Market Power in the Carbonated Soft Drink Industry

William J. Allender and Timothy Richards*
Arizona State University

January 16, 2012

Abstract

We investigate the strategic pricing for leading brands sold in the carbonated soft drink (CSD) market in the context of a flexible demand specification (i.e. random parameter nested logit) and a structural pricing equation. Our approach does not rely upon the often used ad hoc linear approximations to demand and profit-maximizing first-order conditions. We estimate the structural pricing equation using four different estimators (i.e. OLS, LIML, 2SLS, and GMM) and compare the implied deviation from Bertrand-Nash competition. Our results suggest that retailers, on average, price CSD brands below their cost, likely a result of the competitive retailing environment. We also find CSD wholesalers price their brands significantly more cooperatively than Bertrand-Nash would suggest, thus inflating profits.

1 Introduction

The rich content of scanner data enables the estimation of structural econometric models to be used to investigate market power and analyze policy. Recent advances in structural approaches to empirical market power analysis combines estimated demand functions with game-theoretic models of a particular industry to estimate its competitive nature. However, economic and econometric theory are often silent on the specific econometric estimators that should be employed. Each one having similar, but differing restrictions on the assumptions of the underlying sample from which the data is collected. As a result, the objective of this study is to empirically compare several different estimators of a supply model, and look

*Ph.D. Student and Morrison Chair respectively, Morrison School of Management and Agribusiness, Arizona State University, email: William.Allender@asu.edu. Support from the Economic Research Service of the USDA is gratefully acknowledged. The authors would like to thank Dr. Seung C. Ahn for his helpful comments.

at the differences each one implies with respect to the nature of the competitive game the carbonated soft drink (CSD) market plays.

The CSD category is used in our empirical analysis for several reasons. First, the industry is highly concentrated at the manufacturing level, being largely dominated by two manufacturers. However, retail outlets have recently been introducing, and pushing, their own private label brands as a way of expanding category profits. Second, the industry is well known among a wide range of consumers not only in the U.S., but largely throughout the world. Thus many retail outlets carry the same set of products making the competitive nature of the CSD industry an empirical question at not only the manufacturer level, but also at the retail level. Finally, given the long history of the industry, particularly in the U.S., a steady state equilibrium is likely to exist. Thus the CSD market is an ideal category for comparing the competitive nature implied from several different econometric estimators.

The remainder of the paper is organized as follows. In the next section we present the econometric model used to analyze the market power of the CSD market. This section begins with a brief overview of the demand and supply models used, followed by their specification. In section 3 we present the estimation methods used for the demand and supply models, along with a description of our identification strategy. The following section details the data used to estimate the models. Empirical results, as well as the implications of our finding are presented in section 5 and the last section concludes.

2 Econometric Model of the CSD Market

2.1 Overview

We model the CSD market using a structural model of consumer, retailer, and producer behavior. This method ensures that our empirical results are consistent with theoretical expectations of firm behavior at all stages of the supply side. On the demand side we model consumer's discrete purchasing choices using a random parameter nested logit model. Because CSDs are a highly differentiated food product a discrete choice framework is appropriate (Jain, Vilcassim, and Chintagunta 1994; and Nevo, 2001). A random coefficient model captures unobserved household heterogeneity, while also controlling for other exogenous factors that may influence a household's brand decisions.

On the supply side we assume Bertrand-Nash competition among both producers and retailers (Draganska and Klapper 2007). We parameterize deviations from the maintained equilibrium to test both the direction and degree of diversion from Bertrand competition

(Villas-Boas and Zhao 2005; Draganska and Klapper 2007). We then estimate the model using several different estimation techniques to compare the implied market power in the CSD market. We begin by explaining the consumer demand component, and continue on to behavior by both sets of participants in the CSD supply chain.

2.2 Structural Model of CSD Demand

Consumer demand is represented by a random utility model in which consumers are assumed to make a hierarchical decision regarding first the choice of retailer and, once this decision is made, the choice of soft drink from the set of all products being offered, or make no purchase at all. This latter alternative forms the outside option. Indirect utility u_{ijkt} for consumer i obtained from purchasing product $j \in \{1, 2, ..., J\}$, on purchasing occasion $t \in \{1, 2, ..., T\}$, at store $k \in \{1, 2, ..., K\}$, is given by:

$$u_{ijkt} = \boldsymbol{\gamma}^{\intercal}\mathbf{z}_{jkt} - \alpha_i p_{jkt} + \xi_{jkt} + \varsigma_{ikt} + (1 - \sigma_K)\varepsilon_{ijt}, \tag{1}$$

where ξ_{jkt} is an error term that accounts for all product-specific variations in demand that are unobserved by the econometrician such as the amount of shelf space allocated to each product or the amount of national advertising. Product j's price in store k is represented by p_{jkt}. The set of brand specific variables $\mathbf{z}_{jkt}$ contains an intercept term (υ_{jkt}) which represents the product-specific preference parameter. Other brand attributes include an indicator of whether the product is offered on a temporary discount (dc_{jkt}), and an interaction term between the retail price and the discount ($dc_{jkt}p_{jkt}$) (Chintagunta, 2002). By including an interaction term we allow for the possibility that promotions rotate the demand curve in addition to the expected demand-shifting effect. In this way, we allow items on promotion to become less elastic if households perceive discounting as a means of differentiating otherwise similar products.

It is well understood that the simple logit model suffers from the independence of irrelevant alternative (IIA) property, so we extend it in two ways. First, we explicitly account for the hierarchical nature of a consumer's choice process by using a nested logit model (McFadden, 1980). A nested logit model provides both an intuitive way of describing the consumer's decision and analytical solutions for the retailer's profit maximizing pricing decision. Nesting by store represents a natural choice because consumers are likely to substitute among products (in the same category) within a store, rather than comparing products across stores. With a nested logit model, substitution within each store still reflects IIA, while between comparison between stores does not.

In order to incorporate this nesting assumption, we allow the ε_{ijt} terms to follow a Generalized Extreme Value distribution (GEV, McFadden, 1980). With the GEV assumption, we allow for differing degrees of substitution between products across groups. The distribution of ς_{ik} is defined such that the term $\varsigma_{ikt} + (1 - \sigma_K)\varepsilon_{ijt}$ is i.i.d. Type I extreme-value distributed if the consumer specific error term ε_{ijt} is itself extreme-value distributed (Cardell, 1997). The parameter σ_K measures utility-correlation within each nest (store) and, as such, is interpreted as an inverse measure of store heterogeneity. The parameter is bound between 0 and 1. If $\sigma_K = 1$ then the correlation among stores goes to 1 and stores are regarded as perfect substitutes. On the other hand if the parameter is zero the model reverts back to a simple multinomial logit model. We include the no purchase option (outside good) with $j = 0$, which allows us to test whether sales have any general demand-expansion effects.

Based on the indirect utility model in equation (1), and following Train (2003) the level of mean utility that varies over consumers for each choice of product j in store k is: $\delta_{jk} = \boldsymbol{\gamma}^\intercal \mathbf{z}_{jkt} - \alpha_i p_{jkt} + \xi_{jkt}$. So, the probability of consumer i choosing brand j in store k at time t is given by (suppressing the time subscript for clarity):

$$\Pr_{jk} = \Pr_{j|k} \Pr_{k} = \frac{e^{\delta_{jk}/\sigma_K} \left(\sum_{j \in k} e^{\delta_{jk}/\sigma_K} \right)^{\sigma_K - 1}}{\sum_{l=1}^{K} \left(\sum_{j \in l} e^{\delta_{jk}/\sigma_K} \right)^{\sigma_K}} \tag{2}$$

where $\Pr_{j|k}$ is the conditional probability of consumer i choosing brand j conditional on brand j being in store k and $\Pr_k$ is the marginal probability of choosing a brand in store k (with the marginality being over all alternatives in store k) where the utility of the no purchase option (outside option) has been normalized to zero. This equality is exact, since any probability can be written as the product of a marginal and a conditional probability.

Unfortunately, the GEV model still suffers from the IIA property within stores, which means that the substitution elasticities between products depends only on their market shares and not on more fundamental attributes that are likely to influence demand. Thus, our second extension allows the product-preference and marginal utility of income parameters in equation (1) to vary over consumers in a random way (Berry, Levinsohn, and Pakes 1995; and Nevo, 2001). Specifically, the marginal utility of income is normally distributed over consumers such that:

$$\alpha_i = \alpha + \sigma_\alpha v_i, \ v_i \sim N(0, 1), \tag{3}$$

where α is the mean price response across all consumers and v_i is the consumer-specific variation in response with parameter σ_α. Furthermore, consumers are assumed to differ in their preferences for each product attribute such that unobserved consumer heterogeneity is reflected in the distribution of each brand's preference parameter (Erdem, 1996; and Nair, Chintagunta and Dubé, 2004) as follows:

$$v_{ijkt} = v + \sigma_v \mu_i, \ \mu_i \sim N(0,1). \tag{4}$$

In contrast to the IIA property of a simple logit model, the heterogeneity assumption in (3) and (4) creates a general pattern of substitution over the J alternatives through the unobserved, random part of the utility function given in (1). As a result, the utility from different soft drinks is correlated according to their set of attributes included in $\mathbf{z}_{jk}$. Non-IIA substitution is critical in models of differentiated product pricing because parameter estimates would otherwise be entirely confounded by misestimates of the partial elasticity of demand facing each product.

With a discrete choice model of demand, it is assumed that each consumer purchases only one unit of the chosen product. Defining the densities of μ_i and v_i as $f(\mu)$ and $g(v)$, respectively, the probability of product j being purchased in store k is obtained by integrating over the distribution reflecting consumer heterogeneity:

$$\Pr_{jk} = \iint \frac{\exp\left[\frac{\delta_{jk}}{\sigma_K}\right]\left(\sum_{j\in k}\exp\left[\frac{\delta_{jk}}{\sigma_K}\right]\right)^{\sigma_K - 1}}{\sum_{l=1}^{K}\left(\sum_{j\in l}\exp\left[\frac{\delta_{jk}}{\sigma_K}\right]\right)^{\sigma_K}} f(\mu)\, g(v)\, d\mu\, dv, \tag{5}$$

which we estimate using simulated maximum likelihood following Petrin and Train (2010) which is described in section 3 below.

2.3 Structural Model of CSD Supply

In order to measure the degree of market power exercised by wholesalers and retailers, we develop a model of the CSD supply chain that is used to derive equilibrium wholesale and retail margins. We assume a Bertrand-Nash equilibrium such that suppliers quote a wholesale price to retailers, taking into consideration the retailer's response, and then retailers set prices to be paid by consumers (Kadiyali, Vilcassim, and Chintagunta, 1996, 1999). We solve the model using backward induction, first describing the second-stage pricing decision made by retailers, and then the first-stage wholesale pricing decision. This kind of vertical model (often referred to as a conjectural variation model) is found to fit well among the data in a number of categories (Besanko, Dubé, and Gupta, 2003; Villas-Boas and Zhao, 2005; Draganska and Klapper, 2007; Berto Villas-Boas, 2007) and is found to best fit the CSD category (Dhar, Chavas, Cotterill and Gould, 2005). In the remainder of this section we derive the subgame perfect Nash equilibrium in prices to this channel game. To make the notation as clear as possible we suppress the time subscript t.

2.3.1 Retailer Decision

Each retailer, k, chooses prices p_j for all products to maximize category profits. In other words, the retailer solves:

$$\Pi^k = \max_{p_j} Q \sum_{j=1}^{J} (p_j - w_j)\, s_j, \ \forall\, j \tag{6}$$

where Q is the total market, w_j is the wholesale price, p_j is the price retailers charge, and s_j is the market share of brand j. To simplify our derivation, without loss of generality, we assume unit retailing costs are zero. Equation (6) assumes retailers maximize profits across all product categories, and not just on a category by category basis. In other words, retail prices reflect the implicit assumption that retailers internalize all pricing externalities across categories. Assuming retailers behave as Stackelberg followers, our first order condition for product j is given by:

$$s_j + \sum_{l=1}^{J} \frac{\partial s_l}{\partial p_j}(p_l - w_l) = 0, \ \forall\, j. \tag{7}$$

for each retailer k. Stacking the first-order conditions for all brands and solving for retail prices in matrix notation yields:

$$\mathbf{p} = \mathbf{w} - \mathbf{S}_p^{-1}\mathbf{S}, \tag{8}$$

where $\mathbf{p}$ is a $J \times 1$ vector of retail prices, $\mathbf{w}$ is a $J \times 1$ vector of wholesale prices, $\mathbf{S}$ is a $J \times 1$ vector of market shares, and $\mathbf{S}_p$ is a $J \times J$ matrix of share derivatives with respect to all retail prices. Since suppliers are assumed to take retailers' optimal responses into account in setting upstream prices, equation (8) represents the retail decision rule that frames their pricing decisions.

2.3.2 Wholesaler Decision

Wholesalers are assumed to set prices such that the surplus they obtain over production costs is maximized for all the products they supply, while taking into account the retailers' response. The profit maximization problem for the vendor who sells product j is:

$$\Pi_j^{W_n} = \max_{w_j} Q \sum_{j=1}^{J} s_j(w_j - c_j), \ \forall j, \tag{9}$$

where c_j denotes the (constant) production cost of product j incurred by the wholesaler and the other variables are as described above. The first order condition for the supplier is given by:

$$s_j + \sum_{k=1}^{J} \left(\frac{\partial s_j}{\partial pk} \frac{\partial p_k}{\partial w_j} \right) (w_j - c_j) = 0. \, \forall j \tag{10}$$

However, the retail-wholesale pass-pass through term $\partial p_k / \partial w_j$ represents values that are not observable in the data. Therefore, we recover each pass-through rate by totally differentiating the retailer's first order condition as in Villas-Boas and Zhao (2005). In other words, by totally differentiating equation (7) we solve for $\partial p_k / \partial w_j$ with respect to all wholesale prices. Doing so and simplifying gives us:

$$\sum_{k=1}^{J} \frac{\partial s_j}{\partial p_k} \frac{\partial p_k}{\partial w_j} + \sum_{k=1}^{J} \sum_{l=1}^{J} (p_l - w_l) \left(\frac{\partial^2 s_l}{\partial p_j \partial p_k} \right) \frac{\partial p_k}{\partial w_j} + \sum_{l=1}^{J} \frac{\partial s_l}{\partial p_j} \frac{\partial p_l}{\partial w_j} = \frac{\partial s_j}{\partial p_j}, \forall \, j \tag{11}$$

which can be simplified by defining a $J \times J$ matrix $\mathbf{G}$ with typical element $g_{j,k}$ such that:

$$g_{j,k} = \frac{\partial s_j}{\partial p_k} + \sum_{l=1}^{J} (p_l - w_l) \left(\frac{\partial^2 s_l}{\partial p_j \partial p_k} \right) + \frac{\partial s_k}{\partial p_j}, \; \forall \, j,k \tag{12}$$

Using the above expression, we write the wholesale margin in matrix notation as:

$$\mathbf{w} - \mathbf{c} = -((\mathbf{G}^{-1}\mathbf{S}_p)\mathbf{S}_p * \mathbf{I}_N)^{-1}\mathbf{S}, \tag{13}$$

where $\mathbf{I}_N$ is a $J \times J$ identity matrix and * indicates element by element multiplication. Substituting this expression back into the solution for retail prices provides a single expression for the whole - retail plus wholesale - margin in which the wholesale price is equal to the retail price minus the retail margin and wholesale margin:

$$\mathbf{p} = \mathbf{c} - \mathbf{S}_p^{-1}\mathbf{S} - ((\mathbf{G}^{-1}\mathbf{S}_p)\mathbf{S}_p * \mathbf{I}_N)^{-1}\mathbf{S} \tag{14}$$

where the first expression on the right side is the marginal cost, the second expression on the right side is the retail margin and the third represents the wholesale margin. At this point, all of the parameters required to identify the wholesaler's cost are contained in the price and demand side estimates ($\mathbf{p}$, $\mathbf{S}_p$, and $\mathbf{G}$). Marginal costs ($\mathbf{c}$), in turn, are estimated as a linear function of input prices, such that:

$$\mathbf{c} = \boldsymbol{\vartheta}^{\mathsf{T}} \boldsymbol{\iota}, \tag{15}$$

where $\boldsymbol{\iota}$ is a vector of input prices and $\boldsymbol{\vartheta}$ is a vector of parameters to be estimated. This function is estimated after substituting the demand parameters into equation (14) in the three-step procedure described in detail in section 3.

2.3.3 Measuring Market Power

The model of wholesale and retail margins in equation (14) is derived under the assumption of Bertrand-Nash rivalry. However, by parameterizing deviations from Bertrand-Nash (Villas-Boas and Zhao, 2005), we test for the presence of market power among either wholesalers or retailers. We augment the margin expressions in equation (14) by introducing multipliers θ, and ϕ in a manner similar to Villas-Boas and Zhao (2005) and Draganska and Klapper (2007). To clarify our exposition, let $\mathbf{m}^R = -\mathbf{S}_p^{-1}\mathbf{S}$, be the retail margin, and let $\mathbf{m}^W = -((\mathbf{G}^{-1}\mathbf{S}_p)\mathbf{S}_p * \mathbf{I}_N)^{-1}\mathbf{S}$ denote the wholesale margin. Then with the introduction of the multipliers we can rewrite equation (14) as:

$$\mathbf{p} = \mathbf{c} + \phi\mathbf{m}^R + \theta\mathbf{m}^W. \tag{16}$$

Notice that if ϕ or θ are greater than 1 then retailers, or wholesalers, respectively, are pricing more cooperatively than in the Bertrand Nash equilibrium, whereas ϕ, or θ less than 1 would suggest that the firms are pricing more competitively.

The estimating equations for the full model are, therefore, equation (5) for the demand side, and equation (16) for the supply side with equation (15) substituted for the marginal cost expression. In terms of the supply model the estimating equation is written as:

$$\mathbf{p} = \boldsymbol{\vartheta}^{\intercal}\boldsymbol{\iota} + \phi\mathbf{m}^R + \theta\mathbf{m}^W + \epsilon, \tag{17}$$

where ϵ captures variations in price that are not explained by the model and are assumed to be i.i.d. The key hypothesis tests, therefore, involve ϕ and θ.

3 Estimation Method

There are several complications to address when estimating the demand (5) and supply models (17) above. The share equation cannot be estimated using ordinary least squares because prices are likely to be correlated with elements in the error term ε_{ijk}. Promotional activities, in-store merchandising, and other strategies cause price and market share to be jointly endogenous, making our estimates biased (Villas-Boas and Winer, 1999). We therefore use simulated maximum likelihood (SML) estimation following Petrin and Train (2010), who suggest a control function approach whose econometric foundation is found in the sample-selection models of Heckman (1978) and Hausman (1978). The idea of the control function approach is to derive a proxy variable that conditions on the part of the endogenous price variable, thus making the remaining variation in the price variable independent of the error term. Then the standard simulated maximum likelihood approach will be consistent.

Specifically, we regress the endogenous variable, price, against a set of instrumental variables (which are detailed below) via OLS and obtain the residuals, ζ_{ijkt}, from this model. The residuals from this regression are independent of the original error term ε_{ijk}. The control function is then defined as the linear combination: $CF(\zeta_{ijkt}|\lambda) = E(\varepsilon_{ijk}|\zeta_{ijkt}) = \lambda\ \zeta_{ijkt}$, where λ is a parameter to be estimated. We then let $\lambda = \lambda_0 + \sigma_\lambda \eta$, where $\eta \sim N(0, 1)$ to allow for greater flexibility. The control function, $CF(\zeta_{ijkt}|\lambda)$, is then added to the utility function given in equation (1). The full model, including the control function, is then estimated using SML which uses Monte Carlo simulation to solve the integral in equation (5) up to an approximation that is accurate to the number of random draws chosen, R. This method provides consistent parameter estimates under general error assumptions and is readily able to accommodate complex structures regarding consumer heterogeneity.

To identify the endogenous price variable, we require instruments that are correlated with price, but not the unobservables in the demand equation. Because there could be market and time specific price shocks in one particular market or time, we include a set of market and quarter specific binary variables as well as a constant term. Furthermore, because input prices are likely to be correlated with price, but not the error term in the demand equation, we also include retail and production input prices. We then exploit the panel data nature of the model and interact these input prices with a set of binary market specific variables. Finally, because it is possible that the price variable is capturing the state dependence of consumer demand over time due to habit, learning, or inertia of purchasing habits, a lagged share value and lagged price of 1 week are also included. The lagged price and share values are pre-determined from the perspective of current-period demand are thus appropriate instruments. Our identification strategy is well-accepted in the literature (Berto Villas-Boas, 2007; Draganska and Klapper, 2007).

Estimating the models sequentially in this way can cause a compounding error problem (Cameron and Trivedi, 2005). However, we find this unlikely because the parameter estimates of the estimated model are very similar to those that result if the control function is not included in the utility specification.[1] Therefore, we conclude that the compounding error problem is, at most, negligible and do not employ a boot strapping method. To aid in the speed and efficiency of SML estimation, we include a Halton draw sequence. Halton draws can significantly reduce the number of draws with no degradation of simulation performance

[1] These results are available from the authors upon request.

(Bhat, 2003). We find that $R = 200$ draws are sufficient to produce stable estimates without excessive estimation time. Bhat (2003) provides experimental evidence that Halton sequences can reduce the number of draws required to produce estimates at a given accuracy by a factor of 10.

In estimating the supply model, we use several different estimators to investigate the difference in the competitive nature implications from each. First, brand specific price shocks could affect the price of carbonated soft drinks. To account for this, we add binary indicator variables for the 19 most popular brands.[2,3] Second, given the retail and wholesale margins are likely to be endogenous we require a single equation instrumental variable (IV) estimator. Specifically, we estimate equation (17) using 2SLS, LIML, and GMM. For comparison purposes with an inconsistent estimator, we also use OLS. 2SLS is a natural choice for an estimator because the model is linear, and the consistency of the estimator in the presence of endogenous variables is well understood (Theil 1953, and Basmann, 1957). However, the LIML estimator has been shown to have less bias in finite samples, but, due to its lack of finite moments, tends to have a wider sampling distribution (Davidson and MacKinnon, 2004). Both of these estimators, along with OLS, are members of the k-class estimators (Theil, 1961).[4] Finally, we use the GMM estimator because it is the most general estimator and is by far the most widely used in current application of market power estimation (Hansen, 1982).

For the IV estimators of the supply function we require instruments that are correlated with the endogenous margin variables, but not the unobservables in the pricing equation. Therefore, we use a similar set of IVs as those described for the demand model. The difference here is that instead of interacting the input prices with market specific binary variables, they are interacted with brand specific binary variables. While demand shocks are likely to occur at the market level, wholesalers and retailers are likely to decide their margin level brand by brand. This identification is well accepted in the literature (Berto Villas-Boas, 2007).

Though simultaneous estimation of the demand and supply models is preferred, sequential estimation of the demand model and then the supply model is necessary as the estimates of the demand model are used directly in the estimation of supply. Sequential estimation of the demand and margin equations in this way is common in the literature and has been

[2] We had also included market specific binary variables to account for any market specific price shocks but found these were not relevant in explaining price. They were therefore excluded from the final specification.

[3] The 13 least popular brands of our sample were not purchased as often and as a result were not found to be significant in explaining price. They were therefore excluded from the final specification.

[4] Mariano (2001) gives a more recent account of the finite sample properties of 2SLS, LIML, and other k-class estimators.

shown to produce results that differ little from those obtained with simultaneous estimation (Villas-Boas and Zhao, 2005).

4 Data Description

The empirical estimation requires CSD retail level sales data. Consequently, this study uses Nielsen research group's Scantrack data, which measures weekly retail sales for 52 weeks in 2005. The data consist of dollar sales, unit volume (ounces), promotion attributes, and product specific identifiers. The Scantrack data features weekly sales information at the UPC level for participating retailers in 52 markets. Covering all 52 markets would have been intractable so we focus on the 5 largest markets: Chicago, IL; Los Angeles, CA; New York, NY; Atlanta, GA; and Philadelphia, PA. This sample of markets provides a wide variety of supermarkets representative of most retail outlets and prices. In total we have 143 retailers yielding $163,593$ observations.

Because the Scantrack data features sales information for supermarkets only, the outside option consists of the entire potential market for soft drink sales. Following Nevo (2001) and Berry, Levinsohn and Pakes (1995) the size of the whole market is calculated on a weekly basis by multiplying each of the total metropolitan statistical areas' populations by the USDA estimate of per capita consumption of soft drinks. The difference between the inside and outside option is then reduced to CSDs sold through convenience stores, food service outlets, and by retailers that do not participate in the retail-scanner data syndication (Wal-mart, Costco, and other club and super stores).

There are a large number of products available in the CSD category – too large to model in a tractable way and obtain reliable estimates of each products' mark-up. Therefore, we choose thirty three brands that were common among all markets and retailers with the highest market share.[5] The brands used, along with their summary statistics are shown in table 1. The fact that soft drinks are a differentiated food category is evident in the variability of prices among brands.

[Insert table 1]

Wholesale input costs consist of market-specific indices of commercial electricity prices, and the prices of: high fructose corn syrup, sugar, aluminum, and diesel. These input price indices were obtained from the USDA (USDA-NASS) database and their summary statistics are available in table 1. In order to avoid losing the detailed weekly variation in the CSD

[5] Diet sodas were excluded because retail price information wasn't available.

sales data the input prices are smoothed from monthly to weekly observations. This linear filter approach is common in empirical industrial organization research (e.g. Slade, 1995).

5 Empirical Results and Discussion

In this section we first present the demand estimates in order to establish the validity of the demand model. We then present the supply-side model estimates and compare the market power implications across the different estimation methods for both retailers and wholesalers.

5.1 CSD Demand Results

The demand for CSDs is modeled using a random coefficient nested logit model with a control function added to the utility function to alleviate endogeneity in the pricing variable. As a first step to interpreting the results, we test the validity of the random-parameter nested logit model against a simple logit alternative. A simple specification test involves testing the significance of the GEV scale parameter, σ_K. If $\sigma_K = 0$, then the GEV model collapses to a standard mixed logit. In the results shown in table 2, the t-ratio for the null hypothesis that $\sigma_K = 0$ is 815.95, so we easily reject the null hypothesis and conclude that the GEV model is preferred. Furthermore, it is common in the retailing literature to assume that individual stores price as local monopolists (Chintagunta, 2002). However, the GEV scale parameter, σ_K, which represents a measure of the extent to which consumers substitute among stores, does not support this assumption as the stores in the sampled markets are regarded as very good – but not perfect – substitutes for each other ($\sigma_K = 0.925$, while $\sigma_K = 1.0$ implies perfect substitutability).[6]

[Insert table 2]

Next, we compare the random-parameter nested logit to a constant parameter alternative. For this purpose, we use a likelihood-ratio (LR) test where the constant parameter model is the restricted version and the random parameter is the unrestricted version. The LR statistic for this test is 13.177, which is Chi-square distributed with 3 degrees of freedom. The Chi-square value at the 5% level of significance is 7.814, therefore we reject the simpler model and conclude that the random parameter model is a better fit. Further, the significance of the variance parameters in the random coefficients model further supports this specification over a fixed-coefficient alternative. Therefore, we can conclude that the random parameter model

[6] A formal test of the hypothesis: $H_o : \sigma_K = 1$ produces a t-ratio of 65.95, so we reject the null hypothesis and conclude that soft drinks from different stores are not perfect substitutes.

is preferred to the constant-parameter alternative, so we will use this version to interpret the demand results.

As a final specification test, we investigate the endogeneity of the price variable. Following Eichenbaum, Hansen and Singleton (1988) we test the hypothesis that $\lambda_0 = 0$ using a simple t-test. We find the test statistic for this hypothesis is -3.48 so we conclude that $\lambda_0 \neq 0$ at the 5% level of significance and the price variable is indeed endogenous. We test the relevance of the IVs by regressing the price variable on them. These results are presented in table 3. We find the F-statistic for this regression is $71,724.6$, so the IVs are indeed relevant (Staiger and Stock, 1997). As a last step to evaluating the IVs, we test their exogeneity following Hansen (1982) and find a test statistic of 2568.68, which is Chi-square distributed with 10 degrees of freedom. This suggests the IVs are not strictly exogenous to the error term in the demand model. However, the R^2 from the regression of the demand model's error term on the IVs was only 0.016. Furthermore, many different IVs were used, and the ones chosen were found to have the smallest J-statistic. We therefore argue that the IVs used were the best available.

[Insert table 3]

There are a number of results from the CSD demand model that are of substantive interest to CSD retailers and manufacturers. First, the marginal utility of income (price coefficient) is negative and significant as expected. We also find the standard deviation of the price coefficient is greater than 1, and significant at the 5% level. This suggests that there is a significant degree of heterogeneity among consumers when it comes to their marginal utility of income for carbonated soft drinks.[7] Second, the discount and price interaction term shift the demand curve out during a promotion week and rotate it clockwise (more inelastic). Both of these results are as retailers intend.

5.2 Market Power Results

We now present the results of the supply model for the various estimation routines. Before we interpret the results of the parameters of interest, namely ϕ and θ, we test the endogeneity of the two margin measures, followed by tests of relevance and exogeneity of the IVs chosen. In order to test whether or not the two margin measures are indeed endogenous we use the Hausman test (Hausman 1978). From table 5 the Hausman test statistics from the 2SLS and GMM estimators suggest that the margin variables are indeed endogenous while the hypoth-

[7] For comparison, the price coefficient estimate and standard deviation of the model estimated without the control function included is: -.36891 (-24.40), and 4.21742 (423.58) respectively, with the t-ratios is parentheses.

esis is not rejected in the case of the LIML estimator. Given the wide sampling distribution of the LIML estimator we conclude that the margin variables are indeed endogenous. Therefore, an IV estimation routine is appropriate in order to get consistent parameter estimates. We now test the hypothesis that the IVs are not relevant. For the wholesale margin variable we reject the hypothesis and conclude that the IVs are indeed relevant. However, the F-statistic from the regression of the retail margin on the IVs suggests they are not relevant to the retail margin measure. Nevertheless, many different IV combinations were used, and in each case, the F-statistic was very low, always below 1.5. Therefore, we once again argue that the IVs used were the best available. Finally, we test the exogeneity of the IVs following Hansen (1982) and find the test statistics are 30.65, 1158.05, and 2208.11, for the LIML, 2SLS, and GMM estimators respectively, which are all Chi-square distributed with 39 degrees of freedom which yields a critical value of 54.572. In the case of the LIML estimator the IVs are indeed exogenous. However, in the case of 2SLS, and GMM we conclude the IVs are not strictly exogenous. However, we find the uncentered-R^2 is 0.0072, and 0.0138 for the 2SLS and GMM estimators respectively, which are indeed quite small.[8]

[Insert table 4]

In this model, the conduct parameters are interpreted as measuring the extent of deviation from the maintained competitive nature assumed, on the part of either retailers (ϕ), or wholesalers (θ) (Villas-Boas and Zhao, 2005; Draganska and Klapper, 2007). In the retail case, the competitive nature assumed is Bertrand-Nash pricing, so an estimated value of $\phi > 1$ implies retailers, on average, price CSDs more cooperatively than Bertrand-Nash would predict, while $\phi < 1$ suggests retailers price them more competitively. Upstream wholesale deviations are interpreted in the same way. Finally, we note that if $\phi = \theta = 0$, then retailers and wholesalers, respectively, do not take advantage of the differentiated nature of their products and price as purely competitive sellers would.

The results for the four different estimators are presented in table 5. Consistent across all estimators is the sign of the parameter estimates of ϕ and θ, where $\phi < 0$ and $\theta > 0$. However, the OLS and LIML estimator suggest that ϕ is not statistically different from 0 at the 5% or 10% level of significance. On the other hand, both the 2SLS and GMM estimators agree that $\phi < 0$. The results of the four estimators taken together imply that retailers in general price CSDs at (OLS and LIML) or below (2SLS and GMM) their costs. While these results are somewhat counterintuitive, if we take into consideration the myriad of product

[8] The regression results of the pricing equation error term on the IVs are available from the authors upon request.

categories within the food retailing environment it may be profitable for retailers to price one category at or below cost to attract new customers to the store (e.g. loss leader pricing).

Several studies suggest that promoting one product, or category - even below cost - can have a sales expansion effect for a particular retailer as a whole (Hess and Gerstner, 1987; Walters, and MacKenzie, 1988; and Hosken and Reiffen, 2001). Thus in the context of our results, it may be the case that retailers are pricing CSDs such that the average price promotion in the category provides a loss to the retailer in that category, but an increase in profits overall as a result of an increase in sales in all other categories. Extending this research to specifically include the competitive interactions among retailers and retail price promotions at the category level Chintagunta (2002) and Richards (2007) conclude that promotions have their greatest impact on in-store product share, but promotions can increase store share if consumers regard the retailers has highly substitutable. Another possible explanation for observing retailers pricing CSDs at or below cost may have to do with manufacturers offering incentives to retailers to promote their products. Specifically, Agrawal (1996) and Lal and Villas-Boas (1998) develop a theoretical model of manufacturer and retailer competition that suggest manufacturers selling brands with little brand loyalty will incentivize retailers to promote their brand through price promotion, while retailers will frequently offer promotions on brands with a high level of brand loyalty to capture market share from other retailers.

[Insert table 5]

Looking more closely at the results of manufacturer market power (θ), we find once again that all four estimators agree that manufacturers price CSDs above purely competitive levels. However, the magnitude differs significantly among estimators. First, the OLS estimator suggests wholesalers price their brands only slightly above competitive levels. Given the endogeneity of the margin variables, the OLS parameter estimates are likely biased downward. Second, the parameter estimate from the LIML estimator implies wholesalers price their products significantly more cooperatively than Bertrand-Nash competition. The t-statistic from the parameter estimate suggests it is not statistically different from 0 at the 10% level of significance however. This suggests the parameter is not statistically different from zero, which is not surprising given the LIML estimator tends to have a wider sampling distribution (Davidson and MacKinnon, 2004). Finally, the 2SLS and GMM estimates, both having similar parameter estimates as expected, suggest that wholesalers price their brands significantly more cooperative than Bertrand-Nash competition. Given that the industry is dominated by two suppliers, this is not a surprising result. It is also unlikely that wholesalers are offering trade deals to retailers to promote their brands. Rather, it seems that retailers are discounting the brands in the CSD category as a means of competing for retail market share. This likely expands CSD category sales, which greatly increases the profit of the wholesalers.

6 Conclusion and Implications

In this study we estimate the market power of both retailers and wholesalers in the carbonated soft drink market. Pass-through is modeled using a structural model of each retail food outlet. Demand is assumed to be discrete, which we estimate using a random parameter, nested logit model with a control function added to account for the endogeneity of price. We then assume Bertrand-Nash pricing such that the wholesaler quotes a price to the retail, taking into consideration the retailer's response, while the retailer then sets retail prices in a two stage, non-cooperative game theoretic framework. We derive the retail pricing equation directly from the first-order conditions for multi-product retail profit maximization, and the wholesale pricing equations indirectly due to the unobservability of the wholesale pass-through term. For both retailers and wholesalers, pricing conduct is allowed to deviate from either the competitive benchmark or complete collusion through the inclusion of a conduct parameter. This allows us to empirically investigate the degree of market power among either retailers or wholesalers.

The supply model is estimated using several different IV estimators to compare the competitive nature implied by each. Specifically OLS, 2SLS, LIML, and GMM are used with 52 weeks of retail scanner data covering 2005. The OLS and LIML estimators suggest retailers price CSDs at their respective prices, while the 2SLS and GMM estimators suggest retailers are pricing the category below cost. These results are consistent with previous research that investigates the competitive nature of the food retailing market and suggest retailers price specific categories at or below cost as a mean of expanding overall store sales (Hess and Gerstner, 1987; Hosken and Reiffen, 2001; and Richards, 2007).

On the wholesale side, the OLS parameter estimate suggests that margins, while positive, are very narrow. However, the OLS estimate is biased downward as a result of the endogenous margin variables. The magnitude of the conduct parameter from the LIML estimator suggests wholesalers price their products significantly more cooperatively than Bertrand-Nash would imply. Though, due to the wide sampling distribution of the LIML estimator, the hypothesis that the parameter is statistically different from zero is not rejected. Finally, the results of the 2SLS and GMM estimates agree that wholesalers are in fact pricing their products significantly more cooperatively then Bertrand-Nash would suggest. Given the competitive nature of the retailing environment, wholesalers are likely taking advantage of retailer's dependence on CSD price promotions to capture inflated wholesale margins and excess profit. On the other hand, this may be because wholesalers compete for market share through venues other than price competition, such as new product introductions and advertising. The specific nature of wholesalers inflated prices is a matter for future research.

While our results were robust across the different estimators and various specifications of demand, there are several shortcomings of the paper that should be addressed. First, the demand model relies heavily on the assumption that consumers choose one product / brand within the category in a single shopping trip. This assumption is a bit overly restrictive for the CSD market and can generate incorrect consumer responses to marketing mix variables which could effect the margin estimates (Dubé, 2004). Second, if consumers stockpile then static estimates of long-run price sensitivity may be overstated (Hendel and Nevo, 2006). However, we were unable to account for this in the demand model because information on consumer inventory levels was not available. Additional research is needed to account for these weaknesses.

References

[1] Agrawal, D. “Effect of Brand Loyalty on Advertising and Trade Promotions: A Game Theoretic Analysis with Empirical Evidence” *Marketing Science* 15 (1996): 86-108.

[2] Basmann, R. "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation" *Econometrica* 25 (1957): 77-83.

[3] Berry, S., J. Levinsohn, and A. Pakes. “Automobile Prices in Market Equilibrium,” *Econometrica* 63 (1995): 841-890.

[4] Berto Villas-Boas, S. “Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data” *The Review of Economic Studies*. 74 (2007):625-652.

[5] Besanko, D., J Dubé, and S. Gupta. "Competitive Price Discrimination Strategies in a Vertical Channel Using Aggregate Retail Data" *Management Science* 49 (2003): 1121-1138.

[6] Bhat, C. R. "Simulation Estimation of Mixed Discrete Choice Models using Randomized and Scrambled Halton Sequences" *Transportation Research Part B: Methodological* 37 (2003): 837-855.

[7] Cameron, A. and P. Trivedi. *Microeconometrics: Methods and Applications.* New York, Cambridge University Press, 2005.

[8] Cardell, N. S. "Variance components structures for the extreme value and logistic distributions with applications to models of heterogeneity". *Econometric Theory* 13 (1997): 185–213.

[9] Chintagunta, P. K. "Investigating Category Pricing Behavior at a Retail Chain" *Journal of Marketing Research* 39 (2002): 141-154.

[10] Davidson, R and J. MacKinnon. *Econometric Theory and Methods.* New York, Oxford University Press, 2004.

[11] Dhar, T., J. Chavas, R. Cotterill, and B Gould "An Econometric Analysis of Brand-Level Strategic Pricing Between Coca-Cola Company and PepsiCo" *Journal of Economics & Management Strategy* 14 (2005): 905-931.

[12] Draganska, M. and D. Klapper. “Retail Environment and Manufacturer Competitive Intensity” *Journal of Retailing*. 83 (2007):183-198.

[13] Dubé, J. "Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks" *Marketing Science* 23 (2004): 66-81.

[14] Eichenbaum, M., L. Hansen and K. Singleton. "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty" *The Quarterly Journal of Economics* 103 (1988): 51-78.

[15] Erdem, T. "A Dynamic Analysis of Market Structure Based on Panel Data" *Marketing Science* 4 (1996): 359-378.

[16] Hansen, L. "Large Sample Properties of Generalized Method of Moments Estimators" *Econometrica* 50 (1982): 1029-1054.

[17] Hausman, J. A. "Valuation of New Goods Under Perfect and Imperfect Competition" *NBER: The Economics of New Goods* (1997): 207-248.

[18] Heckman J. “Dummy Endogenous Variables in a Simultaneous Equation System,” *Econometrica*. 46 (1978): 931–959.

[19] Hendel, I. and A. Nevo. "Measuring the Implications of Sales and Consumer Inventory Behavior" *Econometrica*. 74 (2006): 1637-1673.

[20] Hess, J., and E Gerstner. "Loss Leader Pricing and Rain Check Policy" *Marketing Science* 6 (1987): 358-374.

[21] Hosken, D., and D. Reiffen. "Multiproduct Retailers and the Sale Phenomenon" *Agribusiness* 17 (2001): 115-137.

[22] Jain, D. C., N. J. Vilcassim, and P. K. Chintagunta. "Random-Coefficients Logit Brand-Choice Model Applied to Panel Data" *Journal of Business & Economic Statistics* 12 (1994): 317-328.

[23] Kadiyali V., N. Vilcassim, and P. Chintagunta. "Empirical Analysis of Competitive Product Line Pricing Decisions: Lead, Follow, or Move Together?" *The Journal of Business*. 69 (1996): 459-487.

[24] —. "Investigating Dynamic Multifirm Market Interactions in Price and Advertising" *Management Science.* 45 (1999): 499-518.

[25] Lal, R., and J. Villas-Boas. “Price Promotions and Trade Deals with Multiproduct Retailers” *Management Science* 44 (1998): 935-949.

[26] Mariano, R. "Simultaneous Equation Model Estimators: Statistical Properties and Practical Implications" Ch. 6 in *A Companion to Econometric Theory,* ed. B. Baltagi, Oxford, Blackwell Publishers, (2001): 122-143.

[27] McFadden, D. L. "Econometric Models for Probabilistic Choice among Products." *Journal of Business* 53 (1980): S13-29.

[28] Nair, H., P. Chintagunta, J. Dubé. "Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants" *Quantitative Marketing and Economics* 2 (2004): 23–58.

[29] Nevo, A. “Measuring Market Power in the Ready-To-Eat Cereal Industry” *Econometrica* 69 (2001): 307-342.

[30] Petrin, A. and K. Train. "A Control Function Approach to Endogeneity in Consumer Choice Models." *Journal of Marketing Research* 47 (2010): 3-13.

[31] Richards, T. "A Nested Logit Model of Strategic Promotion" *Quantitative Marketing and Economics* 6 (2007): 63-91.

[32] Slade, M. "Product Rivalry with multiple Strategic weapons: an Analysis of price and advertising Competition" *Journal of Economics and Management Strategy.* (1995): 445-476.

[33] Staiger, D., and J. Stock. "Instrumental Variables Regression with Weak Instruments" *Econometrica* 65 (1997): 557-586.

[34] Theil, H. "Repeated Least Squares Applied to Complete Equation Systems" The Hague, Central Planning Bureau, mimeo, 1953.

[35] —. *Principles of Econometrics.* New York, A Wiley/Hamilton Publication, 1961.

[36] Train, K. E. *Discrete Choice Methods with Simulation.* New York: Cambridge University Press, 2003.

[37] United States Department of Agriculture. Economic Research Service. 2005. Retrieved June 13 2009, from http://www.ers.usda.gov/Data/

[38] Villas-Boas, J. and R. Winer. "Endogeneity in Brand Choice Models" *Management Science.* 45 (1999):1324-38

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

[40] Walters, R., and S. MacKenzie. "A Structural Equations Analysis of the Impact of Price Promotions on Store Performance" *Journal of Marketing Research* 25 (1988): 51-63.

Table 1: Carbonated Soft Drink Summary Statistics.

Price	Unit	Mean	Std. Dev	Max	Min
Overall Product Price	cents per oz.	*0.0865*	*0.1160*	*0.9626*	*0.0020*
A & W	cents per oz.	0.0166	0.0032	0.0247	0.0047
Barq's	cents per oz.	0.0186	0.0035	0.0249	0.0077
Canada Dry	cents per oz.	0.0176	0.0066	0.0530	0.0074
Coca-Cola Cherry	cents per oz.	0.0214	0.0042	0.0347	0.0081
Coca-Cola Caffeine Free	cents per oz.	0.0229	0.0048	0.0353	0.0117
Coca-Cola	cents per oz.	0.0320	0.0175	0.0725	0.0069
Private Label	cents per oz.	0.0116	0.0039	0.0208	0.0074
Dr. Pepper Caffeine Free	cents per oz.	0.0229	0.0037	0.0284	0.0125
Dr. Pepper	cents per oz.	0.0184	0.0038	0.0328	0.0030
Fanta	cents per oz.	0.0177	0.0041	0.0331	0.0053
I.B.C.	cents per oz.	0.0512	0.0075	0.0554	0.0347
Monster	cents per oz.	0.1337	0.0076	0.1369	0.0994
Mountain Dew Code Red	cents per oz.	0.0154	0.0028	0.0226	0.0086
Mountain Dew	cents per oz.	0.0319	0.0205	0.0740	0.0047
MUG	cents per oz.	0.0171	0.0028	0.0250	0.0089
Pepsi Caffeine Free	cents per oz.	0.0167	0.0036	0.0247	0.0072
Pepsi	cents per oz.	0.0315	0.0197	0.0710	0.0044
Pepsi Vanilla	cents per oz.	0.0164	0.0040	0.0220	0.0055
Pepsi Wild Cherry	cents per oz.	0.0165	0.0033	0.0247	0.0030
Perrier	cents per oz.	0.0601	0.0202	0.0876	0.0296
RC	cents per oz.	0.0149	0.0010	0.0189	0.0102
Red Bull	cents per oz.	0.2399	0.0146	0.2880	0.1955
S. Pellegrino	cents per oz.	0.0663	0.0097	0.0707	0.0296
Schweppes	cents per oz.	0.0225	0.0123	0.0536	0.0058
Seagram's	cents per oz.	0.0185	0.0047	0.0312	0.0108
7 Up	cents per oz.	0.0109	0.0145	0.0526	0.0003
Shasta	cents per oz.	0.0152	0.0038	0.0267	0.0069
Sierra Mist	cents per oz.	0.0265	0.0191	0.0695	0.0038
Sprite	cents per oz.	0.0307	0.0188	0.0730	0.0033
Squirt	cents per oz.	0.0178	0.0033	0.0277	0.0069
Sunkist	cents per oz.	0.0166	0.0050	0.0645	0.0044
Vintage	cents per oz.	0.0152	0.0033	0.0204	0.0074
Welch's	cents per oz.	0.0148	0.0010	0.0185	0.0102
Electricity Cost	$ per 1000 kWh	99.5769	4.7386	106.0000	92.0000
High Fructose Corn Syrup	cents per lb.	21.1117	0.6256	21.4900	19.4800
Sugar	cents per 10 lbs.	9.8419	1.3456	13.7180	8.51000
Aluminum	$ per Metric Ton	187.9300	11.1521	223.1510	173.1940
Diesel	Index	23.9265	3.2885	31.5700	19.3400

Table 2: Random Coefficient Logit Demand Estimates: Retail Level Demand Model.

Variable	Multinomial Logit Model		Random Coef. Logit Model		Random Coef. Para. Distributions	
	Estimate	t-ratio.	Estimate	t-ratio.	Estimate	t-ratio.
Constant (υ_{jkt})	-8.0518*	-1494.33	-8.0517*	-1503.09	0.0018	1.30
Price (α)	-0.3500*	-21.23	-0.3477*	-21.09	5.9848*	588.71
Discount Dummy (dc_{jkt})	0.1592*	32.91	0.1588*	33.63		
Discount*Price ($dc_{jkt}p_{jkt}$)	-0.4757*	-11.93	-0.4913*	-12.26		
Q_1	0.1025*	24.66	0.1021*	24.50		
Q_2	0.0349*	8.56	0.0344*	8.20		
Q_3	0.0723*	17.79	0.0722*	17.69		
Chicago IL	-0.6257*	-151.03	-0.6254*	-129.52		
Los Angeles CA	-0.9172*	-221.99	-0.9174*	-239.03		
New York NY	-0.8277*	-136.21	-0.8277*	-161.31		
Atlanta GA	-0.5487*	-117.09	-0.5483*	-99.65		
σ_K	0.9254*	782.03	0.9254*	815.95		
λ	-0.1299*	-2.55	-0.1744*	-3.48	0.6214*	15.40
Log-Likelihood	-146209.22		-146202.63			
$T * R^2_{err}$	-		2568.68			
LRI	0.5073		0.5073			

* Indicates significance at the 95% level.

Dependent variable: Probability brand j was purchased in store k.

Table 3: OLS regression of the endogenous price variable on IVs..

Variable	Estimate	t-ratio.
Constant	0.5581*	78.96
Lagged Price (1 week)	0.9600*	1213.66
lagged Pr_k (1 week)	0.0000*	-10.35
HFCS*Chicago, IL	-0.0206*	-72.14
HFCS*Los Angeles, CA	-0.0210*	-73.96
HFCS*New York, NY	-0.0211*	-61.21
HFCS*Atlanta, GA	-0.0203*	-67.53
HFCS*Philadelphia, PA	-0.0205*	-72.40
Aluminum*Chicago, IL	0.0001*	4.86
Aluminum*Los Angeles, CA	0.0002*	10.56
Aluminum*New York, NY	0.0002*	6.19
Aluminum*Atlanta, GA	0.0000	0.76
Aluminum*Philadelphia, PA	0.0000*	2.75
Diesel*Chicago, IL	-0.0022*	-34.63
Diesel*Los Angeles, CA	-0.0024*	-39.00
Diesel*New York, NY	-0.0024*	-22.14
Diesel*Atlanta, GA	-0.0021*	-27.18
Diesel*Philadelphia, PA	-0.0022*	-34.90
Corn*Chicago, IL	-0.0377*	-32.45
Corn*Los Angeles, CA	-0.0414*	-36.57
Corn*New York, NY	-0.0417*	-20.43
Corn*Atlanta, GA	-0.0341*	-24.00
Corn*Philadelphia, PA	-0.0345*	-30.55
R^2	0.906	
F-Stat.	71,724.6	

* Indicates significance at the 95% level.

Dependent variable: Price

HFCS - High Fructose Corn Syrup price.

Table 4: OLS regression of the endogenous measures of the margin variable on IVs.

Dependent Variable:	Retail Margin		Wholesale Marg.	
Variable	Estimate	t-ratio.	Estimate	t-ratio.
Chicago IL	-0.0215	-0.31	-0.0116*	-10.13
Los Angeles CA	0.0275	0.41	-0.0068*	-6.11
New York NY	0.0132	0.18	-0.0108*	-8.97
Atlanta GA	0.0228	0.33	0.0044*	3.86
A & W	0.2546	0.01	0.1171	0.40
Barq's	0.4905	0.18	0.0102	0.22
Canada Dry	0.1995	0.07	0.0152	0.32
Coca-Cola Cherry	0.6496	0.20	-0.0103	-0.19
Coca-Cola Caffeine Free	1.3737	0.28	0.0808	1.00
Coca-Cola	-0.0038	0.00	-0.0124	-0.16
Private Label	0.4336	0.07	-0.0055	-0.05
⋮				
HFCS*A & W	-0.0118	-0.02	-0.0029	-0.25
HFCS*Barq's	-0.0149	-0.14	-0.0001	-0.09
HFCS*Canada Dry	-0.0077	-0.07	-0.0001	-0.07
HFCS*Coca-Cola Cherry	-0.0229	-0.18	0.0015	0.73
HFCS*Coca-Cola Caffeine Free	-0.0391	-0.20	-0.0026	-0.80
HFCS*Coca-Cola	0.0026	0.01	0.0004	0.15
HFCS*Private Label	-0.0155	-0.06	0.0023	0.57
⋮				
Sugar*A & W	-0.0032	-0.01	-0.0013	-0.24
Sugar*Barq's	-0.0145	-0.28	-0.0002	-0.29
Sugar*Canada Dry	-0.0023	-0.04	-0.0005	-0.57
Sugar*Coca-Cola Cherry	-0.0148	-0.24	0.0002	0.27
Sugar*Coca-Cola Caffeine Free	-0.0497	-0.54	-0.0018	-1.24
Sugar*Coca-Cola	-0.0042	-0.05	0.0001	0.07
Sugar*Private Label	-0.0085	-0.07	0.0020	1.06
⋮				
Lagged Price (1 week)	-0.1292	-0.74	-0.0040	-1.41
lagged Pr_k (1 week)	-0.0002	-0.82	-0.0000*	-14.77
lagged $Pr_{j\|k}$ (1 week)	0.0293	0.36	0.0108*	8.08
Uncentered R^2	0.00032		0.02906	
F-Statistic	0.8		76.2	

* Indicates significance at the 95% level.

⋮ - For space's sake the rest were omitted, but they were similar to those shown.

Table 5: Supply Side Estimation Results.

Variable	OLS		LIML		2SLS		GMM	
	Estimate	t-ratio.	Estimate	t-ratio.	Estimate	t-ratio.	Estimate	t-ratio.
ϕ	-0.0000	-0.44	-7.2489	-0.81	-0.1181*	-3.81	-0.1341*	-5.08
θ	0.0121*	5.20	18.953	0.73	3.0856*	7.70	3.5014*	10.13
Electricity	-0.0000	-0.74	0.0146	0.30	-0.2756*	-4.90	-0.1584*	-9.93
HFCS	0.0052*	17.94	-0.0512	-0.34	0.7462*	5.01	0.5089*	10.03
Sugar	0.0028*	10.11	-0.0318	-0.23	0.4908*	5.19	0.3823*	10.05
Gasoline	0.0002*	2.08	0.0217	0.32	0.3028*	4.56	0.0657*	9.13
A & W	-0.1256*	-28.97	-1.4822	-0.56	-0.2748*	-5.17	-0.4726*	-9.66
Barq's	-0.0070*	-5.06	-0.2586	-0.36	0.0007	0.03	-0.1757*	-8.19
Canada Dry	-0.0108*	-7.66	-0.4507	-0.53	-0.0170	-0.78	-0.1954*	-8.93
Coca-Cola Cherry	-0.0160*	-11.03	-0.7245	-0.64	-0.0819*	-3.48	-0.2684*	-10.79
Coca-Cola Caffeine Free	-0.0352*	-20.81	-0.0956	-0.12	-0.0410	-1.70	-0.2193*	-9.25
Coca-Cola	-0.0643*	-38.55	-0.2749	-0.34	-0.0444	-1.83	-0.2193*	-9.60
Private Label	-0.1148*	-59.30	-1.5426	-0.73	-0.3042*	-8.58	-0.5074*	-14.01
Dr. Pepper Caffeine Free	-0.0774*	-52.24	-0.7166	-0.69	-0.1152*	-5.05	-0.2979*	-12.68
Dr. Pepper	-0.1086*	-63.37	-0.5550	-0.58	-0.1100*	-4.51	-0.2877*	-12.21
Fanta	-0.1208*	-59.78	-4.8418	-0.83	-0.3977*	-9.18	-0.6127*	-14.36
⋮								
J-Statistic	-		75,871.06		1,582.44		2,809.285	
$T * R^2_{err}$	-		30.647		1,158.047		2,208.106	
Hausman statistic	-		0.680		81.7886		166.643	
Overidentifying test stat.	-		28.731		-		-	

* Indicates significance at the 95% level.

Dependent variable: Price.

HFCS - High Fructose Corn Syrup price.

⋮ - For space's sake the rest were omitted, but they were similar to those shown.