 46, 3, 435-445.

Dube, J. P., Hitsch, G. J., Rossi, P. E. (2010). State Dependence and Alternative Explanations for Consumer Inertia, The RAND Journal of Economics, 41, 3, 417-445.

Dube, J. P., Hitsch, G. J., Rossi, P. E., Vitorino, M. A. (2008). Category Pricing with State-Dependent Utility, Marketing Science, 27, 3, 417-429.

Dukes, A. J., Gal-Or, E., Srinivasan, K. (2006). Channel Bargaining with Retailer Asymmetry, Journal of Marketing Research, 43, 1, 84-97.

Hall, R. E. (1979). Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. NBER Working Paper, (R0015).

Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica: Journal of the Econometric Society, 1269-1286.

Heckman, J., Singer, B. (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, Econometrica, 52, 3, 271-318.

Hotz, V. J., Miller, R. A. (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models, The Review of Economic Studies, 60, 3, 497-529.

Hotz, V. J., Miller, R. A., Sanders, S., Smith, J. (1994). A Simulation Estimator for Dynamic Models of Discrete Choice, The Review of Economic Studies, 61, 2, 265-289.

Iyer, G. (1998). Coordinating Channels Under Price and Nonprice Competition, Marketing Science, 17, 4, 338-355.

Kim, M., Kliger, D., Vale, B. (2003). Estimating Switching Costs, Journal of Financial Intermediation, 12, 1, 25-56.

Klemperer, P. (1987a). Markets with Consumer Switching Costs, The Quarterly Journal of Economics, 102, 4, 375-394.

Klemperer, P. (1987b). The Competitiveness of Markets with Switching Costs, RAND Journal of Economics, 18, 1, 138-150.

Ma, Y., Seetharaman, P. B., Narasimhan, C. (2012). Modeling Dependencies in Brand Choice Outcomes Across Complementary Categories, Journal of Retailing, 88, 1, 47-62.

McGuire, T. W., Staelin, R. (1983). An Industry Equilibrium Analysis of Downstream Vertical Integration, Marketing Science, 2, 2, 161-191.

Pancras, J., Sudhir, K. (2007). Optimal Marketing Strategies for a Customer Data Intermediary, Journal of Marketing Research, 44, 4, 560-578.

Pakes, A., McGuire, P. (1994). Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model, RAND Journal of Economics, 25, 4, 555-589.

Pakes, A., McGuire, P. (2001). Symmetric Markov-Perfect Equilibrium, Stochastic Algorithms and the Curse of Dimensionality, Econometrica, 69, 5, 1261-1281.

Petrin, A., Train, K. (2010). A Control Function Approach to Endogeneity in Consumer Choice Models, Journal of Marketing Research, 47, 1, 3-13.

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

Sudhir, K. (2001). Structural Analysis of Manufacturer Pricing in the Presence of a Strategic Retailer, Marketing Science, 20, 3, 244-264.

Seetharaman, P. B. (2004). Modeling Multiple Sources of State Dependence in Random Utility Models: A Distributed Lag Approach, Marketing Science, 23, 2, 263-271.

Shum, M. (2004). Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market, Journal of Economics & Management Strategy, 13, 2, 241-272.

Su, C. L., Judd, K. L. (2012). Constrained Optimization Approaches to Estimation of Structural Models, Econometrica, 80, 5, 2213-2230.

Viard, B. (2007). Do Switching Costs Make Markets More or Less Competitive? The Case of 800-Number Portability, The RAND Journal of Economics, 38, 1, 146-163.

Villas-Boas, J. M. (2004). Consumer Learning, Brand Loyalty, and Competition, Marketing Science, 23, 1, 134-145.

Villas-Boas, J. M., Zhao, Y. (2005). Retailers, Manufacturers, and Individual Consumers: Modeling the Supply Side in the Ketchup Marketplace, Journal of Marketing Research, 42, 1, 83-95.

Villas-Boas, J. M. (2015). A Short Survey on Switching Costs and Dynamic Competition, International Journal of Research in Marketing, 32, 2, 219-222.

Villas-Boas, S. B. (2007). Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data, The Review of Economic Studies, 74, 2, 625-652.

Villas-Boas, J. M. Winer, R. S. (1999). Endogeneity in brand choice models, Management Science, 45, 10, 1324-1338.Wernerfelt, B. (1991). Brand Loyalty and Market Equilibrium, Marketing Science, 10, 3, 229-245.

Yoffie, D. B. (1994). Cola Wars Continue: Coke vs. Pepsi in the 1990s, Harvard Business School Case 9-794-055.

TABLE 1: DESCRIPTIVE STATISTICS ON COLA DATASET[1]

(JUNE 1991 – JUNE 1993)

Number of Households = 370
Number of Shopping Trips = 31062
Number of Purchases = 5784

Brand	Price ($ / unit)	Display	Feature	Market Share
Coke	$0.7996 ($0.0672)	0. 1525 (0. 1000)	0.2598 (0.1318)	0.3079 (0.1046)
Pepsi	$0.7402 ($0.0580)	0.2283 (0.1861)	0.3362 (0.2184)	0.4569 (0.1179)
Royal Crown	$0.8036 ($0. 0785)	0. 1005 (0.0817)	0.1060 (0.0933)	0.1682 (0.0898)
Private Label	$0.5231 ($0.0724)	0.1119 (0.0825)	0.0708 (0.0950)	0.0671 (0.0552)

Note: A unit is standard size of 32 oz.

TABLE 2: ESTIMATION RESULTS – INERTIAL DEMAND MODEL[2]

	Proposed Demand Model		**One-Period Inertia Model**	
	Segment 1	**Segment 2**	**Segment 1**	**Segment 2**
Coke	1.007 (0.196)	-0.190 (0.189)	1.826 (0.159)	-0.118 (0.180)
Pepsi	1.100 (0.196)	0.016 (0.219)	1.951 (0.157)	0.149 (0.188)
Private Label	-1.833 (0.166)	-1.710 (0.158)	-1.799 (0.155)	-1.500 (0.138)
Royal Crown	0.535 (0.173)	-0.561 (0.194)	1.033 (0.146)	-0.570 (0.177)
Price	-5.055 (0.227)	-6.709 (0.283)	-5.488 (0.204)	-6.445 (0.239)
Display	1.067 (0.068)	1.456 (0.069)	1.078 (0.067)	1.556 (0.079)
Feature	0.179 (0.070)	0.428 (0.057)	0.173 (0.068)	0.431 (0.078)
Inertia	1.393 (0.050)	0.815 (0.057)	1.094 (0.046)	0.962 (0.090)
Summer	-0.129 (0.061)	-0.154 (0.059)	-0.162 (0.060)	-0.118 (0.061)
Segment Size	28.52%	71.48%	31.15%	68.85%
Likelihood	-15450		-15780	
BIC	31179		31840	

[1] Standard Deviations are reported within parentheses.
[2] Standard errors are reported within parentheses in Tables 3, 4 and 5. The standard errors in Table 4 are bootstrapped standard errors.

TABLE 3: ESTIMATION RESULTS – DYNAMIC PRICING MODEL

	Dynamic Model (with Retailer)		Myopic Model (with Retailer)		Dynamic Model (without Retailer)	
Parameters	**Estimates**	**S.E.**	**Estimates**	**S.E.**	**Estimates**	**S.E.**
Pepsi	0.366	0.138	0.254	0.186	0.577	0.117
Coke	0.433	0.158	0.336	0.219	0.636	0.130
Petroleum	0.008	0.012	0.011	0.015	0.005	0.010
Sugar	-0.017	0.017	0.021	0.022	-0.010	0.014

TABLE 4: HARVESTING AND INVESTING INCENTIVES

DECOMPOSITION METHOD 1: SIMULTANEOUS CHANGES FOR ALL PLAYERS

	Proposed Full Dynamic Model	Myopic Model (No Investing Incentive)	Static Model (No Investing and Harvesting Incentive)	Investing Incentive	Harvesting Incentive	Net Effect
Margin of Coke	\$0.17	\$0.22	\$0.19	-22.08%	11.77%	-12.91%
Margin of Pepsi	\$0.20	\$0.24	\$0.20	-16.86%	19.84%	-0.37%
Retailer's Margin of Coke	\$0.23	\$0.22	\$0.20	2.51%	9.64%	12.40%
Retailer's Margin of Pepsi	\$0.23	\$0.22	\$0.20	2.34%	12.51%	15.14%

DECOMPOSITION METHOD 2: OTHER PLAYERS REMAIN DYNAMIC

	Proposed Full Dynamic Model	Myopic Model (No Investing Incentive)	Static Model (No Investing and Harvesting Incentive)	Investing Incentive	Harvesting Incentive	Net Effect
Margin of Coke	\$0.17	\$0.22	\$0.21	-22.87%	6.68%	-17.72%
Margin of Pepsi	\$0.20	\$0.24	\$0.21	-17.29%	11.56%	-7.73%
Retailer's Margin of Coke	\$0.23	\$0.23	\$0.21	0.75%	12.13%	12.96%
Retailer's Margin of Pepsi	\$0.23	\$0.23	\$0.20	0.77%	14.52%	15.40%

FIGURE 1: SHARE OF QUANTITY SALES AND CUSTOMER INSTALLED BASE OF COKE (RELATIVE TO PEPSI)

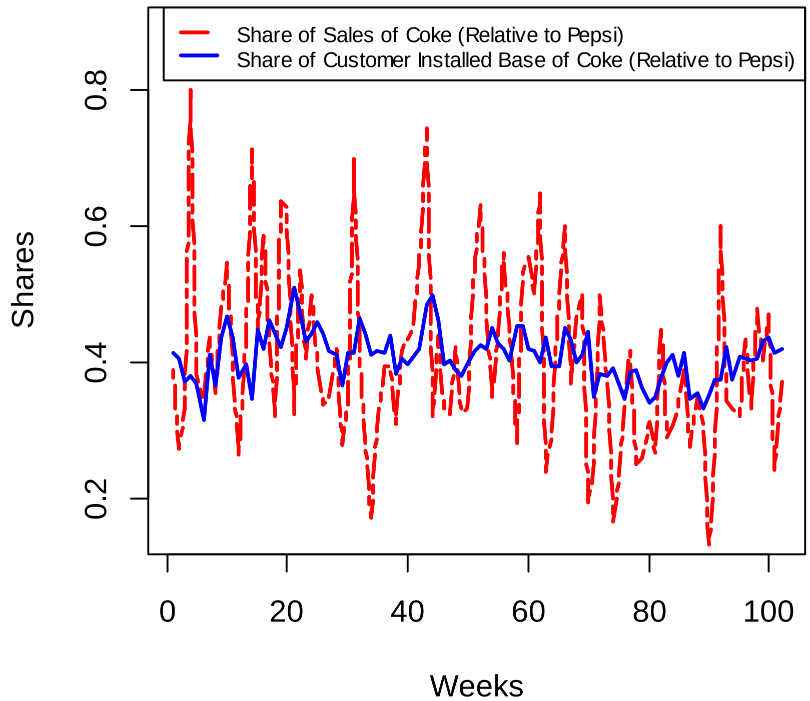
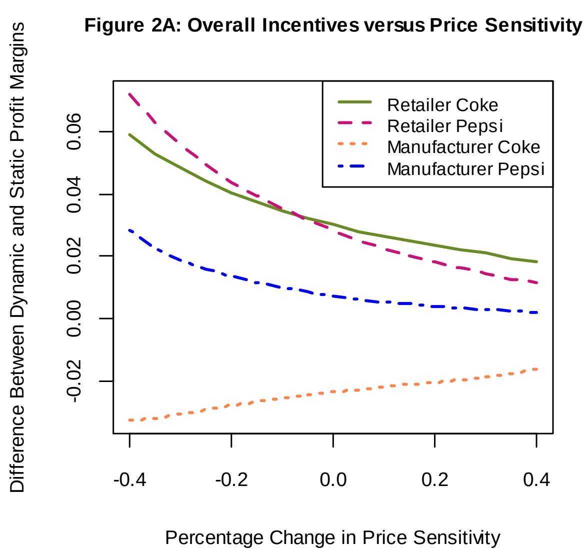


FIGURE 2: INCENTIVES AND PRICES AS A FUNCTION OF CONSUMER PRICE SENSITIVITY



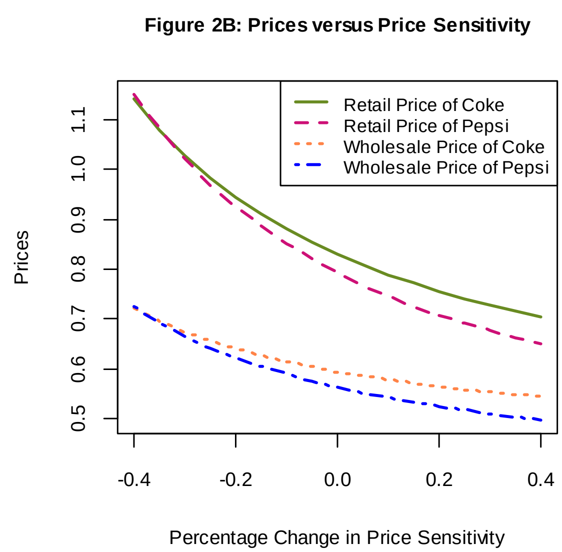


FIGURE 3: PROFIT MARGINS AND PRICES AS A FUNCTION OF CONSUMER INERTIA

Figure 3A: Profit Margins versus Segment 1 Inertia

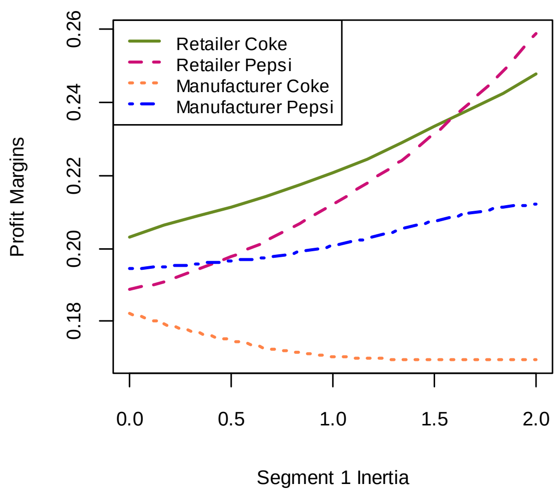


Figure 3B: Prices versus Segment 1 Inertia

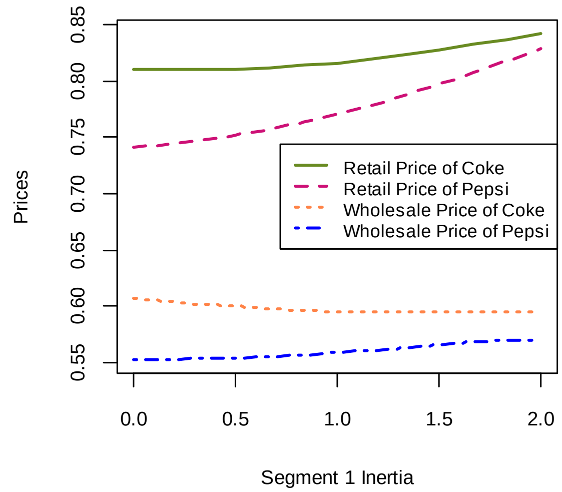


Figure 3C: Profit Margins versus Segment 2 Inertia

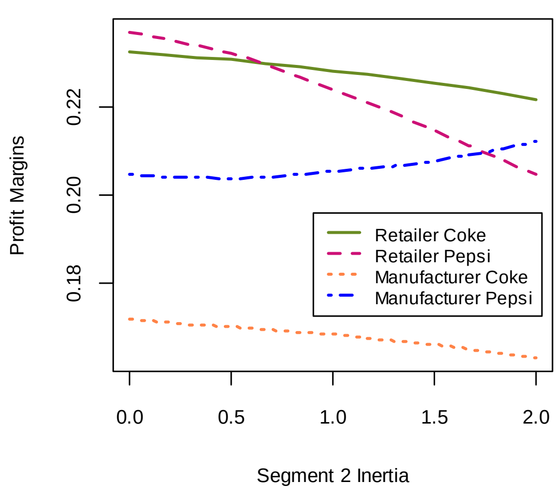
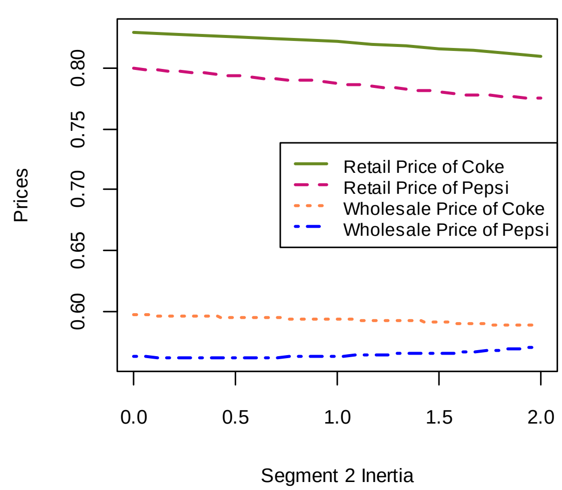


Figure 3D: Prices versus Segment 2 Inertia



[i] This coefficient is more generally referred to as the *state dependence coefficient*, and captures inertia only when it takes positive values; it captures variety seeking when it takes negative values. In this paper, we will refer to the state dependence coefficient as the *inertia parameter* for expositional convenience since it only takes positive values in our cola dataset.

[ii] Unlike the random coefficients logit model, the latent class logit model yields convenient closed-form expressions for aggregate brand market share functions (as will be explained in the next section). Further, Andrews, Ainslie and Currim (2002) show that the latent class logit model yields aggregate estimates of brand demand, as well as holdout demand forecasts, that are just as accurate as those yielded by random coefficients logit models.

[iii] While there are 4 brands – Coke, Pepsi, Royal Crown, and Private Label – in the cola category, we endogenize the prices of only the two major brands – Coke, Pepsi – in the empirical analysis. This is done for computational convenience. The prices of Royal Crown and Private Label are treated as exogenous to the analysis.

[iv] Our demand model ignores the purchase of multiple units of a product (i.e., the quantity decision). In the empirical application on the cola category, which we will describe later, first we find that the correlation between the market share based on purchase incidences and the market share based on quantity outcomes over all weeks in the data is 0.92. Second, we conduct an independent sample t-test in order to check whether the means of the two market shares (based on quantity and incidence outcomes) differ from each other. This t-test yields no statistically significant difference between the means of the two measures. We also conduct a paired sample t-test in order to see whether the two market shares differ in a given week, and again find that there is no statistically significant difference between the weekly measures. These results suggest that ignoring quantity outcomes will not lead to biases in our demand estimation results.

[v] This is a standard assumption in the literature. Other components of marginal costs, such as inventory holding costs, can be considered as relatively minor when compared to the wholesale prices.

[vi] In the empirical application, our model estimation stops at second order with interactions.

[vii] As is commonly done in the literature, price / display / feature of a brand are calculated as a share-weighted average price / display / feature over all SKUs representing that brand.

[viii] We have also included the prices of other raw materials and found that the estimated coefficients are insignificant.

[ix] Substantive insights gleaned from our empirical analysis remain similar when the heterogeneity specification is modified to include additional supports for the heterogeneity distribution. These results are available upon request.

[x] Although we expect the effect to be different in other places that exhibit larger temperature differences across seasons, it does not have any bearing on the measurement of the investing and harvesting incentives of manufacturers and the retailer, which is our main research objective. We compare the investing and harvesting incentives during summer and in other seasons and find almost the same results.

[xi] As discussed in endnote iii, we ignore the strategic aspect of the prices of Royal Crown and the Private Label. This can be rationalized by the observation in Table 1 that they have much smaller market shares than Coke and Pepsi and are, therefore, unlikely to significantly influence the wholesale and retail prices of Coke and Pepsi.

[xii] We calculate the sum of squared errors (SSE) from the two models to compare the model performances. The SSE for the dynamic model is 1.065, significantly lower than the SSE for the myopic model that is 1.434, indicating the former model is able to explain better the variation in retail prices.

[xiii] For the simplicity of analysis we assume zero cost and demand shocks for all brands.

[xiv] In the counterfactual simulations, Pepsi is the more preferred brand (compared to Coke) in both consumer segments since it has higher brand preference (even though the difference is not statistically significant); and it displays and features more frequently.

[xv] In order to study whether the retailer's harvesting incentives, as well as the vertical conflict between the retailer and the manufacturers, are driven by the fact that the retailer is a monopolist, we run a simulation study in which we introduce a second retailer and assume that both brands (Coke and Pepsi) are carried by both retailers. We find that retailers continue to have no investing incentives under retailer competition, whereas manufacturers continue to have both investing and harvesting incentives under retailer competition. The reason for this finding is that customers who are inertial toward a brand, without any inertia toward either retailer, can easily switch between retailers from one period to the next. Therefore, neither retailer has an incentive to "invest" on either brand even under retailer competition. We thank an anonymous reviewer for pushing us to run such a simulation study.

[xvi] Seetharaman (2004) shows that in-store display advertising, as well as newspaper feature advertising, serve this role by increasing consumer inertia toward brands in the long run.