Pricing Behavior in Markets with State Dependence in Demand

HAI CHE
Walter A. Haas School of Business
University of California at Berkeley
545 Student Services Building
Berkeley, CA 94720-1900
Email: haiche@berkeley.edu
Phone: (510)643-8918

P.B. SEETHARAMAN
Jesse H. Jones Graduate School of Management
Rice University
P.O. Box 2932
Houston, TX 77252-2932
Email: seethu@rice.edu
Phone: (713)348-6342

K. SUDHIR
Yale School of Management
135 Prospect St
PO Box 208200
New Haven, CT 06520
Email: k.sudhir@yale.edu
Phone: (203)432-3289

September 2004

*We thank all faculty members at the John M. Olin School of Business, most especially Chakravarthi Narasimhan, for guiding us on the scope and direction of this project over the past couple of years. We also thank Tat Chan and Chuck Moul for their comments on earlier versions of this paper, which is based on chapter 1 of the first author’s doctoral dissertation at Olin. Technical Appendix is available upon request. Please send comments to: Hai Che, Assistant Professor, Haas School of Business, University of California at Berkeley, Berkeley, CA 94720-1900. Email address: haiche@berkeley.edu
Pricing Behavior in Markets with State Dependence in Demand

September 2004

Abstract

Marketing researchers have documented significant biases in estimated price elasticities from brand choice models that ignore state dependence effects. Yet empirical research on firm behavior has thus far ignored the effects of state dependence. In this paper we investigate the supply side consequences of ignoring state dependence using household-level data from the cereal category.

Our key results are as follows: We find evidence of both positive (i.e., inertia) and negative (i.e., variety-seeking) state dependence in the cereal category. Ignoring state dependence in demand biases inferences about firm behavior. Specifically, we attribute the observed prices in the cereal category to tacit collusion between manufacturers, when in fact accounting for the effects of state dependence shows that competitive behavior is indeed non-cooperative. We find that firms are forward looking in that they incorporate the effect of their current prices on future profits. However, the relative “bang for the buck,” in terms of improvements in prediction and fit, is much greater from modeling state dependence in demand than from modeling inter-temporal firm behavior on the supply side.

Keywords: Brand Choice, State Dependence, Inertia, Variety-Seeking, Pricing, Scanner Panel Data, Empirical Industrial Organization.
1 Introduction

A Marketing Decision Support System (MDSS) for pricing requires two critical inputs: (1) a model of demand, and (2) a model of competition. Marketing analysts spend considerable effort in calibrating a comprehensive model of demand, but typically much less effort on calibrating the competition model. Often, they assume that price competition is based on the Bertrand Nash equilibrium. But recent research using empirical industrial organization methods has demonstrated that there are substantial deviations (both positive and negative) in prices from the predictions of the Bertrand Nash equilibrium (Kadiyali (1996); Roy et al. (1994); Sudhir (2001a)). Therefore it is critical that the model of competitive response for an MDSS also be empirically calibrated, rather than assumed.¹

Research on inferring competition typically uses aggregate (i.e., store level) data because they are more easily available. Early research in this area therefore used reduced form linear and log-linear models (e.g., Vilcassim et al. (1999); Kadiyali et al. (2000)). Berry et al. (1995) enabled researchers to estimate the random coefficients multinomial logit (RC-MNL) that accounted for unobserved household heterogeneity with just aggregate store level data. Since then, the RC-MNL model has become very popular in the literature when using aggregate data (Dube et al. (2002)). But the use of aggregate data has restricted researchers from taking into account other well-known drivers of demand that have been discovered using household level data.

A particularly important omission is the effect of state dependence on demand. State dependence arises due to two behavioral phenomena: inertia and variety seeking. While inertia refers to a household repeat-purchasing a brand due to habits formed out of past consumption experiences, variety seeking refers to a household switching from one brand to another on account of satiation with attributes consumed in the past. Using household level data, Allenby and Lenk (1994) and Keane (1997) show that unobserved heterogeneity is overestimated when state dependence is omitted. Moreover, disentangling state dependence and heterogeneity is critical to accurately infer price elasticities (e.g., Roy et al (1996); Seetharaman (forthcoming)) and competitive market structure (Erdem (1996)).

Our goal in this paper is to shed insights on whether the hitherto unaccounted effects of state dependence adversely affect the empirical inference of competitive response, and if so, what aspects of state dependence effects are critical for inclusion in an MDSS. Specifically, we investigate the following

---

¹Some researchers estimate reaction functions to calibrate competitive response (e.g., Leeflang and Wittink (1992)). For a detailed discussion on the pros and cons of using the reduced form reaction function approach versus the structural empirical industrial organization approach, see Kadiyali et al. (2001).
three research questions about firm behavior, given their importance in calibrating a model of competitive response for an MDSS. Since it is not possible in general to disentangle state dependence from heterogeneity with aggregate data, we use household level data to help disentangle state dependence effects from heterogeneity.

1. Are inferences about price competition biased if we ignore state dependence?

   When household choices are state dependent (i.e., past brand choices affect current choice), households are found to be less price sensitive (Seetharaman et al. (1999)). Normatively, firms should raise prices in response to this lower price sensitivity. If indeed firms raise prices for this reason, models that do not account for state dependence will attribute the higher prices to tacit collusion among firms, even when firms are pricing competitively. Thus inferences of competition based on demand models that ignore state dependence could be systematically biased and lead us to use the wrong model of competitive response in an MDSS.

2. To what extent do firms take intertemporal effects of state dependence into account in setting prices?

   When households’ brand choices are state dependent, current brand choices will affect future brand choices. Therefore, firms should normatively take into account the impact on future demand when setting current prices. For example, if there is inertia, firms may lower current prices to lock in customers for the future (Klemperer (1987); Freimer and Horsky (2004)). Past research has found that managers find it difficult to correctly accommodate intertemporal effects in their decision making (e.g., Chakravarti et al. (1979), Meyers and Hutchinson (2001)), but do better when taking into account merely current effects (McIntyre (1982)). It is, therefore, an empirical question as to the extent to which firms account for intertemporal effects of state dependence (i.e., how far do they look ahead?) in setting prices. Answering this question helps us to improve our ability to predict the prices of competitors, which is critical for an MDSS to set a firm’s own optimal prices.

3. Does state dependence impact prices more through its impact on current elasticities or through its intertemporal impact on observed prices?

   If the impact of state dependence on current elasticities is much stronger than intertemporal effects, a simple static model ignoring the intertemporal impact can still be a good approximation
of competitive response. Given the computational difficulties in solving dynamic Nash equilibria, this finding is of practical use in developing an MDSS. This does not mean that a profit maximizing firm should not price in a dynamically optimal fashion to improve its own profits; it only implies that competitor responses can be predicted reasonably accurately without modeling the intertemporal effects. Conditional on static responses by competitors, a firm could still choose its own price in a dynamically optimal fashion.

Apart from addressing these empirical/substantive issues related to calibrating a model of competitive response, this paper contributes to demand side modeling by proposing a new attribute-based specification of state dependence that allows a household’s current brand choice to be a function of product attributes of brands consumed by the household in the past. This specification is in the spirit of stochastic choice models of variety seeking (e.g., McAlister (1982), Lattin and McAlister (1985), Trivedi et al. (1994)), but embeds both inertia and variety seeking. We find that this specification is superior to extant models of state dependence and emphasizes the importance of modeling demand at the SKU level as suggested by Fader and Hardie (1996), rather than by aggregating SKUs with different attributes to a brand level as is often done in the extant literature.

The rest of the paper is organized as follows. Section 2 presents the proposed model of demand and supply and Section 3 describes the estimation procedure. Section 4 discusses the data and Section 5 the empirical results. Section 6 concludes.

2 Model

We model a market with multiple manufacturers who sell to consumers through multiple retailers. We follow the previous literature in assuming that each retailer is a local monopolist (e.g., Besanko et al. (1998); Sudhir (2001b)). This assumption is justified for many packaged goods categories because estimated cross-store elasticities are fairly small in magnitude (Walters (1991)). We present our model in three parts: (1) Demand Model (2) Myopic Supply Model and (3) Forward-looking Supply Model.

Demand

We model demand using a household level logit model. The probability of a household $h$ \( h = \)
purchasing one of \( J \) available brands (denoted by \( j = 1,...J \)) or not purchasing in the category \( (j = 0) \) on a shopping occasion \( t = 1,2,...,T \) is given by:

\[
\theta_{ht} = \frac{\exp(v_{htj})}{\sum_{k=0}^{J-1} \exp(v_{hkt})}
\]

where \( v_{htj} = X_{htj} \cdot \beta_h - \alpha_h \cdot p_{htj} + SD_h \cdot sim_{ij} + \xi_{jt} \) for brand \( j = 1,...J \), and \( v_{h0t} = \delta_h \cdot I_{ht} \) for the outside good. The vector \( X_{htj} \) denotes the set of explanatory variables characterizing brand \( j \), as experienced by household \( h \), on shopping occasion \( t \). This vector includes brand dummies and marketing mix variables – specifically feature and display – excluding price. \( p_{htj} \) denotes the retail price of brand \( j \) at shopping occasion \( t \). \( \xi_{jt} \) denotes a composite (stochastic) measure of unobserved (to the econometrician) characteristics of brand \( j \) at shopping occasion \( t \), that is common across all households. It refers to common demand shocks that affect all households (such as brands’ national advertising campaigns on television, macro-economic conditions etc.) that are not recorded in scanner panel data and, therefore, unobservable to researchers, but are observable by the price-setting firms.

\( sim_{ij} \) is a similarity variable that captures how much perceived similarity brand \( j \) has to the brand \( i \) bought by household \( h \) on its previous purchase occasion and is operationalized as

\[
sim_{ij} = \frac{I_{ij} + \sum_{l=1}^{L} r_l I_{ijl}}{1 + \sum_{l=1}^{L} r_l},
\]

where \( i \) refers to the brand purchased by household \( h \) at its previous purchase occasion, \( L \) stands for the number of product attributes represented among all brands within the product category. \( I_{ij} \) is an indicator variable that takes the value 1 if \( i = j \), i.e., \( i \) and \( j \) are the same brand, and 0 otherwise, \( I_{ijl} \) is an indicator variable that takes the value 1 if brands \( i \) and \( j \) share attribute \( l \) and 0 otherwise, and \( r_l > 0 \) stands for the perceived importance for attribute \( l \) in determining inter-brand similarity.\(^4\) Note that \( sim_{ij} \) is restricted to lie between 0 and 1, and is monotonically increasing in the number of attributes shared by brands \( i \) and \( j \).

The effect of how similarly brand \( j \) is perceived in comparison to the brand \( i \) purchased by household \( h \) at its previous purchase occasion, i.e., \( sim_{ij} \), can be positive or negative, depending on whether state dependence effects are positive or negative respectively. For example, if a household is inertial, i.e., \( SD_h > 0 \), the effect of purchasing a brand similar to the one previously purchased will be to increase utility. Conversely, if a household is variety seeking, i.e., \( SD_h < 0 \), the effect of

\(^3\)For model exposition, we will assume that all \( J \) brands are always available, but brand unavailability is taken into account in the estimation.

\(^4\)The above-mentioned specification for the unobserved variable \( sim_{ij} \) restricts it to lie within a pre-specified range (i.e., [0,1]). This enables us to separately identify it from the unobserved state dependence parameter \( SD_h \), which is unrestricted in magnitude. Since \( sim_{ij} \) is defined relative to the brand \( i \) chosen by the household at its previous purchase occasion, its value remains unchanged from one shopping trip to the next if the household does not make a purchase in the product category.
purchasing a brand similar to the one previously purchased will be to decrease utility. We allow
the state dependence parameter, $SD_h$, to be a function of household-specific demographic variables:
$SD_h = \gamma_{0h} + DEMO_h \cdot \gamma_{1h}$, where $DEMO_h$ is a row vector of demographic variables (such as family
size, income etc.) characterizing household $h$, $\gamma_{1h}$ is the corresponding column vector of parameters,
and $\gamma_{0h}$ is the baseline state dependence level of households in the market. This allows the estimated
degree of inertia or variety seeking to be an explicit function of household characteristics. We accom-
modate unobserved heterogeneity by allowing household specific parameters to follow a joint, discrete
distribution (as in Kamakura and Russell (1989)).

The market share of brand $j$ at period $t$ is given by $S_{jt} = \frac{\sum_{h=1}^{H} \theta_{hjt}}{M}$ where $\theta_{hjt}$ is the unconditional
brand choice probability of household $h$, taking its lagged brand choice $i$ as given and integrating over
the unobserved heterogeneity distribution.

Myopic Supply Model

We develop a supply model where firms set prices “as if” they were myopic taking into account
only current demand (and its dependence on past demand through state dependence) and cost condi-
tions. Since we wish to empirically infer the nature of horizontal strategic interactions between manu-
facturers and vertical strategic interaction between each manufacturer and the retailer, we estimate
a menu of games allowing for different combinations of horizontal and vertical strategic interactions
and pick the best fitting model. We allow for two types of horizontal interactions between manufactur-
ers: (1) Bertrand, and (2) Collusive$^5$; and two types of vertical strategic interactions (1) Manufacturer
Stackelberg, and (2) Vertical Nash.

The retailer sets the vector of retail prices $p_t = (p_{1t},...,p_{Jt})$ in week $t$ to maximize category profits.
i.e., $\max_{p_t} \pi_R = \sum_{j=1}^{J} (p_{jt} - w_{jt})S_{jt}(p_t)M$ where $J$ is the number of brands in the category, $w_{jt}$ is the
wholesale price of brand $j$ at time $t$, $S_{jt}(p_t)$ is the market share of brand $j$ at time $t$ and $M$ is the size
of the retailer’s local market.

Suppose there are $F$ firms, each of which produces some subset, $F_j$, of the $j = 1,...,J$ different
brands in the product category. A manufacturer pricing under Bertrand competition sets wholesale
prices to maximize total profits from its product line as follows: $\pi_f = \sum_{j \in F_j} (w_{jt} - mc_{jt})S_{jt}(p_t)M$ where
$mc_{jt}$ is the marginal cost of producing brand $j$, and the other terms are as explained earlier. In contrast,
a manufacturer pricing under collusion sets wholesale prices to maximize total profits from all products

$^5$We also tested for intermediate levels of cooperative behavior between manufacturers using the weighted objective
function approach (Sudhir (2001a)).
sold as follows. \( \pi_f = \sum_{j=1}^{J} (w_{jt} - mc_{jt})S_{jt}(p_t)M. \) Following the procedures developed in Sudhir (2001b) and Berto Villas-Boas (2004), we solve the first order conditions of the retailer and manufacturer to obtain the econometric pricing equation below. The derivation is provided in the appendix.

\[
p_t = mc_t + (w_t - mc_t) + (p_t - w_t) = mc_t - (\Phi^r_t G^{-1} \Omega_t)^{-1} S(p_t) + \Phi^{-1}_t S(p_t)
\]

(2)

For the Manufacturer Stackelberg Model:

(1) \( \Phi_t \) is a \( J \times J \) matrix with elements \( \Phi_{jrt} = \frac{\partial S_{rt}(p_t)}{\partial p_{jt}}, j, r = 1, ..., J; \)

(2) \( G_t = \left[ \begin{array}{ccc}
2 \frac{\partial S_{1t}}{\partial p_{1t}} + \sum_{k=1}^{J} (p_{kt} - w_{kt}) \frac{\partial^2 S_{kt}}{\partial p_{jt}^2} & \cdots & \left( \frac{\partial S_{1t}}{\partial p_{jt}} + \Phi_{jrt} \right) + \sum_{k=1}^{J} (p_{kt} - w_{kt}) \frac{\partial^2 S_{kt}}{\partial p_{jt} \partial p_{jt}} \\
\vdots & \ddots & \vdots \\
\left( \frac{\partial S_{1t}}{\partial p_{jt}} + \Phi_{jrt} \right) + \sum_{k=1}^{J} (p_{kt} - w_{kt}) \frac{\partial^2 S_{kt}}{\partial p_{jt} \partial p_{jt}} & \cdots & \left( 2 \frac{\partial S_{Jt}}{\partial p_{Jt}} + \sum_{k=1}^{J} (p_{kt} - w_{kt}) \frac{\partial^2 S_{kt}}{\partial p_{jt}^2} \right)
\end{array} \right] \)

and (3) \( \Omega_t \) is a \( J \times J \) matrix with \( \Omega_{jrt} = \Omega^*_{jr} \times \Phi_{jrt} \), where \( \Omega^*_{jr} = \begin{cases} 1, & \text{if } \exists f : \{r, j\} \subset F_f \\ 0, & \text{otherwise} \end{cases} \) for the Bertrand Equilibrium and \( \Omega^* \) is a \( J \times J \) square matrix with all the elements being 1 for the collusive equilibrium.

For the Vertical Nash Model, we replace \( \Phi^r_t G^{-1}_t = -I. \) and retain the rest of the equations as above where \( I \) is a \( J \) dimensional identity matrix.

Forward Looking Supply Model

We estimate the same set of models of horizontal and vertical strategic interactions as under the myopic model. But in the forward looking model, retailers and manufacturers look ahead and base their pricing decisions not only based on their impact on current sales, but also on the future. To assess how far firms look ahead, we sequentially estimate models where firms look ahead \( T \) periods, \( T = 1, 2, 3... \) and stop when there is no improvement in model fit going from \( T = t \) to \( T = t + 1. \)

We show the estimation equations for a 1 period look ahead model and provide the derivations in the appendix. Derivations and generalizations for \( T > 1 \) periods are in a technical appendix available from the authors.

The retailer’s objective in the one period look ahead model is to set prices such that they maximize category profits from the current period and the next period (discounted by a factor \( \delta \)) i.e., \( \max_{p_{1t}, p_{jt}} V_R = \pi_R(p_t) + \delta \pi_R(p_{t+1}) \). The manufacturer \( f \)'s objective similarly is to maximize \( \max_{w_{1t}, w_{jt}} V_f = \pi_f(w_t) + \delta \pi_f(w_{t+1}) \). Since at this stage, the next period prices are set without look ahead the next period profit
margin will be the same as derived in the previous section. But the current period margins will change to reflect the look-ahead. The supply side pricing equation can then be decomposed as before:

\[
p_t = mct + PCM^\text{myopic}_M(t = 1) - \delta \Delta_1 [PCM^\text{myopic}_M(t = 2)] \\
+ PCM^\text{myopic}_R(t = 1) - \delta \Delta_1 [PCM^\text{myopic}_R(t = 2)]
\]

where \( \Delta_1 = \bar{\theta}_{r2r1} - \sum_{k=1,k \neq r}^J \bar{\theta}_{r2r1} \bar{\theta}_{r2r1} \) is the unconditional (over the unobserved heterogeneity distribution) transition probability for a household that bought brand \( r \) in period 1 continuing to buy the same brand \( r \) during period 2, and \( \bar{\theta}_{r2r1} \) is the unconditional (over the unobserved heterogeneity distribution) transition probability for a household that did not buy brand \( r \) in period 1 switching to buy brand \( r \) during period 2.

If households are inertial (\( SD > 0 \)) and \( \bar{\theta}_{r2r1} - \sum_{k=1,k \neq r}^J \bar{\theta}_{r2r1} > 0 \), it can be shown that retailers and manufacturers will price lower in the first-period than in the myopic case. The intuition for this analytical result (similar to Klemperer (1987)), is that firms have an incentive to lower their current prices in order to lock-in customers for the second period. This incentive to lower price will be moderated by the number of customers who are already locked in to the firm’s brand at the beginning of the first period, i.e., the inertials who bought the firm’s brand in their previous purchase. 6 On the other hand, if households are variety seeking (\( SD < 0 \)) and \( \bar{\theta}_{r2r1} - \sum_{k=1,k \neq r}^J \bar{\theta}_{r2r1} < 0 \), forward looking manufacturers and retailers will price higher than in the myopic case. Not only do firms not have lock-in incentives to lower prices, but they can raise prices because variety seeking customers who switch due to higher current prices are more likely to switch back.7 Since Klemperer (1987) did not model variety seeking in his work, this is a new analytical insight in the literature on state dependence.

When there are heterogeneous households in the market, the actual prices offered by firms in each period will be based on the composition of state-dependent households. In other words, the prices will respond to not only whether inertial or variety seeking households are expected to arrive at the store on a given week, but also the degree of heterogeneity across these households in terms of the

6However, there is another effect. If households are inertial (\( SD > 0 \)) but \( \bar{\theta}_{r2r1} - \sum_{k=1,k \neq r}^J \bar{\theta}_{r2r1} < 0 \), retailers and manufacturers will price higher in the first-period than in the myopic case. This will happen when the effects of inertia are small relative to effects of marketing activities. This does not invalidate Klemperer’s result, since he assumes a duopoly pricing game using a Hotelling line, where the effects of switching cost on consumer demand are always high enough for the two firms to compete more aggressively in the first period. While in the case of oligopoly pricing and multinomial logit demand, the effects of marketing activities on households’ brand choice could dominate the effects of inertia, i.e., \( \bar{\theta}_{r2r1} < \sum_{k=1,k \neq r}^J \bar{\theta}_{r2r1} \).

7The competitive effects, through marketing mix activities, only strengthen the effects of variety seeking on pricing in this case, and the first-period prices will be strictly higher than under the myopic case.
brands bought in the past. This can be one explanation for time-varying prices observed in this product category, and is behaviorally consistent with Freimer and Horsky’s (2004) normative predictions about alternating price promotions by manufacturers in consecutive periods when there is a high degree of inertia.

3 Estimation

We employ a two-stage estimation approach (Newey and McFadden 1994). In the first step, we estimate the demand model using a likelihood-based approach, correcting for potential price endogeneity. While these estimates are unbiased and consistent, they are inefficient since the information in supply equations are not exploited. However the advantage is that demand estimates are unbiased by any potential mis-specification in supply equations (especially, when we wish to choose the best supply specification). Further, due to the large number of observations in household level data, efficiency concerns are not as high as with aggregate data. Conditional on the demand estimates, we compute the margins under alternative supply side specifications. We then estimate a cost function, conditional on the margins under alternative supply models and choose the best fitting supply model on the basis of fit.

Demand-Side Estimation

Recall that a household $h$’s probability of choosing brand $j$ at shopping occasion $t$, $\theta_{hjt}$ is a function of $\xi_{jt}$, which is the common shock among households that is unobserved by the econometrician. Villas-Boas and Winer (1999) argue that since profit maximizing firms take $\xi_{jt}$ into account when setting prices, there is a price endogeneity problem. We therefore need instruments for price to obtain an unbiased estimate of the price coefficient. Denoting $P_{jt}^{\text{Instrument}}$ as instruments for price, we can estimate the pricing equation: $P_{jt} = \gamma_0 + \gamma_1 \cdot P_{jt}^{\text{Instrument}} + \eta_{jt}$. The random errors $\eta_{jt}$ could arise from both cost shocks and demand shocks. So when prices are endogenous, $\xi_{jt}$ and $\eta_{jt}$ should be correlated. Assuming a joint normal distribution of $\xi_{jt}$ and $\eta_{jt}$, and $E(\xi_{jt}|\eta_{jt}) = \rho_j \sigma_{\xi_j} \sigma_{\eta_j}$, testing for price endogeneity is equivalent to estimating whether $\rho_j = 0$ or not.

We estimate the demand model using Limited Information Maximum Likelihood (LIML). The joint likelihood of the household $h$ over the demand and pricing equation is:

$$L_h(\alpha_h, \beta_h, \delta_h, \gamma_0, \gamma_1, \sigma_\xi, \sigma_\eta) = \prod_{t=1}^{T_h} \prod_{j=0}^{J} \left[\theta_{hjt}(\xi_{jt})\right]^{y_{hjt}} \cdot f(\xi_{jt}|\eta_{jt}) \cdot f(\eta_{jt})d\xi_{jt}$$  \hspace{1cm} (4)
where \( y_{hjt} \) \( \begin{align*} &= 1 \text{ if alternative } j \text{ is chosen} \\
&= 0 \text{ otherwise} \end{align*} \) and \( f(\eta_{jt}) \) refers to the density of the distribution of \( \eta_{jt} \) evaluated at the estimated residual from the pricing equation estimation. We note that since scanner panel data are used in the estimation, one explicitly observes the number of no purchases (i.e., purchases of the outside good) during each week instead of having to impute them using some assumed market size.

During estimation, we integrate \( \xi \) out from the likelihood using simulation. Previous studies applying LIML (such as Villas-Boas and Winer 1999) do not allow for heterogeneous demand parameters due to the computational burden involved. In contrast, we allow for different segments of households that differ from each other not only in their propensity for variety seeking or inertia, but also in their responsiveness to the marketing mix and intrinsic preferences for the large number of choice alternatives (>100). We estimate a semi-parametric mixture model of unobserved heterogeneity (Kamakura and Russell 1989), under which the sample likelihood function becomes

\[
L = \prod_{h=1}^{H} \left\{ \prod_{k=1}^{K} [L_{hk} \ast \Pr(k)] \right\} \tag{5}
\]

where \( K \) stands for the number of supports of the discrete heterogeneity distribution, \( L_{hk} \) stands for the likelihood function computed for household \( h \) under the assumption that \( h \) belongs to segment \( k \), and \( \Pr(k) \) stands for the probability of household \( h \) belonging to segment \( k \).

**Supply-Side Estimation**

For estimation we assume that the marginal cost for each brand \( j \), \( mc_{jt} \), is a brand-specific linear function of observable cost shifters, as shown below.

\[
mc_{jt} = \sum_{k}^{} input_{jkt} \cdot \omega_{k} \tag{6}
\]

where \( input_{jkt} \) is the input price\(^8\) at time \( t \) of attribute \( k \) contained in brand \( j \) (including the manufacturer-specific wage rate), \( \omega = (\omega_1, \omega_2, \ldots, \omega_K) \) is a vector of coefficients associated with the \( K \) attributes. We estimate the manufacturer and retailer margins conditional on alternative supply specifications, based on the demand estimates. Subtracting out these margins from prices, we then estimate the cost equation as follows:

\(^8\)In addition to input prices, we also include a manufacturer dummy. We will discuss in detail later how we construct brand-specific cost data.
\[ p_{sjt} - \hat{PCM}_{M,sjt} - \hat{PCM}_{R,sjt} = \gamma_0 + \gamma_s \cdot \text{Store}_s + \gamma_m \cdot \text{Manufacturer}_j + \sum_k \text{input}_{jkt} \cdot \omega_k + \eta_{sjt} \]  

(7)

where \( \eta_{sjt} \) is a random error which is assumed to be normally distributed with mean zero and covariance matrix \( V_{\eta} \). \( \hat{PCM}_{R,sjt} \) and \( \hat{PCM}_{M,sjt} \) are retailer and manufacturer margins computed based on the estimates of demand parameters conditional on the supply-side pricing game.

In order to estimate the supply model, we write the empirical likelihood function as the likelihood of observed prices given the distribution assumptions on \( \eta_{sjt} \). Denoting \( g(.) \) as the marginal (normal) density of \( \eta_{sjt} \), this likelihood function is written as:

\[ L = \prod_{t=1}^{T} \left[ \prod_{j=1}^{J} g(\eta_{sjt}) \right] \]  

(8)

The two-period supply-side model is estimated as follows (the same backward-induction procedures are used to estimate three- and four-period supply models): 1) Estimate the second-period margins using the demand-side estimates (calculate the margins for each brand at each week between weeks 2 and 104); 2) Plug the estimated second-period margins into the first-period pricing equation to obtain the estimates of first-period margins (between weeks 1 and 103); 3) Estimate first-period pricing equations. We use a simulation study (details available in a technical appendix), to check that our proposed estimation approach estimates both the demand and supply models quite well.

### 4 Data

We need household level data, as opposed to aggregate data, in order to identify state dependence. We use IRI scanner panel data in the breakfast cereal category, drawn from urban and suburban areas in a large U.S. city, covering a 104 week period from June 1991 to June 1993. In the suburban market 483 households purchase 24,971 units\(^9\) of cereals in 68,432 shopping trips, while in the urban market 480 households purchase 23,011 units of cereals in 71,247 shopping trips\(^10\). In all, 117 brands are purchased by these households. If households buy cereal on different days of the same week, we treat each purchase separately for estimation. Some descriptive statistics about marketing variables are

\(^9\)A unit refers to a regular package size in the product category.

\(^10\)We estimate the demand models separately for these two markets. We report only the estimation results for the suburban market data. Results for the urban market are available upon request from the authors.
reported in Table ???. Also included in the dataset are households’ demographics, such as family size, income, household head age, kids’ age, etc.

Since we model state dependence in attribute space, we require data on product attributes. We collected such information by visiting supermarkets and reading product labels, and also from manufacturers’ websites. When brands have been discontinued, we collected attribute information from cereal reviews found in various sources.

We collected weekly cost data for different ingredients (corn, wheat, rice and oatmeal), sugar, fruits (apple, grape) from the Department of Agriculture and weekly labor wage data for 5 different states (corresponding to locations of cereal manufacturer plants) from the U.S. Current Population Survey Annual Earnings File. A brand’s cost is based on its ingredients and state of production. For example, Rice Krispies’ input cost is based on the cost of rice and sugar and Michigan’s labor wage rate.

5 Empirical Results

We report the results of the empirical analysis in two parts: (1) Tables 1 - 2 report the demand-side results.\textsuperscript{11} and (2) Tables 3 - 5 report the supply-side results.

5.1 Demand-Side Results

We estimate heterogeneity in a latent class framework, by increasing the number of supports for the heterogeneity distribution until there is no improvement in fit. To assess the role of state dependence, we evaluate three models: (1) a multinomial logit model with no state dependence (McFadden (1980))\textsuperscript{12}, (2) a multinomial logit model with a lagged brand choice dummy to capture state dependence (Seetharaman et al. 1999), and (3) a multinomial logit model with the proposed attribute-based similarity specification of state dependence. We find that the three-support heterogeneous version of the proposed attribute based state dependence model has the best fit (see Table 1), in terms of both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). This demonstrates that the proposed state dependence model is superior to extant approaches of modeling state dependence. It also suggests the importance of modeling demand at an SKU level in categories where a brand has a number of SKUs with widely varying attributes (as argued by Fader and Hardie 1996).

\textsuperscript{11}The estimation was done in the C++ programming environment on a personal computer employing a Pentium 2 GHz chip. Convergence times varied from one day (for the static models without endogeneity correction) to one week (for intertemporal models with endogeneity correction).

\textsuperscript{12}This model is also called the zero-order Multinomial Logit model.
As found in previous research (Keane 1997; Seetharaman et al. 1999), we note that the magnitude of the estimated price coefficient, $\alpha$, is overstated when state dependence effects are ignored, i.e., $|\alpha_{no\ SD}| > |\alpha_{with\ SD}|$ (see Table 1). This suggests that the optimal Bertrand-Nash prices will be systematically lower if one bases it on demand estimates that ignore state dependence. We revisit this issue when we discuss the supply side estimates.

We further discuss only the results of the three support heterogeneous version of the attribute-based state dependence model, since it has the best fit. Table 2 contains estimates of this model with and without price endogeneity accounted for. We use prices from another market as instruments, as in Nevo ((2001)) to account for price endogeneity. The justification for the use of these instruments are based on arguments developed by Hausman (1997). However, these instruments will only partially correct the endogeneity bias if national advertising creates common demand shocks across markets. A Hausman (1987) test rejects the null hypothesis that prices are exogenous. As expected, the magnitude of the estimated price coefficient increases with endogeneity correction. Note that including state dependence reduces the magnitude of the estimated price coefficient, while correcting for endogeneity increases it. The remaining coefficients remain fairly similar between the two models.

We find three segments of size 27%, 16% and 56%. The largest support (with a mass of 56.8%) is observed to be extremely price sensitive, while the remaining two supports are relatively price insensitive. We report aggregate summary statistics for $SD$ (e.g., mean, median) within each of the three supports. To the extent that supports 1, 2 and 3 are observed to have negative, positive and insignificant state dependence coefficients respectively, we will refer to them as the variety seeking, inertial and zero-order segments respectively. Support 2, accounting for about 16% of the market, is strongly inertial (second column, with the mean of $SD$ being 5.08). Support 3, accounting for about 56% of the market, is relatively zero-order (third column, with the mean of $SD$ being 0.79). The remaining 27% of the consumers in Support 1 are variety seeking (first column, with the mean of $SD$ being -1.35).

Thus we empirically document the existence of a variety seeking segment in the market using scanner

---

13 Given a discrete number of segments, the comparison is based on the market-level price coefficient, which is computed as a weighted average of support-specific price coefficients. We also find that the price elasticities are systematically lower (discussed later) when we account for state dependence.

14 Unfortunately, instruments for prices used by other researchers do not work in our context. Lagged prices are not valid instruments because unobservable demand effects affect both current prices and future prices through state dependence. Attribute-based BLP type instruments do not work because the set of products and their attributes remain virtually unchanged during the period of analysis, providing no variation in these instruments. Note that displays and features are also potentially endogenous. Other market features and displays are not effective instruments; in the absence of effective instruments, and since our focus is on prices, we abstract away from this issue in the paper. Note that Sudhir (2001b) rejects feature endogeneity once price endogeneity has been accounted for.

12
panel data for the first time. This is in contrast to previous research (e.g., Erdem 1996, Seetharaman, et al. 1999), who find that, almost without exception, households tend to be inertial.

We profiled the demographic characteristics of members belonging to the different segments. From the profiling results, it is clear that variety-seeking households are, on average, larger than inertial households. This is consistent with the intuition that variety-seeking is partly an artifact of aggregating brand choices made by household members who are heterogeneous in their tastes.\(^\text{15}\) However, we find some differences between the relationship between demographics and state dependence across the different segments. The sign on the family size variable is positive within segment 1, but negative within segment 2, which suggests an interesting second-order effect, i.e., within each segment, larger families show less state dependence (i.e., more zero-order behavior) than smaller families. To the extent that previous studies have restricted demographic effects to be equal across segments (not expecting them, a priori, to be different), such second-order effects have not been estimated before.\(^\text{16}\) The profiling also showed that, on average, variety-seeking households have higher-income than inertial households, but as income increases within each segment, households are found to become less state dependent in their brand choices.\(^\text{17}\) Further research (possibly using laboratory experiments) is required to understand the behavioral rationale of these curious second-order effects, which indicate non-linear dependencies of state dependence on demographics. This also raises the intriguing question of whether insignificant demographic effects in extant brand choice models (that restrict demographic effects to be equal across segments) could be an artifact of such opposing effects that ‘cancel each other out’.

To explore whether variety seeking effects survive beyond intra-household heterogeneity, we re-estimate our proposed brand choice model using purchase data of single-member households only (205 out of 963 households). We still uncover a segment of variety-seekers, although the size of the variety seeking segment is smaller (10.2%). To the extent that the estimated degree of variety seeking in this market becomes only mitigated, but not eliminated, in an analysis of single-member households, we believe that variety seeking indeed characterizes brand choices of some households in the breakfast cereal category.

\(^{15}\)Technically, this should be considered as intra-household heterogeneity in tastes and not variety seeking (Kahn et al. 1986). However, since we do not observe consumption, it is impossible to disentangle intra-household heterogeneity and variety seeking (as with all previous research in this area). We explore this issue in more detail by restricting attention to single member households.

\(^{16}\)Furthermore, we find that zero-order households, i.e., those in segment 3, become more inertial as their family size increases, which is puzzling, given the signs of these effects for segments 1 and 2.

\(^{17}\)Again, zero-order households of segment 3 behave differently compared to this, by becoming more inertial as their income increases.
A key element of our proposed brand choice model is that state dependence in households’ brand choices over time is driven by observed product attributes. In other words, inertial households choose brands with similar attributes, while variety seeking households choose brands with dis-similar attributes, on successive purchase occasions. We see that sugar-level, fruit-nut and product-type (kids, family, adult/family) all have positive and significant effects on perceived inter-product similarity between alternatives. Ingredients (corn, wheat, rice and oatmeal) and fiber do not significantly affect perceived inter-product similarity. These findings suggest that promotional messages by marketers using easily communicable taste characteristics, such as sugar level and the presence of fruits and nuts, will be more effective in influencing perceptions and choices of households than using less easy-to-discriminate characteristics, such as ingredients or fiber. For example, Kellogg’s could target variety seeking households (if they can be identified using appropriate data sources) with a portfolio of coupons across brands that are diverse on discernible product attributes to ensure that variety seeking households stay within Kellogg’s franchise.

In terms of household preferences for attributes, we see that variety-seekers prefer high-fiber and fruit-nut cereals to sweetened cereals, while inertials prefer high-fiber, and zero-order households prefer sweetened and fruit-nut cereals. Overall, product attributes are observed to significantly explain households’ brand preferences in this product category. Both inertial and variety seeking segments of households exhibit lower price sensitivity than the zero-order segment, and also have positive coefficients for feature and display. Product inventory increases the utility of the outside good and reduces purchase incidence only for variety seekers; for the other two segments it has no significant effect. The behavioral literature (e.g., McAlister (1982)) has shown that variety seeking households satiate on the product attributes in existing product inventory, and switch brands accordingly; our findings suggest that variety seeking households also respond strongly to product inventory in their category purchase decisions.

We calculate price elasticities of different brands based on the estimated demand parameters. Cheerios, a large share brand in the product category, has an own-price elasticity of -1.91. However, Quaker Oats, a similar brand in terms of product attributes, but with much smaller market share, has a much higher own-price elasticity of -7.8. In fact, large share brands are observed to have smaller own-price as well as cross-price elasticities compared to small share brands, which is consistent with their strong market positions.\(^{18}\) We uncover a systematic reduction in price elasticities after we account for state

\(^{18}\)This shows that the estimated extent of unobserved heterogeneity across households is strong enough to relax the
dependence. On average, the two elasticities differ by 10-20\%. For example, Cheerios’s own price elasticity is estimated to be -2.15 if one ignored state dependence, compared to -1.91 after accounting for state dependence; the difference is statistically significant at the 5% level. We elaborate further on the supply side effects of this bias in the next section.

5.2 Supply-Side Results

We estimate three types of pricing games: static (i.e., ignoring state dependence in demand), myopic (i.e., accommodating state dependence in demand, but ignoring inter-temporal behavior of firms), and inter-temporal (i.e., accommodating both state dependence in demand and one-period look ahead). For each type of pricing game, we estimate [2 VSIs (Vertical Nash, Stackelberg) \times 2 HSIs (Bertrand, Collusive)] = 4 types of interactions between firms. Therefore, we estimate 4 (interaction types per pricing game) \times 3 (pricing games) = 12 supply-side specifications, each of which would generate different predicted values of manufacturer and retailer margins. Then, based on the best-fitting model among these 12, we also estimate two period and three period look ahead versions of the games to investigate how many periods firms look ahead into the future when pricing.

Table 3 shows the maximized log-likelihood values associated with the various supply-side specifications. We report the results for the three types of pricing games under three columns: (1) static, (2) myopic, and (3) inter-temporal (one period look-ahead). The Vuong test for non-nested models is used to select the best-fitting supply-side specification. The Vuong test statistics are reported below the log-likelihood values. We now use the supply side results to address the three research questions that we set out to answer.

Are inferences about price competition biased if we ignore state dependence?

Consider the differences between the results for the static and myopic models. For the static model (that ignores state dependence), the collusive pricing interactions between manufacturers outperform Bertrand pricing interactions, whereas the result reverses for the myopic model (that accounts for state dependence). Thus, one would wrongly attribute the seemingly high observed prices in the cereal category to collusion between manufacturers if one ignored the effects of state dependence.\footnote{This result obtains even if we predict market shares of brands based on steady-state probabilities of the Markov chain characterizing households’ brand choices, instead of basing them on households’ lagged choices that are observed in the data.}

\footnote{This result obtains even if we predict market shares of brands based on steady-state probabilities of the Markov chain characterizing households’ brand choices, instead of basing them on households’ lagged choices that are observed in the data.}
Prices of breakfast cereals were known to be very high, vis-a-vis their marginal costs of production, during our period of study. Recall that ignoring state dependence leads us to exaggerate the level of household price sensitivity. Thus, the normative Bertrand prices without accounting for state dependence are lower than the observed high prices of cereals in the data, leading us to conclude that firms are tacitly colluding in this market. However, after accounting for state dependence on the demand side, the observed prices are consistent with Bertrand competition. Thus we demonstrate that the inferred model of competitive response in an MDSS can be biased by the omission of state dependence in demand.

This finding also has implications for regulatory agencies such as the FTC that routinely try to determine the nature of competition in product markets, while investigating anti-trust charges. The cereals market was subject to charges of anti-competitive practices during our period of study, but our analysis suggests that without modeling state dependence, we may wrongly interpret that firms behaved collusively.20

To what extent do firms take intertemporal effects of state dependence into account in setting prices?

Our second research question is related to the issue of whether firms are forward looking and if so, to what extent when setting prices. As discussed earlier, while normatively firms should be forward looking (e.g., Klemperer (1995)), behavioral research suggests that managers have difficulty taking into account intertemporal effects (e.g., Chakravarthi et al. (1979)). We use \( \chi^2 \) tests to compare the goodness-of-fit between the myopic pricing model and the one period look-ahead inter-temporal pricing model, since the former model is nested within the latter (Table 3). We find that the one period look-ahead model fits the data better than the myopic pricing model, suggesting that firms (manufacturers and retailers) account for the effects of their current pricing decisions on next-period profits. However there is little improvement in the two-period and three-period look ahead models. In fact the fit reduces when we use three-period look ahead. We conclude, therefore, that the one-period look ahead pricing model is an adequate characterization of inter-temporal pricing behavior of firms in this product category. Recent studies in experimental economics (Camerer et al. (2004)) suggest that even strategic players endowed with highest thinking ability rarely think more than two steps ahead. Our findings provide

---

20Consumers kept complaining about high cereals prices throughout the 1980s and in the early 1990s. Critics charged the industry with possible cooperation among manufacturers in their pricing strategies leading to a shared monopoly. The Federal Trade Commission (FTC) investigated, but eventually dropped, the allegation of cooperation between firms. Interestingly, the top four cereals manufacturers then slowly engaged in a “price war,” beginning with General Mills’ initial price cuts in 1994 (Scott Bruce and Bill Crawford Cerealizing America: The Unsweetened Story of American Breakfast Cereal, Faber & Faber, 1995 and also PBS, August 8, 1996).
some support to this finding using market data.

*Does state dependence impact prices more through its impact on current elasticities or its intertemporal impact on observed prices?*

Given that we find that state dependence affects current elasticities and that firms take into account intertemporal effects, we address the question of which of these two effects has greater relative impact in being able to predict observed prices accurately. This question is of interest from a marketing engineering perspective. If approximations to complex behavior are needed from a computational perspective, what is the best approximation? We address this question in two ways: First, we assess how much relative percentage improvement in in-sample likelihood is obtained by including the two effects of state dependence. Second, we address the relative gains in accuracy that can be obtained in predicting prices through the two effects of state dependence.

As far as the fit in the estimation sample is concerned, the log-likelihood for (1) the best fitting myopic model without state dependence is 30,231, (2) the best fitting myopic model with state dependence is 32,196 and (3) the best fitting inter-temporal model is 32,401. The total percentage improvement in likelihood by including both state dependence and inter-temporal effects is 7.2%, of which 6.5% is due to modeling state dependence, and 0.7% is due to inter-temporal effects. Thus inter-temporal effects contribute to only about 9.7% of the improvement in likelihood, while 90.3% of the improvement is through the modeling of state dependence.

Second, we investigate the improvement in predicting observed prices. We use the demand and cost estimates from the supply model to predict prices for the 117 brands in the dataset. We report the mean squared errors in the price predictions in Table 4. Relative to the one-period model without state dependence, the one-period pricing model with state dependence reduces the mean squared errors (MSE) of prediction by 22.34%. Relative to the one-period model without state dependence, the one-period look ahead pricing model with state dependence and allowing for forward looking behavior reduces the MSE of prediction by 23.67%. The incremental reduction in MSE by allowing for forward looking behavior over and above state dependence is just 1.71%. Therefore, 94% (22.34/23.67) of the reduction in MSE is obtained by accounting for state dependence, with only 6% of the reduction obtained by incorporating forward looking behavior. This suggests that state dependence effects on current elasticity are relatively more important than inter-temporal effects in predicting observed prices. Thus, based on both model fit and prediction, our results show that modeling *household purchase dynamics*

---
21This result obtains for both in-sample and out-of price prediction.
through state dependence is more critical than modeling firm pricing dynamics through forward looking behavior.

From a practical point of view, in developing decision support systems, this implies that there is much greater “bang-for-the-buck” in computing “near-optimal” prices by modeling state dependence than by modeling pricing dynamics. This is a useful empirical insight given that the computational complexity of solving dynamic oligopolistic pricing games (with 117 brands involving 6 manufacturers and strategic retailer as in our application) are far greater compared to solving a static game.

**Costs**

Table 5 presents the full estimation results for the best-fitting pricing models: one period look ahead pricing models accounting for state dependence. The Bertrand Nash models fits the data best. General Mills has a relative cost advantage compared to other manufacturers. The retailer that stocks the largest number of cereal brands in the store (92 brands per week on average) has a relative cost disadvantage compared to the other two retailers (83 and 87 brands per week on average respectively). Prices of corn, rice and oats are significant in the cost equation. A variance decomposition of observed prices shows that raw materials contribute to about 9.7% and price-cost margins contributes to 66.5% of the variance in observed retail prices\(^ {22} \).

It is of interest that cost estimates are fairly identical between the one-period and two-period pricing models. This finding runs counter to one of the implications of the illustrative example in Berry and Pakes (2000), that cost estimates will be biased toward zero if we ignore the inter-temporal effect. Investigating a monopolist’s pricing problem for a single experience good, Berry and Pakes (2000) show, through Monte Carlo simulation, that there is a negative correlation between cost variables and the inter-temporal effect that is omitted in the myopic model. The intuition for this is that since customers with experience are more likely to buy again in the future, a monopolist taking into account the benefit of future purchases would not raise prices as much as a myopic firm would in response to a cost increase. Since the cereal market is inertial overall, and an inertial good has the same demand properties as an experience good, one would also expect a negative correlation between the cost variables and the omitted inter-temporal effect. However in the context of oligopolistic multi-product price competition as in our case, the effects of inertia may not be large enough to dominate the effects of marketing activities on household’s brand choice. The net effect of inertia and marketing mix is given by

\[
\Delta_1 = \bar{\theta}_r R_r - \sum_{k=1,k \neq r}^J \bar{\theta}_r R_k \quad (\text{as in equation (3)}) \quad \text{In the case of extreme inertia, in which consumers do}
\]

\(^ {22} \)The remaining 24% of the variance is due to unobserved cost shocks.
not respond to marketing mix, $\Delta_1 = 1$, because $\vartheta_{r2r1} = 1 > \sum_{k=1,k\neq r}^{J} \vartheta_{r2k1} = 0$. In most cases, consumers do respond to marketing activities and the sign of $\Delta_1$ suggests whether inertia effects or competitive marketing mix effects dominate. When inertia effects dominate, $\vartheta_{r2r1} > \sum_{k=1,k\neq r}^{J} \vartheta_{r2k1}$ and therefore $\Delta_1 > 0$; while when marketing-mix effects dominate inertia effects, $\vartheta_{r2r1} < \sum_{k=1,k\neq r}^{J} \vartheta_{r2k1}$ and therefore $\Delta_1 < 0$. In our analysis, we find that the estimate of $\Delta_1$ is positive in some periods and negative in other periods, suggesting that neither the inertia effects nor the competitive marketing mix effects dominate in this market. This leads to very low correlation between the inter-temporal effect and cost variables. Therefore we do not find any significant bias in the factor cost estimates between the myopic and inter-temporal models.

6 Conclusion

Marketing researchers spend considerable effort on calibrating demand models, but pay far less attention to calibrating firm side behavior. Yet recent research has shown that the common assumption of Bertrand competition that is used in many marketing decision support systems may be violated in many markets. While a rich literature in marketing, based on demand-side analyses, has documented the effects of state dependence on the estimated market structure among brands within a product category, the implications of such state dependence effects on supply-side inferences are not well understood.

In this paper, we investigate the effects of state dependence on firm pricing decisions and the pricing equilibrium. Towards this end, we propose a structural model of demand and supply accounting for state dependence effects. Our demand model accounts the effects of both inertia and variety seeking, allowing them to be functions of perceived inter-product similarities between brands along observed product attributes. On the supply side, we formulate and estimate an inter-temporal model of pricing that takes into account the effects of state dependence in demand. We estimate the proposed model using scanner panel data in the breakfast cereal category.

Our four key empirical findings are summarized below.\textsuperscript{23} The first is an insight on modeling state dependence itself, and the next three are related to insights about firm behavior.

1. Previous studies using scanner panel data only uncover inertial behavior, but we find both inertial and variety seeking segments in the breakfast cereal category. Our study is the first to uncover

\textsuperscript{23}We also estimated this model using data on the ketchup category and the substantive/empirical claims that we make in this paper generalize to this category as well. The key difference is that the ketchup category has no variety seeking segment. Details of this analysis are available from the authors.
variety seeking using scanner panel data. We conjecture that previous studies’ inability to estimate variety seeking is because they aggregate SKUs to the brand-level and use the top five or so brands only in the estimation. In the breakfast cereal category, since SKUs are typically marketed as separate brands, we do not face the aggregation issue. Previous research has also ignored the influence of product attributes on perceived inter-product similarities (which we found to be important) because they aggregated disparate SKUs with different attributes into a single brand. Although, estimating the model with a large number of brands (SKUs) is computationally challenging, our study demonstrates that this can be important in making correct inferences about consumer demand.

2. We find that the collusive pricing model outperforms the Bertrand pricing model if state dependence effects in demand are ignored, while the Bertrand model performs better if state dependence effects are accommodated. This demonstrates that the inferred model of competitive response can be seriously biased when state dependence is ignored, with clear implications for developing an empirically validated MDSS. The result is also of importance to antitrust regulators, because our analysis argues that the high prices in this market are an outcome of non-cooperative pricing in the face of low price elasticities, rather than being a collusive outcome.

3. We find that an inter-temporal pricing model, that allows firms to look one-period ahead while making their pricing decisions, fits the data better than does a myopic pricing model. Thus, firms appear to take into account the impact on future demand of setting prices for their brands in the current period, even though behavioral work has suggested that decision makers have difficulty in accounting for intertemporal effects in their decision making. Nevertheless, the gains from employing two- or three-periods look-ahead are minimal or negative. To accurately take into account the dynamics from several periods ahead, firms have to predict both household and competitor behavior far into the future, a very difficult task. Our finding is consistent with the hypothesis of boundedly rational firms. In this sense, it is indeed plausible that there are gains from looking one period ahead, but not much farther into the future.

4. Modeling state dependence considerably reduces the errors in predicting observed prices of brands and improves the fit of the model. The gains from modeling inter-temporal pricing are

---

24 Trivedi, Bass and Rao (1994) uncover variety seeking in the video rental market using survey data.
25 Further research is necessary to investigate this issue further.
far more limited. This suggests that empirical industrial organization applications can benefit significantly from using household level data to model state dependence in markets where state dependence is likely to exist. However, modeling inter-temporal pricing seems to offer limited benefits to warrant the complexity of solving dynamic games in this context.

In this paper, we take an important first step in investigating the role of demand side mis-specification on the inference of supply-side behavior. Specifically, we investigate the role of state dependence. However, data limitations as well as computational difficulties lead us to abstract away from many issues that should be addressed in future work. For example, since state dependence effects are likely to be strong only in short intervals (such as weekly data), it would be useful to investigate the implications for the inferred model of competitive response given the periodicity of data. For example, Nevo (2001) uses aggregate data on breakfast cereals at quarterly intervals and finds that pricing behavior is consistent with Bertrand pricing even though he does not model state dependence. One reason for this could be that state dependence effects become weak at quarterly frequency. A systematic investigation of the effects of periodicity of the available data on inference of pricing behavior is necessary to resolve this issue. This would require the availability of weekly data on multiple geographic markets in a single category, so that we have sufficient degrees of freedom to still estimate the model even after aggregating the data to bi-weekly, monthly and quarterly levels. It is also possible that while retailers set their prices weekly, manufacturers have a different periodicity of pricing. Modeling the pricing decision at the frequency at which those decisions are made in the real world would be important in future work.

In the absence of data on wholesale prices, we followed the approach developed by Sudhir (2001b) and Berto Villas-Boas (2004) in modeling manufacturer-retailer interactions. With wholesale price data, it would be possible to investigate more flexible models of interactions between manufacturers and retailers (e.g., Kadiyali et al. 2000). One possibility would be to empirically estimate a model of bargaining between manufacturers and retailers.

More broadly, several other demand phenomena have also been ignored in extant research on firm-side behavior: these include stockpiling (e.g., Erdem et al. (2003), Hendel and Nevo, (2004)), consideration sets (e.g, Andrews and Srinivasan (1995)) etc. The analytical framework that we develop in this paper using household level data (as opposed to the extant research using aggregate data) should enable researchers to investigate the importance of accounting for these factors. Researchers have also discovered changes in consumers’ price and promotion sensitivity over time, as a function of past pro-
motions and advertising activities (e.g., Kopalle et al. (1999)). An investigation of whether and how firms respond to such dynamic demand changes would also be of interest in future research.

A key finding in our paper is that firms do only one period look-ahead consistent with a bounded rationality perspective. Even here, taking into account the current period effects of state dependence is far more critical in predicting competitive response than intertemporal effects. But it is possible that firms take into account intertemporal effects reasonably well in the context of inventory management systems etc., where computer based solutions are extensively used. Ailawadi et al. (forthcoming) find that their predictions about competitive response are improved when inventory dynamics (such as forward buying etc.) are included. However they do not investigate the relative benefits of current period effects and intertemporal effects. A systematic investigation about the relative importance of dynamics for predicting observed prices in other contexts where dynamics are important (e.g., consumer stockpiling, retailer forward buying) would be insightful in building a MDSS.

**Appendix:**

**Myopic Pricing Model**

Suppose that each manufacturer and the retailer are myopic in pricing and only maximizes current-period profits. Here, we will derive the myopic model with the assumptions of a Bertrand-Nash game among multiple manufacturers and a Stackelberg game between the manufacturers and the single retailer.

**The Retailer**

The retailer takes the wholesale prices as given in the Stackelberg game, and acts like a monopolist in pricing the whole category.

At any period $t$, the retailer’s profit maximization problem is, 
$$
\max_{p_{1t}, \ldots, p_{Jt}} \pi_{Rt} = \sum_{j=1}^{J} (p_{jt} - w_{jt})s_{jt}(p),
$$
where $s_{j}(p)$ is the market share of brand $j$, which is a function of the prices of all brands, $M$ is the size of the market. The first-order condition is $rac{\partial \pi_{Rt}}{\partial p_{j}} = 0$, and in details, $s_{jt}(p) + \sum_{r=1}^{J} (p_{Rt} - w_{Rt}) \frac{\partial s_{rt}(p)}{\partial p_{j}} = 0$.

Defining $\Phi_{jrt} = -\frac{\partial s_{rt}(p)}{\partial p_{jt}}$, $j, r = 1, \ldots, J$, and using the matrix form, the retailer’s mark-up is $p_{t} - w_{t} = \Phi_{t}^{-1}s_{t}(p)$.

**The Manufacturer**

Manufacturers’ markups can be solved for explicitly by defining $\Phi_{jrt} = -\frac{\partial s_{rt}(p)}{\partial p_{jt}}$, $j, r = 1, \ldots, J$, $\Omega_{jr}^{*} = \begin{cases} 1, & \text{if } \exists f : \{r, j\} \subset F_f \\
0, & \text{otherwise} \end{cases}$ and $\Omega$ is a $J \times J$ matrix with $\Omega_{jrt} = \Omega_{jr}^{*} \times \Phi_{jrt}$. Following the same derivation as in the retailer pricing case, manufacturer’s mark-up is $w_{t} - mc_{t} = -(\Phi_{t}G_{t}^{-1}\Omega_{t})^{-1}s_{t}(p)$.
Price-Cost Margin: Summary of Myopic Model

The price-cost margin is defined as \( p_t - mc = (p_t - w_t) + (w_t - mc) \). Based on the above discussion, we see that \( p_t - mc = \Phi_t^{-1} s_t(p) - (\Phi_t' G_t^{-1} \Omega_t)^{-1} s_t(p) \).

One-Period Look Ahead Pricing Model

Suppose that each manufacturer and the retailer look ahead one-period in pricing to maximize inter-temporal profits. Also, we only derive the two-period model with the assumptions of a Bertrand-Nash game among multiple manufacturers and a Stackelberg game between the manufacturers and the single retailer.

The Retailer

The retailer takes the wholesale prices as given in the Stackelberg game, and acts like a monopolist in pricing the whole category.

In the one-period model, at period \( t (t = 1, 2) \), the retailer’s profit maximization problem is,

\[
\max_{p_{t1}, \ldots, p_{tJ}} \pi_{R_t} = \sum_{j=1}^{J} (p_{jt} - w_{jt}) s_{jt}(p),
\]

where \( s_{jt}(p) \) is the market share of brand \( j \), which is a function of the prices of all brands, \( M \) is the size of the market.

If the monopolist retailer, instead of being myopic, looks one-period ahead and maximizes two-period profits, his objective function will become, \( V_R = \pi_{R_1}(p_{j1}) + \delta \pi_{R_2}(p_{j2}) \), where \( j = 1, \ldots, J \).

The first-order conditions now become,

\[
\frac{\partial \pi_{R_1}}{\partial p_{j1}} + \delta \sum_{r=1}^{J} \frac{\partial \pi_{R_2}}{\partial s_{r2}(p)} \frac{\partial s_{r2}(p)}{\partial s_{r1}(p)} \frac{\partial s_{r1}(p)}{\partial p_{j1}} = 0
\]

\[
\frac{\partial \pi_{R_2}}{\partial p_{j2}} = 0
\]

It is obvious that \( \frac{\partial \pi_{R_2}}{\partial s_{r2}(p)} = p_{r2} - w_{r2} \). In order to calculate \( \frac{\partial s_{r2}(p)}{\partial s_{r1}(p)} \), we assume the following relationship holds between \( s_{r1} \) and \( s_{r2} \),

\[
s_{r2} = \theta_{r2r1} * s_{r1} + \sum_{k=1,k\neq r}^{J} \theta_{r2k1} * s_{k1},
\]

where, as defined in the demand models, \( \theta_{r2r1} \) is the transition probability that consumers who bought brand \( r \) in period 1 and continue to buy it during period 2, and \( \theta_{r2k1} \) is the transition probability that consumers who did not buy brand \( r \) in period 1 but switched to it during period 2.

The following relationship holds, \( s_{r2} = \theta_{r2r1} * s_{r1} + \sum_{k=1,k\neq r}^{J} \theta_{r2k1} * s_{k1} = \theta_{r2r1} * s_{r1} + \sum_{k=1,k\neq r}^{J} \theta_{r2k1} * (1 - s_{r1} - s_{r2} - \ldots - s_{s_{k-1,k \neq r} - \ldots - s_{j1}) \) and therefore,

\[
\frac{\partial s_{r2}(p)}{\partial s_{r1}(p)} = \theta_{r2r1} - \sum_{k=1,k\neq r}^{J} \theta_{r2k1} \begin{cases} > 0 & \text{if } SD > 0 \text{ and } \theta_{r2r1} - \sum_{k=1,k\neq r}^{J} \theta_{r2k1} > 0 \\ < 0 & \text{if } SD < 0 \text{ and } \theta_{r2r1} - \sum_{k=1,k\neq r}^{J} \theta_{r2k1} < 0 \end{cases}
\]
since \( \sum_{k=1,k \neq r}^J \theta_{r2k1} = 1 - \theta_{r2r1} \). With these two terms derived, we could now write down the retailer’s price functions for both period 1 and period 2.

The detailed first-order conditions could be written as, \( s_j(p) + \sum_{r=1}^J (p_r - w_r) \frac{\partial s_r(p)}{\partial p_{j1}} + \delta \sum_{r=1}^J (p_r - w_r) \frac{\partial s_r(p)}{\partial p_{j2}} = 0 \) for period 1; and \( s_j(p) + \sum_{r=1}^J (p_r - w_r) \frac{\partial s_r(p)}{\partial p_{j1}} + \delta \sum_{r=1}^J (p_r - w_r) \frac{\partial s_r(p)}{\partial p_{j2}} = 0 \) for period 2, where \( \Delta_1 = \theta_{r2r1} - \sum_{k=1,k \neq r}^J \theta_{r2k1} \).

Defining \( \Phi_{jrt} = -\frac{\partial s_r(p)}{\partial p_{j}} \), \( j, r = 1, ..., J \), and using the matrix form, the retailer’s second-period pricing function can be written as \( p_2 - w_2 = \Phi_2^{-1} s_2(p) \) and the retailer’s first-period pricing function can be written as \( p_1 - w_1 = \Phi_1^{-1} \left\{ s_1(p) - \Phi_1 \left[ \delta \Phi_2^{-1} s_2(p) (\Delta_1) \right] \right\} \).

**The Manufacturer**

Manufacturers’ markups can be solved for explicitly by defining \( \Omega_{j}^r = -\frac{\partial s_r(p)}{\partial p_{j}} \), \( j, r = 1, ..., J \), and \( \Omega \) is a \( J \times J \) matrix with \( \Omega_{jrt} = \Omega_{j}^r \times \Phi_{jrt} \). Following the same derivation as in the retailer pricing case, in period 2, manufacturer’s mark-up is \( w_2 - mc = -(\Phi_2^* G_2^{-1} \Omega_2)^{-1} s_2(p) \) and in the first-period, manufacturer’s price-marginal cost margin is, \( w_1 - mc = \Omega_1^{-1} \left\{ s_1(p) + \Omega_1 \left[ \delta (\Phi_2^* G_2^{-1} \Omega_2)^{-1} s_2(p) (\Delta_1) \right] \right\} \).

**Price-Cost Margin: Summary of The One-Period Look Ahead Model**

The price-cost margin is defined as \( p_t - mc = (p_t - w_t) + (w_t - mc) \). Based on the above discussion, we see that

\[
p_t - mc = \begin{cases} 
\Phi_1^{-1} s_1(p) - \delta \Phi_2^{-1} s_2(p) (\Delta_1) + \Omega_1^{-1} s_1(p) + \delta (\Phi_2^* G_2^{-1} \Omega_2)^{-1} s_2(p) (\Delta_1) & \text{when } t = 1 \\
\Phi_2^{-1} s_2(p) - (\Phi_2^* G_2^{-1} \Omega_2)^{-1} s_2(p) = PC M^{Myopic}(t = 2) & \text{when } t = 2
\end{cases}
\]

where \( \Delta_1 = \theta_{r2r1} - \sum_{k=1,k \neq r}^J \theta_{r2k1} \).

**References**


Tables And Figures

![Figure 1: A Schematic Model of the Market](image-url)
Table 1: Demand-Side Estimation: Fit Results

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>Multinomial Logit</th>
<th>State-Dependence with Lag Dummies</th>
<th>State-Dependence with Similarity Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Without Heterogeneity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>8</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Price Coefficient</td>
<td>-25.50</td>
<td>-23.23</td>
<td>-23.09</td>
</tr>
<tr>
<td>State-dependence Parameter</td>
<td>2.87</td>
<td>3.15</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-146949</td>
<td>-137997</td>
<td>-137407</td>
</tr>
<tr>
<td>AIC</td>
<td>293914</td>
<td>276020</td>
<td>274850</td>
</tr>
<tr>
<td>BIC</td>
<td>293937</td>
<td>276057</td>
<td>274901</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td></td>
<td>17904</td>
<td>19084 (1180)</td>
</tr>
<tr>
<td>2. With Heterogeneity (2-Support Case)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>17</td>
<td>27</td>
<td>32</td>
</tr>
<tr>
<td>Price Coefficient</td>
<td>-91.28</td>
<td>-74.45</td>
<td>-66.35</td>
</tr>
<tr>
<td>State-dependence Parameter</td>
<td>-3.25</td>
<td>-0.49</td>
<td>0.32</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-141525</td>
<td>-132171</td>
<td>-131388</td>
</tr>
<tr>
<td>AIC</td>
<td>283084</td>
<td>264396</td>
<td>262840</td>
</tr>
<tr>
<td>BIC</td>
<td>283034</td>
<td>264340</td>
<td>262778</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td></td>
<td>18708</td>
<td>20274 (1566)</td>
</tr>
<tr>
<td>(3-Support Case)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>26</td>
<td>41</td>
<td>46</td>
</tr>
<tr>
<td>Price Coefficient</td>
<td>-90.65</td>
<td>-89.87</td>
<td>-85.46</td>
</tr>
<tr>
<td>State-dependence Parameter</td>
<td>-2.19</td>
<td>-1.92</td>
<td>-1.37</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-139959</td>
<td>-130756</td>
<td>-129930</td>
</tr>
<tr>
<td>AIC</td>
<td>279970</td>
<td>261594</td>
<td>259952</td>
</tr>
<tr>
<td>BIC</td>
<td>279918</td>
<td>261515</td>
<td>259863</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td></td>
<td>18406</td>
<td>20058 (1652)</td>
</tr>
<tr>
<td>Number of Observations:</td>
<td>68,312</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Demand-Side Estimation:  
Parameter Estimates for Three-support Demand Model  
With State-dependence

<table>
<thead>
<tr>
<th>Variables</th>
<th>No Endogeneity Correction</th>
<th>With Endogeneity Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Support 1 Estimates</td>
<td>Support 2 Estimates</td>
</tr>
<tr>
<td></td>
<td>(Estimates)</td>
<td>(Estimates)</td>
</tr>
<tr>
<td>Constant</td>
<td>-6.85 (0.26)</td>
<td>-7.00 (0.30)</td>
</tr>
<tr>
<td>Attributes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sweetened</td>
<td>-0.12 (0.07)</td>
<td>-0.03 (0.08)</td>
</tr>
<tr>
<td></td>
<td>-0.33 (0.04)</td>
<td>-0.33 (0.04)</td>
</tr>
<tr>
<td>fiber</td>
<td>0.28 (0.07)</td>
<td>0.28 (0.06)</td>
</tr>
<tr>
<td></td>
<td>0.08 (0.04)</td>
<td>0.08 (0.04)</td>
</tr>
<tr>
<td>fruit-nuts</td>
<td>0.24 (0.06)</td>
<td>0.14 (0.06)</td>
</tr>
<tr>
<td></td>
<td>0.02 (0.04)</td>
<td>0.60 (0.04)</td>
</tr>
<tr>
<td>Marketing Mix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>price</td>
<td>-3.50 (0.90)</td>
<td>-4.08 (1.02)</td>
</tr>
<tr>
<td></td>
<td>-1.37 (0.42)</td>
<td>-2.06 (0.48)</td>
</tr>
<tr>
<td>feature</td>
<td>1.27 (0.16)</td>
<td>1.36 (0.16)</td>
</tr>
<tr>
<td></td>
<td>0.92 (0.10)</td>
<td>0.90 (0.10)</td>
</tr>
<tr>
<td>display</td>
<td>1.28 (0.17)</td>
<td>1.25 (0.17)</td>
</tr>
<tr>
<td></td>
<td>1.22 (0.09)</td>
<td>1.25 (0.10)</td>
</tr>
<tr>
<td>Inventory</td>
<td>0.96 (0.06)</td>
<td>0.99 (0.07)</td>
</tr>
<tr>
<td></td>
<td>-0.04 (0.10)</td>
<td>-0.04 (0.10)</td>
</tr>
<tr>
<td>State-dependence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>intercept</td>
<td>-1.35 (0.39)</td>
<td>-1.23 (0.41)</td>
</tr>
<tr>
<td></td>
<td>4.77 (0.05)</td>
<td>5.02 (0.05)</td>
</tr>
<tr>
<td>family size</td>
<td>0.25 (0.12)</td>
<td>0.23 (0.13)</td>
</tr>
<tr>
<td></td>
<td>-0.16 (0.03)</td>
<td>-0.16 (0.03)</td>
</tr>
<tr>
<td>income</td>
<td>0.76 (0.06)</td>
<td>0.82 (0.07)</td>
</tr>
<tr>
<td></td>
<td>-0.10 (0.01)</td>
<td>-0.12 (0.01)</td>
</tr>
<tr>
<td>age</td>
<td>-0.72 (0.13)</td>
<td>-0.59 (0.14)</td>
</tr>
<tr>
<td></td>
<td>0.05 (0.03)</td>
<td>0.07 (0.03)</td>
</tr>
<tr>
<td>child age</td>
<td>-0.67 (0.09)</td>
<td>-0.70 (0.10)</td>
</tr>
<tr>
<td></td>
<td>-0.27 (0.03)</td>
<td>-0.31 (0.03)</td>
</tr>
<tr>
<td>State-dependence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>-1.46 (0.33)</td>
<td>-1.35 (0.34)</td>
</tr>
<tr>
<td>standard error</td>
<td>4.83 (0.10)</td>
<td>5.08 (0.10)</td>
</tr>
<tr>
<td>Similarity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ingredient</td>
<td>0.09 (1.58)</td>
<td>0.09 (1.57)</td>
</tr>
<tr>
<td>sweetened</td>
<td>0.17 (0.01)</td>
<td>0.16 (0.01)</td>
</tr>
<tr>
<td>fiber</td>
<td>0.09 (1.58)</td>
<td>0.09 (1.57)</td>
</tr>
<tr>
<td>fruit-nuts</td>
<td>0.08 (0.01)</td>
<td>0.09 (0.01)</td>
</tr>
<tr>
<td>product type</td>
<td>0.11 (0.01)</td>
<td>0.11 (0.01)</td>
</tr>
<tr>
<td>Segments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>29.17% (0.01)</td>
<td>27.33% (0.01)</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>46</td>
<td>66</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>68,312</td>
<td>68,312</td>
</tr>
<tr>
<td>L1</td>
<td>-129,930</td>
<td>-56,529</td>
</tr>
<tr>
<td>ch^2</td>
<td>20058</td>
<td>647.5</td>
</tr>
<tr>
<td>AIC</td>
<td>259952</td>
<td>113190</td>
</tr>
<tr>
<td>BIC</td>
<td>260082</td>
<td>113377</td>
</tr>
</tbody>
</table>
Table 3: Supply-Side Estimation: Fit Results

<table>
<thead>
<tr>
<th>Manufacturer Interaction</th>
<th>Retailer Objective</th>
<th>Log-Likelihoods (Vuong Test Statistics)</th>
<th>Log-Likelihoods (χ^2 Test Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Myopic Pricing Without State-Dependence</td>
<td>Myopic Pricing With State-Dependence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One-Period Look Ahead Pricing With State-Dependence</td>
<td></td>
</tr>
<tr>
<td>Tacit Collusion</td>
<td>Category Profit Maximization</td>
<td>30231</td>
<td>31772</td>
</tr>
<tr>
<td></td>
<td>Manufacturer Vertical Nash</td>
<td>(—)</td>
<td>(4.76**)</td>
</tr>
<tr>
<td></td>
<td>Category Profit Maximization</td>
<td>29302</td>
<td>31015</td>
</tr>
<tr>
<td></td>
<td>Manufacturer Stackelberg</td>
<td>(4.63**)</td>
<td>(5.31**)</td>
</tr>
<tr>
<td>Bertrand Competition</td>
<td>Category Profit Maximization</td>
<td>29216</td>
<td>31949</td>
</tr>
<tr>
<td></td>
<td>Manufacturer Vertical Nash</td>
<td>(3.97**)</td>
<td>(2.39*)</td>
</tr>
<tr>
<td></td>
<td>Category Profit Maximization</td>
<td>28967</td>
<td>32196</td>
</tr>
<tr>
<td></td>
<td>Manufacturer Stackelberg</td>
<td>(5.13**)</td>
<td>(—)</td>
</tr>
</tbody>
</table>

(—): the best fitting model

(**): >95% confidence of accepting the best fitting model and rejecting the alternative model based on Vuong test statistics

(*): >90% confidence of accepting the best fitting model and rejecting the alternative model based on Vuong test statistics

(**): >95% confidence of accepting the intertemporal pricing model based on χ^2 test statistics
Table 4: Prediction of Prices for Brands

<table>
<thead>
<tr>
<th>Predicting Price</th>
<th>Mean Squared Errors (MSE) of Price Predictions</th>
<th>Variance of MSE</th>
<th>Reduction in (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicting Price of Brands Using Myopic Model without State-dependence</td>
<td>0.053</td>
<td>0.230</td>
<td></td>
</tr>
<tr>
<td>Predicting Price of Brands Using Myopic Model with State-dependence</td>
<td>0.041</td>
<td>0.193</td>
<td>Compared with Myopic No SD 22.34</td>
</tr>
<tr>
<td>Predicting Price of Brands Using One-period look ahead Model with State-dependence</td>
<td>0.041</td>
<td>0.190</td>
<td>Compared with Myopic No SD 23.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Compared with Myopic With SD 1.710</td>
</tr>
</tbody>
</table>
Table 5: Supply-Side Estimation:
Parameter Estimates for One-period Look Ahead Pricing Models
Using Demand Estimates With State-Dependence

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Bertrand (Vertical Nash)</th>
<th>Collusive (Vertical Nash)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimates</td>
<td>S.E.</td>
</tr>
<tr>
<td>Manufacturer Dummies</td>
<td>Nabisco</td>
<td>0.0105</td>
</tr>
<tr>
<td></td>
<td>General Mills</td>
<td>-0.0159</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Quaker</td>
<td>0.0109</td>
</tr>
<tr>
<td></td>
<td>Kellogg</td>
<td>0.0350</td>
</tr>
<tr>
<td>Store Dummies</td>
<td>Store 2 Dummy</td>
<td>-0.0086</td>
</tr>
<tr>
<td></td>
<td>Store 3 Dummy</td>
<td>-0.0075</td>
</tr>
<tr>
<td>Cost</td>
<td>Constant</td>
<td>0.1354</td>
</tr>
<tr>
<td></td>
<td>Corn</td>
<td>0.0187</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>0.0134</td>
</tr>
<tr>
<td></td>
<td>Oats</td>
<td>0.0227</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>0.0050</td>
</tr>
<tr>
<td></td>
<td>Wage</td>
<td>4.17E-05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Bertrand (Stackelberg)</th>
<th>Collusive (Stackelberg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimates</td>
<td>S.E.</td>
</tr>
<tr>
<td>Manufacturer Dummies</td>
<td>Nabisco</td>
<td>0.0107</td>
</tr>
<tr>
<td></td>
<td>General Mills</td>
<td>-0.0178</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>-0.0088</td>
</tr>
<tr>
<td></td>
<td>Quaker</td>
<td>0.0114</td>
</tr>
<tr>
<td></td>
<td>Kellogg</td>
<td>0.0429</td>
</tr>
<tr>
<td>Store Dummies</td>
<td>Store 2 Dummy</td>
<td>-0.0085</td>
</tr>
<tr>
<td></td>
<td>Store 3 Dummy</td>
<td>-0.0080</td>
</tr>
<tr>
<td>Cost</td>
<td>Constant</td>
<td>0.1354</td>
</tr>
<tr>
<td></td>
<td>Corn</td>
<td>0.0086</td>
</tr>
<tr>
<td></td>
<td>Wheat</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>0.0070</td>
</tr>
<tr>
<td></td>
<td>Oats</td>
<td>0.0237</td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>0.0038</td>
</tr>
<tr>
<td></td>
<td>Wage</td>
<td>4.53E-05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32269.3</td>
</tr>
<tr>
<td></td>
<td>31975.2</td>
</tr>
</tbody>
</table>

Note: For identification, one of the manufacturer dummies (Ralston) and one of the store dummies are set to 0. Store 2 and 3 on average carry a smaller number of brands every week.