Vertical Contracts between Manufacturers and Retailers: An Empirical Analysis*

Sofia Berto Villas-Boas
(University of California, Berkeley)

October, 2002

ABSTRACT

This paper tests different models of vertical contracting between manufacturers and retailers in the supermarket industry. I estimate demand and use the estimates to compute price-cost margins for retailers and manufacturers under different supply models without observing wholesale prices. I then test which set of margins seems to be compatible with the margins obtained from direct estimates of cost and select the best among the non-nested competing models. The models considered are: (1) a double marginalization pricing model; (2) a vertically integrated model; and (3) a variety of alternative (strategic) supply scenarios, allowing for collusion, non-linear pricing and strategic behavior with respect to private label products. Using data on yogurt sold at several stores in a large urban area of the United States, I find that wholesale prices are close to marginal cost and that retailers have pricing power in the vertical chain. This is consistent with non-linear pricing by the manufacturers or with high bargaining power of the retailers.

Keywords: Market power, vertical contracts, multiple manufacturers and retailers, non-nested tests, yogurt local market, non-linear pricing, bargaining power.

* I wish to thank Richard Gilbert and Aviv Nevo as well as Daniel Ackerberg, Severin Borenstein, Kenneth Chay, Bronwyn Hall, Rene Kamita, Michael Katz, Daniel McFadden, Thomas Rothenberg, Miguel Villas-Boas and Catherine Wolfram for many helpful discussions and advice. I also wish to thank seminar participants at Boston University, Columbia University, Northwestern University, NYU, Princeton University, Stanford University, UC Berkeley, UCLA, Universidade Católica Portuguesa, University of Chicago and Universidade Nova de Lisboa for helpful comments and suggestions. Fellowship *PRAXIS BD/9128/96* from the Fundação para a Ciência e Tecnologia is gratefully acknowledged. Address: Department of Agricultural and Resource Economics, University of California at Berkeley, 310 Giannini Hall, Berkeley, CA 94720-3310; e-mail: sberto@are.berkeley.edu.

1. Introduction

Vertical contracts are multidimensional agreements between manufacturers and retailers with terms that reflect the relative bargaining power of the parties involved and that are responses to moral hazard and adverse selection problems and to the need of risk sharing between the parties. There are several wide-ranging reasons why we should care about analyzing vertical contracts. First, vertical contracts may promote efficiency in the vertical channel. This efficiency is a result from the departure from the simple uniform pricing scheme that results in double marginalization. The problem of double marginalization arises when the only contractual instrument used is the wholesale price. As a consequence, the sum of profits for the manufacturer and retailer may be less than if they could have coordinated their decisions. Second, vertical contracts may impair competition through their horizontal effects on the upstream (manufacturer) and downstream (retail) markets by increasing the possibility for oligopolistic coordination (increasing market power) or by excluding rivals (and hence diminishing product variety and choices). Finally, the vertical structure in a particular market can significantly affect downstream prices (see Hastings, 2002) and price dynamics (see, for example, Chevalier, Kashyap and Rossi, 2000) and condition the assessment of merger activities in the upstream and downstream markets. While there is extended theoretical work on vertical contracts (for a survey, see Katz (1989)), vertical contracts are especially difficult to examine empirically due to their infra-marginal components and due to limited data availability. In particular, wholesale price data are typically unavailable and retailers' and manufacturers' marginal costs are difficult to measure separately. This paper presents a method to analyze vertical contracting that does not require data on wholesale prices or marginal cost (of either the retailers or the manufacturers).

In this paper I focus on whether the contracting between manufacturers and retailers in the supermarket industry follows the double marginalization model or something more efficient. The research plan of this paper is as follows: First, I estimate demand and use the estimates to compute price-cost margins for retailers and manufacturers under different supply models, without observing wholesale prices. I then assess the fit of these different vertical models and select the best among the competing non-nested models.

The first supply model I consider is the double marginalization model. The implied price-cost margins are inefficient from the perspective of the joint profit of retailers and manufacturers. The second model I consider is vertically integrated pricing, which will maximize joint profits and therefore is the efficient outcome from the retailers and manufacturers point of view. The implied price-cost margins correspond to those of a vertically integrated monopolist who sells all the products in the market. I also examine intermediate scenarios incorporating the role of private labels, collusion and fixed fees in their design. In two of the models, either the retailers or the manufacturers are allowed to use non-linear pricing contracts (involving fixed fees). In another model, the retailers are assumed to behave as if they were vertically integrated with respect to the private labels. Finally, collusion at the manufacturer level or at the retailer level is examined. I empirically focus on the yogurt market, defined by two zip code areas, in a large Midwestern city. This paper uses a panel data set of quantities sold and retail prices for 43 yogurt products over a period of 104 weeks collected by scanning technology at three retailers in the market. I choose this product category because it has substantial retail price variability that is not solely due to promotional retail activity, which in turn is important for input price changes to be reflected in changes of retail prices. Another reason for choosing this product has to do with the potential wholesale price variability due to its short shelf life. Consequently manufacturers can adjust wholesale prices more often to respond to significant marginal cost changes.

The results do not provide support for models that imply double marginalization pricing in the vertical structure. The supply model that fits the data best assumes that wholesale prices are close to marginal cost and that the retailers have pricing power in the vertical chain. This is consistent with high bargaining power of the retailers or with non-linear pricing by the manufacturers. In the optimal non-linear pricing contract, the manufacturer sets the marginal wholesale price close to the manufacturer's marginal cost in order for the retailer to have the right incentives when setting the retail prices. Then the manufacturer transfers revenue from the retailers via a fixed fee or by selling the non-marginal units at higher wholesale prices.

The two main contributions of this paper are that, given demand assumptions, I am able to (1) estimate, without observing wholesale prices, the price-cost margins for all manufacturers and all retailers in a certain local market given different supply models; and (2) test the validity of each one of the models by comparing the computed margins with the price-cost margins estimated using components of marginal cost. Previous work, typically, does not model the retailers' decisions (for example, BLP (1995) and Nevo (2001)). In these papers the implied price cost margins are determined by the manufacturers and by maximizing the profits from the set of products that each of them sells. My results suggest that, at least for the market I study, the model that is more consistent with the data has retailers making the pricing decisions not the manufacturers. This model implies different price-cost margins, since the retailers and manufacturers will be maximizing their profits over a different set of products.

In terms of the techniques used in this paper, the estimation of firm's (implied) price-cost margins without observing actual costs follows Bresnahan (1981, 1987) (see Bresnahan (1989) for a survey). The starting point is the estimation of a demand system and the elasticities of substitution between the different products. In the context of oligopoly markets with differentiated products, two problems may arise: the high dimensionality of elasticities to be estimated (equal to the square of the number of products) and the endogeneity of prices. To solve the dimensionality problem I follow the discrete choice literature (see, e.g., McFadden (1973,1984), Cardell (1989), Berry (1994), Berry, Levinsohn and Pakes (1995) and Nevo (2001)) by projecting consumer choices on a set of product characteristics, with smaller dimension than the square of the number of products. To account for the fact that prices set by retailers and manufacturers can be correlated with unobserved product characteristics I use, as instruments for prices, direct components of marginal cost, namely input prices, interacted with product-specific fixed effects. The intuition for interacting input prices with product dummies is to allow for each input to enter the production function of each product differently. This is a new approach to instrument for prices and, given the good first-stage fit, appears to generate robust results.

Several recent papers examine retailer and manufacturer vertical relationships in different industries (see e.g., Bresnahan and Reiss (1985) in the automobile market, Corts (2001) in the U.S. motion picture industry and Mortimer (2002) for video rentals). More closely related to this paper, Chintagunta, Bonfrer and Song (2000) estimate the impact of the introduction of a private label by one retailer on the relative market power of the retailer and the manufacturers and Kadiyali, Chintagunta and Vilcassim (2000) measure the share of profits to retailers and manufacturers. Two key distinguishing features of this paper relative to the two previous ones is that they use data on wholesale prices reported by the retailer and that they use a conduct parameter approach (that measures deviations from Bertrand pricing behavior) in their analysis. Finally, Villas-Boas and Zhao (2001) evaluate the degree of manufacturer competition and the retailer and manufacturers interactions in the ketchup market in a certain city and Sudhir (2001) studies competition among manufacturers under alternative assumptions of vertical interactions with one retailer. One innovation of the paper is to allow for multiple retailers when analyzing the vertical interactions between manufacturers and retailers.

The rest of this paper is organized as follows. The next section presents the model. Section 3 describes with more detail the method of estimation, the instruments and the testing procedures used. In section 4, I describe the yogurt market and the data being used. Finally, section 5 looks at the results, and section 6 presents conclusions and extensions indicating, in particular, how the methodology proposed here can be used in different settings.

2. The Model

The model consists of a standard discrete choice demand model and of different alternative scenarios of vertical relationships between manufacturers and retailers. For each supply scenario the price-cost margins for the retailers and for the manufacturers are expressed solely as functions of demand substitution patterns.

2.1. Demand Side

Let the consumer choose each period t among N_t different products[1] sold by three retailers. Using the typical notation for discrete choice models of demand, the indirect latent utility of consumer i from buying product j during week t is given by

$$U_{ijt} = d_j + x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ijt} \tag{1}$$

where d_j are product (brand-store) fixed effects capturing time invariant product characteristics, x_{jt} are the observed product characteristics, p_{jt} is the price of product j, ξ_{jt} are the mean across consumers of unobserved (by the econometrician) product characteristics (or better, changes in unobserved product characteristics since a product dummy is included in equation (1)) and ϵ_{ijt} represents the distribution of consumer preferences about this mean with density $f(\epsilon)$. The random coefficients β_i are consumer taste parameters for the different product characteristics and the term α_i represents the marginal utility of price. They are allowed to vary across consumers according to

$$[\alpha_i, \beta_i]' = [\alpha, \beta]' + \Gamma\ D_i + \Upsilon\ v_i \tag{2}$$

where the variable D_i has observed consumer characteristics such as demographics, while unobserved consumer characteristics are picked up by v_i.

The matrices of non-linear demand parameters to be estimated are Γ and Υ. Unobserved consumer characteristics v_i are assumed to be normally distributed $N(0, I)$, and the observed consumer characteristics D_i have an empirical distribution $\hat{\mathrm{F}}(\mathrm{D})$ from the demographic data. Additionally an outside good is included in the model, called good zero, allowing for the possibility of consumer i not buying one of the N_t marketed goods. Its price is not set in response to the prices of the other N_t products. In the outside good I include yogurts sold by smaller retail stores or grocery stores not considered in the analysis and also yogurts of small manufacturers sold in the three retail stores

[1] The same physical product sold at two different retailers is defined as two different products.

studied. As usual, the mean utility of the outside good, δ_{0t}, is normalized to be constant over time and equal to zero.[2] Given a measure M of the market size, assumed proportional to the population in the contiguous zip code areas where the stores are located, then the observed market share of product j is given by $s_j = q_j/M$, where q_j are the units sold. [3]

I make the usual assumption that consumers purchase one unit of that product among all the possible products available at a certain time t that maximizes their indirect utility.[4] Then the market share of product j during week t is given by the probability that good j is chosen, that is,

$$s_{jt} = \int 1[(D_i, v_i, \epsilon_{it}) \mid U_{ijt} \geq U_{iht} \; \forall \; h = 0, ...N_t] \; dF(\epsilon) \; dF(v) \; dF(D). \tag{3}$$

If consumer heterogeneity enters only through the random shock (that is, assuming that both D and v are fixed) and ϵ_{ijt} is distributed i.i.d. with an extreme value type I density, then (3) becomes

$$s_{jt} = \frac{e^{\delta_{jt}}}{e^{\delta_{0t}} + \sum_{k=1}^{N_t} e^{\delta_{kt}}} = \frac{e^{\delta_{jt}}}{1 + \sum_{k=1}^{N_t} e^{\delta_{kt}}} \tag{4}$$

which is the Multinomial Logit model. Assuming still that ϵ_{ijt} is distributed i.i.d. extreme value, but now allowing for consumer heterogeneity to affect the taste parameters for the different product characteristics corresponds to the full random coefficients model or mixed Logit model.[5] The market share of product j will no longer have a closed form expression.

2.2. Supply Side

In what follows, each supply model is described in detail and then solved to obtain an expression for both the retailer's and the manufacturer's implied price-cost margins just as a function of demand side parameters.[6]

---

[2]Without making any additional assumptions it would not be identified. The alternative would be to normalize any one of the N_t goods.

[3]In this case, q_j are the servings sold of yogurt. One serving corresponds to a cup of 6 ounces. Accordingly, p_j is the price per serving of product j.

[4]The studies that explicitly model multiple-discrete choices (e.g., Dubin and McFadden, 1984; Hanemann, 1984; Hausman, Leonard and McFadden, 1995; Hendel, 1999 and Dubé, 2001) need individual level data for estimation. Since this paper uses only market-level data the techniques proposed could not be directly applied here.

[5]This is a very general model. As shown in McFadden and Train (2000), any discrete choice model derived from random utility maximization can be approximated, to any degree of accuracy, to a Mixed Logit.

[6]The (Logit and random coefficients) expressions of the price-cost margins in a simplified model with only two retailers, two wholesalers and two products are available at http://socrates.berkeley.edu/∼villas/homepage.html.

2.2.1. Scenario 1: Double Marginalization Model

In this model manufacturers set their prices first, and retailers follow, setting the retail prices given the wholesale prices. The margins that result from their behavior correspond to the pure double marginalization price-cost margins in the context of linear pricing in oligopoly markets at the manufacturer and at the retail level.

Let there be N_r Nash-Bertrand retailers competing in the retail market and let their be N_w Nash-Bertrand manufacturers competing in the wholesale market. To solve this vertical model one starts, as usual, by looking at the retailer's problem. Each retailer r's profit function in week t is given by

$$\pi_{rt} = \sum_{j \epsilon S_{rt}} \left[p_{jt} - p_{jt}^{w} - c_{jt}^{r} \right] s_{jt}(p) \tag{5}$$

where S_{rt} is the set of products sold by retailer r during week t, p_{jt}^{w} is the wholesale price he pays for product j, c_j^r is the retailer's marginal cost of product j and $s_{jt}(p)$ is the share of product j. The first order conditions, assuming a pure strategy Nash-equilibrium in prices, are

$$s_{jt} + \sum_{m \epsilon S_{rt}} [p_{mt} - p_{mt}^{w} - c_{mt}^{r}] \frac{\partial s_{mt}}{\partial p_{jt}} = 0 \quad \text{for} \quad j = 1, \cdots, N_t \tag{6}$$

where N_t is the number of products in the market.

Define T_r as the retailer's ownership matrix with the general element $T_r(i, j)$ equal to one when both products i and j are sold by the same retailer and zero otherwise. Let Δ_{rt} be the retailer's response matrix, containing the first derivatives of all the shares with respect to all retail prices, with element $(i, j) = \frac{\partial s_{jt}}{\partial p_{it}}$. Stacking up the first order conditions given by (6) for all N_t products and rearranging terms, the following vector expression for the retailers' implied price-cost margins just as a function of the demand side for each week t is obtained

$$p_t - p_t^{w} - c_t^{r} = -(T_r * \Delta_{rt})^{-1} s_t(p), \tag{7}$$

where $T_r * \Delta_{rt}$ is the element by element multiplication of the two matrices. If the equilibrium is unique, equation (7) implicitly defines the retail prices as a function of the wholesale prices.

Looking now at the manufacturer, each of them maximizes his profit choosing the wholesale prices p^w, knowing that the retailers behave according to (7). The manufacturer's profit function is given by

$$\pi_{wt} = \sum_{j \epsilon S_{wt}} [p_{jt}^{w} - c_{jt}^{w}] \; s_{jt}(p(p^{w})), \tag{8}$$

where S_{wt} is the set of products sold by manufacturer w during week t and c^w_{jt} is the marginal cost of the manufacturer that produces product j. The first-order conditions are, assuming again a pure strategy Nash-Equilibrium in the wholesale prices,

$$s_{jt} + \sum_{m \epsilon S_{wt}} [p^w_{mt} - c^w_{mt}] \frac{\partial s_{mt}}{\partial p^w_{jt}} = 0 \text{ for } \ j = 1, \cdots, N_t. \tag{9}$$

Let T_w be a matrix of ownership for the manufacturers, analogously defined as the matrix T_r above. In particular, element (j, m) of T_w is equal to one if manufacturer who sells product j also sells product m and is equal to zero otherwise. Let Δ_{wt} be the manufacturer's response matrix, with element $(j, m) = \frac{\partial s_{mt}}{\partial p^w_{jt}}$, containing the derivatives of the market shares of all products with respect to all wholesale prices, which in turn depends on the curvature of demand. In other words, this matrix has the cross-price elasticities of the derived demand and the effect of cost pass-through.[7] Collecting terms and solving for the manufacturers' implied price-cost margins yields

$$p^w_t - c^w_t = -(T_w * \Delta_{wt})^{-1} s_t(p). \tag{10}$$

To obtain Δ_{wt}, first note that $\Delta_{wt} = \Delta'_{pt} \Delta_{rt}$, where Δ_{pt} is a matrix of derivatives of all the retail prices with respect to all the wholesale prices. So all that is needed is to find expressions for, and compute, Δ_{pt} and pre-multiply Δ_{rt}, from the retailer's problem, by the transpose of Δ_{pt} to get the manufacturer's response matrix Δ_{wt}. From now on, the time subscript is dropped to simplify notation. To get the expression for Δ_p, let us start by totally differentiating for a given j equation (6) with respect to all prices $(dp_k, k = 1, \cdots, N)$ and a wholesale price p^w_f, with variation dp^w_f:

$$\sum_{k=1}^{N} \underbrace{[\frac{\partial s_j}{\partial p_k} + \sum_{i=1}^{N} (T_r(i,j) \frac{\partial^2 s_i}{\partial p_j \partial p_k} (p_i - p^w_i - c^r_i)) + T_r(k,j) \frac{\partial s_k}{\partial p_j}]}_{g(j,k)} \ dp_k \ - \ \underbrace{T_r(f,j) \frac{\partial s_f}{\partial p_j}}_{h(j,f)} dp^w_f = 0. \tag{11}$$

Putting all $j = 1,...N$ products together, let G be the matrix with general element $g(j, k)$ and let H_f be the N dimensional vector with general element $h(j, f)$. Then

$$G \ dp - H_f \ dp^w_f = 0. \tag{12}$$

Solving for the derivatives of all prices with respect to the wholesale price f the f-th column of Δ_p

[7] This matrix becomes very complicated with multiple products in the context of multiple retailers and manufacturers. Please refer to supplement available at http://socrates.berkeley.edu/∼villas/homepage.html for details.

is obtained:

$$\frac{dp}{dp_f^w} = G^{-1}H_f. \tag{13}$$

Stacking all N columns together, $\Delta_p = G^{-1}H$, which has the derivatives of all prices with respect to all wholesale prices. The general element of Δ_p is $(i,j) = \frac{\partial p_j}{\partial p_i^w}$. Finally, the sum of the implied price-cost margins for the retailers and the manufacturers is by definition obtained by adding up (7) and (10)

$$p_t - c_t^r - c_t^w = -(T_r * \Delta_{rt})^{-1} s_t(p) - (T_w * \Delta_{wt})^{-1} s_t(p). \tag{14}$$

2.2.2. Scenario 2: Non-Linear Pricing Models

In a one manufacturer and one retailer case, in the classical nonlinear optimal (two-part tariff) pricing model, the manufacturer sets the wholesale price equal to marginal cost and lets the retailer be the residual claimant. Then the manufacturer is able to extract part or the full "monopoly" (or vertically integrated firm's) surplus in the form of a fixed fee that the retailer has to pay. Two-part tariffs are seen as optimal contracts whenever there is downstream market power in the retail market[8] and under fairly general market assumptions. Two-part tariff as the optimal contract design has been shown to be optimal in the simple double marginalization model where retailers follow manufacturers in a price setting game with a certain demand (Tirole, 1988, page 176), an uncertain demand (Rey and Tirole, 1986) or under asymmetric information (Tirole, 1988, page 177).[9] It is no longer true that the optimal two-part tariff in the context of multiple retailers yields marginal cost pricing by the manufacturers (Mathewson and Winter, 1984 and Schmalensee, 1981). However two-part tariffs are still optimal in the context of multiple manufacturers and a single retailer (Shaffer and O'Brien, 1997 and Tirole, 1988, page 180). In the one manufacturer, one retailer model, resale price maintenance implies that the manufacturer sets a wholesale price equal to the monopoly price and then imposes resale price at the monopoly price. The retailer makes zero profit, and the manufacturer recovers the monopoly profit.

Scenario 2 considers the existence of these non-linear pricing contracts in the context of multiple retailers and multiple manufacturers. In particular, two sub-cases are considered to test the validity of two solutions to the double marginalization problem. In the first case, the wholesale margins are assumed to be zero, which means that the retailers have the pricing decision given that wholesale prices equal marginal costs. In the second case, the retail margins are assumed to be zero. Given

[8] This is not to be confused with retailers having vertical power in the vertical structure. If there is retail market power, retailers impose an externality by adding a second margin to the wholesale margin.

[9] However, in the presence of uncertainty, two-part tariffs have poor properties in terms of risk sharing.

that, it is as if the manufacturers were setting the final price consumers are facing, like in resale price maintenance. In both of these sub-cases, the profits of the vertical structure may be greater than the sum of the profits of retailers and manufacturers in the first scenario of double marginalization. The potential increase in the whole channel's profits is due to the elimination of the first or the second margin in the vertical structure and the fact that the retailers have some retail power, i.e., face downward sloping demand curves.

Case 1: "Wholesale margins are zero and retailers have pricing decisions." In this case, retailers maximize their profits, given that wholesale prices are equal to marginal costs. The manufacturers' implied price-cost margins are zero for all products. The implied price-cost margins for the retailers are given by equation (7) subject to $p_t^w = c_t^w$, that is,

$$p_t - c_t^r - c_t^w = -(T_r * \Delta_{rt})^{-1} s_t(p). \tag{15}$$

This means that the retailer gets from his optimization the profits corresponding to the downstream vertically integrated structure for each of the j products.

Case 2: "Zero retail margins and manufacturers have pricing decisions." In this case, the retail implied price-cost margins are zero for all products since the retailers add to the wholesale prices only the retail costs, i.e. $p_{jt} = p_{jt}^w + c_{jt}^r \ \ \forall j$. This means that the manufacturers get from their optimization the profits corresponding to the downstream vertically integrated structure for each of the j products. The manufacturers' implied price-cost margins are given by

$$p_t^w - c_t^r - c_t^w = -(T_w * \Delta_{rt})^{-1} s_t(p). \tag{16}$$

It is worth noting that the implied price-cost margins in equation (16) are different from equation (15) because the retail ownership T_r differs from the manufacturer ownership T_w or, in other words, because the manufacturers and the retailers are maximizing their profits over a different set of products. In BLP (1995) and Nevo (2001) the (manufacturer) implied price-cost margins computed are given by expressions similar to (16) and the retailers' decisions are not modeled.

2.2.3. Scenario 3: The Hybrid Model

Each retailer behaves as a vertically integrated firm with respect to its own private label products and plays the vertical Nash-Bertrand game in the other products (the national brands). This scenario's implied price-cost margins have bits and pieces of the ones from scenario 1 and scenario 2's first case (for the expressions in a simple model, please refer to the appendix). In particular, the retail margins will be the same as in scenario 1 given by equation (7). However, the wholesale margins change: When vertically integrating into the upstream market, the retailers affect the price-cost margins of the national brands' manufacturers. By vertically integrating, the retailers eliminate the wholesale margins in the private labeled products, and the final retail price of the private labels falls. Demand for the products sold by the manufacturers of national brands changes (decreases), and consequently the national brand manufacturers need to adjust their wholesale prices. For this particular market, at the manufacturer level, the wholesale margins for the private label products are zero and thus not optimized over. The implied manufacturers' price-cost margins for the national brands are given by

$$p_t^w - c_t^r - c_t^w = -(T_w^* * \Delta_{wt}^*)^{-1} s_t^*(p), \tag{17}$$

where T_w^* is the manufacturers' ownership matrix as before without the rows and columns that correspond to the private label products in the sample. The expression for Δ_{wt}^* is equal to Δ_{wt} but without rows and columns of the private label products. In $s_t^*(p)$ are the shares of the national brands, namely $s_t(p)$ without rows for the private label products.

2.2.4. Scenario 4: Manufacturer-Level Collusion Model

This scenario corresponds to manufacturers choosing wholesale prices that maximize the sum of the manufacturers' profits. Because manufacturers are assumed to be colluding, it is as if one single upstream firm owned the full set of products. Thus the manufacturers' ownership matrix T_w is a matrix full of ones, henceforth called T_1. Manufacturers' price-cost margins are given by equation (10) subject to $T_w = T_1$, which results in

$$p_t^w - c_t^w = -(T_1 * \Delta_{wt})^{-1} s_t(p). \tag{18}$$

The implied price-cost margins of the retailers, which are assumed to set their retail prices given the wholesale prices, are given by

$$p_t - p_t^w - c_t^r = -(T_r * \Delta_{rt})^{-1} s_t(p). \tag{19}$$

Finally, the sum of the implied price-cost margins for the retailers and the manufacturers is obtained by adding up (18) and (19)

$$p_t - c_t^r - c_t^w = -(T_r * \Delta_{rt})^{-1} s_t(p) - (T_1 * \Delta_{wt})^{-1} s_t(p). \tag{20}$$

2.2.5. Scenario 5: Retail Level Collusion Model

Assuming collusion at the retail level corresponds to assuming that $T_r = T_1$. Retail price-cost margins are given by

$$p_t - p_t^w - c_t^r = -(T_1 * \Delta_{rt})^{-1} s_t(p), \tag{21}$$

while manufacturer price-cost margins have the following expression

$$p_t^w - c_t^w = -(T_w * \Delta_{wt})^{-1} s_t(p). \tag{22}$$

Finally, the sum of the implied price-cost margins for the retailers and the manufacturers is, by definition, obtained by adding up (21) and (22)

$$p_t - c_t^r - c_t^w = -(T_1 * \Delta_{rt})^{-1} s_t(p) - (T_w * \Delta_{wt})^{-1} s_t(p). \tag{23}$$

2.2.6. Scenario 6: Monopolist Model

This last scenario examines the question of whether this industry is jointly profit maximizing. The way scenario 6 relates to scenario 2 is that one can view scenario 2 as asking the question of whether this industry has efficient pricing subject to horizontal constraints. Scenario 6 implies, in a world where there are several manufacturers and several retailers, that they not only vertically integrate but that they coordinate their horizontal pricing decisions (meaning that they are colluding). The fact that they cannot collude is what is meant by horizontal constraints in scenario 2. In this present scenario, wholesale margins are zero. Furthermore, $T_r = T_w = T_1$. Consequently the implied price-cost margins of the full vertically and horizontally integrated structure are given by

$$p_t - c_t^r - c_t^w = -(T_1 * \Delta_{rt})^{-1} s_t(p). \tag{24}$$

3. Estimation and Testing Procedures

3.1. Demand Estimation using Generalized Method of Moments (GMM)

When estimating demand the idea is to estimate the parameters that produce product market shares close to the observed ones. This procedure is non-linear in the demand parameters and suffers from the fact that prices are endogenous variables. The key step is therefore to construct a demand side equation to be estimated linear in the parameters associated with the endogenous variables so that instrumental variables estimation (or, in other words, GMM estimation) can be directly applied. This follows from equating the estimated product market shares[10] to the observed shares and solving for the mean utility across all consumers that is defined as

$$\delta_{jt}(\Gamma, \Upsilon) = d_j + x_{jt}\beta - \alpha p_{jt} + \xi_{jt}. \tag{26}$$

For the Logit model the mean utility δ_{jt} can be recovered analytically, following Berry (1994)'s inversion technique, by $log(s_{jt}) - log(s_{0t}) = \delta_{jt}$. However, in the full model, solving for the mean utility has to be done numerically (see BLP, 1995). Finally, once this inversion has been made, one obtains equation (26) which is linear in the parameter associated with price. Let θ be the demand side parameters to be estimated. In the Logit case $\theta = \theta_L = (\alpha, \beta, d)$ and in the full model $\theta = (\theta_L, \Gamma, \Upsilon)$ where Γ and Υ are the non-linear parameters. For the Logit case, θ_L is obtained directly from estimating (26) by 2SLS.[11] In the full random coefficients model, θ is obtained by GMM following Nevo's (2000) estimation algorithm, where equation (26) enters in one of the steps.[12] To ensure finding a global minimum, I start by using a gradient method (providing an analytical gradient) with different starting values of the non-linear parameters to find a minimum of the GMM objective function. Then I use that minimum as a starting value for the Nelder-Mead (1965) simplex search method (which is a direct search method that does not use numerical or analytical gradients) to check if the results coincide.[13]

Finally, robust standard errors of the parameters are obtained. For the random coefficients model, the White (1980) estimate of the covariance matrix of the demand side parameters estimated and defined in the presence of heteroscedastic demand residuals, is given by

[10]For the Logit model the expression for the estimated market share is given by (4). For the random coefficient model the product market share in equation (3) is approximated by the Logit smoothed accept-reject simulator given by

$$s_{jt} = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{e^{\delta_{jt} + [x_{jt}, p_{jt}](\Gamma\ D_i + \Upsilon\ v_i)}}{1 + \sum_{k=1}^{N_t} e^{\delta_{kt} + [x_{kt}, p_{kt}](\Gamma\ D_i + \Upsilon\ v_i)}}, \tag{25}$$

where N_s are the random draws from the distribution of v and D. This simulator is continuous in the data and in the parameters to be estimated, so gradient-based methods are applied to estimate Γ and Υ.

[11]This is optimal in the presence of homoscedastic errors. The 2SLS estimators are unbiased, consistent and asymptotically normally distributed even in the presence of heteroscedasticity. However, one needs to obtain an appropriate White (1980) estimate of the 2SLS estimators' variance covariance matrix.

[12]The main idea is to concentrate the GMM objective function such that it will be only a function of the non-linear parameters. By expressing the optimal vector of linear parameters as a function of the non-linear parameters and then substituting back into the GMM objective function it can be optimized with respect to the non-linear parameters alone.

[13]The Nelder-Mead search method is generally less efficient and slower to converge than the gradient methods but is more robust specially if the objective function is highly discontinuous.

$$
\begin{aligned}
V_{rc} = \; & ([X \vdots F_{(\Gamma,\Upsilon)}]' Z_d (Z_d' Z_d)^{-1} Z_d' [X \vdots F_{(\Gamma,\Upsilon)}])^{-1} [X \vdots F_{(\Gamma,\Upsilon)}]' Z_d (Z_d' Z_d)^{-1} (\sum u_{d,i}^2 Z_d Z_d') (Z_d' Z_d)^{-1} Z_d' [X \vdots F_{(\Gamma,\Upsilon)}] \cdot \\
& ([X \vdots F_{(\Gamma,\Upsilon)}]' Z_d (Z_d' Z_d)^{-1} Z_d' [X \vdots F_{(\Gamma,\Upsilon)}])^{-1},
\end{aligned}
$$

where $[X \vdots F_{(\Gamma,\Upsilon)}]$ is a matrix that has in columns the regressors X associated with the linear demand parameters and the gradient $F_{(\Gamma,\Upsilon)}$ of the mean utility with respect to the non-linear parameters Γ and Υ.[14]

3.2. Instruments and Identification of Demand

The need to use instrumental variables in the estimation of demand results from the fact that when retailers decide retail prices, they take into account all the product characteristics, not only the ones that are observed, the x_{jt}, but also those characteristics that are not observed by the analyst, the ξ_{jt}. The retailers also take into account any changes in their products' characteristics when setting the retail prices. Since a product fixed effect is included, it will capture the product characteristics that are constant over time, both observed and unobserved. The econometric error that remains in ξ_{jt} will therefore only include the changes in unobserved product characteristics such as unobserved promotions, changes in shelf display and/or changes in unobserved consumer preferences. This implies that the prices in (26) are correlated with the changes in unobserved product characteristics affecting demand. Hence, to obtain a precise estimate of the price coefficients, those endogenous variables need to be instrumented for.

Recall that a valid instrument has to satisfy two requirements. It has to be both uncorrelated with the residual and correlated with the endogenous variable that one is instrumenting for. The price decision takes into account cost-side variables, such as input prices. It is reasonable to assume that the prices of inputs are uncorrelated with the changes in unobserved product characteristics, the ξ_{jt}. For example, changes in shelf display are most likely not correlated with input prices such as the prices of milk and sugar. One problem is that there is not brand or product level variation for the prices of the inputs in the data. This can be solved by interacting those input prices with product dummy variables. The idea behind these instruments is that some yogurts are more sugary, fruit yogurts use more fruit, and so on. Since the exact content is not observed, it is estimated this way. Thus these cost data multiplied by product fixed effects are the instruments for the endogenous retail prices. In particular, two sets of instruments are considered when estimating demand to examine the effects of the two alternative instrumental variables specifications. In the first specification, by allowing marginal cost of a given yogurt brand sold at two different retailers

[14] For the Logit model the gradient does not enter the expression.

to vary, I interact 43 product dummies (where product is defined as brand-store) with the input prices. In the alternative specification, I assume that the marginal cost of the same brand sold at two different retailers is the same. This results in 21 brand dummies interacted with the input prices. One last note on why these instruments (input prices multiplied by product fixed effects) are valid. The residual in the demand equation has only the part that is not explained by "store-brand" level fixed effects. If I had not included brand dummies in the demand, I would have the problem that the instruments would be correlated with constant unobserved product characteristics. Since I account for those by estimating the brand-store fixed effects, I don't have that problem.

The remainder of the paper relies heavily on having consistently estimated the demand parameters or, in other words, the demand substitution patterns and the derived demand substitution patterns (determined by the curvature of demand). The standard question is what is the exogenous variation that identifies these substitution patterns? There are basically two sources of identification in the data. One is the relative price variation over time. In this paper the experiment is to ask consumers to choose between different products over time, where a product is perceived as a bundle of attributes (among which are prices). Since prices are not randomly assigned, I use input price changes over time that are significant and exogenous to unobserved changes in product characteristics as instruments for prices. Regarding identification of elasticities for the same product sold at different retailers, presumably their input prices are identical (and the above argument for identification does not apply) so the source of identification there is the fact that the choice set of consumers changed due to one of the stores being closed for renovation during some weeks.

3.3. Supply Model Testing Procedures

After having estimated the demand model in a first step I access the validity of the supply models in a second step.[15] I basically follow two approaches: one more formal non-nested hypothesis testing approach and another more informal, but maybe more intuitive, specification testing procedure.

[15] Alternatively, both supply and demand models could have jointly been estimated. Besides from technical simplicity in doing this two step procedure, an additional advantage is that the model testing procedure described next reduces to simply estimating one equation and provides an elegant way to compare the non-nested models.

3.3.1. Informal Model Specification Test

Starting with the later, the idea is to, given demand estimates θ, estimate a supply equation that is derived from the firms profit maximizing decisions and is given by

$$p = f(c\ \gamma) + SIPCM_r(\theta)\ \lambda_r + SIPCM_w(\theta)\ \lambda_w + u^s, \tag{27}$$

where c is a matrix of cost side variables such as input prices, γ is a vector of coefficients associated with cost-side variables, $SIPCM_r(\theta)$ and $SIPCM_w(\theta)$ are the retail and manufacturer price-cost margins, respectively, implied by the different scenarios and λ_r and λ_w are parameters associated to the implied margins. It is assumed that $f(c\ \gamma) = e^{c\ \gamma}$ to ensure that all products have positive estimated marginal costs.[16] For each model separately, the specification test is to test the null hypothesis that all the coefficients in λ_r and λ_w are not significantly different from one.[17]

3.3.2. Formal Ranking of Supply Models: Non-nested Hypothesis Testing

This paper follows a "menu approach" (as in Bresnahan (1987) or in Gasmi, Laffont and Vuong (1992)) presenting different models of vertical relationships and the objective is to determine which model fits the data better. Because most of the models cannot be nested in another proposed model, pairwise non-nested testing procedures proposed by Smith (1992) are applied here. The basic is that, if a certain model is true, and building on the previously described informal testing procedure, that means that all parameters associated with the price-cost margins variables in equation (27) are not different from one. Then subtract the implied price-margins from the retail price and recover what the marginal costs would be for that vertical supply model in question. Let y be the difference in retail price and price-cost margins. Then, for each pairwise comparison there are two competing regression models M_g and M_h defined, respectively, as $M_g : y_g = X\beta + u_g$ and $H_h : y_h = X\gamma + u_h$. In the model, X is a matrix of input prices and β and γ are parameters to be estimated by GMM. The Cox-type statistic to compare each pair of models is then constructed by examining the behavior under M_h of the difference of the estimated GMM criterion functions for model M_h and for the

[16] An even more informal and intuitive test for the supply models is whether implied marginal costs are negative for some models.

[17] See the supplement to this paper available at http://socrates.berkeley.edu/∼villas/homepage.html on supply estimation, testing procedures and on deriving the corrected variance covariance matrix of the parameters of interest in λ_r and λ_w. The need for the variance correction has to do with the fact that the price-cost margins that are regressors in the supply equation, have embedded estimators (as in Newey and McFadden, 1994) from the first step of estimating demand.

alternative model M_g.[18] The intuition behind the non-nested tests is to see how the price-cost margins of alternative models explain the residual (unobserved determinants of price) of the null model. Normalized, standardized and compared to a standard normal critical value, a large positive statistic in this one-sided goodness of fit test leads to a rejection of the null model M_h against M_g.

3.4. Identifying Power of Non-Nested Tests

How I separate empirically the different models that I am comparing is due to non-linearity of demand together with variation in product ownership structure.[19] The retailers and manufacturers in the data sell different combinations of products and that gives me the variation to estimate a menu of different models of retail pricing and manufacturer pricing (that all implicitly involve a change in the retail and manufacturer ownership structure). It is also crucial that demand is non-linear. As an illustrative example, if demand is linear, vertical models that assume changes in retail ownership, keeping manufacturer ownership constant are all indistinguishable (this would be the case when comparing the double marginalization model with the model of retail collusion). This is because, manufacturers' price-cost margins are unaffected by retail price-cost margins, because the Hessian in equations (10) and (18) is equal to zero.

4. The Market and the Data

4.1. The Market

The analysis focuses on the yogurt market in a Midwestern metropolitan area, more precisely in two zip code areas.[20] The choice of this product, to analyze vertical relationships between manufacturers and retailers has to do with the pattern of yogurt retail price variability and with the potential wholesale price variability. Yogurt has to be consumed within twenty-eight days of its production, so its shelf life is short. In case there are some significant marginal cost changes, there exists, in principle, the possibility for the manufacturers to adjust the wholesale prices accordingly. Some wholesale price variability can therefore be expected. Looking now at the retail price variability of yogurt it is fairly large and does not seem to be uniquely related to promotional activities. This is particularly important since if price promotional aspects were the drivers of retail price variability

[18] See the supplement to this paper available at http://socrates.berkeley.edu/∼villas/homepage.html for the detailed derivation of the test statistic.

[19] However, models that have different wholesale and retail margins but have the same total margins are not distinguishable.

[20] For confidentiality reasons the city's name is not revealed, nor are the retail store names.

then input price changes may not have been reflected in changes of retail prices. This could imply that input prices would be poor instruments for the retail prices. Figure 1 plots the price series for one of the large selling brands of yogurt. Temporary price reductions are characterized by a decrease in price during a number of successive weeks, and after that period price rises to its original level. Such a pattern is not present for the yogurt price series. Looking instead at the price series for a heavily promoted brand, such as a large selling soft drink brand, (see Figure 2) the pattern of price promotions is evident.

Yogurt is produced by a few leading national yogurt manufacturers: Dannon and General Mills, who together account for almost 62% of the total U.S. yogurt sales. Private label brands from retail stores are in third place with 15% of the market and Kraft comes next. All other manufacturers have individual shares of less than 2% (Frozen Food Digest, October 1995: 38). This industry is fairly concentrated at the manufacturer level. Therefore, in principle, it is interesting to confront supply models of upstream price collusion (for example, scenario 4) with the data. Scenario 3 was inspired by the importance that private labels seem to have in the yogurt market. One of the most important characteristics of the yogurt market is that yogurt sales are mostly driven by new product introductions.[21] In 1994, there were over one thousand new dairy product introductions and over one hundred yogurt introductions alone and for the sample period and market considered in this analysis, there are five new product introductions. In terms of product variety, each store sells an average of 150 yogurts from seven manufacturers. Product variety together with successful advertising (influencing consumers' evaluations of the different products) can result in positive price-cost margins for the manufacturers due to product differentiation alone and this would be reflected in the estimates for the price-cost margins in the non-collusive supply scenarios considered in this paper.

At the retail level there is a small number of large retailers (or retail chains) competing directly with each other and who have jointly 75% of total sales to final consumers in the whole metropolitan area. All other retailers not considered have individual shares less than 5% (figures for 1992). Three retail stores are considered in the data, where store 1 is a smaller store than stores 2 and 3. The last two retail stores belong to two retail chains, while store 1 is unique in the whole metropolitan area. The retail stores in the data are located within less than two miles from one another, and in fact two retailers are located at both sides of a street intersection (see Figure 3). Some smaller

[21] On the competitive effects of product line extensions in the U.S. yogurt market by the two leading manufacturers, see Kadiyali, Vilcassim and Chintagunta (1999). In another paper, Draganska and Jain (2000) use store-level data for the yogurt category to derive recommendations for effective product-line extension decisions based on what-if experiments.

grocery stores are located within the two zip code areas considered, but the closest large retail store nearby is located in a different zip area.

4.2. Data

The analysis is done using a data set on retail prices, advertising, aggregate market shares[22] and product characteristics for 43 products produced by five manufacturers. In particular, the number of products are equal to 43 for all weeks but for the weeks during which retailer 2 closed due to remodeling, when the number of products in the sample is 25. Information on consumer demographics, wages by state and input prices is also used.

The price, feature (advertising) and market share data come from an Information Resources Inc. (IRI) scanner data set that covers the purchases in three retail stores in a Midwestern urban area during 104 weeks.[23] Summary statistics for prices, feature, quantity sold and shares are presented in Table 1. Feature is a dummy variable that takes the value of one when the product was featured during that week. Table 1 also presents summary statistics on combined shares for all the products, combined shares for the products sold by each manufacturer and combined shares for the products sold by each retailer. The combined shares for the products analyzed are on average 34%. Quantity sold is defined as servings sold, where one serving corresponds to a 6-ounce yogurt cup. Price and servings sold series for the 43 products in the sample were obtained by aggregation.[24]

Market shares are defined by converting quantity sold to servings sold and then dividing by the total potential servings in the market. The potential market, in terms of servings, is assumed to be half of a serving per capita a week. Hence the potential market in terms of servings is equal to half of the resident population in the two zip code areas. This assumption is consistent with U.S. consumption patterns. According to U.S. Department of Agriculture, Americans consume on average 9 pounds of yogurt a year, which in terms of servings corresponds to approximately half of a serving per capita a week. Table 3 provides the average U.S. per capita consumption for the years 1991-94 and compares it with international patterns. The highest per capita consumption average

---

[22]The household sample is not used in this analysis because it does not seem to be representative. In fact, twenty five products with substantial market shares in the aggregate sample have zero market shares in the individual sample during more than eight weeks. Combining market shares by store there are again discrepancies when using the individual sample. Store one has a 21% combined market share in the individual sample while stores two and three have 18% and 68%, respectively. However in the aggregate data store one has only 6%, store two has 57% and store three has 37%.

[23]I thank David Bell for letting me have access to this data set.

[24]For a particular retailer, a product is defined such that when sold in different sizes would be aggregated as the same product. Also products with the same brand name and with price correlation close to one and with similar product characteristics were aggregated. For the list of the products in the sample, please refer to Table 2.

for the countries considered is in Bulgaria with 3.39 servings per week[25] and the lowest in Russia with less than 25% of a serving per week. The estimates of the marginal utility of price are robust to small variations of the weekly per capita consumption assumption,[26] as can be seen in Table 4.

In the potential market, Dannon comes first in terms of local market shares of its products with an average of 17%. Next comes General Mills with 9%. The private labels come third with 4%. Kraft comes last among the products analyzed with 3%. Furthermore, in this local market, combined market shares for the 10 products sold by retailer 1 are on average 2%, and for store 2, which sells 18 products, they are 20%. Store 3 has average combined shares of its 15 products of about 14%.

The product characteristics data were collected by inspection of the label reads and for those products currently unavailable in any supermarket because they were discontinued, from manufacturers' descriptions. Table 5 describes the following product characteristics: calories, total fat, cholesterol, carbohydrates, dummy for vitamins, dummy for calcium above 30% daily value, Aspartame dummy, Fruit on the Bottom dummy, available in different sizes dummy and store dummy variables. These are most of the characteristics found to be more relevant to consumers when purchasing yogurt, according to Frozen Food Digest (1995), and also according to manufacturer's yearly market surveys and brand-name comparison articles (e.g., Nutrition Action, 1998, Center for Science in the Public Interest).

The cost data set is described in Table 6, with reference about the different sources. For most of the input price data series there is considerable time variation as can be seen in Figures 4 and 5 where the weekly input price series normalized by the correspondent weekly average are presented.

Using information on consumer demographics, such as family size, income and age, allows us to consider consumer heterogeneity in the taste parameters for the different products. A sample from the joint distribution of income and age of the resident population was obtained from the 1990 Census at the zip code level, for the zip codes of interest (see Table 7). The population in the market considered is about 76% white with median household income of about thirty thousand dollars a year and on average with 2.5 persons in the household.

[25] The primary yogurt culturing bacteria, Lactobacillus bulgaricus, was named by 1908's Physiology and Medicine Nobel laureate Dr. Metchnikoff, in honor of the yogurt-loving Bulgarians.

[26] The ranking of the different supply models is also invariant to small changes in weekly per capita assumption.

5. Results

5.1. *Demand Estimation*

The Logit model for demand is considered first to get a feel for what is going on in the data. It also allows to compare and choose between two different instrumental variable specifications and illustrate the need to instrument for prices when estimating demand. Understanding the drawback of having poor substitution patterns, I then estimate a full random coefficients discrete choice model of demand for differentiated products.

5.1.1. *Logit Demand*

Table 8 presents the results from regressing the mean utility δ_j, which for the Logit case is given by $ln(s_{jt}) - ln(s_{0t})$, on prices and product dummy variables in equation (26). The second column displays the estimate of ordinary least squares for the mean price coefficient α and columns three and four have estimates of α for two different instrumental variables (IV) specifications, using input prices as instruments for the prices. In the first IV specification, assuming that marginal cost for the same product sold at different retailers is different, prices are instrumented by input prices interacted with 43 product dummy variables. In the second IV specification, prices are instrumented by input prices interacted with 21 product dummies. This last specification corresponds to assuming that marginal cost for the same product sold at different retailers is constant. Regarding the need to instrument for prices, the Hausman (1978) test for exogeneity suggests that there is a gain from using instrumental variables versus ordinary least squares when estimating demand. The last columns of Table 8 present the results from including feature, i.e., from the regressing the mean utility δ_j on price, feature dummy variable and product dummy variables. The coefficient of feature is not significant for OLS as well as for the two IV specifications, and the effects of including feature on the price coefficients and on the product characteristic coefficients are insignificant both statistically and economically. Furthermore, I cannot reject the exogeneity of the feature variable.

The bottom of Table 8 reports that the first stage R-squared and F-Statistic of both IV specifications are high and the Wald test for zero coefficients associated with the instruments is clearly rejected, suggesting that the demand instruments have some power.[27] Estimates of first-stage coefficients have in general the expected positive sign and are significant for plastic, sugar and milk. Estimates for the average effect of strawberry price on the price of yogurt are positive and in general

[27] First-stage results for both are available at http://socrates.berkeley.edu/~villas/homepage.html.

significant for fruit yogurts. Coefficients for the wages in the states where plants of the different products are located are significant and positive. To choose between the two different specifications, the assumption of constant marginal cost (across same "physical" product sold in different retailers) is tested. In particular, I test the assumption that the coefficients associated with the same input for the same "physical" product are equal to each other. This is a much stronger assumption than what is needed but in case of not being able to reject it this makes the choice for the specification that assumes constant marginal cost (Specification 2) stronger. This assumption cannot be rejected, so I choose to proceed with the second specification.[28]

5.1.2. Random Coefficients Demand

Results from estimating equation (26) for the full model are presented in Table 9 considering consumer heterogeneity by allowing the coefficients on price, calories, calcium and store-specific dummy variables to vary across consumers as a function of their income, their age and other unobserved consumer characteristics. Interpreting the estimates, the mean price coefficient is similar to the Logit estimate for the mean of the marginal utility of price. From the coefficient on the interaction of price with income one interprets that consumers with higher income are less price sensitive. Age does not significantly seem to affect the mean price sensitivity however unobservable characteristics in the population seem to affect it significantly. The coefficients associated with the store dummies are to be interpreted as relative to the smaller store 1. For example, unobservable characteristics in the population do not seem to explain why people choose stores 2 and 3 over store 1. In fact, older people seem to significantly prefer store 1 over both the two other stores, given the negative and significant coefficient associated to the interaction between the store dummies and age of the population. The preferences for the larger stores 2 and 3 rise with an increase in income. Higher calcium content seems to be preferred by older consumers. The estimates for the interactions of demographics with the constant term (that captures consumers' valuation for the outside option) suggest that older consumers and consumers with less income are less likely to buy yogurt.

Product fixed effects are estimated, the d_js, capturing the part of the mean utility level that is constant over time and associated with product characteristics that don't change in the sample period, improving the fit of the model. However, given that the product characteristics used are time invariant, the estimates of their coefficients cannot be directly obtained when product fixed

[28] The estimate of the Wald test statistic is 160.9, which is less than 326.3, the 95% the critical value for a chi-square with 286 degrees of freedom. The critical value C for a chi-square with large R degrees of freedom for the significance level a can be approximated (Greene, 1997, p.70) as $C \approx 0.5[\Phi^{-1}(a) + \sqrt{2R-1}]^2$.

effects are included. It is easy to see that the coefficient β of a constant product characteristic x_j is not identified, i.e. it is indistinguishable from the coefficient d_j. Nevertheless, the taste coefficients for the product characteristics can be indirectly obtained by estimating the following regression $\hat{d} = X\beta + \zeta$ using Generalized Least Squares, where $\hat{d}$ are the estimated product fixed effects and assuming $E[\zeta|X] = 0$. Table 9 displays the estimates of the consumer taste parameters for the different product characteristics. The R-squared of 0.78 suggests that the fit obtained is good. All coefficients are statistically significant, and the signs are, in general, consistent with consumer surveys. For example, for the average consumer, calories and Aspartame (an artificial sweetener that was not 100% FDA approved) have a negative marginal utility. The availability of different sizes of yogurt for a certain brand is, on average, positively valued, as is less sugar content, the most calcium content and the possibility of trying new yogurt flavors. Finally, on average, consumers tend to prefer stores 2 and 3 over the smaller store 1.

Additional specifications are presented in Table 10. Column 3 presents the results for the GMM estimation of the full model while column 4 presents the NLLS estimates of the full model. The coefficients change considerably as do the estimated price-cost margins. Column 5 presents the estimates from the specification that sets the unobserved shocks v_i to zero for all the product characteristics. Comparing columns 3 and 5, the estimates are essentially unchanged, as are the average estimated price-cost margins for the different scenarios considered (as listed in the bottom of Table 10) and also the ranking of the different supply models. This suggests that the heterogeneity is driven by demographics and not by random shocks. Finally, columns 6 through 8 present the results for a full model of demand including feature: Column 6 is equal to 3 but adds feature, column 7 has the NLLS results with feature and, finally, column 8 does not allow for random shocks v_i and includes feature. Comparing the estimates of columns 6 and 7, there is still a considerable change when instrumenting for prices after including feature. The heterogeneity is mostly explained by demographics since the estimates from column 6 are similar to those in column 8. The coefficients for feature are overall statistically insignificant and comparing column 3 with column 6, there is not a significant effect on the estimates, on the price-cost margins and also on the ranking of the models from including feature.

5.2. Elasticities and Price-Cost Margins

5.2.1. Elasticities and Price-Cost Margins: Logit Demand

For the estimated own and cross elasticities for the Logit model, see columns three and four of Table 11. The elasticities vary by brand, where the mean of the distribution of own-price elasticities is

-3.95 with a standard deviation of 0.61. In terms of cross-price elasticities, they are on average 0.026 with a standard deviation of 0.02. Summary statistics for the price-cost margin estimates, given a Logit demand model, are provided in the top part of Table 13. In each line are the price-cost margins for the different models. For the models that estimate both retail and wholesale margins those are added up to have an estimate of the whole vertical margin for each product. Comparing the whole vertical margins for the different models, one notes that for four of the models there are some products, during some weeks that exhibit estimated price-cost margins greater than 100%, which implies negative marginal cost estimates. This happens in particular for the double marginalization model, for the hybrid model, for the wholesale (scenario 4) and for the retail (scenario 5) collusion models being considered. When retailers decide the prices (scenario 2, case 1) the price-cost margins estimated are, on average, slightly higher than the price-cost margins that result when manufacturers decide the prices (scenario 2, case 2). Finally, the monopolist model in scenario 6 predicts, on average, larger margins than the previous two models. Additionally, columns 3 through 11 of Table 2 present the average estimated price-cost margins by product for the retailers and the wholesalers for each different scenario. One fact that is evident is that private labels have the largest estimated price-cost margins under all scenarios considered. This is an implication of the Logit demand specification. The lower the price, the lower the elasticity (in absolute value). Since relative price-cost margins are negatively related to elasticities, the lower in absolute value are the elasticities, the higher the price-cost margins. Private labels with lower prices than national brands exhibit therefore, not surprisingly, the highest implied price-cost margins. One last limitation of the Logit demand specification are the implied cross-price elasticities. Products with similar market shares and prices have similar cross-price elasticities.

5.2.2. Elasticities and Price-Cost Margins: Random Coefficients Demand

In the full model, the above described and other limitations in terms of elasticities disappear. On the one hand, own price elasticities are no longer uniquely driven by functional form specifications, such as above. In particular, the marginal utility of price α will now vary by product, in the sense that it is obtained as the average of all the price sensitivities for all the consumers of that particular product. On other hand cross-price substitution patterns are richer. The elasticities for the full model are given by

$$\frac{\partial s_{jt}}{\partial p_{kt}}\frac{p_{kt}}{s_{jt}} = \begin{cases} \frac{p_{jt}}{s_{jt}} \int \alpha_i s_{jt,i}(1 - s_{jt,i}) \; dF(v) \; dF(D) & \text{if j=k} \\ -\frac{p_{kt}}{s_{jt}} \int \alpha_i s_{jt,i} s_{kt,i} \; dF(v) \; dF(D) & \text{otherwise,} \end{cases}$$

where

$$s_{ijt} = \frac{e^{\delta_{jt}+\mu_{ijt}}}{1+\sum_{k=1}^{N_t} e^{\delta_{kt}+\mu_{ikt}}}$$

is now the individual probability of i purchasing product j during week t. In particular α_i is the marginal utility of price for consumer i. The integral over the unobserved and observed consumer demographics is simulated by drawing N_s random pairs (v, D). The simulated elasticities are then given by

$$\frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \frac{p_{jt}}{s_{jt}} \sum_{n=1}^{N_s} \alpha_n s_{jt,n}(1-s_{jt,n}) & \text{if j=k} \\ -\frac{p_{kt}}{s_{jt}} \sum_{n=1}^{N_s} \alpha_n s_{jt,n} s_{kt,n} & \text{otherwise,} \end{cases}$$

Table 11 presents the own price elasticities as well as the mean and the standard deviations of the cross-price elasticities for the random coefficients demand model. In terms of own price elasticities, they are on average slightly lower than the ones estimated in the Logit demand model. The mean of the own price elasticities is now -3.69 with a standard deviation of 0.55. In terms of the cross-price elasticities, their summary statistics are presented in the last columns of Table 11. They vary significantly by product, ranging, on average, from 0.001 to 0.194. Due to its large dimension, detailed information on estimated cross-price elasticities for the 43 products in the sample is not presented. Let me however summarize that, overall, the results seem reasonable and intuitive. As an illustrative example, Dannon Low fat Plain Yogurt sold in store 3 is less sensitive to a change in price of Dannon Low fat Plain Yogurt sold at any of the two other stores (elasticities of 0.005 and 0.01) than to a change in the price of Dannon Light Vanilla Yogurt sold at the same store (elasticity of 0.023). Also, when looking within the same store, the effect on Dannon Light fruit yogurt from changes in the price of Dannon Classic Flavor Fruit yogurt (elasticity of 0.222) seems to be larger than the effect from changes in the price of Dannon Low fat plain yogurt (elasticity of 0.063). These yogurts are maybe used for different purposes (plain yogurt is sometimes used for cooking) and therefore purchased by consumers with different characteristics.

Overall, the products seem to be less sensitive to changes in prices of products in other stores than to changes in prices in the same store. To summarize this, Table 12 presents the mean cross-price elasticities for products within a store and contrasts it with the mean cross-price elasticities between products of different stores. If one defines a measure of the relative substitution, as the mean cross-price elasticities within the same store divided by the mean cross-price elasticities across stores, on average, the substitution within store is four times the substitution across stores. In particular, breaking up this analysis by store, it is interesting to verify that the smaller store 1 seems to have the most loyal customers, versus the larger stores. Nevertheless, the substitution within stores 2 and 3 is still larger than across stores.

As a diagnostic of how far the results are from the Logit demand model, the variance of the cross-price elasticities in the random coefficients model is computed. The last columns of Table 11 present the mean and the standard deviations of the cross-price elasticities with respect to a certain price (which should be zero, according to Logit assumptions). For all products, the standard deviation, relative to the mean, is fairly large and the Logit restrictions seem less reasonable, especially for products of store 1.

Looking now at the estimated price-cost margins they differ by retail store. Retail store 1, the smallest retailer in the sample, exhibits the largest variability of price-cost margins across time for all scenarios estimated and also the largest average price-cost margins. Summary statistics for the price-cost margin estimates in the random coefficients demand specification are presented in the bottom part of Table 13. In each line are the price-cost margins for the different models. For the models that estimate both retail and wholesale margins, those are added up to have an estimate of the whole vertical margin for each product. Comparing the total vertical margins for the different models, one notes again that for some of the models there are certain products, during some weeks, that exhibit estimated price-cost margins greater than 100% (implying negative marginal costs). This is a first indicator of a possible problem in those models fitting the data. Whether this problem is going to be statistically significant is tested in the next section.

5.3. Model Testing

5.3.1. Informal Specification Test

Results from informally testing the validity of each supply model are presented in Table 14. Looking at the second and third columns of Table 14 all models cannot be individually rejected when testing the null hypothesis that all the parameters λ are jointly equal to one.[29] These null hypotheses imply that, given the assumptions for demand, the price-cost margins estimated under the scenario in question are consistent with the price-cost margins obtained from supply-side estimates.[30] Next I look at each of the parameters λ in equation (27) individually and test for each to be equal to one. Results are presented in the last column of Table 14. The price-cost margins implied by the double marginalization model are the ones that seem the least consistent with the data, since the null hypothesis of the parameter associated with each individual margin being one is rejected in 86%

[29] I use a distance metric statistic, which is the GMM analog of the Likelihood ratio statistic (Newey and West, 1987) to test each supply model.

[30] The purpose and interpretation of the supply parameters λ here is different from the Conduct Parameter (CP) Models where, for some values of the estimate of the CP, inferences (subject to Corts' (1999) cautions) are drawn about market power in a certain industry.

of the cases. Looking at more efficient contracting solutions between manufacturers and retailers, I reject the null hypothesis of each individual margin being consistent with the manufacturer having the pricing decision and zero retail margins 84% of the times. In contrast, the retail pricing decision and zero wholesale margins hypothesis is only rejected 24% of the times. The hybrid model is rejected 81% of the times. Both collusion models are rejected more than 80% of the times, and this can be possibly due to product differentiation in this market. This is because it is more difficult to coordinate on a price (see Scherer, 1980) and also to penalize (and hence to sustain collusion) in the context of differentiated products.[31] Finally, the efficient pricing model is only rejected 33% of the times. As a preliminary conclusion, there seems to be informal evidence that the contracting solution may follow something between zero manufacturer margins and retail pricing decisions and the fully efficient model. This would mean that not only retailers are deciding prices but the margins could be higher than the ones predicted by Nash-pricing behavior of the retailers.

5.3.2. Ranking of Supply Models: Non-nested Hypothesis Testing

To formally rank the

models Table 15 presents the estimates for the test statistics for pair-wise comparisons of all models, given a random coefficients demand specification. Once again, the intuition behind the non-nested pairwise comparisons is to see how the price-cost margins of alternative models explain the residual (unobserved determinants of price) of the null model. This residual is obtained by subtracting the computed price-cost margins and estimated marginal costs from retail prices, under the null model being considered. In each row is the (null) model being tested and in each column is the alternative being used to test it. If the alternative model is performing too well, then the null model is rejected by a large and significant test statistic. After doing the pair-wise comparison, the model that assumes zero wholesale margin and in which retailers have pricing decisions provides the best fit. It outperforms other models at 5% significance since all elements of the column correspondent to this model (labeled 2.1) are larger than the critical value. This leads to a rejection of the null models in each row against model 2.1. The best model also escapes rejection against any alternative specified since all the elements in the row correspondent to model 2.1 are less than the critical value for all alternatives considered. The "second place" model seems to be the monopolist model (model 6). It is only rejected by the best model and by the model that assumes that there are no retail margins and that manufacturers are setting the prices.

[31]Firms may in fact want to choose minimum differentiation to support collusion, as in Friedman and Thisse (1993).

Regarding some evidence in favor of robustness with respect to demand assumptions, for the Logit demand specification, the ranking in terms of the two best models is invariant. The ranking of the different models is also invariant for the additional demand specifications that are presented in Table 10.

6. Discussion

The purpose of this paper is to present a method to analyze vertical contracts. Alternative models of competing manufacturers' and retailers' decision-making are used to determine whether contracting in the supermarket industry follows the double marginalization pricing model or whether more efficient contracting solutions are in place. This paper extends the literature in analyzing vertical contracts in as much as it considers multiple retailers and does not require the need to observe data on wholesale prices. The approach, given demand estimates, is to compute price-cost margins for retailers and manufacturers implied by alternative vertical contracting models and to confront those with price-cost margins obtained from direct estimates of cost. In the more efficient contracts considered, via vertical integration, collusion or bargaining power, the double marginalization externality imposed by the retailers disappears. Consequently the sum of retailers' and manufacturers' profits may increase.[32]

For the market I study, the results rule out double marginalization. In particular, they suggest that, on the margin, manufacturers are pricing at marginal cost and that retail prices are the unconstrained profit maximizing prices. This result is consistent with several scenarios. For example, this result is consistent with non-linear pricing by manufacturers, via quantity discounts or two-part tariff contracts. In the optimal non-linear pricing contract, the manufacturer sets the marginal wholesale price close to the manufacturer's marginal cost for the retailer to have the right incentives when setting the retail prices. Then the manufacturer extracts revenue from the retailers via a fixed fee or by selling the non-marginal units at higher wholesale prices. The existence of quantity discounts is common practice in this industry while anecdotal evidence suggests that retail supermarkets do not often pay fixed fees to their manufacturers, and if they do, these fees are not close to the retail profits. Instead, there seem to be substantial fees paid by the manufacturers to the retailers (the so-called slotting allowances). The non-existence of (or the small) fixed fees from the retailers to the manufacturers could be explained by the fact that there are multiple manufacturers in this market with whom the retailers can bargain more aggressively the fixed fee down,

[32]In certain cases, profits may decrease and the manufacturers may not choose the vertically integrated solution, as e.g. in Mc Guire and Staelin (1983) and Coughlan and Wernerfelt (1989).

threatening to buy from another manufacturer. This result is also consistent with high bargaining power of the retailers that are able to force the wholesale prices down to marginal cost. In fact, in the last few decades, arguments have been made that retailers have acquired greater bargaining power relative to manufacturers (Progressive Grocer, April 1992) suggesting a possible departure from the double marginalization model in the industry. Among the several reasons for this that have been pointed out by industry participants and by researchers, private labels that compete directly with the national brands (e.g., Narasimhan and Wilcox, 1998) provide a new bargaining tool for retailers when negotiating with manufacturers.[33] Another reason is the increase of concentration at the retail level. As a result retailers have market power which they can use to bargain more aggressively with the manufacturers.[34] An indication of retailer market power is the increase in competition for shelf space implying that manufacturers have to pay retailers slotting allowances (e.g., Chu, 1992 and Shaffer, 1991) to get their products displayed. The bottom line is that, without information on fixed fees, the above theoretical and anecdotal predictions cannot be tested and one cannot formally identify which interpretation of the results applies.

One implication of the results is whether we should care about the efficiency gain from solving the vertical coordination problem associated with double marginalization. For the market that I study, the magnitude of the deadweight loss associated with the double marginalization model in comparison with the "best" model is roughly one thousand and six hundred dollars a week, which represents four percent of the sum of the three retailers' revenues. Extrapolating to an United-States/yearly basis (given the US consumption patterns of half a serving a week, total US population and the average price of a yogurt serving being forty-five cents) then national yogurt-retail revenues are about two billion dollars, and four percent of that is about ninety million dollars, which is a considerably big number.

Another implication of the results relates to the pricing decision-makers in a particular industry. In the related literature, traditionally, the retailers' pricing decisions have been assumed away. For the market I study, this model is outperformed by the alternative model of retailers having the pricing decisions. Estimating the price-cost margins under the assumption that manufacturers are setting the prices and retailers are neutral pass-through intermediaries, when in fact retailers are deciding the prices, could lead to bias and affect the conclusions when accessing market power and or merger activities in a certain industry.[35] Furthermore, the bias is expected to be more serious the

[33]Retailers are able to sell products that can be purchased at a potentially lower wholesale price, that carry their store brand and are displayed next to the national brands. At a 1995 convention, Douglas Ivester, then-president and CEO of Coca Cola, called private labels *"parasites"* and said they were responsible for *"eroding category profits."*

[34]For example, see New York Times, November 13, 1998, page C1.

[35]This issue is addressed in a companion study in progress.

more the sets of products that retailers sell and the sets of products that manufacturers sell do not coincide. More broadly, and since retailers may not be a neutral pass-through intermediaries, when analyzing price dynamics in the economy as a whole, retail behavior and retail market conditions should also be considered in addition to manufacturer behavior.

Future research considers the fact that looking at just one category may be restrictive since manufacturers, retailers and consumers make their pricing and purchase decisions in the context of multiple categories.[36] Given that consumers purchase a basket of goods during a shopping trip, a multiple category demand may be a more realistic framework to consider (see Part IV). In terms of pricing decisions, the fact that one manufacturer sells products in different product categories affects not only its pricing strategy but may possibly benefit its bargaining flexibility with the retailers. Also retailers use strategic category pricing to drive consumers into the store and increase sales.

Finally, and to motivate future empirical research on vertical contracts, I illustrate how the methodology proposed in this paper can be applied to address two questions. First, given the estimates of demand and a model of a pre- and post-vertical merger supply behavior, one can predict whether a potential vertical merger affects horizontal competition in the upstream and downstream markets involved.[37] Second, and related to pass-through effects of foreign trade policy, given the estimates of demand in a certain country for a particular good that involves a vertical trading supply model across different countries, one can analyze the effect of an increase of a tariff (depreciation of the exchange rate) on domestic or foreign margins. Trade policy makers are particularly interested in who absorbs most of the effects of a particular trade policy (foreign margins or domestic margins). That is in turn determined by the vertical relationships between domestic and foreign upstream or downstream firms.[38]

7. References

Berry, Steven T., 1994. "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, No. 2, pp.242-262.

[36] For the retailers analyzed, yogurt sales represent on average only 2% of total retail sales in contrast to the two largest dollar sales categories: soft drinks 17 % and cereal 12 %.

[37] See Manuszak (2001) for an analysis on how upstream (horizontal) mergers affect market power in the retail gasoline markets.

[38] For example, if import prices do not rise as much as the dollar depreciation (i.e. the pass-through effect is less than one), then foreign profit margins are being diminished (see, for example, Feenstra (1989)).

Berry, S., J.Levinsohn and A.Pakes, 1995. "Automobile Prices in Market Equilibrium," *Econometrica*, 63, No. 4, pp.841-890.

Bresnahan, T., 1981. "Departures from Marginal-Cost Pricing in the American Automobile Industry," *Journal of Econometrics*, 17, pp.201-227.

Bresnahan, T. and P. C. Reiss, 1985. "Dealer and Manufacturer Margins," *RAND Journal of Economics*, 26, No. 2, pp.253-268.

Bresnahan, T., 1987. "Competition and Collusion in the American Automobile Oligopoly: the 1955 Price War," *Journal of Industrial Economics*, 35, pp.457-482.

Bresnahan, T., 1989. "Empirical Studies of Industries with Market Power," in Schmalansee, R., R. D. Willig eds., *Handbook of Industrial Organization, Volume II*, Amsterdam: North Holland, pp.1011-1057.

Cardell, N. Scott,1997. "Variance Components Structures for the Extreme Value and Logistic Distributions With Application to Models of Heterogeneity," *Econometric Theory*, 13 (2), pp.185-213.

Chevalier, J.A., A. K. Kashyap and P.E. Rossi. "Why don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *NBER Working Paper*, No. W7981.

Chintagunta, P., A. Bonfrer and I. Song, 2000. "Studying the Impact of Store Brand Entry: An application of a Random Coefficients Logit Model with Aggregate Data," *working paper* , University of Chicago.

Chu, W., 1992. "Demand Signaling and Screening in Channels of Distribution," *Marketing Science*, 11, No. 4, pp.327-347.

Corts, K. S., 1999. "Conduct Parameters and the Measurement of Market Power," *Journal of Econometrics*, 88, pp.227-250.

Corts, K. S., 2001. "The Strategic Effects of Vertical Market Structure: Common Agency and Divisionalization in the U. S. Motion Picture Industry," *Journal of Economics and Management Strategy*, 10, No. 4, pp.509-528.

Coughlan, A.T. and B. Wernerfelt, 1989. "On Credible Delegation by Oligopolists: A Discussion of Distribution Channel Management," *Management Science*, 35, No. 2, pp.226-239.

Cox, D.R., 1961. "Tests of Separate Families of Hypotheses," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics*, 1, pp.105-123.

Dixit, A. and J. E. Stiglitz, 1977. "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67, pp.297-308.

Draganska, M. and D. Jain, 2000. "Product-Line Length and Competitive Pricing," *working paper*, Northwestern University.

Dubé, J.P., 2001. "Multiple Discreteness and Product Differentiation: Strategy and Demand for Carbonated Softdrinks," *working paper*, University of Chicago.

Dubin, J. A. and D. McFadden, 1984. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica*, 52, No. 2, pp.345-362.

Dubin, J. and D. Rivers, 1986. "Statistical Software Tools," Pasadena: *California Institute of Technology.*

Feenstra, R.C., 1989. "Pass-through of tariffs and exchange rates," *Journal of International Economics*, 27, pp.25-45.

Friedman, J.W. and J-F Thisse, 1993. "Partial Collusion Fosters Minimum Product Differentiation," *RAND Journal of Economics*, 24, No. 4, pp.631-645.

Gasmi, F., J.J. Laffont and Q. Vuong, 1992. "Econometric Analysis of Collusive Behavior in a Soft-Drink Market," *Journal of Economics and Management Strategy*, 1, pp.277-311.

Gourieroux, C. and A. Monfort, 1994. "Testing Non-nested Hypotheses," in R.F.Engle and D. L. McFadden, eds., *Handbook of Econometrics, Volume IV*, Amsterdam: North Holland, pp.2583-2637.

Greene, W.H., 1997. "Econometric Analysis," 3rd Edition, New Jersey, Prentice-Hall.

Hanemann, W. M., 1984. "Discrete/Continuous Models of Consumer Demand," *Econometrica*, 52, No. 3, pp.541-561.

Hansen, L.P., J. Heaton and A. Yaron, 1996. "Finite Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economic Statistics*, 14 (3), pp.262-280.

Hastings, J. S., 2002. "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in California," University of California Energy Institute POWER Working Paper no. 84.

Hausman, J., 1978. "Specification Tests in Econometrics," *Econometrica*, 46, pp.1251-1271.

Hausman, J., 1983. "Specification and Estimation of Simultaneous Equations Models," in Z. Griliches and M. Intilligator, eds., *Handbook of Econometrics, Volume I*, Amsterdam: North-Holland, pp.391-448.

Hausman, J., G. Leonard and D. McFadden, 1995. "A Utility Consistent, Combined Discrete Choice and Count Data Model," *Journal of Public Economics*, 56 (1), pp.1-30.

Hausman, J., 1996. "Valuation of New Goods Under Perfect and Imperfect Competition," in T. Bresnahan and R. Gordon, eds., *The Economics of New Goods*, Studies in Income and Wealth, Vol 58, Chicago: National Bureau of Economic Research.

Hendel, I., 1999. “Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns,” *Review of Economic Studies*, 66, pp.423-446.

Kadiyali, V., P. Chintagunta and N. Vilcassim, 2000. “Manufacturer-Retailer Channel Interactions and Implications for Channel Power: An empirical investigation of pricing in a local market,” *Marketing Science*,v.19 n.2 , pp.127-148.

Kadiyali, V., N. Vilcassim and P. Chintagunta, 1999. “Product line extensions and competitive market interactions: An empirical Analysis,” *Journal of Econometrics*, 88, pp.339-363.

Katz, M., 1989. “Vertical Contractual Relations,” in Schmalansee, R., R. D. Willig eds., *Handbook of Industrial Organization, Volume I*, Amsterdam: North Holland, pp.655-721.

Manuszak, M. D., 2001. “The Impact of Upstream Mergers on Retail Gasoline Markets,” *working paper*, Carnegie Mellon University.

Mathewson, G. F. and R. A. Winter, 1984. “An Economic Theory of Vertical Restraints,” *Rand Journal of Economics*, 15, No. 1, pp.27-38.

McFadden, D., 1973. “Conditional Logit Analysis of Qualitative Choice Behavior,” *Frontiers of Econometrics*, P. Zarembka, eds., New York, Academic Press, pp.105-142.

McFadden, D., 1984. “Econometric Analysis of Qualitative Response Models,” in Z. Griliches and M. Intilligator, eds., *Handbook of Econometrics, Volume II*, Amsterdam: North-Holland, pp.1396-1456.

McFadden, D. and K. Train, 2000. “Mixed MNL Models of Discrete Response,” *Journal of Applied Econometrics*, 15, No. 5, pp.447-470.

McGuire, T.W. and R. Staelin, 1983. “An Industry Analysis of Downstream Vertical Integration,” *Marketing Science*, 2, No. 2, pp.161-191.

Messinger, P. R. and C. Narasimhan, 1995. " Has Power Shifted in the Grocery Channel?," *Marketing Science*, Vol. 14, No. 2, pp.189-223.

Mizon, G. and J-F. Richard, 1986. "The Encompassing Principle and its Application to Non-nested Hypotheses," *Econometrica*, 54, pp.657-678.

Mortimer, J. H., 2002. "The Effects of Revenue-Sharing Contracts on Welfare in Vertically-Separated Markets: Evidence from the Video Rental Industry," *working paper*, Harvard University.

Narasimhan, C. and R.T. Wilcox, 1998. "Private Labels and the Channel Relationship: A Cross-Category Analysis," *Journal of Business*, 71, No. 4, pp.573-600.

Nevo, A. 1998. "Identification of the Oligopoly Solution Concept in a Differentiated-Products Industry", *Economics Letters*, 59, pp.391-395.

Nevo, A., 2000. " A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, No. 4, pp.513-548.

Nevo, A., 2001. " Measuring Market Power in the Ready-To-Eat Cereal Industry ," *Econometrica*, 69, No. 2, pp.307-342.

Newey, W. and D. McFadden, 1994."Large Sample Estimation and

Hypothesis Testing," in R.F. Engle and D. McFadden, eds., *Handbook of Econometrics, Volume IV*,Amsterdam: North-Holland, pp.2211-2245.

Newey W. and K. West, 1987. " Hypothesis Testing with Efficient Method of Moments Estimation," *International Economic Review*, 28, pp.777-787.

Rey, P. and J. Tirole, 1986. "The Logic of Vertical Restraints," *American Economic Review*, 76, pp.921-939.

Scherer, F., 1980. “Industrial Market Structure and Economic Performance,” second edition. Chicago: Rand-McNally.

Schmalensee, R., 1981. “Monopolistic Two-Part Pricing Arrangements,” *RAND Journal of Economics*, 12, No. 2, pp.445-466.

Shaffer, G., 1991. “Slotting Allowances and Resale Price Maintenance: A Comparison of Facilitating Practices,” *Rand Journal of Economics*, 22, No. 1, pp.120-135.

Shaffer, G. and D. P. O’Brien, 1997. “Nonlinear Supply Contracts, Exclusive Dealing, and Equilibrium Market Foreclosure,” *Journal of Economics & Management Strategy*, 6, pp.755-785.

Smith, R. J., 1992. “Non-nested Tests for Competing Models Estimated by Generalized Method of Moments,” *Econometrica*, 60, No. 4, pp.973-980.

Sudhir, K., 2001. “Structural Analysis of Manufacturer Pricing in the Presence of a Strategic Retailer,” *Marketing Science*, 20, No. 3, pp.244-264.

Tirole, J., 1988. *The Theory of Industrial Organization*, Cambridge: The MIT Press.

Villas-Boas, J.M. and Y. Zao, 2000. “The Ketchup Marketplace: Retailer, Manufacturers and Individual Consumers,” *working paper*, University of California, Berkeley.

Vuong, Q.H., 1989. “Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses,” *Econometrica*, 57, No. 2, pp.307-333.

White, H. 1980. “A Heteroskedasticity-consistent covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” *Econometrica*, 48, No. 4, pp.817-838.

8. Tables and Figures

Description	Mean	Median	Standard Deviation	Max	Min	Brand Variation	Week Variation
Prices (cents per serving)	49	48	9.2	72	24	68.3%	2.4%
Feature (=1 if featured)	0.03	0	0.15	1	0	10.8%	5.3%
Servings sold (1 serving=6 ounces)	246	132	393.3	9538	1	43.6%	4.1%
Share of product within market (%)	0.8	0.4	1.3	32	0.03	43.6%	4.1%
Combined Shares of products (%)	34	37	12.7	75	12		
Combined Shares by Manufacturer (%)	Mean	Median	Standard Deviation	Max	Min		
Dannon	16.8	16.4	7.6	50.0	4.7		
General Mills	8.8	9.0	3.6	31.1	4		
Private Label of Retailer 2	4.1	3.3	4.2	38.5	0.6		
Kraft	3.4	3.1	1.6	13.6	1.1		
Private Label of Retailer 3	1.3	1.2	0.5	3.7	0.6		
Combined Shares by Retailer (%)	Mean	Median	Standard Deviation	Max	Min		
Retailer 1	2.3	2.3	1.0	9.2	1		
Retailer 2	19.8	20.5	9.2	57.6	1.2		
Retailer 3	13.6	13.5	3.4	24.3	6.7		

Table 1: Prices, Feature, Servings Sold and Market Shares of Products in Sample: Summary Statistics.
Source: IRI.

Product ID #	Manufacturer	Retailer	Product Name	Price	
				Mean	Std
1	Kraft	1	Breyer Light Fruit Yogurt	38.94	4.73
2	Dannon	2	Dannon Light Vanilla Yogurt	47.62	3.48
3	Dannon	3	Dannon Light Vanilla Yogurt	42.06	3.07
4	Dannon	1	Dannon Lowfat Plain Yogurt	52.56	3.97
5	Dannon	2	Dannon Lowfat Plain Yogurt	48.19	4.75
6	Dannon	3	Dannon Lowfat Plain Yogurt	46.90	2.48
7	Dannon	1	Dannon Light Fruit Yogurt	57.87	5.01
8	Dannon	2	Dannon Light Fruit Yogurt	54.69	5.09
9	Dannon	3	Dannon Light Fruit Yogurt	47.08	2.33
10	Dannon	2	Dannon Nonfat Plain Yogurt	48.69	4.54
11	Dannon	3	Dannon Nonfat Plain Yogurt	46.56	2.58
12	Dannon	1	Dannon Classic Flavor Fruit Yogurt	52.50	5.53
13	Dannon	2	Dannon Classic Flavor Fruit Yogurt	53.68	7.55
14	Dannon	3	Dannon Classic Flavor Fruit Yogurt	46.96	3.29
15	Dannon	1	Dannon Classic Flavor Vanilla Yogurt	53.31	3.27
16	Dannon	2	Dannon Classic Flavor Vanilla Yogurt	48.82	4.68
17	Dannon	3	Dannon Classic Flavor Vanilla Yogurt	46.38	3.04
18	Dannon	1	Dannon Fruit on the Bottom Yogurt	51.12	6.48
19	Dannon	2	Dannon Fruit on the Bottom Yogurt	53.18	6.47
20	Dannon	3	Dannon Fruit on the Bottom Yogurt	47.31	2.41
21	Store 2	2	Private Label 2 Lowfat Fruit Yogurt	52.17	7.43
22	Store 2	2	Private Label 2 Lowfat Plain Yogurt	30.76	2.00
23	Store 2	2	Private Label 2 Lowfat Vanilla Yogurt	30.13	0.87
24	Store 2	2	Private Label 2 Nonfat Fruit Yogurt	54.63	7.29
25	Store 2	2	Private Label 2 Nonfat Plain Yogurt	54.82	7.35
26	Store 3	3	Private Label 3 Lowfat Fruit Yogurt	35.83	1.01
27	Store 3	3	Private Label 3 Lowfat Plain Yogurt	30.52	2.07
28	Kraft	1	Light N'Lively Nonfat Fruit Yogurt	48.40	4.52
29	Kraft	2	Light N'Lively Nonfat Fruit Yogurt	46.93	4.71
30	Kraft	3	Light N'Lively Nonfat Fruit Yogurt	46.44	3.24
31	Kraft	1	Light N'Lively Lowfat Fruit Yogurt	49.38	4.28
32	Kraft	2	Light N'Lively Lowfat Fruit Yogurt	46.67	5.04
33	Kraft	3	Light N'Lively Lowfat Fruit Yogurt	45.23	4.26
34	General Mills	2	Yoplait Custard Style Lowfat Fruit Yogurt	60.69	5.86
35	General Mills	3	Yoplait Custard Style Lowfat Fruit Yogurt	57.52	4.77
36	General Mills	2	Yoplait Custard Style Lowfat Vanilla Yogurt	63.54	6.58
37	General Mills	3	Yoplait Custard Style Lowfat Vanilla Yogurt	57.06	5.48
38	General Mills	1	Yoplait Fruit Yogurt	57.69	9.47
39	General Mills	2	Yoplait Fruit Yogurt	58.67	4.73
40	General Mills	3	Yoplait Fruit Yogurt	52.62	4.67
41	General Mills	1	Yoplait Light Fruit Yogurt	52.10	10.65
42	General Mills	2	Yoplait Light Fruit Yogurt	56.21	5.61
43	General Mills	3	Yoplait Light Fruit Yogurt	49.15	4.10

Table 2: Information about the 43 Products in Sample - Prices.

Price in cents per serving. One serving is equivalent to 6 ounces of yogurt. Product ID #: First column has the product identification number. Source: IRI.

Product ID #	PCM$^{(*)}$ S-1,R	PCM$^{(*)}$ S-1,W	PCM$^{(*)}$ S-2,C 1	PCM$^{(*)}$ S-2,C 2	PCM$^{(*)}$ S-3,R	PCM$^{(*)}$ S-3,W	PCM$^{(*)}$ S-4,R	PCM$^{(*)}$ S-4,W	PCM$^{(*)}$ S-5,R	PCM$^{(*)}$ S-5,W	PCM$^{(*)}$ S-6
1	33%	33%	33%	33%	33%	33%	33%	41%	39%	34%	39%
2	29%	30%	29%	28%	29%	30%	29%	36%	31%	32%	31%
3	32%	34%	32%	32%	32%	33%	32%	39%	35%	36%	35%
4	24%	26%	24%	26%	24%	26%	24%	30%	28%	29%	28%
5	29%	30%	29%	28%	29%	29%	29%	35%	31%	32%	31%
6	28%	30%	28%	29%	28%	30%	28%	35%	31%	32%	31%
7	22%	24%	22%	23%	22%	24%	22%	27%	26%	26%	26%
8	25%	26%	25%	25%	25%	26%	25%	31%	27%	28%	27%
9	28%	30%	28%	29%	28%	30%	28%	35%	31%	32%	31%
10	28%	30%	28%	28%	28%	29%	28%	35%	31%	31%	31%
11	29%	30%	29%	29%	29%	30%	29%	35%	32%	32%	32%
12	24%	27%	24%	26%	24%	27%	24%	30%	29%	29%	29%
13	26%	28%	26%	26%	26%	27%	26%	33%	29%	29%	29%
14	28%	30%	28%	29%	28%	30%	28%	35%	32%	32%	32%
15	24%	26%	24%	25%	24%	26%	24%	29%	28%	28%	28%
16	28%	30%	28%	28%	28%	29%	28%	35%	31%	31%	31%
17	29%	31%	29%	29%	29%	30%	29%	35%	32%	33%	32%
18	25%	27%	25%	27%	25%	27%	25%	31%	29%	30%	29%
19	26%	27%	26%	26%	26%	27%	26%	32%	28%	29%	28%
20	28%	30%	28%	29%	28%	30%	28%	35%	31%	32%	31%
21	27%	25%	27%	25%	27%	0%	27%	33%	29%	25%	29%
22	44%	42%	44%	41%	44%	0%	44%	55%	48%	42%	48%
23	45%	43%	45%	42%	45%	0%	45%	56%	49%	43%	49%
24	26%	24%	26%	24%	26%	0%	26%	32%	28%	24%	28%
25	26%	24%	26%	24%	26%	0%	26%	32%	28%	24%	28%
26	37%	35%	37%	35%	37%	0%	37%	46%	41%	35%	41%
27	44%	41%	44%	41%	44%	0%	44%	54%	48%	42%	48%
28	26%	27%	26%	26%	26%	27%	26%	32%	31%	27%	31%
29	30%	28%	30%	27%	30%	27%	30%	37%	32%	28%	32%
30	29%	28%	29%	27%	29%	27%	29%	35%	32%	28%	32%
31	26%	26%	26%	26%	26%	26%	26%	32%	30%	26%	30%
32	30%	28%	30%	28%	30%	27%	30%	37%	32%	28%	32%
33	30%	28%	30%	28%	30%	28%	30%	36%	33%	29%	33%
34	23%	22%	23%	22%	23%	22%	23%	28%	25%	23%	25%
35	23%	23%	23%	23%	23%	23%	23%	29%	26%	24%	26%
36	22%	21%	22%	21%	22%	21%	22%	27%	24%	22%	24%
37	24%	24%	24%	23%	24%	23%	24%	29%	26%	24%	26%
38	23%	23%	23%	23%	23%	23%	23%	28%	26%	25%	26%
39	23%	23%	23%	22%	23%	22%	23%	29%	25%	23%	25%
40	25%	25%	25%	25%	25%	25%	25%	31%	28%	26%	28%
41	25%	26%	25%	26%	25%	26%	25%	31%	30%	28%	30%
42	24%	24%	24%	23%	24%	23%	24%	30%	27%	25%	27%
43	27%	27%	27%	27%	27%	27%	27%	33%	30%	28%	30%

Table 2: (cont.) Information about the 43 Products in Sample - Price-Cost Margins (PCM).
(*): PCM= Average Price-cost Margin across weeks. Price-Cost Margin=$(p-c)/p$, where p is price and c is marginal cost. S=Scenario; R=retail; W=wholesale. **S-1**: Double Marginalization; **S-2, C 1**: Wholesale Price at marginal cost, Retailer pricing decisions; **S-2, C 2**: Zero Retail Margin, Manufacturer pricing decisions; **S-3**: Hybrid model; **S-4**: Wholesale collusion; **S-5**: Retail collusion; **S-6**: Monopolist. Source: My calculations, Logit Demand. For product names correspondent to each product identification number (**Product ID #**) please refer to previous page in the fourth column of Table 2.

	USA	Germany	France	Bulgaria	Russia	Canada
$(*)$	0.48	1.24	1.92	3.39	0.22	0.34
$(**)$	4.7	11	17	30	2	3.1

Table 3: US Per Capita Consumption compared to selected Countries.
Row **(*)**: Per capita weekly consumption in servings (1 serving = 6 ounces). Row **(**)**: Per capita yearly consumption in kg. 1990's. Sources **(**)**: USA: Economic Research Service, USDA, Statistical Series, Food Consumption Prices & Expenditures 1991-94. Figures for Germany, France, Bulgaria and Russia- Boston Consulting Group , 1998. Canada: Statistics Canada, Matrix 5666. Per capita consumption of Dairy Products. Conversion **(*)**: 1kg = 1000/170 servings. 1 year = 52 weeks.

	Specification 1				Specification 2			
Servings per capita per week	0.25	0.5	0.75	1	0.25	0.5	0.75	1
α	-8.89	-7.14	-7.01	-6.99	-9.31	-8.42	-8.17	-8.04
	(0.72)	(0.59)	(0.58)	(0.59)	(0.81)	(0.64)	(0.63)	(0.63)

Table 4: Sensitivity Analysis to Yogurt weekly per capita consumption assumption. Dependent variable in all columns 2 through 9 is $ln(s_{jt}) - ln(s_{0t})$. Regressions include brand dummy variables. 4310 observations. White standard errors are in parenthesis. α**:** Estimate of marginal utility of price. Instrumental Variable Specification 1 assumes that marginal cost of the same product sold at different retailers is different, while Specification 2 assumes that it is constant.

Description	Mean	Median	Std	Max	Min
Store 2 (=1 if product from store 2)	0.42	–	–	–	–
Store 3 (=1 if product from store 3)	0.35	–	–	–	–
Total Calories	170.6	150	59.15	253	0
Fat Calories	16.4	20	14.45	39	0
Total Fat (g)	1.94	2.5	1.66	4.6	0
Cholesterol (mg)	11.67	13	5.62	20	5
Total Carbohydrates (g)	29.32	25	11.32	48	13
Dietary Fiber (g)	0.09	0	0.29	1	0
Sugar (g)	25.2	22	10.89	42	10
Protein (g)	8.67	9	2.55	13	4
Vitamin (=1 if product has vitamin A or C)	0.51	–	–	–	–
Calcium (=1 if product has more calcium than 30% Daily Value)	0.63	–	–	–	–
Aspartame (=1 if product has Aspartame)	0.11	–	–	–	–
Fruit on the Bottom (=1 if yes)	0.09	–	–	–	–
Available in Different Sizes (=1 is yes)	0.53	–	–	–	–
New Fruit Flavors (=1 if yes)	0.26	–	–	–	–

Table 5: Characteristics of Products in Sample.
Source: Yogurt cups' label-reads.

Description	Mean	Median	Std	Max	Min
Citric Acid ($/Lb)	1.9	1.3	0.84	3	1.23
Plastic (cents/Lb)	32.6	33	3.26	3.8	27
Sugar (cents/Lb)	9	8.6	1.14	14.4	8.2
Non-fat Grade A milk ($/Lb)	1	1.1	0.08	1.2	0.86
Whey Protein ($/Lb)	0.5	0.5	0.09	0.6	0.31
Corn ($/Bushel)	2.3	2.3	0.16	2.5	1.98
Strawberry ($/CWT)	0.8	0.7	0.29	1.4	0.35
Wages Ohio (weekly earnings/number hours a week - $/hour)	11.2	11	0.56	12.6	10.4
Location of plant for Dannon Yogurts: Minster OH.					
Wages Illinois (weekly earnings/number hours a week - $/hour)	12.1	12.1	0.3	12.8	11.5
Location of plant for Breyers, Light N'Lively (Kraft): Moleena, IL; location of plant for Private Label of Store 3 and location of the three retailers.					
Wages Michigan (weekly earnings/number hours a week - $/hour)	12	11.8	0.61	14.4	10.9
Location of plant for Yoplait Yogurts: Kalamazoo, MI.					
Wages Oregon (weekly earnings/number hours a week - $/hour)	12.9	13	0.37	13.8	12.1
Location of plant for Private Label of Store 2: Clackamas, OR.					
Interest Rate (Federal Funds Effective Rate - %)	4	3.7	1	6.3	2.9
Interest Rate (Commercial Paper 3 months - %)	4.1	3.9	0.96	6.2	3.1

Table 6: Input Prices.

Sources: Citric Acid (Chemical Week); Plastic (Chemical Marketing Reporter); Sugar (Coffee, Sugar and Cocoa Exchange); Non-fat Grade A milk, Whey protein (Cheese Market News, US. Dep. Agriculture); Corn, Strawberry (National Agriculture Statistics Service, US. Dep. Agriculture); Wages (CPS Annual Earning File - NBER 50); Interest Rates (Federal Reserve).

ZIP Area 1				Age						
				< 25	$25-34$	$35-44$	$45-54$	$55-64$	$65-74$	> 75
Income	Less	than	$5,000	190	355	263	161	194	188	83
	$5,000	to	$9,999	114	281	210	154	148	407	411
	$10,000	to	$14,999	85	263	197	213	155	424	298
	$15,000	to	$24,999	201	735	551	327	386	567	285
	$25,000	to	$34,999	160	943	751	407	568	424	146
	$35,000	to	$49,999	90	1016	1109	816	652	325	93
	$50,000	to	$74,999	30	483	926	878	609	166	82
	$75,000	to	$99,999	6	37	212	271	156	74	7
	$100,000	or	more	0	46	71	142	97	17	19
ZIP Area 2				Age						
				< 25	$25-34$	$35-44$	$45-54$	$55-64$	$65-74$	> 75
Income	Less	than	$5,000	41	157	167	163	118	149	193
	$5,000	to	$9,999	84	74	128	171	177	481	756
	$10,000	to	$14,999	106	253	251	160	227	479	627
	$15,000	to	$24,999	162	1126	697	540	725	764	542
	$25,000	to	$34,999	230	1049	939	525	485	628	305
	$35,000	to	$49,999	127	1258	1255	826	748	494	279
	$50,000	to	$74,999	62	699	1032	816	705	278	110
	$75,000	to	$99,999	32	201	228	262	242	89	72
	$100,000	or	more	0	15	93	108	84	31	36
Two ZIPS				Age						
				< 25	$25-34$	$35-44$	$45-54$	$55-64$	$65-74$	> 75
Income	Less	than	$5,000	231	512	430	324	312	337	276
	$5,000	to	$9,999	198	355	338	325	325	888	1167
	$10,000	to	$14,999	191	516	448	373	382	903	925
	$15,000	to	$24,999	363	1861	1248	867	1111	1331	827
	$25,000	to	$34,999	390	1992	1690	932	1053	1052	451
	$35,000	to	$49,999	217	2274	2364	1642	1400	819	372
	$50,000	to	$74,999	92	1182	1958	1694	1314	444	192
	$75,000	to	$99,999	38	238	440	533	398	163	79
	$100,000	or	more	0	61	164	250	181	48	55

Table 7: Demographics by Zip Code Areas

Age of Householder by Household Income. Source: 1990 Census at the Zip Code level.

Variable	No Feature OLS	No Feature IV1	No Feature IV2	With Feature OLS	With Feature IV1	With Feature IV2	With Feature IV1$^{(*)}$	With Feature IV2$^{(*)}$
Price	−5.54	−7.14	−8.42	−5.27	−7.00	−8.29	−6.96	−8.10
	(0.34)	(0.59)	(0.64)	(0.35)	(0.63)	(0.66)	(0.37)	(0.69)
Feature				0.32	0.16	0.04	0.36	0.60
				(0.22)	(0.14)	(0.14)	(0.36)	(0.47)
Measures of Fit								
R^2	0.72			0.72				
Price Exogeneity Test		10.58	28.38		11.17	28.20	10.30	22.64
Feature Exogeneity Test							0.005	0.35
95% critical value		(3.84)	(3.84)		(3.84)	(3.84)	(3.84)	(3.84)
Test of Overidentification		1139	728		1138	727	1137	2173
95% critical value		(613)	(312)		(613)	(312)	(618)	(311)
First Stage								
R^2		0.78	0.74		0.81	0.76		
F-Statistic		13.64	12.95		14.61	13.18		
Wald Test: cost coefficients =0		5134	2106		5157	2173		
95% critical value		(615)	(311)		(615)	(311)		

Table 8: Results from Logit Demand.

Dependent variable in all columns is $ln(s_{jt}) - ln(s_{0t})$. Regressions include brand dummy variables. 4310 observations. White standard errors are in parenthesis. Instrumental Variables (IV1) for prices in this column are input prices multiplied by 43 product dummy variables, assuming that marginal cost differs for the same product sold at different retailers (Specification 1). Instrumental Variables (IV2) for prices in this column are input prices multiplied by 21 product dummy variables, assuming that marginal cost for the same product sold at different retailers is constant (Specification 2). IV(*): Specifications that also instrument for feature are in these last two columns. Source: My calculations.

Variable	Mean in population	Interaction with		
		Unobservables	Income	Age
Constant*	−74.262	−0.270	−1.823	10.811
	(9.600)	(0.186)	(0.410)	(2.436)
Price	−7.884	1.0116	3.212	0.091
	(0.975)	(0.351)	(1.355)	(0.072)
Store 2 (=1 if product from store 2)*	55.846	0.375	2.009	−10.283
	(9.786)	(0.283)	(0.654)	(2.479)
Store 3 (=1 if product from store 3)*	58.851	0.813	1.263	−9.885
	(9.343)	(0.518)	(0.413)	(2.202)
Total Calories*	−0.162	0.002	0.002	−0.006
	(0.011)	(0.002)	(0.002)	(0.007)
Fat Calories*	−2.719			
	(0.078)			
Cholesterol (mg)*	−0.056			
	(0.012)			
Total Carbohydrates (g)*	2.237			
	(0.065)			
Dietary Fiber (g)*	0.769			
	(0.096)			
Sugar (g)*	−1.201			
	(0.036)			
Protein (g)*	0.441			
	(0.027)			
Vitamin (=1 if product has vitamin A or C)*	0.619			
	(0.057)			
Calcium (=1 if more than 30% Daily Value)*	5.314	0.190	0.246	0.236
	(0.946)	(0.218)	(0.212)	(0.086)
Aspartame (=1 if product has Aspartame)*	−5.719			
	(0.206)			
Fruit on the Bottom (=1 yes)*	−4.044			
	(0.146)			
Available in Different Sizes (=1 is yes)*	4.651			
	(0.127)			
New Fruit Flavors (=1 if yes)*	−13.339			
	(0.394)			
GMM	566.83			
R^2 of GLS regression	0.56			
Weighted R^2 of GLS regression	0.78			

Table 9: Results from the Random Coefficients Model of Demand.
(*) were obtained from a (**GLS**) regression of estimated product dummy variables on product characteristics, with 43 observations. **Std**: Standard errors are in parenthesis. Source: My calculations.

	Variable	No Feature GMM Est. (s.e)	No Feature NLLS Est. (s.e)	No Feature $v_i = 0$ Est. (s.e)	With Feature GMM Est. (s.e)	With Feature NLLS Est. (s.e)	With Feature $v_i = 0$ Est. (s.e)
Mean	Constant	−74.26(9.6)	−8.07(1.14)	−64.8(8.1)	−2.09(0.45)	−1.91(0.28)	−1.5(0.41)
	Price	−7.88(0.9)	−5.59(0.34)	−8.0(0.88)	−7.61(1.0)	−5.32(0.36)	−7.83(0.9)
	Feature				0.44(0.29)	0.35(0.12)	0.28(0.21)
	Store 2	55.9(9.8)	0.92(0.88)	52.1(8.5)	12.3(0.76)	1.76(0.05)	12.3(0.6)
	Store 3	58.9(9.3)	−2.29(0.83)	47.9(8.29)	13.2(0.87)	2.82(0.06)	13.1(0.8)
	Calories	−0.2(0.01)	−0.22(0.01)	−0.18(0.01)	−0.2(0.01)	−0.07(0.01)	−0.01(0.01)
	Calcium	5.31(0.9)	−0.57(0.68)	5.12(0.80)	2.36(0.6)	3.4(0.07)	2.61(0.23)
Std dev	Constant	0.27(0.18)	0.09(0.04)		0.26(0.19)	0.08(0.04)	
	Price	1.01(0.4)	0.47(0.13)		1.01(0.35)	0.48(0.12)	
	Store 2	0.4(0.3)	0.07(0.08)		0.37(0.28)	0.06(0.08)	
	Store 3	0.8(0.5)	0.13(0.06)		0.82(0.51)	0.12(0.06)	
	Calories	0(0.002)	0(0.003)		0(0.0015)	0(0.0003)	
	Calcium	0.2(0.22)	0.16(0.05)		0.19(0.22)	0.16(0.05)	
Interact	Constant	−1.82(0.41)	−0.2(0.15)	−1.5(0.31)	−1.85(0.42)	−0.23(0.16)	−1.5(0.31)
With	Price	3.21(1.36)	1.23(0.45)	2.97(1.37)	3.18(1.40)	1.24(0.55)	2.80(1.34)
Income	Store 2	2.0(0.65)	−0.05(0.1)	1.76(0.61)	2.01(0.65)	−0.02(0.1)	1.75(0.61)
	Store 3	1.26(0.4)	0.14(0.1)	1.25(0.36)	1.28(0.42)	0.17(0.09)	1.24(0.36)
	Calories	0(0.002)	0(0.001)	0(0.001)	0(0.002)	0(0.0006)	0(0.001)
	Calcium	0.25(0.2)	0.08(0.08)	0.27(0.15)	0.26(0.21)	0.08(0.08)	0.27(0.16)
Interact	Constant	10.81(2.4)	−0.57(0.33)	10.43(2.1)	10.74(2.4)	−0.52(0.33)	10.4(2.1)
With	Price	0.09(0.07)	0.13(0.11)	0.12(0.13)	0.09(0.06)	0.08(0.08)	0.09(0.05)
Age	Store 2	−10.3(2.5)	0.37(0.25)	−9.91(2.2)	−10.15(2.5)	0.34(0.24)	−9.88(2.18)
	Store 3	−9.9(2.2)	0.38(0.23)	−10.1(1.9)	−9.76(2.2)	0.36(0.22)	−10.05(1.9)
	Calories	−0.01(0.01)	0(0.001)	−0.01(0.01)	−0.01(0.01)	0(0.002)	0(0.01)
	Calcium	0.24(0.09)	0.4(0.19)	0.15(0.69)	0.21(0.86)	0.38(0.18)	0.27(0.7)
DoubleMg							
PCMw		31.6%	41.6%	30.6%	32.8%	29.1%	30.4%
PCMr		34.5%	41.5%	32.8%	35.8%	28.8%	31.7%
Monopolist		39.2%	47.1%	38.3%	40.6%	32.8%	35.5%
First Stage							
R2		0.74		0.74	0.76		0.76
Wald (cr.val)		2166(312)		2166(312)	2173(311)		2173(311)
GMM/NLLS		566.8	2157	644.7	565.1	2146	644.1
R2 min. dist.		0.56	0.68	0.69	0.83	0.65	0.81

Table 10: Additional Specifications - Random Coefficient Model of Demand.
Estimates (Est.) and standard errors in parenthesis (s.e) for different specifications in each column. Column 3 presents the GMM estimates without feature, column 4 the NLLS for the same specification and column 5 presents the GMM estimates with $v_i = 0$. Columns 6 through 8 are analogous to 3 through 5 but include feature. Source: My calculations.

		Logit Demand		Random Coefficients Demand			
Product ID #	Store	Own Price Elasticity	Cross-Price Elasticities (varying price of product in row)	Own Price Elasticity	Mean (a)	Std (b)	(b)/(a)
1	1	−3.126	0.012	−2.972	0.036	0.044	1.227
2	2	−3.821	0.018	−3.586	0.022	0.011	0.512
3	3	−3.374	0.019	−3.184	0.029	0.018	0.606
4	1	−4.225	0.002	−4.037	0.006	0.008	1.257
5	2	−3.868	0.014	−3.621	0.013	0.007	0.548
6	3	−3.767	0.007	−3.530	0.012	0.008	0.625
7	1	−4.648	0.009	−4.380	0.028	0.034	1.238
8	2	−4.327	0.129	−3.974	0.134	0.066	0.492
9	3	−3.731	0.115	−3.388	0.194	0.122	0.626
10	2	−3.897	0.037	−3.637	0.039	0.019	0.49
11	3	−3.733	0.022	−3.492	0.039	0.024	0.62
12	1	−4.219	0.003	−4.033	0.008	0.009	1.208
13	2	−4.303	0.024	−3.998	0.026	0.017	0.646
14	3	−3.766	0.02	−3.520	0.028	0.016	0.557
15	1	−4.285	0.003	−4.092	0.007	0.009	1.22
16	2	−3.916	0.02	−3.674	0.018	0.012	0.687
17	3	−3.723	0.013	−3.481	0.021	0.012	0.561
18	1	−4.108	0.004	−3.908	0.01	0.012	1.22
19	2	−4.246	0.056	−3.935	0.059	0.037	0.627
20	3	−3.783	0.044	−3.514	0.06	0.034	0.562
21	2	−4.165	0.057	−3.854	0.053	0.033	0.629
22	2	−2.471	0.008	−2.361	0.008	0.004	0.542
23	2	−2.420	0.007	−2.329	0.007	0.005	0.664
24	2	−4.379	0.025	−4.053	0.034	0.016	0.468
25	2	−4.408	0.001	−4.098	0.001	0	0.48
26	3	−2.864	0.046	−2.694	0.066	0.036	0.55
27	3	−2.453	0.004	−2.348	0.006	0.003	0.583
28	1	−3.887	0.008	−3.706	0.024	0.029	1.246
29	2	−3.765	0.017	−3.529	0.018	0.006	0.321
30	3	−3.729	0.011	−3.496	0.02	0.015	0.752
31	1	−3.966	0.01	−3.745	0.021	0.027	1.26
32	2	−3.738	0.03	−3.501	0.032	0.014	0.448
33	3	−3.627	0.02	−3.391	0.037	0.028	0.753
34	2	−4.863	0.028	−4.486	0.026	0.017	0.635
35	3	−4.610	0.026	−4.231	0.042	0.028	0.664
36	2	−5.107	0.003	−4.722	0.004	0.002	0.628
37	3	−4.586	0.003	−4.255	0.006	0.004	0.708
38	1	−4.634	0.008	−4.341	0.025	0.032	1.245
39	2	−4.680	0.063	−4.302	0.072	0.043	0.592
40	3	−4.204	0.051	−3.836	0.088	0.062	0.708
41	1	−4.187	0.005	−3.991	0.02	0.027	1.335
42	2	−4.492	0.049	−4.149	0.049	0.023	0.464
43	3	−3.924	0.057	−3.597	0.092	0.071	0.769
Average					0.036	0.033	0.923

Table 11: Diagnostic of Logit Assumption.
Mean (column (a)) and Standard deviations (column (b)) of the cross-price elasticities for the different products under a random coefficients demand specification. For names equivalent to the Product Identification Numbers **ID#** see Table 2.

	Mean Cross Price Elasticities		Relative Substitution
	Same Store Products	Other Store Products	
	(a)	(b)	(a)/(b)
Average across Products	0.070	0.020	3.559
Average by Store	(c)	(d)	(c)/(d)
Store 1	0.062	0.007	9.397
Store 2	0.036	0.019	1.868
Store 3	0.091	0.029	3.185

Table 12: Relative Substitution for products within and across stores.

Description	Mean	Median	Std	Min	Max
Given a Logit Demand					
PCM Model 1: Double Marginalization - Wholesale Margin (%)	28.3	26.7	6.2	17.7	73.5
PCM Model 1: Double Marginalization - Retail Margin (%)	28.0	27.5	6.5	17.4	62.4
PCM Model 1: Retail+Wholesale Margin (%)	56.3	54.6	12.6	35.1	135.4
PCM Model 2 Case 1 : Zero Wholesale Margin, Retailer Decision (%)	28.0	27.5	6.5	17.4	62.4
PCM Model 2 Case 2: Zero Retail Margin, Manufacturer Decision (%)	27.6	26.4	5.9	17.6	61.9
PCM Model 3: Hybrid Model - Wholesale Margin (%)	22.7	25.8	10.8	0	72.9
PCM Model 3: Hybrid Model - Retail Margin (%)	28.0	27.5	6.5	17.4	62.4
PCM Model 3: Retail+Wholesale Margin (%)	50.8	50.8	11.4	21.8	134.8
PCM Model 4: Wholesale Collusion - Wholesale Margin (%)	34.5	33.0	8.4	20.1	88.0
PCM Model 4: Wholesale Collusion - Retail Margin (%)	28.0	27.5	6.5	17.4	62.4
PCM Model 4: Retail+Wholesale Margin (%)	62.5	60.4	14.9	37.5	149.8
PCM Model 5: Retail Collusion - Wholesale Margin (%)	29.5	28.1	6.4	18.1	78.2
PCM Model 5: Retail Collusion - Retail Margin (%)	31.2	30.2	7.1	19.5	67.5
PCM Model 5: Retail+Wholesale Margin (%)	60.6	57.9	13.2	37.7	145.6
PCM Model 6: Monopolist (%)	31.2	30.2	7.1	19.5	67.5
Given a Random Coefficients Demand					
PCM Model 1: Double Marginalization - Wholesale Margin (%)	31.6	29.8	8.4	16.6	131.3
PCM Model 1: Double Marginalization - Retail Margin (%)	34.5	32.0	10.0	17.8	103.5
PCM Model 1: Retail+Wholesale Margin (%)	66.2	62.5	17.3	34.4	216.8
PCM Model 2 Case 1 : Zero Wholesale Margin, Retailer Decision (%)	34.5	32.0	10.0	17.8	103.5
PCM Model 2 Case 2: Zero Retail Margin, Manufacturer Decision (%)	30.7	29.2	6.9	16.9	80.9
PCM Model 3: Hybrid Model - Wholesale Margin (%)	24.6	8.1	48.5	0	527.8
PCM Model 3: Hybrid Model - Retail Margin (%)	34.5	32.0	10.0	17.8	103.5
PCM Model 3: Retail+Wholesale Margin (%)	59.11	46.2	49.2	17.8	562.2
PCM Model 4: Wholesale Collusion - Wholesale Margin (%)	51.6	42.0	32.4	19.9	409.1
PCM Model 4: Wholesale Collusion - Retail Margin (%)	34.5	32.0	10.0	17.8	103.5
PCM Model 4: Retail+Wholesale Margin (%)	86.1	75.4	39.8	35.7	481.5
PCM Model 5: Retail Collusion - Wholesale Margin (%)	39.9	31.0	9.1	16.9	136.8
PCM Model 5: Retail Collusion - Retail Margin (%)	39.2	36.2	12.4	18.6	138.1
PCM Model 5: Retail+Wholesale Margin (%)	72.1	67.7	19.7	35.8	226.1
PCM Model 6: Monopolist (%)	39.2	36.2	12.4	18.6	138.1

Table 13: Price-Cost Margins (PCM) by Scenario.
PCM=$(p-c)/p$ where p is price and c is marginal cost. Std: Standard deviation. Source: My calculations.

Model	DM(*)	Critical value (**C**)	Cases of Rejection(**)
1. Double Marginalization	9.08	103.06	86%
2.1 Zero Wholesale Margin, Retailer Pricing decisions	18.24	59.02	24%
2.2 Zero Retail Margin, Manufacturer Pricing decisions	20.10	59.02	84%
3. Hybrid Model	6.52	95.36	81%
4. Wholesale Collusion	16.28	103.06	85%
5. Retail Collusion	7.26	103.06	83%
6. Monopolist	14.29	59.02	33%

Table 14: Validity of Different Supply Models.

(*): Distance Metric (DM) Tests for Validity of Different Supply Models based on Generalized Method of Moments (GMM) Estimation. $DM = R[GMM^r(\hat{\theta}_T^r) - GMM^u(\hat{\theta}_T^u)]$ which is distributed Chi-square with degrees of freedom equal to the number of restrictions R. GMM_r is the estimated GMM criterion function for the restricted model and GMM_u is the estimated GMM criterion function for the unrestricted model. Critical values C for a Chi-square χ^2_R with large degrees of freedom R for the significance level a can be approximated (Greene, 1997, p.70) as follows: $C \approx 0.5[\Phi^{-1}(a) + \sqrt{2R-1}]^2$. Significance level above is 10 %. Null hypothesis is that all the coefficients associated with the price-cost margins are equal to one. **(**)**: Percentage cases of rejection from testing individually if each product exhibits price-cost margins consistent with the ones implied by the model in each row (parametrically this is done by testing individually whether each λ is equal to one). Source: My calculations.

	Alternative Models						
H_0 Model	1	2.1	2.2	3	4	5	6
1. Double Marginalization	–	2.04	1.39	0.13	1.82	2.61	1.93
2.1 Wholesale Price at marginal cost	1.00	–	1.29	0.14	0.61	0.71	1.42
2.2 Zero Retail Margin	1.18	3.87	–	0.18	0.60	0.80	3.31
3. Hybrid model	0.14	2.29	1.03	–	0.23	0.21	2.10
4. Wholesale Collusion	0.41	2.28	0.18	0.06	–	0.41	2.29
5. Retail Collusion	2.10	2.94	0.59	0.13	1.18	–	2.13
6. Monopolist	1.06	4.24	2.56	0.14	0.61	0.86	–

Table 15: Pair-wise statistics to determine which model most adequately explains the data.

In each row is the null model being tested and in each column the alternative model being used to test the null model. Source: My calculations. Non-nested Cox-type test statistic (Smith, 1992) for strictly non-nested hypothesis (SNN) and for overlapping models (OVE and AEV) that can be discriminated in the Vuong (1989) two step procedure are distributed standard normal. One-sided test at 5% with critical value 1.65. Random coefficients demand.

Figure 1

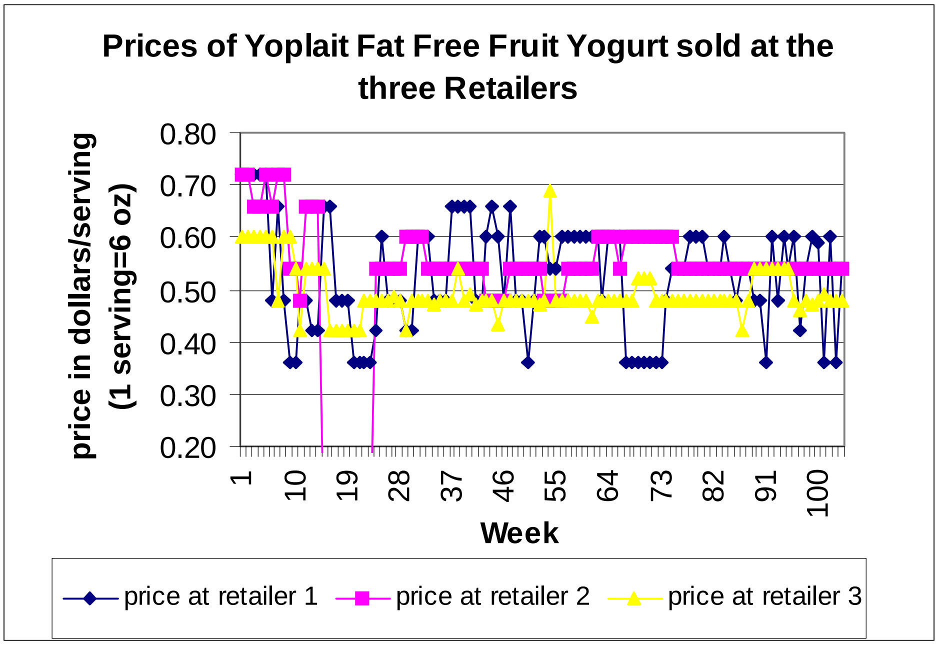


Figure 2

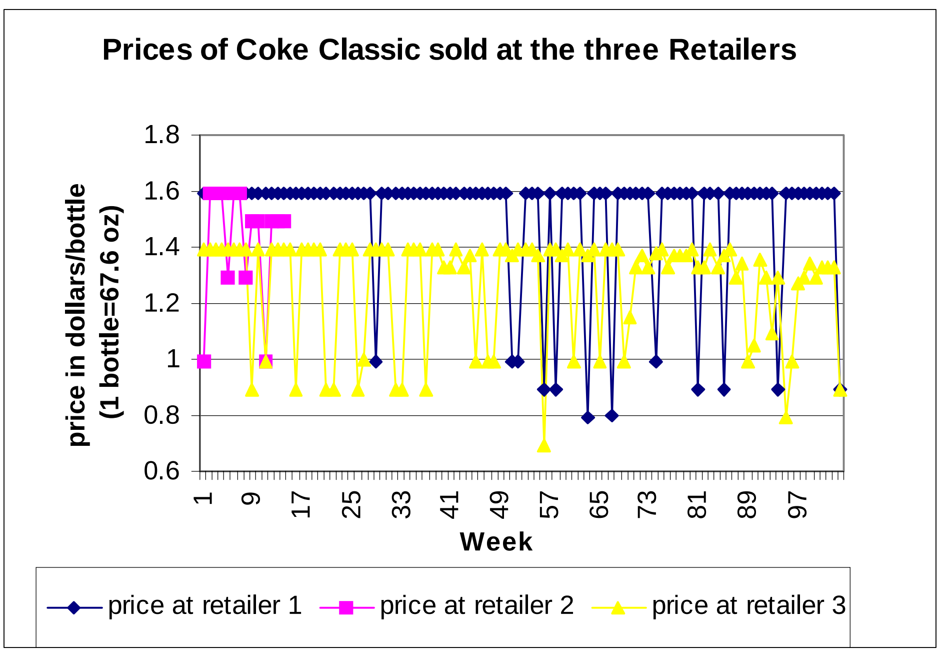
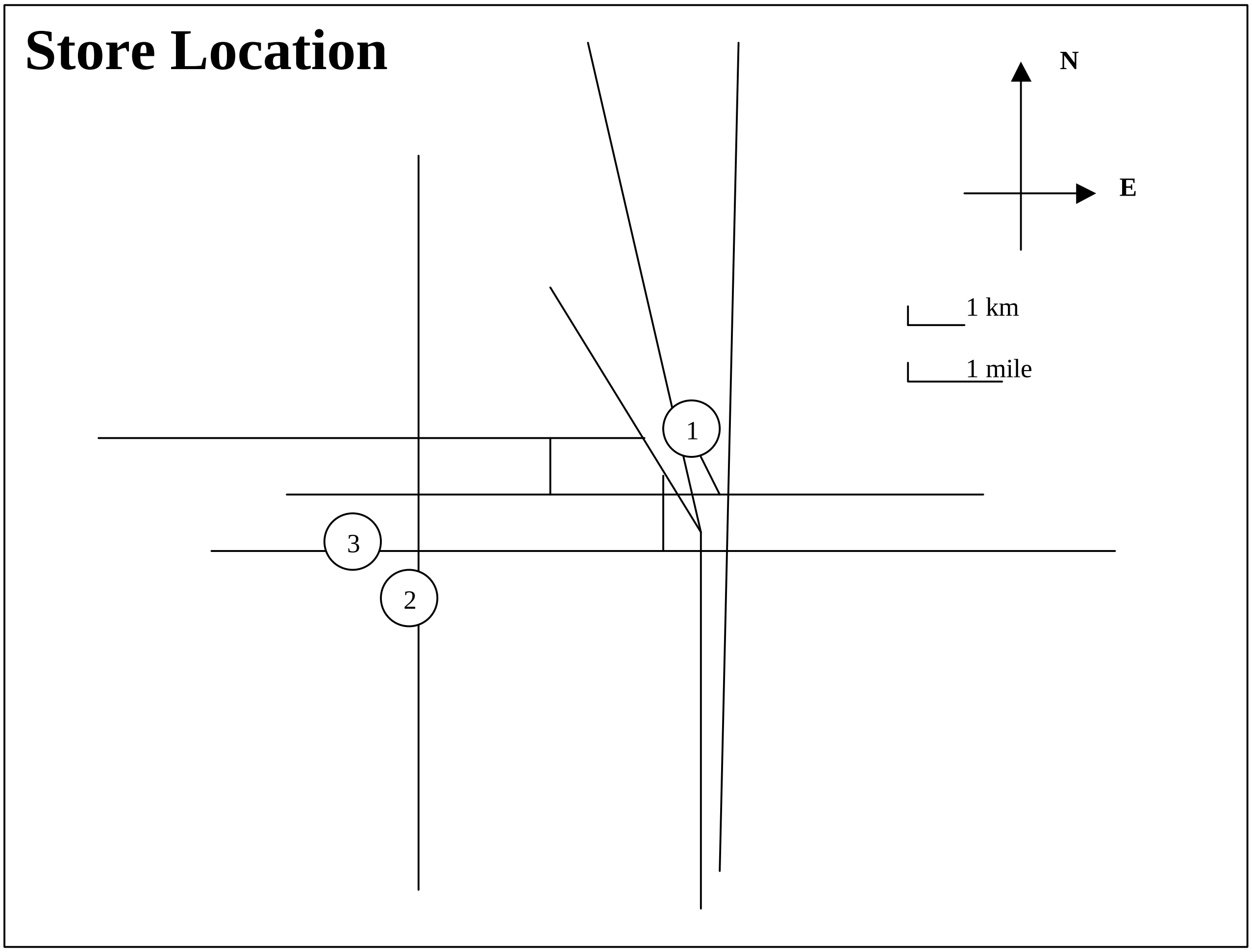


Figure 3



Figures 4 and 5