Product Differentiation and Mergers in the Carbonated Soft Drink Industry

Jean-Pierre Dubé[1]
University of Chicago Graduate School of Business
First Version September 1999.

November 29, 2000

[1] I am very grateful to Arie Beresteanu, David Besanko, Judy Chevalier, Tim Conley, Anne Gron, Sachin Gupta, Rob Porter and John Woodruff for their comments. I also benefitted from the comments of seminar participants at Northwestern University and the Econometric Society 2000 World Congress. The scanner data used in this study were provided by AC Nielsen. Financial support from a Northwestern University Dissertation Grant and from the Center for the Study of Industrial Organization are gratefully acknowledged. Address for correspondence: University of Chicago, Graduate School of Business, 1101 E. 58th Street, Chicago Il 60637. E-mail: jean.dube@gsb.uchicago.edu.

Abstract

Economists increasingly rely on structural methods to assess mergers in differentiated products industries. They typically estimate aggregate demand models using market level data, requiring a trade-off between the scope of products considered and the implied underlying consumer behavior.

Instead, I use a unique micro data set containing detailed transaction-level information for a panel of household purchases of carbonated soft drinks. I estimate household demand for soft drinks, allowing for more realistic multi-item purchase behavior for a large set of product alternatives. Integrating across the predicted household purchases, I combine aggregate demand with a model of static oligopoly. Finally, I simulate the competitive impact of several soft drink mergers on equilibrium prices and quantities. I compare my results with the popular aggregate mixed logit model. (JEL L15, C5, D4)

keywords: mergers, demand estimation, multiple-discreteness, structural modeling

1 Introduction

In 1986, following a wave of consolidation in the carbonated soft drink (CSD) industry, the Federal Trade Commission (FTC) contested the proposed acquisitions of Dr. Pepper by Coca Cola Co., and of SevenUp by PepsiCo. Pepsi immediately called off its merger; however, Coke persisted in bringing its case to trial[1]. In theory, a merger raises legal concern if the joint management of the merged firms' products leads to higher prices[2]. In practice, policy analysts typically lack cost data for computing Lerner indices, the price-cost margins, resorting instead to such market share-based concentration measures as the Herfindahl index (HHI) to screen for market power.[3] A similar lack of empirical support for economic arguments led the court to rule against Coke based on the illegality of the post-merger increase in market share (White 1989). The HHI seems inadequate for evaluating industries in which, like CSDs, firms naturally possess market power by producing differentiated products. Moreover, with differentiation, a firm may be able to generate unilateral market power, independent of its market share, by merging with a competitor producing a highly-substitutible good. Unfortunately, the inability to determine the degree of substitutibility - the cross-price elasticities of demand- of the large number of CSD products differing primarily by flavor prevented a more compelling economic basis for the ruling.

Since the early 1980s, the increasing availability of aggregate brand level data collected by AC Nielsen and IRI from supermarket checkout scanners has permitted improved demand estimation in industries with differentiated products. Using such aggregate data, economists typically combine estimated demand with a game-theoretic model of the competitive industry structure to simulate the impact of a merger on equilibrium prices (Baker and Bresnahan 1985, Hausman, Leonard and Zona 1994, Werden and Froeb 1994 and Nevo 1999). In the same spirit, I revisit the CSD case, combining estimated demand with a model of static Bertrand oligopoly to simulate the impact of the blocked CSD mergers. In contrast with the existing literature, I use a micro scanner data set as existing aggregate models are unable to capture both the complex consumer purchase behavior and the large scope of relevant products.

In using aggregate data to estimate differentiated product demand, researchers are

[1] F.T.C. v. Coca-Cola Co., 641 F. Supp. 1128 (1986).

[2] The 1992 merger guidelines specifically set a 5% price increase as the objectionable limit.

[3] Theoretically, dividing the HHI by the elasticitiy of demand measures the Cournot equilibrium Lerner Index for homogeneous product industries with constant marginal costs and no capacity constraints (Tirole 1988).

often faced with a trade-off between a meaningful scope of products and assumptions implying restrictive, unverifiable underlying consumer behavior. For instance, a typical household shopping trip for CSDs may involve the purchase multiple units of an assortment of brands from a broad set of more than 1000 SKUs.[4] The traditional residual demand approach reduces analysis to the merged firms' products without assuming specific consumer behavior (Baker and Bresnahan 1985), but its feasibility diminishes for mergers between firms with extensive product lines[5]. The multi-level demand approach reduces the number of estimated parameters by grouping products a priori into segments and assuming consumers make sequential budgetting decisions (Hausman, Leonard and Zona 1994 and Hausman 1996, Cotterill, Franklin and Ma 1996). In addition to imposing a segmentation scheme, its feasibility diminishes as the number of products in a segment or the number of segments grows. Finally, the parsimonious discrete choice model (DCM) of demand accommodates more products by projecting consumer preferences onto product attributes (McFadden 1975,1981, Berry and Pakes 1993, Werden and Froeb 1994, Berry, Levinsohn and Pakes [BLP] 1995 and Nevo 2000). However, the DCM restricts consumer behavior to single-unit purchases, an assumption that is typically violated in CSD shopping data.[6] The large number of CSD alternatives makes the DCM the most appealing aggregate model, however the restrictive implied consumer behavior could have an adverse effect on estimated substitution patterns and merger predictions.

Given the lack of an appealing aggregate model, I use an unusually-detailed disaggregate household-level scanner data set to model consumer demand for CSDs using point-of-purchase information. I recast Hendel (1999)'s *multiple-discreteness* model into a random utility context that exploits the panel structure of the data. Using the characteristics approach, as in the standard DCM, the model allows for a large number of products. The use of household information not only allows for more flexible taste heterogeneity (Goldberg 1995, BLP 1998 and Petrin 1999), it also enables modeling the complex multiple-product purchase behavior explicitly. I use the panel structure to incorporate lagged choice variables to account for product loyalty, a form of dynamic heterogeneity that is not typically included in studies using aggregate data. The econometric procedure also corrects for the remaining within-panel unobserved serial dependence. The point-

[4] In Dubé (2000), I present comparable evidence for other differentiated products industries such as ready-to-eat cereals, canned soups and cookies.

[5] One needs to estimate at least as many parameters as the square of the number of merged products.

[6] Nevo (2000) circumvents this problem for ready-to-eat cereals by modeling a single daily consumption decision of a fixed serving size. Unlike cereals, CSDs are not a staple item and, therefore, are not likely to be consumed in a systematic way.

of-purchase data includes state variables reflecting in-store marketing conditions, such as feature advertising and in-aisle displays, that are not typically included in studies using aggregate city-level or national data. The data also provides a much richer set of product alternatives, disaggregating brands to the individual package sizes.

Aggregating the individual estimates to compute market demand, I compute the manufacturer margins and marginal costs that prevail in Bertrand Nash equilibrium. I then combine the estimates of aggregate demand and marginal cost to simulate the effects of several hypothetical CSD mergers. The simulations provide evidence supporting Coke's claim that the merger with Dr. Pepper would not have been anticompetitive in terms of price increases. In contrast, the results support the FTC's opposition to the merger between Pepsi and SevenUP, which internalizes the competitive pressure of Pepsi on SevenUP and allows the latter to increase its prices substantially. Finally, the merger between Coke and Pepsi results in very large price increases, as expected. In an appendix, I report analogous results using a DCM with store-level data for the largest chain in the same market. Given its popularity, the DCM provides an interesting comparison model. The DCM yields much lower cross-elasticities, leading to much lower gains in unilateral market power as a result of the mergers. The DCM's small simulated impact of mergers on prices seems unrealistic, especially for the merger between Coke and Pepsi.

The paper is organized as follows. Section 2 provides a brief outline of the CSD industry and a summary of the relevant empirical literature. Section 3 develops the proposed model of demand along with the model of static multiproduct oligopoly used to describe the CSD manufacturers. Section 4 outlines the estimation procedure. Section 5 describes the scanner panel data and the market. Section 6 presents the empirical results for demand. Section 7 describes the proposed model's predictions for measured market power and the analysis of mergers. Finally, Section 8 concludes and outlines possibilities for further research. An appendix summarizes the DCM, providing demand estimates and merger results.

2 Mergers in the CSD Industry

Pepsi's William C. Munro once confessed, "The soft drink is not a serious thing. No one needs it."[7] However, a recent study by the national soft drink association (NSDA) estimates that the industry currently employs over 175,000 people, generating $8 billion

[7] J.C. Louis and Harvey Z. Yazijian, *The Cola Wars.* (New York: Everest House, Publishers, 1980), p. 150.

per year in wages and salaries[8]. In 1998, CSDs accounted for 49% of total US beverage gallonage, generating over $54 billion in revenues with roughly 56.1 gallons consumed per capita. In contrast, the second largest beverage, beer, accounts for only 19.4%, roughly 22.1 gallons per capita in 1998.[9]. A.C. Nielsen estimates that the CSD category is the largest in the Dry Grocery Department at US Food Stores, accounting for roughly one tenth of the department's national sales revenue. Clearly, CSDs play an important role in the US economy.

During the past decade, three companies, Coca-Cola Co, PepsiCo and Dr. Pepper-SevenUp (currently owned by Cadbury-Schweppes) have controlled most of the CSD market; but their respective shares are spread across a fairly large number of brands, flavors and package sizes. The NSDA reports that Coke manages 20 different CSD brands, whereas Pepsi and Cadbury-Schweppes each manage 10. By the end of 1997, 10 of the three firms' brands accounted for just over 58% of category sales, with 22 different brand/sizes each holding over 1% of supermarket sales for the top 25 US chains[10]. The high concentration of the leading brand shares among 3 major firms during the 1990s reflects a decade of consolidation during the 1980s.

During the 1980s as the CSD category reached maturity, the huge advertising outlays required for new product entry and the high risks of product failure led to dramatic industry consolidation. By 1989, Cadbury-Schweppes had acquired Canada Dry, Hires Root Beer and Crush; and Hicks and Haas had acquired SevenUP, Dr. Pepper, A&W Rootbeer and Squirt. Hicks and Haas was eventually acquired by Cadbury-Schweppes in 1995. In 1986, at the height of the merger phase, Coke (the number one firm) announced plans to acquire Dr. Pepper (the number three firm) and Pepsi (the number two firm) announced plans to acquire SevenUP (the number four firm). In 1986, the brands associated with these four firms accounted for over 75% of the volume sales in the CSD market. Fearing a dramatic rise in industry prices, the FTC contested both mergers. Pepsi and SevenUP immediately canceled their merger plans. However, Coke persisted, bringing the case to the Federal District Court.

While the Court ultimately rejected the merger on the grounds that it would give Coca-Cola too much market share, the decision was controversial. From a legal standpoint, the FTC's estimate that the merger would increase the Herfindahl index by 341 points to a level of 2646 violated the limits of the Merger Guidelines (White 1989)[11]. In

[8] *Economic Impact Of The Soft Drink Industry*, The national soft drink association, www.nsda.org.
[9] *Beverage World*; East Stroudsburg; May 15, 1999; Greg W Prince.
[10] *Beverage World*; East Stroudsburg; Dec 15, 1997; Kent Phillips.
[11] The court did not accept Coke's claim that the relevant market was the national beverage industry,

addition to the concentration argument, both the FTC and Coca-Cola presented several interesting empirical economic arguments dealing with the extent of the market boundaries, efficiency in the distribution chain, joint production efficiency and differentiated product competition. Due to a lack of appropriate empirical tools at the time, many of these arguments were not taken into consideration for the Court's final decision and the claims still remain open questions.

I focus on the economic claims regarding market power put forth by both the FTC and Coke[12]. The FTC downplayed the importance of differentiation, claiming that margins were high due to tacit collusion. In contrast, Coke argued differentiation was sufficiently strong that coordination would be virtually impossible even with a merger.[13] Furthermore, intense competition between Coke and Pepsi's colas would keep prices low regardless of the merger. Coke also predicted Dr. Pepper would benefit from more efficient production, reducing the latter's production costs and, thus, its price. Finally, Coke argued that only the merger between Coke and Pepsi would lead to an objectionable decrease in competition. As discussed previously, the large number of differentiated CSD products complicates demand estimation, an issue which appears to have prevented valid empirical tests of the arguments put forth by the FTC and Coke.

Since the trial, several authors have attempted to determine the magnitude of CSD margins and whether they reflect collusion or product differentiation. Gasmi, Laffont and Vuong (1992) simplify the scope of products by assuming a cola duopoly between Coke and Pepsi colas, treating the remaining products as a competitive fringe. Using non-nested tests of several structural models and 19 years of accounting data, they find evidence of long-run advertising collusion, not price collusion between 1968 and 1986.

Langan and Cotterill (1994) and Cotterill, Franklin and Ma (1996) use a more elaborate multilevel demand system to extend the scope of products to 9 brands; although their data is averaged across product sizes and package types. Although neither study finds strong evidence supporting price collusion, the former shows the potential for profit-increasing collusive pricing between several brands. The latter includes price reaction functions and finds that market power may in fact come as much from product differen-

including milk, juice and coffee. With this market definition, the postmerger HHI would have only been 739.

[12] Since the case, Higgins, Kaplan and Tollison (1995) have investigated the extent of the CSD market and Muris, Scheffman and Spiller(1992,1993) have provided a comprehensive treatment of the distribution networks and the increased efficiency of vertical integration.

[13] The FTC based their opinion on the relatively high return on stockholder equity for the major producers. Coca-Cola used reduced form regressions to show an inverse relationship between prices and concentration. (see White1989).

tiation as from collusive pricing. The use of multi-stage-budgeting assumptions requires a priori assumptions about the segmentation of products and the sequential process by which consumers make choices amongst these segments. For CSDs, one might prefer to allow the data to reveal any potential segmentation structure, rather than impose one. The approach is feasible for the limited set of 9 brands considered above, but the parameter dimensionality problem resurfaces when I disaggregate to the UPC level. Finally, the simulation of equilibrium prices with a multilevel demand specification is complicated due to the highly non-linear way in which prices, in the form of indices, enter the various levels[14]. Typically, equilibrium prices are approximated, rather than simulated. Nevo (2000) finds that the approximations tend to decrease in accuracy for mergers with a large impact on prices. Given the lack of an appealing aggregate model of demand for the CSD industry, I use disaggregate data to estimate a consumer-level model.

3 The model

3.1 Individual CSD demand

One of the predominant features of the CSD purchase data is the frequent incidence of multiple-item purchases. In contrast with the behavior implied by the typical DCM, households do not always select a single unit of a single CSD product on a given shopping trip. Table(8) , in the Appendix, breaks down the distribution of shopping trips by the total number of CSD products purchased and the total number of units. In fact, only 39% of these trips result in a single-unit purchase, implying that 69% of the trips involving a CSD purchase violate the behavioral assumption of the DCM. Households appear to be seeking variety in their purchases by purchasing an assortment of CSDs. A meaningful model of CSD demand must, therefore, allow households to choose an integer number of products and an integer quantity of each. Hendel (1999) refers to this form of decision-making as the *multiple discreteness* problem.[15]

Hendel (1999) proposes a static random profit model to predict firms' cross-sectional differences in computer holdings. He explains this variety of computer holdings by unobserved (to the econometrician) differences in firms' computing tasks. Within a firm,

[14] The use of price indices may also be the driving force of some of the counterintuitive price elasticities. Several cross-elasticities are negative, implying complementarity between such products as Coke and Sprite. This property suggests that increasing the price of Sprite lowers Coke demand.

[15] For product categories in which households typically purchase a single brand on a given trip, a separate line of research has examined the quantity purchase decision using the Hanneman (1984) random utility model (Chiang 1991 and Chintagunta 1993). These models do not account for purchases of multiple brands on the same trip.

heterogeneity across the tasks themselves generates a need for multiple types of computers as well as multiple units.

In the context of a panel of household CSD purchases, the need for assortment arises due to the separation between the time of purchase and the time of consumption. Typically, a consumer makes supermarket purchase decisions in anticipation of a stream of future consumption occasions before the next shopping trip. The shopper must select a product to satisfy each of these anticipated needs. Differences in tastes across these expected consumption occasions leads to the purchase of several alternatives. For instance, a single shopper may be purchasing for several members of a household with varying tastes, such as adults versus children. Alternatively, if consumers are uncertain of their own future tastes, they may purchase a variety to ensure they have the right product on hand (Hauser and Wernerfelt 1991, Walsh 1995). Over time, tastes may also reflect household-specific loyalties towards certain products or brands. The number of anticipated needs may depend on a household's existing CSD inventories. In practice, the anticipated consumption occasions are unobserved, so I simulate them. I assume the actual number of expected consumption occasions is drawn from a Poisson process whose mean is a function of household demographics and CSD inventories. The estimation procedure yields the expected total CSD purchase vector for a shopping trip, aggregated across all the expected needs.[16]

Kim, Allenby and Rossi (1999) also study households purchasing assortments. Their imperfect substitutes specification assumes households maximize a separate subutility function for each product alternative, rather than for each expected consumption occasion. This specification would require a large number of parameters to accommodate all of the CSD alternatives. The approach also estimates demand conditional on purchase, whereas the current model allows consumers to purchase zero units for a given shopping trip. The ability to model unconditional demand will be vital for studying the impact of mergers on pricing since increased prices may cause some consumers to stop purchasing CSDs altogether. Without the no-purchase option, firms could collude (merge) and charge infinite prices.

Formally, on a given shopping trip, a household h purchases a basket of various alternatives to satisfy J_h different anticipated consumption occasions until the next trip. In fact, the actual number J_h is not observed by the econometrician. Instead, I assume that J_h derives from a distribution characterized by household demographics and its purchase

[16] For the special case in which the consumption occasions are observed, Hausman, Leonard and McFadden (1995) develop a less complicated two-stage model.

history (inventory). Each household has quasilinear preferences that are separable in its purchases of the I softdrink products available and a composite commodity of other goods, z. Conditional on J_h, the total utility of household h at the time of a shopping trip is given by (I suppress the time index to simplify notation):

$$U^h = \sum_{j=1}^{J_h} u_j^h(\sum_{i=1}^{I} \Psi_{ij}^h Q_{ij}^h, D_h) + z, \tag{1}$$

where D_h is a $(d \times 1)$ vector of household characteristics, $Q_{ij}^h$is the quantity purchased of alternative i and Ψ_{ij}^h captures the household's valuation of alternative i 's attributes on consumption occasion j. The random component of the utility function comes from the treatment of Ψ_{ij}^h as a random variable. The additive separability of the J_h subutility functions does not allow for valuation spillovers between consumption occasions. For a given consumption occasion, the perfect substitutes structure combined with curvature assumptions for $u_j^h(\cdot, \cdot)$ ensure that households select a non-negative quantity of a single alternative. Since the perceived product qualities, Ψ_{ij}^h, vary across the J_h consumption occasions, households may purchase several products on a given trip.

The household's expenditure constraint is given by:

$$\sum_{j=1}^{J_h} \sum_{i=1}^{I} p_i Q_{ij}^h + z \leq y_h.$$

where p_i is the price of product i and y_h is the household's total shopping budget. So long as the subutility functions satisfy the correct shape and continuity properties, the expenditure equation will be binding and may be substituted into the original utility function to give:

$$U^h = \sum_{j=1}^{J_h} u_j^h(\sum_{i=1}^{I} \Psi_{ij}^h Q_{ij}^h, D_h) - \sum_{j=1}^{J_h} \sum_{i=1}^{I} p_i Q_{ij}^h + \mathrm{y}_h. \tag{2}$$

Conditional on the number of anticipated consumption occasions, J_h, the household's problem will be to pick a matrix with columns $Q_j (j = 1, ..., J_h)$ to maximize 2.

Assuming a specific functional form for the subutility functions, the household decision is broken into J_h separate problems, with a subutility for each expected consumption occasion j:

$$\begin{aligned} u_j^h(\beta_j^h, D_h, X) &= \left(\sum_{i=1}^{I} \Psi_{ij}^h Q_{ij}^h \right)^{\alpha} S(D_h) - \sum_{i=1}^{I} p_i Q_{ij}^h \\ \Psi_{ij}^h &= \max(\ 0\ ,\ X_i \beta_j^h + \xi_i)^{m(D_h)} \end{aligned} \tag{3}$$

where X_i is a $(1 \times k)$ vector of brand i's observable attributes, β_j^h is a $(k \times 1)$ vector of random tastes for attributes for consumption need j, and ξ_i is an unobserved attribute. Below I discuss the potential endogeneity that could arise due to correlation between ξ_i and the price. The term $m(D_h)$ captures the taste for quality as function of the household's characteristics, permitting a vertical dimension in consumer tastes. Households with a larger value of $m(D_h)$, will perceive a greater distance between the qualities of goods. $S(D_h)$ captures the effect of household characteristics on the scale of purchases. The α determines the curvature of the utility function. So long as the estimated value of α lies between 0 and 1, the model maintains the concavity property needed for an interior solution.

The model captures household-level heterogeneity in several fashions. The tastes for quality, the scale of purchases and the expected number of consumption needs (mean of the Poisson) are functions of observed household characteristics. In addition, unobserved heterogeneity enters the specification of attribute tastes in the form of random coefficients:

$$\beta_j^h = \widetilde{\beta} + \gamma D_h + \quad \sigma_j^h$$

where $\widetilde{\beta}$ captures the component of tastes for attributes that is common to all households and consumption needs. The $(k \times d)$ matrix of coefficients, γ, captures the interaction of demographics and tastes. Finally, is a diagonal matrix whose elements are standard deviations and σ_j^h is a $(k \times 1)$ vector of independent standard normal deviates. For each household, the taste vector will be distributed normally with, conditional on demographics, mean $\widetilde{\beta} + \gamma D_h$ and variance $\quad '$.

The household's problem consists of maximizing (2). For a given expected consumption occasion, the household can compute the optimal quantity of each of the I products. Each of these optimal quantities has a corresponding utility, which is unobserved to the econometrician. Thus, for each household, there exists a vector of latent utilities, $u_j^* = (u_{j1}^*, ..., u_{jI}^*)$, where $\mathrm{u}_{ji}^* = \max_Q u_j^h(\Psi_{ij}^h Q_{ij}, D_h)$ represents the utility from consuming the optimal quantity of product i for need j. For a given expected consumption occasion j, the perfect substitutability ensures that a household selects brand i if $u_{ji}^* = \max(u_{j1}^*, ..., u_{jI}^*)$. The optimal quantity of brand i for occasion j solves the first order condition:

$$\alpha \left(\Psi_{ij}^h\right)^{\alpha} \left(Q_{ij}^h\right)^{\alpha - 1} S_h - p_i = 0.$$

Rewriting the first order condition in terms of Q_{ij}^h:

$$Q_{ij}^{h*} = \left(\frac{\alpha \left(\Psi_{ij}^{h} \right)^{\alpha} S_h}{p_i} \right)^{\frac{1}{1-\alpha}} \tag{4}$$

which is the optimal quantity of product i for consumption occasion j. The fact that consumers must purchase integer quantities does not pose a problem since the subutility functions are concave and monotonically increasing in Q_{ij}. These properties ensure that I only need to consider the two contiguous integers to Q_{ij}^{h*}. This specification also allows for zero demand (no purchase) depending on the values of the product valuations Ψ. I then compare the $2 \cdot I$ potential quantities, picking the one yielding the highest utility. Households carry out this decision for each expected consumption occasion, selecting an optimal quantity for each. For each trip, I observe the sum of all of these optimal quantities in the form of an aggregate purchase vector.

Given the distributional assumptions for the attribute tastes and the process determining the expected number of consumption occasions, the overall expected total purchase vector for a given trip can be estimated conditional on the observable information and summed over the J_h consumption occasions:

$$EQ_h(D_h, X) = \sum_{J_h=1}^{\infty} \sum_{j=1}^{J_h} \int_{-\infty}^{\infty} Q_j^{h*}(D_h, \beta_j^h, \Theta) \Phi(d\beta | D_h, \Theta) P(dJ_h(D_h)). \tag{5}$$

Estimation requires specifying functional forms for the mean of the Poisson, $\Gamma(D_h)$, the vertical aspect of tastes, $m(D_h)$, and the scale of purchases, $S(D_h)$.

In assessing the impact of mergers on consumer well-being, I compute the change in consumer surplus associated with the change in prices. A popular measure of consumer surplus is the Hicksian compensating variation, which captures the amount by which consumer incomes must be compensated to equalize the pre and post merger levels of utility. I find Δy_h such that optimal true and counterfactual utilities are equal:

$$U^h(\mathbf{p}, y_h) = U^h(\mathbf{p}^*, y_h + \Delta y_h).$$

Since utility function (2) is linear in y_h, I can compute the compensating variation for each household's shopping trip as:

$$\Delta y_h = U^h(\mathbf{p}, y_h) - U^h(\mathbf{p}^*, y_h).$$

3.2 Comparison with the Standard DCM

The proposed random utility model has the particularly interesting feature that it is a generalization of the typical DCM. Disregarding the expected consumption occasions and assuming that consumers are restricted to single-unit purchases, then α no longer plays any role and (2)reduces to:

$$u_{hi} = X_i\beta^h S(D_h) - p_i, i = 1, ..., I.$$

Dividing through by $S(D_h)$ gives:

$$\begin{aligned} \widetilde{u_{hi}} &= X_i\beta^h - \frac{1}{S(D_h)}p_i \qquad (6) \\ &= X_i\beta^h - \phi^h p_i \\ &= \left(X_i\widetilde{\beta} - \widetilde{\phi}p_i\right) + \left(X_i \quad \sigma^h - \omega\sigma^h p_i\right) \end{aligned}$$

where the inverse of $S(D_h)$, ϕ^h, is the price-response parameter. Adding a random disturbance term directly in (6) gives the standard random utility DCM (Manski and McFadden 1981).[17] The recent popularity of aggregate models for which the underlying consumer behavior reflects (6) makes them an interesting comparison model. By comparing the measures of market power and the effects of mergers in the proposed model to those of the aggregate DCM, I compare the results of modeling multiple-discreteness explicitly as opposed to imposing single-unit purchase behavior, which eases aggregation. In the appendix I outline the estimation of an aggregate mixed logit formulation.

3.3 Endogeneity of Prices

A well-documented shortcoming of the characteristics approach is the potential for unobserved (by the econometrician) attributes which may be correlated with the price. For instance, unmeasured physical attributes, advertising or intangibles such as brand reputation could influence pricing. I alleviate the potential endogeneity of prices by including alternative-specific dummy variables that enter the quality function, Ψ_{hjt}. Given the short time span of the data (9 quarters), a single dummy is estimated for each product. Any unobserved attribute that could be correlated with price is assumed to be stable over time. Unlike recent studies using the characteristics approach with aggregate data, my inclusion of transaction-specific features and display activity allows me to proxy for the time-varying, in-store attributes that could influence consumer perceptions of quality.

Nonetheless, unobserved changes in package design, television advertising and shelf space could still introduce variations in households' perceptions of a product's quality during the sample period. These unobserved attributes bias estimation if they are

[17] The mixture of a random disturbance drawn from the extreme value distribution with the normally-distributed random coefficients gives rise to the mixed logit (McFadden and Train 1996).

correlated with any of the observed attributes. For instance, Besanko, Gupta and Jain (1998) find evidence of such high-frequency price endogeneity with weekly store-level data. Villas-Boas and Winer (1999) document that such price endogeneity can even contaminate estimation with individual data. In the aggregate DCM benchmark model, I am able to resolve such high-frequency endogeneity by using the inversion procedure proposed by Berry (1994) and by using marginal cost-shifting instruments (factor prices). However, the highly non-linear specification of the multiple-discreteness model makes it difficult for me to extract much information from supply-side instruments. Although not reported, my results do not change substantially once I include factor prices in the instrument matrix.

3.4 Supply

The softdrink industry is an oligopoly with multiproduct firms. Given the previous empirical findings that prices are not collusive, I use a short-run model in which firms choose profit-maximizing prices quarterly, treating product attributes as fixed. Given that a single distributor typically bottles and distributes the entire product line occur jointly, this model seems more realistic than assuming that brand managers independently set prices for the products under the umbrella of a given brand name. Given the increasingly vertically-integrated nature of manufacturing, bottling and distribution in CSDs (Muris, Scheffman and Spiller 1992), I do not model the channel structure. I use a static model for technical simplicity (see Fershtman and Pakes 2000 for a dynamic model of price wars with differentiated products).[18] The retailers' weekly pricing decisions are also treated as exogenous primarily as a technical convenience; although in practice CSD margins are very low in foodstores[19]. Similarly, the retailer's advertising and display decisions are also treated as exogenous. The retailer is assumed to make a weekly exogenous draw from a distribution of store-wide products to advertise and display. This assumption is consistent with the findings of Slade (1995), whereby retailers compete for overall offerings, rather than on a product-by-product basis. In evaluating mergers, I assume that the large sunk costs associated with a new brand are prohibitively high to expect entry, even if a merger raises overall prices. I also make the standard assumption of the existence of

[18] The incorporation of lagged choice variables imply that current pricing decisions could influence future demand. The size of the effect depends on whether a decrease in current prices increases the number of customers who purchase the good or whether the same number of customers simply purchase more units. For instance, a dynamic model might use consumer penetration as the firms' state variable.

[19] In separate data for national supermarket chains, soft drinks are found to have margins very close to zero and 33% lower than the cross-category average.

a Bertrand-Nash equilibrium with strictly positive prices.

Each of the F firms are assumed to produce some subset, B_f, of the $i = 1, ..., I$ CSD products, making quantity and price decisions at a quarterly frequency based on expected demand. Dropping the time subscript, each firm f has the following cost function:

$$C_f(\{Q_i\}_{i \in B_f}) = C_f + \sum_{i \in B_f} mc_i Q_i$$

where Q_i is the quantity produced of product i and C_f measures the overall fixed costs incurred by firm f. Thus, firm f earns expected profits:

$$\pi_f = \sum_{i \in B_f} (p_i - mc_i) \, Q_i(p^w) - C_f.$$

Assuming the existence of a pure-strategy static Bertrand-Nash price equilibrium with strictly positive prices, each of the prices, p_i $i \in B_f$, satisfies the following first-order conditions:

$$Q_i(p) + \sum_{k \in B_f} (p_k - mc_k) \frac{\partial Q_k(p)}{\partial p_i} = 0, i \in B_f, f = 1, ..., F.$$

I construct the following $(J \times J)$ matrix $\mathbf{\Delta}$ with entries as follows:

$$\widetilde{\Delta_{jk}} = \begin{cases} -\frac{\partial Q_{jk}}{\partial p_i}, if \; \exists \text{ f s.t. } \; \{i, k\} \subset B_f \\ 0, else \end{cases}.$$

Stacking the prices, marginal costs and expected quantities into $(J \times 1)$ vectors, $\mathbf{Q}$, $\mathbf{p}$ and $\mathbf{mc}$ respectively, the first-order conditions can be written in matrix form:

$$\mathbf{Q} - \mathbf{\Delta}(\mathbf{p} - \mathbf{mc}) = 0.$$

From the first-order conditions, I derive the mark-up equation:

$$\mathbf{p} - \mathbf{mc} = \mathbf{\Delta}^{-1}\mathbf{Q} \tag{7}$$

As is typical in the literature, I estimate these mark-ups directly from the estimated demand parameters, without using information on costs (Bresnahan 1989).

Blattberg and Neslin (1990) describe a retail margin-planning strategy whereby supermarket managers set a long-run total average margin that embodies a fixed mark-up over wholesale costs and an occasional promotional discount. Since soft drinks margins are typically found to be close to zero (see footnote above), I treat the fixed margin as zero and the occasional weekly promotion as a random, mean zero disturbance. Since soft drinks are used to generate store traffic, the randomness of promotions reflects the fact that the retailer's decision likely embodies a maximization problem at the store level. Thus, I assume the retail price in store r (where there are R stores) for product j in week t has the following form:

$$p_{jt}^r = p_j^w + \varepsilon_{jt}^r, r = 1, ..., R, j = 1, ..., J, t = 1, ..., T$$

where p_{jt}^r is the shelf price of product j in week t and ε_{jt}^r is the retail mark-up. The quarterly wholesale price is computed by taking the average price across store weeks for each product. Using the average quarterly price is analogous to previous studies using aggregate scanner data that is typically averaged across weeks and stores, to model manufacturer competition (Baker and Bresnahan 1985, Hausman, Leonard and Zona 1994, Hausman 1996 and Nevo 2000). I recover $\mathbf{mc}$ by solving the expression in (7) above.

The estimates of the demand parameters and marginal costs provide a simulator of the CSD industry structure. The highly non-linear demand and the multiproduct firms make it difficult to work out comparative statics analytically. Instead, the analysis of the hypothetical mergers involves using the first-order condition (**??**) to simulate the post-merger equilibrium prices. I solve the equation:

$$\mathbf{p}^* = \mathbf{mc} + \mathbf{\Delta}(\mathbf{p}^*)^{-1}\mathbf{Q}(\mathbf{p}^*)$$

for $\mathbf{p}^*$ numerically.

4 Model Estimation

In section 3, I derive equation (5), the expected purchase vector for each household. Using this formulation, I define the prediction error:

$$\varepsilon_{ht}(D_{ht}, \Theta) = Q_{ht}(D_{ht}, \Theta) - q_{ht} \tag{8}$$

where q_{ht} is the vector of actual purchases of each of the alternatives by household h at time t. If the model represents the true purchasing process, then at the true parameter values, Θ_0:

$$E\left\{\varepsilon_{ht}\left(D_{ht}, \Theta_0\right)\right\} = \overrightarrow{0_I} \text{ for } h = 1, ..., H \text{ and } t = 1, ..., T_h. \tag{9}$$

I also assume that:

$$E\left\{\varepsilon_{ht}\left(D_h, \Theta_0\right)\varepsilon_{hk}\left(D_{ht}, \Theta_0\right)'\right\} = \;_{|t-k|}, \tag{10}$$

where ${}_{|t-k|}$ is a finite $(I \times I)$ matrix. Any function of the observable data, D_{ht}, that is independent of the unobservables must be conditionally uncorrelated with ε_{ht} at $\Theta =$

Θ_0 (Hansen 1982 and Chamberlain 1987). Given such a function, $Z_{ht} = f(D_{ht})$, I can construct conditional moments:

$$E\left\{Z_{ht} * \varepsilon_{ht}\left(D_{ht}, \Theta_0\right) | Z_{ht}\right\} = \overrightarrow{0_I}. \tag{11}$$

For estimation, I use the sample analogue of these moments:

$$g(D_{HT}, \Theta) = \frac{1}{HT} \sum_{h=1}^{H} \sum_{t=1}^{T_h} Z_{ht} * \varepsilon_{ht}\left(D_{ht}, \Theta\right). \tag{12}$$

where $D_{HT} \equiv (D'_{1T1}, ..., D'_{HT_H})$ denotes the matrix containing all of the household/trip information for the sample of H households, where household h makes T_h shopping trips and $T = \frac{1}{H} \sum_{h=1}^{H} T_h$. I choose a value Θ_{GMM} that minimizes the function J_{HT} given by:

$$J_{HT}(\Theta) = \left[g\left(D_{HT}, \Theta\right)\right]' W_{HT}\left[g\left(D_{HT}, \Theta\right)\right] \tag{13}$$

where I specify W_{HT} as the asymptotic variance of g for efficiency (Hansen 1982). The estimation of W is discussed below. This framework gives estimates with the following asymptotic distribution:

$$\sqrt{N}(\Theta_{GMM} - \Theta_0) \implies N(0, \Xi) \tag{14}$$

$$\Xi = \left(plim\left\{\frac{dg(D_{ht}, \Theta_0)}{d\Theta}\right\} W plim\left\{\frac{dg(D_{ht}, \Theta_0)}{d\Theta}\right\}'\right)^{-1}. \tag{15}$$

In order to compute the sample moment conditions, I must evaluate an infeasibly large dimensional integral. Instead, I simulate the integrals using Monte Carlo methods (McFadden 1989 and Pakes and Pollard 1989). For each household trip, I take R independent draws from the Poisson distribution to simulate the number of expected consumption occasions. For each of these R draws, $(N + I - 1) \times K$ draws are taken from the normal distribution to simulate the taste coefficients for these occasions, where K is a sufficiently large number to place an upper bound on the number of occasions simulated for each household. These draws are then used to construct R simulations of the expected purchase vector at each trip, $Q^r_{ht}(D_{ht}, \Theta)$ $r = 1, ..., R$. The R simulations are combined to form an unbiased simulator of the expected purchase vector:

$$\widehat{Q_{ht}}(D_{ht}, \Theta) = \frac{1}{R} \sum_{r=1}^{R} Q^r_{ht}(D_{ht}, \Theta).$$

Since Q^r_{ht} derives from the same distribution as q_{ht},the variance $\widehat{Q_{ht}}(D_{ht}, \Theta)$ is $\frac{1}{R} var(q_{ht})$, which goes to zero as $R \to \infty$. Simulated method of moments consists of substituting

$\widehat{Q_{ht}}(D_{ht}, \Theta)$ into (13). So long as H is sufficiently large, the resulting method of simulated moments estimate, $\Theta_{MSM,}$ will be consistent and will have asymptotic variance $\Xi = \left(\frac{dg^s(\Theta_0)}{d\Theta}' W_{HT} \frac{dg^s(\Theta_0)}{d\Theta}\right)^{-1}$. In the next section, I discuss the estimation of the weight, W_{HT}.

4.1 Estimation of the Weight Matrix, W:

Hansen (1982) shows that, under certain regularity conditions, the efficient weighting matrix W_{HT} is the inverse of S, the variance of the sample moments. In the current context, S must account for the panel structure of the data. To deal with the cross-sectional aspect of the data, I include several state variables, such as temperature and seasonal dummies, to capture contemporaneous aggregate demand shocks that could affect households in a similar fashion. Most households also have fairly long purchase histories, which could exhibit persistent unobserved shocks (McCulloch and Rossi 1994 and Seetharaman 1999 provide parametric time-series methods for multiperiod probit models). The source of these shocks could be measurement error. For instance, household-specific reporting errors in the scanning process could generate unobserved serial dependence. By including observed time-varying factors in the mean of the Poisson function, I assume that this serial dependence is independent of the process generating the number of consumption needs. The variance of the moments has the following form:

$$\begin{aligned} S &= \lim_{HT\to\infty} HT \cdot E\left\{E\left([g(D_{HT}, \Theta_0)][g(D_{HT}, \Theta_0)]' \,|D_{HT}\right)\right\} \\ &= \lim_{HT\to\infty} \frac{1}{HT} \sum_{h=1}^{H} \sum_{t=1}^{T_h} \sum_{k=1}^{T_h} E\left[\left(1+\frac{1}{R}\right) Z_{ht} \quad {}_{tk}Z'_{hk}\right]. \end{aligned}$$

Similar to the discussion in McFadden (1989), the added simulation "noise" will not affect the consistency of the estimator, but it will reduce the efficiency by a factor of $\left(1+\frac{1}{R}\right)$. As $R \to \infty$, the estimator approaches asymptotic efficiency. In this paper, I use thirty simulation draws ($R = 30$) and assume that this will be sufficient to eliminate any added simulation noise.

More formally, I assume the values of a given household's prediction error on a given trip are determined by the values of an underlying random field, ε_s, at location s_{ht} on a lattice H. I index each observation's location by both time and household. I then allow for serial dependence between observations depending on their relative locations on the lattice H. Technically, I could allow for dependence both across households and over time. As discussed above, I only treat intertemporal dependence to simplify the estimation procedure.[20] Conley (1999) provides limiting distributions and covariance estimation techniques for this more general setting. I use Conley's non-parametric, positive semi-definite covariance estimator which is analogous to the spectral time-series estimator of Newey and West (1987). Given a consistent estimate $\widehat{\Theta}$ and a predetermined time L after which the unobserved household-specific shocks die out, the estimator for S is:

$$\begin{aligned}\widehat{S}_{HT} &= \frac{1}{HT}\sum_{h=1}^{H}\sum_{t=1}^{L}\sum_{k=t+1}^{T_h}\omega(t)\left[h^s\left(D_{h,k},\widehat{\Theta}\right)h^s\left(D_{h,k-t},\widehat{\Theta}\right)'+h^s\left(D_{h,k-t},\widehat{\Theta}\right)h^s\left(D_{h,k},\widehat{\Theta}\right)'\right]\\ &\quad -\frac{1}{HT}\sum_{h=1}^{H}\sum_{k=1}^{T_h}h^s\left(D_{h,k},\widehat{\Theta}\right)h^s\left(D_{h,k},\widehat{\Theta}\right)'\end{aligned}$$

where $h^s\left(D_{h,k},\widehat{\Theta}\right) = Z_{hk} * \varepsilon_{hk}\left(D_{hk},\Theta\right)$. I use the Bartlett weight for $\omega(t)$ to assign decreasing weight to the correlation between a given household's purchases as they grow further apart in time:

$$\omega(t) = \begin{cases} 1\text{-}\frac{|t|}{1+L} \text{ if } 1\leq \text{L} \\ 0\text{, else}\end{cases} .$$

4.2 Identification:

I now discuss several data identification issues for the proposed econometric procedure. First, I explain how the data identify the joint distribution of the total number of products and the total number of CSD units purchased on a given trip. Then, I explain how I identify the residual process and the GMM weights in the presence of a large number of moment conditions.

Since I do not observe individual expected needs on a given trip, I estimate aggregate demand per trip. Despite the fact that I do not observe specific needs, I am still able to identify the process that generates them. The main identification problem involves the distinction between a household purchasing 5 units of CSDs to satisfy five needs versus 5 CSDs to satisfy one single need. Since the random tastes are independent across consumption needs, a household with several needs will tend to purchase several different types of CSD. Alternatively, a household with a single consumption need will only purchase one type of CSD. Thus, the number of consumption needs will determine

[20] Intuitively, we do not expect the unobservables generating a given household's choice process to affect other "close" households' choice processes for a given product category. However, we do expect some such "spatial" dependence for the overall shopping choice. For instance, households' store choices may be affected by local convenience stores. This form of dependence is the subject of work in progress.

the joint distribution of the total number of units of CSDs purchased and the number of different brands.

For example, I find in the data that both the total number of CSDs and the number of different types of CSDs purchased on a trip increases with the size of the household. Therefore, household size enters both the scale function, $S(D)$, and the mean of the Poisson, $\lambda(D)$. Since the function $S(D)$ enters the per-task optimal quantity choice in (4), it will be instrumental in the identification of $\lambda(D)$ and total quantity per consumption need. Similarly, the use of demographic variables in determining $m(D)$ in (3) enables the joint identification of $\lambda(D)$ and the taste parameters, β. Although several different sets of parameter values could give the same likelihood for expected total purchases, they will not have the same likelihood for the joint distribution of total products and total units purchased. Since the sample households tend to purchase bundles containing several different CSD brands, the data will identify this joint distribution.

The assumed independence of tastes across consumption needs rules out potential shopping externalities. Purchasing a 12-pack of colas for one expected need does not influence the choice for another need. This assumption seems less of a problem for CSDs than for the purchase of computers, for which there could be obvious shared software-related externalities. Nonetheless, the fact that a consumer has already purchased a cola to satisfy one need might increase the likelihood of purchasing a non-cola to satisfy another. One way in which I could link the choices made during a given trip would be to introduce interaction dummy variables in the utility function. For instance, I could classify all the CSDs in the sample into five flavor groups. While simulating the contemporaneous choices, I would introduce flavor interaction terms that would reflect which flavor combinations have been selected across needs. In addition to providing a link across the consumption needs, these flavor interaction terms would also provide a statistical test for complementarities between flavors. The test would be a simple significance test for whether a given pairwise flavor combination has a positive, negative or zero effect on utility.

With regards to the estimated residual process, I find that the correction for within-panel serial dependence has a noticeable effect on the standard errors of my parameters. However, given the large number of instruments and products, I could run into some trouble with identification if I estimate the underlying covariance matrix freely. For now, the only restrictions I impose are the second moment independence of the instruments and the errors. Even so, with 26 products I still estimate the (26×26) residual covariance matrix, , and a $(K \times K)$ instrument covariance matrix, $E(Z_{h,t}Z'_{h,t+l})$, for each lag l. For precision, I may need to impose some additional restrictions on subsequent estimations. One way to think about valid restrictions is to consider the source of these shocks. For instance, households may randomly shop at a non-sample store, such as a convenience store. I expect this sort of measurement error to have some persistence. However, the persistence may only be for products of the same size. So, the fact that you purchase a 67.6 ounce bottle in a convenience store may only affect the prediction of other 67.6 ounce bottles. In this case, I could set some of the off-diagonal terms between different size products in the autocovariance matrices to zero to improve the identification.

5 Data

During the trial against Coca-Cola, the FTC claimed that the CSD market definition consisted of a national market as well as local markets, approximated by metropolitan areas (as measured by AC Nielsen)[21]. Thus, I focus on a single Nielsen city-market, Denver. Although the use of a single city cannot capture the effects of national CSD competition perfectly, the Denver market presents a very interesting basis of study as its demographic base is perfect for CSD consumption. According to a recent article the population of Denver is unusually young, athletic and outdoors-oriented[22]. For the year ending in January of 1995, Denver had a booming economy and the median age was 33.5, one of the lowest in the country. The Consumer Expenditure Survey claims that Americans between the ages of 35 and 44 hold the largest share of soft drink consumption.[23] Interestingly, Denver has been one of the few city-markets in which Pepsi outperforms Coca-Cola both in the cola segment and for overall CSDs.[24]

Using the Denver data, I limit my analysis to CSD products only. In ruling against Coke, the court defined the extent of the market as carbonated beverages since the prices of other beverages, such as juice and coffee, did not exhibit competitive response to CSD pricing and vice versa.[25] Given the lack of information concerning the fountain

[21] This claim was supported by Nielsen data which showed substantial price dispersion across cities. Moreover, Coke's bottling contracts guaranteed bottlers exclusive rights to metropolitan areas, offsetting any potential for arbitrage across city markets.

[22] Hilary S. Miller [1995], "Rocky Mountain Fever." *Beverage Industry*, 86, 47-51.

[23] Sarah Theodore (1997), "Soft drink demographics hinge on age and demographics," *Beverage Industry*, 88, 48-50.

[24] "15 U.S. Markets: Coke Leads in 11. Pepsi Leads in 4. Biggest Gains: Pepsi in NY; Coke in Minneapolis/St. Paul." *Beverage Digest*, July 18, 1997.

[25] Higgins, Kaplan, McDonald and Tollison (1995) do find evidence of competitive responses from other beverage categories using residual demand analysis. However, they rely on highly aggregated data, using a price index for each product category.

market, I focus on the take-home market in supermarkets, which accounts for the bulk of revenues[26]. The take-home market should provide a meaningful basis for merger analysis since fountain outlets generally offer exclusive contracts to a single manufacturer, limiting the degree to which competition influences prices.

I use scanner data collected by A.C. Nielsen, covering the Denver area between January of 1993 and March of 1995. These data include consumer information for a sample of 2108 households as well as weekly store level information for 58 supermarkets with over \$2 million in "all commodity volume". The ability of household scanner panels to approximate population behavior has been demonstrated in previous studies in which panel estimates of price sensitivities have been shown to be very close to those obtained from aggregate data (Gupta, Chintagunta, Kaul and Wittink 1996). The store level information consists of weekly prices, sales, feature and display activity for all CSD products carried in these stores. Restricting the scope of analysis to those UPC code products with at least one percent of the aggregate sales share yields 26 diet and regular products with a combined share of 51% of the household-level category sales[27]. The household level data covers all shopping trips for these items. For each trip, I know the date, the store chosen and the quantities purchased. For each alternative available within the store, I know the prices and whether the product was featured in a newspaper or as an in-aisle display. Combining the store and purchase data sets, I observe the full set of prices and the in-store marketing environment for all the alternatives on a given trip.

For each shopping trip, I construct a perceived quality measure for each product. The quality consists of three components: fixed physical attributes, time-varying attributes and household-specific loyalty. The fixed physical attributes consist of the ingredients of the product, which I collected from the nutritional information printed on the product packages. These characteristics include total calories, total carbohydrates, sodium content (in mg), and a set of dummy variables that indicate the presence of caffeine, phosphoric acid, citric acid, caramel color and no color. I report these attributes as per-12-ounce-serving, using 4 additional dummy variables to distinguish between package sizes: 6-pack of 12 oz cans, 12-pack of 12 oz cans, 6-pack of 16 oz bottles and 67.6 oz bottles.

The time-varying attributes are the prices and the marketing mix variables, feature

[26] In 1998, total take-home CSD volume was roughly 3.5 times the size of the fountain market. *Beverage World*; East Stroudsburg; May 15, 1999; Greg W Prince.

[27] A recent list of the top 25 supermarket products (in the top 10 national chains), yields almost the identical list of brand/sizes, accounting for 58% of national share. The main difference is the increasing popularity of the 20 oz bottle size. *Beverage World*; East Stroudsburg; Dec 15, 1997; Kent Phillips.

and display. Finally, the household-specific loyalty variables are two dummy variables indicating whether the same brand and same UPC respectively were chosen during the most recent shopping trip on which a purchase occurred. While such loyalty variables are typical in the marketing literature, most empirical IO studies have not had sufficient data to include them. Studies that omit these loyalty terms when they matter will suffer from strong unobserved persistence in the residual process. So long as I control for heterogeneity sufficiently, my estimated loyalty coefficients will not be spurious.[28]

In the Appendix, I provide summary statistics of the demographic variables and the product attributes used in my estimation. I also provide descriptive statistics for the data in the 22 stores used for the aggregate DCM.

6 Results

6.1 Parameter estimates

I now present parameter estimates for four specifications of the proposed model. These specifications differ mainly in their inclusion of random coefficients and interaction terms between demographic variables and product attributes. The second model includes a random intercept in the mean of the Poisson process. Adding a random intercept implies that unobserved household-specific random effects also drive expected consumption needs. The third model also includes interaction terms between some of the demographic variables and certain attributes. In the fourth specification, I make the valuation of citric and caramel random, to allow for more heterogeneity in tastes. In general, I find the most striking differences between these models to come from the addition of the random intercept in the Poisson (models 2, 3 and 4), which changes the relative magnitudes of several variables. All four models are estimated with a full set of product-specific fixed-effects, which I do not report to conserve space. I only report the parameter estimates from the GMM procedure. In the current context, I only focus on the random coefficients of certain attributes since they increase the flexibility of substitution parameters. The mean tastes for attributes have no baring on estimated price elasticities. In Dubé(2000), I project the product fixed-effects onto the physical attribute space to estimate the mean tastes. I then use these mean tastes to provide intuition for substitution patterns and to experiment with shifts in product-positioning in the attribute space.

[28] In general, these lagged purchase indicators are not structural. Chintagunta, Kyriazidou and Perktold (2000) provide a structural interpretation of such lagged terms in the context of the Hanneman model.

Table (1) reports the taste coefficients that enter the quality function, ψ. While the inclusion of a random intercept in the Poisson process (models 2 and 3) changes a few of the parameters, the addition of demographic interactions (model 3) does not lead to substantial qualitative differences. Similarly, the addition of the random terms on citric acid and caramel (model 4) does not seem to change the results; although the attributes are significant. The addition of the random intercept in the Poisson process (models 2, 3 and 4) causes both the mean and variance of the taste for feature ads to decrease, while those of in-aisle displays increase. These changes suggest that some of the random response to marketing variables in the first model was proxying for random needs. Despite these changes, marketing variables appear to have a strong positive influence on purchasing, although households differ substantially in their tastes for these terms. I also find that controlling for both the brand and the specific product chosen on the previous trip seems to explain a lot of the perceived quality. The results suggest that loyalty to a specific brand might be stronger than loyalty to a given UPC. For instance, consumers are slightly more loyal to Coca-Cola in general than to a specific package size of Coca-Cola. Of course, the ability to interpret these parameters as loyalty, as opposed to spurious correlations, depends on the accuracy with which I model heterogeneity. The demographic interaction terms and the additional random coefficient on citric are significant, suggesting that the first two models do not pick up all of the heterogeneity.

The models predict significant unobserved heterogeneity in consumer perceptions of product-specific quality (the standard deviation of the product fixed effects). Ideally, I would interact the product dummies with demographic variables to try to characterize these differences in perception. However, these interactions would require too many additional parameters. Instead, I focus on specific product attributes to explain some of these differences.

As expected, households with a female head under 35 years old tend to have higher preferences for diet products, a well-documented fact in the CSD industry[29]. In fact, I might find additional explanatory power in dummies such as female head with a college degree.[30] Similarly, larger households place slightly more weight on products with more 12-ounce servings, such as the 12-pack. Unexpectedly, households with kids place a higher weight on products with caffeine than without. Part of this effect may be due to the limited scope of products included. In particular, many of the caffeine-free products,

[29] In Europe, Diet Pepsi was reintroduced as Pepsi Max, with twice the caffeine, to overcome its "feminine" image.

[30] This fact is documented in "Just who's buying all these soft drinks, anyway?" *Beverage Industry*, 84(3), 1993.

such as SevenUP and Sprite, tend to appeal more to adults. In contrast with Nevo(2000) who finds little additional random tastes after including demographics variables, I still find evidence for unobserved heterogeneity in tastes for package size (number of 12 ounce servings) and diet, despite controlling for demographic interactions.

Now I present the terms that help determine the other features of the model. For now, I assume a simple linear form for these terms:

$$\begin{aligned}\lambda_h &= \lambda_0^h + \lambda_1 kids + \lambda_2(family\ size) + \lambda_3(last\ trip) \\ &\quad + \lambda_4(last\ csd\ trip) + \lambda_5 temperature + \lambda_6 holiday \\ scale &= s_0 + s_1(family\ size) + s_2(last\ trip) + s_3(last\ csd\ trip) \\ m &= 1 + m_1 income.\end{aligned}$$

Table (2) presents the estimated coefficients. Beginning with the mean of the Poisson process, λ, I find heterogeneity in the expected number of household needs. I find that the expected number of needs depends on the presence of kids and, to a lesser extent, on family size. The inclusion of a random intercept increases the importance of kids, while decreasing the role of family size in determining the expected number of consumption needs. Similarly, temperature no longer has much effect on expected needs. In contrast, the second, third and fourth models both exhibit strong positive effects from holidays. Surprisingly, the time since last trip and since last CSD purchase do not appear to explain much of the expected needs, especially in the second and third models. I anticipated that these terms would proxy for inventory effects. In a previous version of the model, I found a similar insignificantly small effect from an explicit measure of inventory.

The scale of purchases is also increasing in the number of people in the household, especially in the second, third and fourth models. Once again, the effects of time since last trip and time since last CSD purchase are very small (and insignificant). The vertical component is increasing in income, so that households with higher income perceive more distance between products, although this effect diminishes with the addition of the random intercept. Finally, the estimated values of α are positive and below one, which is consistent with the notion that utility is concave.

The reported standard errors have been corrected to account for potential serial-dependence. I attempt to control for many of the observed dynamic factors such as timing of trips, loyalty and inventories. Despite these controls, I still find unexplained persistence in the residuals. Accounting for time-series effects nearly doubles several of the standard errors. Nonetheless, almost all the parameters remain significant after this correction, probably due to my extremely large sample. As an experiment, I recompute

Quality Function	Model 1	Model 2	Model 3	Model 4
ad	1.13	0.66	0.74	0.67
	(0.02)	(0.02)	(0.03)	(0.04)
s.d. ad	0.53	0.03	0.04	0.03
	(0.02)	(0.02)	(0.03)	(0.09)
display	0.95	3.29	3.12	3.35
	(0.02)	(0.07)	(0.05)	(0.07)
s.d. display	0.19	0.57	0.62	0.57
	(0.01)	(0.07)	(0.03)	(0.04)
brand loyalty	2.28	3.56	5.55	3.58
	(0.04)	(0.06)	(0.06)	(0.08)
prod. loyalty	0.94	1.25	1.19	1.21
	(0.07)	(0.31)	(0.14)	(0.15)
s.d. product	1.47	3.19	3.37	3.15
	(0.02)	(0.04)	(0.03)	(0.04)
s.d. diet	0.79	0.57	0.63	0.10
	(0.02)	(0.02)	(0.03)	(0.13)
s.d. citric				0.15
				(0.14)
s.d. caramel				0.58
				(0.03)
s.d. 6-pack	0.99	1.09	1.02	1.09
	(0.02)	(0.06)	(0.03)	(0.04)
s.d. 12-pack	0.58	0.04	0.04	0.04
	(0.01)	(0.01)	(0.03)	(0.01)
s.d. 16oz	1.68	0.15	0.14	0.10
	(0.08)	(0.05)	(0.11)	(0.04)
$kid * caffeine$			0.25	
			(0.01)	
$(family\ size) * servings$			0.02	
			(0.00)	
$(female < 35) * diet$			0.44	
			(0.03)	
Obs	169,788	169,788	169,788	169,788

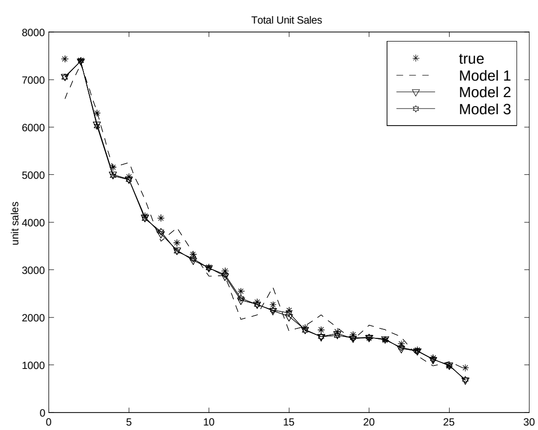
Table 1: Taste Coefficients for Time-Varying Attributes in the Quality Function (standard errors in parentheses)

variable	Model 1	Model 2	Model 3	Model 4
lambda: constant		0.078	0.083	0.078
		(0.002)	(0.006)	(0.002)
lambda: kids	0.076	0.139	0.134	0.140
	(0.005)	(0.002)	(0.005)	(0.002)
lambda: family size	0.060	0.001	0.001	0.001
	(0.002)	(0.000)	(0.000)	(0.000)
lambda: time since last csd	0.001	-0.001	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)
lambda: time since last trip	-0.001	-0.001	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)
lambda: temperature	0.001	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
lambda: holiday	0.005	0.170	0.163	0.170
	(0.002)	(0.003)	(0.006)	(0.003)
lambda: random term		0.052	0.050	0.052
		(0.001)	(0.003)	(0.001)
scale: constant	1.825	-1.008	-0.876	-1.006
	(0.069)	(0.066)	(0.022)	(0.069)
scale: family size	1.292	4.690	4.643	4.699
	(0.077)	(0.154)	(0.105)	(0.152)
scale: time since last trip	0.000	0.000	0.000	0.000
	(0.002)	(0.003)	(0.001)	(0.000)
scale: time since last csd	0.000	0.000	0.000	0.000
	(0.002)	(0.004)	(0.005)	(0.002)
vertical: income	2.059	0.751	0.731	0.746
	(0.129)	(0.019)	(0.059)	(0.022)
alpha	0.031	0.034	0.034	0.034
	(0.001)	(0.001)	(0.001)	(0.001)

Table 2: Non-Linear Coefficients (standard errors in parentheses)

the residuals after setting all of the coefficients for the dynamic factors to zero. I find that the standard errors rise by about 50% on average, some almost double. Evidently, my dynamic controls are already picking up a fair bit of unobserved persistence. Next, I take the actual residuals and average them by product for each household over time. If the model is failing to pick up some of the heterogeneity, I should see non-zero values of these averages, much like a household-specific random effect for each product. In fact, for the top 6 products, I observe about 60% of these random effects lying between $(-.01, .01)$. I also observe about 20% of the random effects lying in (-.2,.2) in a bell-curve like fashion. Thus, I suspect that at least some of the intertemporal persistence I pick up is from mismeasured heterogeneity.

Figure (16) provides a rough idea of how well the proposed models fit the aggregate data. Each point along the horizontal axis corresponds to a product, ordered by its actual share of total sample unit sales. At each point, I plot the corresponding total sales prediction for the first three specifications of the model. In general, the first model does not fit the data quite as well as the second and third. Since the GMM procedure does not provide an estimate of the joint distribution of the data, as in maximum likelihood, I am not able to provide a statistical test for the fit of the model.



(16)

7 Measuring Market Power and Mergers

7.1 Price Elasticity

Before studying mergers, I first report the price-elasticities of demand. As a comparison, I also report elasticities and margins for the DCM. I expect the DCM to be less sensitive to prices for two reasons. First, the implied consumer behavior only allows for single-unit purchases. Thus, individual consumers are forced to respond to a price change by switching to an alternative product or by not responding at all. The DCM cannot treat cases for which a consumer responds to a price change by consuming a different quantity of the same product. As shown in the appendix, the DCM's elasticities are a function of market shares. I expect the small computed CSD market shares to put additional downward pressure on predicted elasticities. I discuss the impact of the lower elasticities below.

Table (7.1) presents estimated own price elasticities for the proposed model and the DCM. Both models yield own-elasticities that are all well above one in magnitude, which is consistent with the static oligopoly model used for the merger analysis in section 6. However, almost all of the DCM's own-price elasticities are markedly lower than those of the proposed model. Since a firm's ability to set a price above marginal cost depends on consumer price sensitivity, the DCM's lower own-elasticties imply higher market power than the proposed model.

In Tables (13, 14 and 15) in the appendix, I report cross-price elasticities from the proposed model. Interestingly, the elasticities demonstrate that consumers substitute between products of the same size in response to price changes. Also, Coke and Pepsi are clearly the primary substitutes of almost every brand, but the reverse is not true. Thus, a merger between either of the colas and a non-cola will likely lead to much lower increases in cola versus non-cola prices. Although not reported, I find the cross-price sensitivities to be much lower for the DCM than for the proposed model. On average, the proposed model's cross-elasticities are over 5 times higher. The DCM's lower degree of product substitutibility implies lower unilateral market power from joint-pricing than the proposed model. Merger analysis with the DCM should therefore lead to lower price increases than the proposed model.

7.2 Margins and Marginal Costs

I use the price equation (7) to compute the wholesale markups implied by the model of static oligopoly described in section 3. For each product, I report the Lerner index, $\frac{p-MC}{p}$. The Lerner index measures market power as the degree to which manufacturers raise prices above marginal costs in equilibrium. The third column of table (4) presents the quarterly medians of the estimated margins for each product. The second column reports the corresponding marginal costs on a dollar-per-12-ounce basis.

Looking at the estimated Lerner indices, Pepsi clearly has a pricing advantage over

Product	Model 3	DCM
PEPSI 6P	-2.59	-1.74
PEPSI DT 6P	-3.46	-1.85
PEPSI DT CF 6P	-3.60	-1.89
PEPSI 16oz	-2.46	-2.99
MT DW 6P	-3.43	-1.90
COKE CLS 6P	-3.13	-1.99
COKE DT 6P	-2.98	-2.04
A and W CF 6P	-3.83	-1.92
DR PR 6P	-4.17	-1.85
PEPSI 12P	-2.97	-2.55
PEPSI DT 12P	-3.26	-2.57
PEPSI DT CF 12P	-3.27	-2.60
MT DW 12P	-4.37	-2.56
COKE CLS 12P	-3.10	-2.73
COKE DT 12P	-3.45	-2.75
COKE DT CF 12P	-4.41	-2.75
SP CF 12P	-6.09	-2.76
DR PR 12P	-3.49	-2.68
PEPSI 67.6oz	-2.49	-1.54
PEPSI DT 67.6oz	-3.10	-1.56
MT DW 67.6oz	-3.55	-1.58
COKE CLS 67.6oz	-3.04	-1.55
COKE DT 67.6oz	-3.61	-1.57
7UP R CF 67.6oz	-3.02	-1.65
7UP DT CF 67.6oz	-2.86	-1.66
DR PR 67.6oz	-3.85	-1.59

Table 3: Own Price Elasticities

its rivals, setting margins of about 50 to 60%. In comparison, Coke margins are quite a bit lower, around 40%. This fact is not surprising since, as described in the data section, Pepsi leads in market share in the Denver CSD market. The higher Pepsi margin reflects both a slightly lower marginal cost as well as a slightly higher shelf price. SevenUP margins are also close to 40%, whereas Dr. Pepper is closer to 30%. The cost of a 67.6 ounce bottle is markedly lower than that of a 6-pack of cans. One would expect to observe such differences in packaging costs associated with plastic bottles versus aluminum cans. Interestingly, for many products, the margins are quite similar for these two size alternatives. The fact that 12-packs of cans have slightly higher marginal costs than 6-packs may reflect the more sophisticated carboard box used to bundle the cans, as opposed to the simple plastic ring used to bind the 6-packs. I find that Dr. Pepper has noticeably higher marginal costs than other products. This result most likely reflects the differences in distribution costs. For instance, SevenUP has substantially lower costs than Dr. Pepper, but the former is typically distributed via a Pepsi bottler since it is not considered to be a direct competitor with colas. I have no explanation for the relatively high marginal cost of Sprite, driven by its very low implied margin. Given that the price of Sprite has virtually no competitive impact on other CSDs, as seen by the cross-elasticities in table (14), I do not expect its high marginal cost to affect the merger analysis in the following section.[31]

In general, there are no systematic differences between the margins predicted by the proposed model and those of the DCM. Both specifications yield similar ranges of predicted market power. However, the sources of market power differ in that the DCM's margins derive much more from own-price elasticities than from the effects of cross-elasticities on joint-pricing. Although not reported, the implied margins from a model treating each product as a separate firm are much higher for the DCM.

7.3 Mergers

Using the marginal costs from the previous section, I simulate the equilibrium prices and quantities for the proposed 1986 mergers between Coke and Dr. Pepper, and Pepsi and SevenUP as well as the hypothetical merger between Coke and Pepsi. Since I am using 1993 data, I must first account for changes in the market structure since 1986. The

[31] Although not reported, regressions using the estimated quarterly per-ounce marginal cost of each product show that costs rise with soft drink industry wages as well as calories (diet versus regular). Also, marginal costs are higher for cans than bottles and the 6-pack is slightly cheaper than the 12-pack. This latter find may be due to the simpler bundling materials used for the 6-pack. Experimentation with multiple specifications typically yielded an adjusted R-squared value of about 0.3.

Product	median price per 12 oz ($/12 oz)	median MC ($/12 oz)	PCM (%)
PEPSI 12P	0.30	0.15	51.41
PEPSI 6P	0.27	0.11	58.77
PEPSI 67.6oz	0.18	0.08	58.96
PEPSI DT 12P	0.30	0.13	55.79
PEPSI DT 6P	0.26	0.14	48.82
PEPSI DT 67.6oz	0.18	0.09	48.11
PEPSI DT CF 12P	0.30	0.12	60.82
PEPSI DT CF 6P	0.26	0.09	65.85
PEPSI 16oz	0.33	0.15	57.02
MT DW 12P	0.30	0.17	41.92
MT DW 6P	0.27	0.12	55.28
MT DW 67.6oz	0.18	0.08	53.72
COKE CLS 12P	0.29	0.18	41.01
COKE CLS 6P	0.27	0.17	37.48
COKE CLS 67.6oz	0.19	0.11	41.39
COKE DT 12P	0.29	0.18	36.85
COKE DT 6P	0.27	0.16	39.56
COKE DT CF 12P	0.29	0.17	46.81
COKE DT 67.6oz	0.19	0.11	38.95
SP CF 12P	0.29	0.23	23.82
7UP R CF 67.6oz	0.18	0.11	37.37
7UP DT CF 67.6oz	0.18	0.09	46.53
DR PR 12P	0.31	0.22	30.31
DR PR 6P	0.28	0.20	26.50
DR PR 67.6oz	0.19	0.12	32.83
A and W CF 6P	0.28	0.18	30.71

Table 4: Predicted Mark-UPs

relevant set of brands has not changed since the time of the trial. However, in 1989, Dr. Pepper and SevenUP were merged into a single CSD firm - both were acquired by Hicks and Haas. Therefore, I first simulate the prices that would prevail without the joint-pricing of Dr. Pepper and SevenUP. Then, I evaluate the mergers relative to these simulated 1986 outcomes, rather than the 1993-1995 outcomes.[32]

To evaluate the mergers, I look at both the legal implications, the impact on prices, as well as the economic implications, the impact on welfare. Legally, a price is not allowed to increase by more than 5%. During the case against Coke, the FTC was even more conservative, using 10% as the upper limit. Table (5) reports the median predicted quarterly price changes, in percent, for each merger across the 9 quarters in the sample[33]. The first column reports the changes in prices associated with breaking apart Dr. Pepper and SevenUP, to replicate the 1986 market structure. This break-up appears to have a very small impact on industry prices. From an economic standpoint, the change in prices associated with a merger does not capture the impact on the well-being of economic agents: consumers and producers. As in Werden and Froeb (1994) and Nevo (2000), table (6) reports the corresponding changes in producer and consumer surplus for each merger. I measure producer surplus as variable profits and consumer surplus as the compensating variation. To compute the aggregate measures, I project the sample estimates of surplus onto the size of the entire market using the number of households in Denver as the size of the population.

The first merger is between Coke and Dr. Pepper. In general, this merger does not seem to have a legally objectionable effect on prices. Coke prices never rise by more than 2% and the prices of Dr. Pepper increase by between 4 and 6%, a borderline violation of the 5% threshold of the merger guidelines; but well below the 10% limit used by the FTC against Coke. The competition between Coke and Pepsi is still sufficient to keep both Coke and Dr. Pepper prices reasonably low. Interestingly, the rise in prices appears to have a large benefit for Pepsi, whose quarterly variable profits rise by $188 thousand (about 2%). Overall, the joint quarterly profitability of Coke and Dr. Pepper increases slightly, on average. However, this represents a 17% increase in total quarterly variable

[32] Since the joining of Dr. Pepper and SevenUP would not have occurred had the FTC not successfully blocked the 1986 mergers, it might not have been completely vacuous to compare the simulated mergers to the 1993-1995 data. One could think of this alternative exercise as a long-run assessment of the consequences of the merger trial. In fact, the comparison of the merger outcomes with the the actual data does not lead to qualitatively different assessments.

[33] I also experimented with approximating the price changes, as in Hausman, Leonard and Zona (1994). I found the approximation method tended to overstate the price increases compared to those computed numerically.

profits for Coke, on average. Quarterly consumer surplus falls, on average, by about \$28 thousand. However, this loss in consumer surplus is still too small to outweigh the large gains in producer surplus, resulting in a net gain to the Denver economy.

Next, I look at the merger between Pepsi and SevenUP. Again, cola prices do not rise by much more than 2%. However, the price of SevenUP rises by almost 14%, for regular, and 15% for diet. The internalized competition allows Pepsi to raise the price of SevenUP by much more than either the legal 5% standard or the 10% FTC standard used against Coke. While the joint profits of both Pepsi and SevenUP rise, the gains only represent a roughly 4.5% increase in total quarterly variable profits for Pepsi, on average. The combination of small profitable gains and the clearly objectionable price increases likely explain why Pepsi did not fight the FTC's decision. While the merger leads to a loss in quarterly consumer surplus of over \$34 thousand, this decrease is still insufficient to lower aggregate Denver welfare.

In the final column, I consider the extreme case of a merger between Coke and Pepsi, which Coke insisted would be the only merger of anticompetitive consequence. I now find substantial price increases, as expected. With the exception of Sprite, all of Pepsi and Coke's products' prices increase by well over 10%. In fact, many rise by more than 20%, especially the diet colas. This drastic reduction in cola competition has the indirect effect of allowing SevenUP, Dr. Pepper and A&W Rootbeer to each raise their prices substantially. The large increase in prices results in well over a \$1 million average increase in the quarterly joint variable profits of Coke and Pepsi (roughly 15%). These gains come at a huge cost to consumers, who lose an average of \$12 million in quarterly consumer surplus. Overall, the Denver economy suffers an almost \$10 million loss in aggregate surplus, on average.

As discussed in Nevo (2000), this analysis is limited by the static nature of the model of producers. The static model does not allow for gains in variable profits, especially for the merger of Coke and Pepsi, to stimulate entry of new firms into the market. Moreover, the current analysis does not account for the potential efficiency gains from joint production. During the 1986 trial, Coke argued that Dr. Pepper would benefit from increased production efficiency and scale economies, both of which would lower production costs substantially. Given the modest predictions for the impact of the merger in the absence of efficiency gains (above), this additional exercise seems unnecessary. However, in the case of Pepsi and SevenUP, it is possible that the latter would benefit from Pepsi's production facilities. In terms of distribution technology, SevenUP already piggybacks off of Pepsi's bottling network in most markets. Thus, any potential efficiency

Product	1986	Coke/Dr. Pep	Pepsi/7UP	Coke/Pepsi
PEPSI 12P	0.08	0.18	0.82	12.55
PEPSI 6P	0.05	0.21	0.68	19.28
PEPSI 67.6oz	0.03	0.25	2.08	12.94
PEPSI DT 12P	-0.01	0.04	1.05	23.72
PEPSI DT 6P	0.02	0.14	0.16	14.55
PEPSI DT 67.6oz	0.00	0.14	1.61	16.88
PEPSI DT CF 12P	-0.47	-0.83	-1.66	21.04
PEPSI DT CF 6P	0.05	-0.38	0.45	19.96
PEPSI 16oz	-0.01	-0.13	0.29	12.66
MT DW 12P	-0.18	-0.70	0.07	18.90
MT DW 6P	-0.03	-0.23	0.23	15.00
MT DW 67.6oz	0.02	-0.09	2.31	12.58
COKE CLS 12P	0.01	0.97	-0.07	16.13
COKE CLS 6P	0.03	1.87	0.12	21.31
COKE CLS 67.6oz	0.02	1.81	0.43	25.00
COKE DT 12P	-0.01	1.16	0.21	21.52
COKE DT 6P	-0.04	0.84	-0.26	22.26
COKE DT CF 12P	-0.38	1.80	-0.58	16.01
COKE DT 67.6oz	-0.02	0.90	0.11	21.08
SP CF 12P	-0.56	-2.17	-6.19	7.73
7UP R CF 67.6oz	0.98	0.28	14.29	8.37
7UP DT CF 67.6oz	2.07	0.18	15.07	11.43
DR PR 12P	0.01	4.32	-0.42	6.73
DR PR 6P	0.80	5.28	0.44	7.30
DR PR 67.6oz	1.19	6.05	-0.06	10.21
A and W CF 6P	1.71	0.41	1.24	11.05

Table 5: Median simulated percent change in price from Mergers

gains for SevenUP would need to reflect the production of concentrated syrup.

Table (7) reports the changes in prices from the analogous merger simulations using the aggregate DCM. As expected given the relatively low estimated price responses, the DCM predicts much lower price increases for the hypothetical mergers. As with the proposed model, the Coke and Dr. Pepper merger leads to very small changes. However, the DCM also predicts fairly small changes for the Pepsi and SevenUP merger. While the 6% increase in SevenUP prices is just above the legal limit, it is still well below the 10% limit used during the case against Coke. In contrast, the proposed model predicted SevenUP price changes of close to 15%. Using the DCM, even the extreme merger between Coke and Pepsi leads to much lower price changes than the proposed model. Only Coke's 67.6 ounce bottles change by as much as 8%. Most prices increase by just

Product	Merger 1	Merger 2	Merger 3
PepsiCo.	188.78	96.61	1125.98
Coca-Cola Co.	17.32	99.23	605.05
Dr. Pepper	-14.70	21.71	384.38
SevenUP	6.25	-28.10	150.78
consumer surplus	-28.96	-34.58	-12280.65
aggregate welfare	172.52	157.77	-9939.30

Table 6: Mean change in surplus (thousands of dollars per quarter)

under 5%, putting the merger right on the legal border of acceptance or rejection. This final prediction has little credibility given that Coke specifically claimed that this merger would dampen industry competition substantially. In general, a merger should internalize some of the competition, allowing firms to raise prices. The low cross-elasticities of the DCM imply much lower substitutability between products and, thus, less competition. Consequently, mergers do not have a very large impact on prices, according to the DCM specification.

8 Conclusions

Developing a viable structural model of demand to assess mergers in industries with a large number of differentiated products turns out to be a formidable task. Only recently have full-scale analyses been performed on complex industries like beer and cereals using aggregate data. Even these new techniques have their limitations, relying on strong assumptions either regarding product segmentation or consumer purchase behavior. For many products, such as CSDs, the behavioral assumptions may be inappropriate and the product groupings may be overly restrictive. Therefore, I use more disaggregate household-level point-of-purchase data to estimate demand. The demand model allows for more sophisticated purchase behavior, reflecting the tendency for consumers to purchase assortments. Allowing for consumers to purchase assortments of alternatives yields lower estimated price sensitivity than an aggregate DCM, which implicitly assumes single-unit purchase behavior. I also find the proposed model predicts much higher levels of substitutability between products, which imply higher manufacturer market power due to joint-pricing of products in the product line. This difference in market power leads to important differences in the merger analysis. The DCM results are not directly comparable as they use aggregate, rather than household, data. Nonetheless, I expect the DCM to be misspecified simply due to the observed multiple-discreteness in the household panel.

Product	1986	Cole/Dr. Pep	Pepsi/7UP	Coke/Pepsi
PEPSI 12P	-1.66	0.24	0.71	4.88
PEPSI 6P	0.07	0.05	0.61	4.32
PEPSI 67.6oz	-1.48	0.17	0.74	4.58
PEPSI DT 12P	-1.76	0.24	0.81	4.71
PEPSI DT 6P	0.06	0.08	0.54	3.99
PEPSI DT 67.6oz	-0.94	0.12	0.81	4.60
PEPSI DT CF 12P	-0.93	0.22	0.80	4.53
PEPSI DT CF 6P	0.52	0.04	0.80	3.50
PEPSI 16oz	0.45	0.04	0.40	2.44
MT DW 12P	-0.55	0.17	0.75	4.80
MT DW 6P	-0.38	0.08	0.93	4.09
MT DW 67.6oz	-0.84	0.15	0.73	4.21
COKE CLS 12P	0.10	0.35	0.06	6.94
COKE CLS 6P	3.20	0.62	0.08	4.94
COKE CLS 67.6oz	-0.60	0.97	0.08	7.96
COKE DT 12P	3.00	0.54	0.01	5.10
COKE DT 6P	3.12	0.74	0.12	4.71
COKE DT CF 12P	-0.42	0.73	0.06	5.88
COKE DT 67.6oz	-0.46	0.95	0.07	7.50
SP CF 12P	-0.83	0.90	0.09	5.37
7UP R CF 67.6oz	1.54	0.15	6.33	1.51
7UP DT CF 67.6oz	1.68	0.22	6.52	1.60
DR PR 12P	2.19	3.07	0.15	1.11
DR PR 6P	1.83	3.32	0.36	1.47
DR PR 67.6oz	0.33	3.66	0.30	1.41
A and W CF 6P	0.08	0.83	1.04	2.30

Table 7: Median simulated percent change in price from Mergers (DCM)

I use the demand estimates to revisit the 1986 merger trial against Coca-Cola's bid to acquire Dr. Pepper.

More than ten years have passed since the proposed mergers between Coke and Dr. Pepper, and Pepsi and SevenUP were successfully opposed by the FTC. During the case against Coke, several sophisticated economic arguments were put forth by both sides. In particular, Coke argued the merger would not increase its ability to raise prices due to the existing differentiation of products, whereas the FTC argued that the merger would lessen competition substantially, leading to more collusive prices. Coke also claimed that only the extreme merger between Coke and Pepsi would have a noticeable effect on industry competition.

The evidence appears to support Coke's claim that the merger with Dr. Pepper would not have a sizeable effect on industry prices. However, I do find large increases in the prices of SevenUP when it is merged with Pepsi. I also find that the hypothetical merger between Coke and Pepsi would have a substantial effect on industry competition. In fact, the merger simulation predicts extremely large price increases for Coke and Pepsi's entire product lines. In contrast, the aggregate DCM predicts much lower price changes for all three mergers, each of which is found to generate price changes below the 5% limit. The fact that the aggregate DCM does not reject the merger between Coke and Pepsi seems unrealistic, especially given that Coke argued it should be anticompetitive.

prods/units	1	2	3	4	5	6	7	8	9	10+	Total
1	20652	11238	1447	2454	245	454	33	282	19	215	37039
2	0	6928	2215	1817	436	464	146	259	45	166	12476
3	0	0	1322	768	302	247	114	109	45	130	3037
4	0	0	0	335	165	109	63	77	28	69	846
5	0	0	0	0	51	69	27	18	16	41	222
6	0	0	0	0	0	7	16	9	8	19	59
7	0	0	0	0	0	0	4	2	2	3	11
8	0	0	0	0	0	0	0	1	2	7	10
9	0	0	0	0	0	0	0	0	0	1	1
10	0	0	0	0	0	0	0	0	0	3	3
Total	20652	18166	4984	5374	1199	1350	403	757	165	654	53704

Table 8: Distribution of total CSD products and total units purchased on a given shopping trip (Conditional on a CSD purchase)

9 Appendix

A Data

I provide summary statistics of the demographic variables and the time-varying product attributes for the panel data in table (10). I compute the time between trips in days. Family size consists of the number of individuals reported for a given household. Temperature is the daily maximum, reported in degrees Fahrenheit. Income bracket is divided into 9 groups: 1 indicates less than \$10,000, 2 indicates between \$10,000 and \$20,000, 3 indicates between \$20,000 and \$30,000, 4 indicates between \$30,000 and \$40,000, 5 indicates between \$40,000 and \$50,000, 6 indicates between \$50,000 and \$60,000, 7 indicates between \$60,000 and \$70,000, 8 indicates between \$70,000 - \$100,000, and 9 indicates over \$100,000.

B Estimating the Aggregate DCM

The following approach for estimating demand is analogous to that of BLP(1995) and Nevo(2000) and I refer the more interested reader to either of these sources for a more thorough discussion. Referring back to section 3, imposing single-unit purchases yields the following indirect utility from purchasing brand j:

$$\begin{aligned} \widetilde{u_{hi}} &= \left(X_i\widetilde{\beta} - \widetilde{\phi}p_i\right) + \left(X_i \quad \sigma^h - \omega\sigma^h p_i\right) \\ &= u_j\left(\theta_1\right) + v_{ij}\left(\theta_2\right) \end{aligned} \tag{17}$$

where u_j is the mean-utility of consuming brand j and v_{ij} is consumer $i's$ idiosyncratic component of utility from consuming brand j. To ease estimation, the parameters are partitioned into two groups, θ_1 and θ_2. Adding a Type I extreme value disturbance yields the following *mixed-logit* probability that consumer i purchases brand j:

$$P_{ij} = \frac{exp\left(u_j + v_{ij}\right)}{1 + \sum exp\left(u_j + v_{ij}\right)}.$$

product	abbreviation
PEPSI COLA REGULAR 12 cans	PEPSI 12P
COKE CLASSIC 12 cans	COKE CLS 12P
PEPSI REGULAR 6 cans	PEPSI 6P
COKE DIET 12 cans	COKE DT 12P
PEPSI REGULAR 67.6oz	PEPSI 67.6oz
PEPSI DIET 12 cans	PEPSI DT 12P
COKE CLASSIC 6 cans	COKE CLS 6P
PEPSI DIET 6 cans	PEPSI DT 6P
COKE CLASSIC 67.6oz	COKE CLS 67.6oz
PEPSI DIET CL 67.6oz	PEPSI DT CL 67.6oz
COKE DIET 6 cans	COKE DT 6P
DR PEPPER 12 cans	DR PR 12P
MOUNTAIN DEW 12 cans	MT DW 12P
DR PEPPER 6 cans	DR PR 6P
7UP CAFFEINE-FREE 67.6oz	7UP R CF 67.6oz
COKE DIET CAFFEINE-FREE 12 cans	COKE DT CF 12P
COKE DIET 67.6oz	COKE DT 67.6oz
7UP DIET CAFFEINE-FREE 67.6oz	7UP DT CF 67.6oz
MOUNTAIN DEW 6 cans	MT DW 6P
SPRITE CAFFEINE-FREE 12 cans	SP CF 12P
PEPSI DIET CAFFEINE-FREE 12 cans	PEPSI DT CF 12P
DR PEPPER 67.6oz	DR PR 67.6oz
MOUNTAIN DEW 67.6oz	MT DW 67.6oz
PEPSI REGULAR 6 16oz bottles	PEPSI 16oz
PEPSI DIET CAFFEINE-FREE 6 cans	PEPSI DT CF 6P
A&W CAFFEINE-FREE 6 cans	A&W CF 6P

Table 9: CSD Products Used for Estimation

Variable	mean	standard deviation
kids	0.3865	0.4870
family size	2.6976	1.4034
income bracket	4.2470	1.9616
female under 35	0.1964	0.3973
time between trips	6.8498	13.7602
inventory	57.6103	86.7435
max. temperature	64.6149	19.8264
spring	0.2411	0.4278
summer	0.2785	0.4483
winter	0.2487	0.4322
price ($)	2.1515	0.3782
ad	0.3203	0.0579
display	0.4174	0.0503

Table 10: Descriptive Statistics (averaged across trips)

		continuous variable		
Flavor		calories	sodium (mg)	carbohydrates
cola	regular	150 (7.5)	40.5 (7.4)	41 (1.51)
	diet	0 (0)	34.7 (8.7)	0 (0)
lem/lime	regular	143.3 (5)	61.7 (16.4)	38.333 (0.5)
	diet	0 (0)	35 (0)	0 (0)
rootbeer	regular	168.3 (4.1)	44.2 (14.6)	44.8 (1.5)
citrus	regular	170 (0)	70 (0)	46 (0)
pepper	regular	148.6 (3.8)	45.7 (8.9)	35.1 (15.5)

Table 11: Continuous Attributes by flavor and diet vs. regular (averages)

		indicators					
Flavor		caffeine	phos.	citric	caramel	clear	#
cola	regular	7	7	4	7	0	7
	diet	6	9	9	9	0	9
lemon\lime	regular	0	0	2	0	2	2
	diet	0	0	1	0	1	1
rootbeer	regular	0	0	0	1	0	1
citrus	regular	3	0	3	0	3	3
pepper	regular	3	3	0	3	0	3

Table 12: Indicator Attributes by flavor and diet vs. regular (counts)

Product	Pep	Pep DT	Pep DT CF	Pep	Mt Dw	Coke	Coke Dt	A&W	Dr P
PEPSI 6P	-2.59	0.15	0.07	0.04	0.17	0.30	0.20	0.10	0.18
PEPSI DT 6P	0.61	-3.46	0.10	0.07	0.09	0.32	0.20	0.04	0.32
PEPSI DT CF 6P	1.30	0.78	-3.60	0.00	0.00	0.17	0.00	0.18	0.00
PEPSI 16oz	0.19	0.00	0.00	-2.46	0.00	0.13	0.00	0.00	0.00
MT DW 6P	0.48	0.13	0.29	0.00	-3.43	0.37	0.09	0.14	0.33
COKE CLS 6P	0.25	0.16	0.16	0.03	0.25	-3.13	0.28	0.12	0.28
COKE DT 6P	0.24	0.23	0.04	0.05	0.00	0.53	-2.98	0.04	0.04
A and W CF 6P	0.69	0.21	0.00	0.00	0.16	0.19	0.19	-3.83	0.32
DR PR 6P	0.40	0.23	0.18	0.00	0.20	0.33	0.15	0.10	-4.17
PEPSI 12P	0.07	0.04	0.00	0.07	0.00	0.05	0.00	0.00	0.03
PEPSI DT 12P	0.11	0.05	0.00	0.00	0.00	0.09	0.05	0.00	0.08
PEPSI DT CF 12P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MT DW 12P	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00
COKE CLS 12P	0.15	0.08	0.00	0.06	0.00	0.00	0.03	0.00	0.00
COKE DT 12P	0.04	0.08	0.00	0.10	0.06	0.00	0.00	0.00	0.07
COKE DT CF 12P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SP CF 12P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DR PR 12P	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08
PEPSI 67.6oz	0.19	0.06	0.02	0.00	0.07	0.13	0.01	0.02	0.08
PEPSI DT 67.6oz	0.12	0.10	0.00	0.04	0.00	0.11	0.04	0.00	0.04
MT DW 67.6oz	0.09	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.00
COKE CLS 67.6oz	0.15	0.05	0.00	0.08	0.00	0.12	0.00	0.00	0.11
COKE DT 67.6oz	0.07	0.12	0.00	0.00	0.00	0.14	0.26	0.07	0.00
7UP R CF 67.6oz	0.12	0.08	0.00	0.00	0.07	0.08	0.00	0.00	0.00
7UP DT CF 67.6oz	0.19	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.14
DR PR 67.6oz	0.15	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00

Table 13: Price Elasticities for 6-packs of cans (quarterly median using Proposed Model)

Product	Pep	Pep DT	Pep DT CF	Mt Dw	Coke	Coke DT	Coke DT CF	SP	Dr P
PEPSI 6P	0.07	0.03	0.00	0.00	0.04	0.02	0.02	0.00	0.03
PEPSI DT 6P	0.11	0.00	0.00	0.00	0.08	0.06	0.00	0.00	0.00
PEPSI DT CF 6P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PEPSI 16oz	0.11	0.07	0.00	0.00	0.00	0.12	0.00	0.00	0.00
MT DW 6P	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
COKE CLS 6P	0.07	0.06	0.00	0.00	0.05	0.00	0.00	0.00	0.00
COKE DT 6P	0.05	0.04	0.00	0.00	0.00	0.04	0.00	0.00	0.00
A and W CF 6P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DR PR 6P	0.08	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00
PEPSI 12P	-2.97	0.17	0.07	0.19	0.18	0.12	0.09	0.00	0.06
PEPSI DT 12P	0.59	-3.26	0.18	0.13	0.24	0.36	0.10	0.00	0.13
PEPSI DT CF 12P	0.40	0.47	-3.27	0.00	0.23	0.49	0.00	0.00	0.23
MT DW 12P	0.15	0.38	0.00	-4.37	0.00	0.26	0.00	0.00	0.00
COKE CLS 12P	0.39	0.10	0.06	0.06	-3.10	0.17	0.14	0.06	0.04
COKE DT 12P	0.35	0.23	0.00	0.13	0.56	-3.45	0.09	0.00	0.08
COKE DT CF 12P	0.32	0.00	0.00	0.24	1.27	0.22	-4.41	0.00	0.11
SP CF 12P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-6.09	0.99
DR PR 12P	0.50	0.26	0.00	0.19	0.20	0.20	0.00	0.00	-3.49
PEPSI 67.6oz	0.21	0.05	0.04	0.08	0.07	0.05	0.02	0.00	0.05
PEPSI DT 67.6oz	0.08	0.12	0.04	0.00	0.14	0.17	0.04	0.00	0.05
MT DW 67.6oz	0.16	0.00	0.00	0.00	0.10	0.08	0.00	0.00	0.00
COKE CLS 67.6oz	0.24	0.12	0.00	0.06	0.07	0.12	0.00	0.00	0.10
COKE DT 67.6oz	0.07	0.17	0.00	0.00	0.14	0.00	0.00	0.00	0.00
7UP R CF 67.6oz	0.34	0.00	0.00	0.00	0.11	0.10	0.00	0.00	0.00
7UP DT CF 67.6oz	0.00	0.00	0.00	0.00	0.17	0.17	0.00	0.00	0.00
DR PR 67.6oz	0.08	0.10	0.00	0.00	0.17	0.16	0.00	0.00	0.00

Table 14: Price Elasticities for 12-packs of cans (quarterly median using Proposed Model)

Product	Pep	Pep DT	Mt Dw	Coke	Coke DT	7UP	7UP DT	Dr P
PEPSI 6P	0.07	0.05	0.02	0.07	0.06	0.04	0.02	0.03
PEPSI DT 6P	0.07	0.07	0.00	0.00	0.03	0.04	0.00	0.03
PEPSI DT CF 6P	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PEPSI 16oz	0.12	0.00	0.00	0.08	0.00	0.05	0.00	0.09
MT DW 6P	0.07	0.00	0.05	0.00	0.00	0.09	0.00	0.00
COKE CLS 6P	0.11	0.09	0.03	0.07	0.04	0.03	0.02	0.00
COKE DT 6P	0.06	0.08	0.00	0.05	0.03	0.00	0.00	0.00
A and W CF 6P	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00
DR PR 6P	0.14	0.05	0.00	0.07	0.00	0.04	0.04	0.05
PEPSI 12P	0.19	0.09	0.02	0.06	0.04	0.03	0.02	0.00
PEPSI DT 12P	0.08	0.08	0.00	0.08	0.04	0.04	0.03	0.00
PEPSI DT CF 12P	0.14	0.00	0.08	0.12	0.00	0.00	0.00	0.00
MT DW 12P	0.20	0.00	0.00	0.16	0.08	0.08	0.00	0.00
COKE CLS 12P	0.09	0.07	0.02	0.09	0.04	0.05	0.03	0.03
COKE DT 12P	0.14	0.06	0.06	0.05	0.06	0.05	0.05	0.00
COKE DT CF 12P	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00
SP CF 12P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DR PR 12P	0.24	0.06	0.00	0.16	0.00	0.00	0.00	0.00
PEPSI 67.6oz	-2.49	0.14	0.08	0.20	0.08	0.11	0.03	0.08
PEPSI DT 67.6oz	0.27	-3.10	0.06	0.12	0.08	0.06	0.05	0.08
MT DW 67.6oz	0.42	0.21	-3.55	0.08	0.06	0.15	0.07	0.12
COKE CLS 67.6oz	0.28	0.14	0.09	-3.04	0.23	0.07	0.09	0.12
COKE DT 67.6oz	0.12	0.25	0.05	0.18	-3.61	0.00	0.06	0.09
7UP R CF 67.6oz	0.31	0.06	0.07	0.09	0.05	-3.02	0.14	0.00
7UP DT CF 67.6oz	0.21	0.11	0.00	0.10	0.09	0.39	-2.86	0.00
DR PR 67.6oz	0.20	0.27	0.30	0.10	0.11	0.11	0.07	-3.85

Table 15: Price Elasticities for 67.6 oz bottles (quarterly median using Proposed Model)

Aggregating these probabilities over the set of consumers, A_j, that purchase brand j gives the market share:

$$S_j = \int_{Aj} \frac{exp\left(u_j + v_{ij}\right)}{1 + \sum exp\left(u_j + v_{ij}\right)} \partial\Phi(v_{ij}).$$

Estimating the share equation directly introduces several computational problems. Instrumenting becomes quite complicated due to the non-linear fashion in which the explanatory variables enter the model. For CSDs, I also expect the prediction error of this system of equations to be highly correlated within a given store-week and across weeks for a given store. To alleviate these issues, Berry (1994) suggests working with the mean utility, u_j, numerically inverting the share equation in terms of the mean utility:

$$u_j\left(\theta_2\right) = X_j\widetilde{\beta} - \widetilde{\phi}p_j.$$

This equation is much simpler to manipulate econometrically since the mean taste parameters now enter linearly. If an attribute is missing with mean taste, ζ_j, then I can construct a Generalized Method of Moments estimator based on the assumption $E\left(X_j\zeta_j|X_j\right) = 0$. Using a matrix of exogenous instruments, Z, which contains X_j as well as supply-side prices of the factors of production, I form the conditional moments $E\left(Z_j\zeta_j|Z_j\right) = 0$. Estimation amounts to finding the vector $(\theta_1^*, \theta_2^*)'$ that minimizes:

$$G\left(\theta_1, \theta_2\right) = \zeta' ZWZ'\zeta$$

where the weight matrix, W, is the inverse of the estimated variance of the conditional moments (Hansen 1982).

Having estimated the demand parameters of the DCM, I compute elasticities using the following formula:

$$\varepsilon^j_{p_k} = \frac{p_k}{\int_{A_j} P_{ij}\Phi(dv)} \int_{A_j} \frac{\partial P_{ij}}{\partial p_k} \Phi(dv).$$

For estimation, I use weekly data for 22 stores from the largest supermarket chain in my data. I measure market shares as the share of total weekly store traffic. Implicitly, I assume that consumers purchase a single unit of one of the CSD products on each trip, where one of the products is the no-purchase alternative. This assumption is clearly inconsistent with the shopping behavior observed in the consumer panel data, thus providing the contrast with the proposed model. Summary statistics of this data are in table (16). Since I use aggregate data from a single chain, the results are not directly comparable to those of the multiple-discreteness model, which uses individual transaction data at all the supermarkets. Moreover, I am unable to provide a statistical test for the validity of the discrete choice assumption, despite the similarities between the two models. Instead, I compare the qualitative outcomes, substitution patterns and merger predictions, as a crude assessment of the discrete choice assumption for data exhibiting multiple-unit purchases.

In Table (17), I report the parameter estimates for two specifications of the aggregate DCM. In the second specification, I add random coefficients on some of the product attributes to be more consistent with the proposed model. For each model, I report both the mean and the standard deviation of each random taste parameter. For both models, I find substantial heterogeneity in the price response coefficient. However, I do not find the standard deviations of the ad and display coefficients to be significantly different from zero in the first specification. Contrary to the findings of the proposed

Variable	67.6 oz bottle		6-pack cans		16 oz bottle		12-pack cans	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
share of unit sales	0.045	0.04	0.05	0.06	0.02	0.05	0.02	0.03
price (cents per 12 oz)	17.83	5.16	21.81	10.96	33.44	10.89	29.83	6.47
feature	0.43	0.50	0.53	0.50	0.14	0.35	0.46	0.50
display	0.46	0.50	0.55	0.50	0.12	0.31	0.50	0.50
temperature (degrees F)	50.67	16.44	50.67	16.44	50.67	16.44	50.67	16.44

Table 16: Data Used in Aggregate DCM

model, the DCM predicts that consumers do not vary in their responses to marketing instruments. Unexpectedly, the influence of temperature is negative, on average, with a relatively large amount of variation. I suspect this result may be due to the fact that some of the winter holidays generate spikes in sales. Including a holiday dummy in the specification (to distinguish the purchase versus no purchase decision) might absorb this negative temperature effect. I also find variation in the degree to which consumers value diet products as well as the various package sizes. Adding these additional random coefficients reduces the mean response of prices as well as the degree of heterogeneity.

Product	DCM 1		DCM 2	
	mean	st.dev.	mean	st.dev.
price	-13.60	2.39	-10.08	0.70
	(1.56)	(0.16)	(1.99)	(0.02)
ad	0.28	0.09	0.39	0.54
	(0.05)	(0.25)	(0.06)	(0.08)
display	0.43	0.13	0.56	0.29
	(0.04)	(0.25)	(0.05)	(0.17)
temperature	-0.02	0.02	-0.02	0.02
	(0.00)	(0.00)	(0.00)	(0.00)
diet				1.20
				(0.03)
6-pack				3.09
				(0.06)
12-pack				0.06
				(1.04)
16-oz bottle				1.77
				(1.04)

Table 17: Aggregate Random Coefficients Discrete Choice Model (standard errors in parentheses)

References

[1] Baker J.B. and T.F. Bresnahan [1985]. "The Gains from Merger or Collusion in Product-Differentiated Industries." *The Journal of Industrial Economics*, 33, 427-444.

[2] Berry, S. [1994]."Estimating Discrete-Choice Models of Product Differentiation." *Rand Journal of Economics,* 25, 242-62.

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

[4] Berry, S., J. Levinsohn, and A. Pakes [1998]. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." Yale University, Working Paper.

[5] Berry, S. and A. Pakes [1993]. "Some Applications and Limitations of Recent Advances in Empirical Industrial Organization: Merger Analysis." *American Economic Review*, 83, 247-252.

[6] Besanko, D., S. Gupta and D. Jain [1998]. "Logit Demand Estimation Under Competitive Pricing Behavior: An Equilibrium Framework." *Management Science*, November.

[7] Blattberg, R.C. and S.A. Neslin [1990]. *Sales Promotion: Concepts, Methods, and Strategies*. Englewood Cliffs, New Jersey: Prentice Hall.

[8] Bresnahan, T.F. [1989]."Empirical Methods for Industries with Market Power", in *Handbook of Industrial Organization*, ed. R. Schmalensee and R. Willig, North Holland.

[9] Chamberlain, G. [1987]. "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions." *Journal of Econometrics*, 34, 305-334.

[10] Chiang J. [1991]."The Simultaneous Approach to the Whether, What, and How much to Buy Questions" Marketing Science vol.10 p297-315.

[11] Chintagunta, P. [1993]. "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households." *Marketing Science,* 12, 184-208.

[12] Conley, T.G. [1999]. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92, 1-45.

[13] Cotterill, R. W., A. W. Franklin, and L. Y. Ma [1996] "Measuring Market Power Effects in Differentiated Product Industries: An Application to the Soft Drink Industry." Food Marketing Policy Center, University of Connecticut, Storrs, CT.

[14] Langan, G. E. and R. W. Cotterill [1994]. "Estimating Brand Level Demand Elasticities and Measuring Market Power for Regular Carbonated Soft Drinks." University of Connecticut, Storrs, CT.

[15] Dubé, J.P. [2000]. "Multiple Discreteness and Differentiated Products: Strategy and Demand for Carbonated Soft Drinks." Northwestern University, *Working Paper.*

[16] Fershtman C. and A. Pakes [2000]. "A Dynamic Game with Collusion and Price Wars." *The Rand Journal of Economics*, 207-36.

[17] 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, 277-311.

[18] Goldberg, P. [1995]. "Product Differentiation and Oligopoly in International Markets: The case of the US automobile industry." *Econometrica*, 63, 891-951.

[19] Gupta, Sachin, P. Chintagunta, A. Kaul and D.R. Wittink [1996]. "Do household scanner data provide representative inferences from brand choices: A comparison with store data." Journal of Marketing Research, 33, 383-441.

[20] Hanemann W. M. [1984]."Discrete/Continuous Models of Consumer Demand" *Econometrica* 52:541-561.

[21] Hansen, L.P. [1982]. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica,* 50, 1092-1054.

[22] Hauser, J.R. and B. Wernerfelt [1990], "An Evaluation Cost Model of Consideration Sets." *Journal of Consumer Research*, 16, 393-408.

[23] Hausman, J.A. [1994]. "Valuation of New Goods Under Perfect and Imperfect Competition." NBER Working Paper #4970.

[24] Hausman, J. A., G.K. Leonard and D. McFadden [1995]. "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses due to Natural Resource Damage." *Journal of Public Economics*,56, 1-30.

[25] Hausman, J.A., G.K.and J.D. Zona [1994]. "Competitive Analysis with Differenciated Products." *Annales d'Economie et de Statistique*, 0, 159-80.

[26] Hendel, I. [1999]. "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns." *Review of Economic Studies, 66, 423-446.*

[27] Higgins, R. S., D. P. Kaplan, M. J. McDonald and R. D. Tollison [1995]. "Residual Demand Analysis of the Carbonated Soft Drink Industry." *Empirica*, 22, 115-26.

[28] Kim, J., G.M. Allenby and P.E. Rossi [1999]. "Modeling Consumer Demand for Variety." University of Chicago, *Working Paper.*

[29] McCulloch R. and P.E. Rossi [1994], "An Exact Likelihood Analysis of the Multinomial Probit Model." *Journal of Econometrics,* 64, 207-240.

[30] McFadden, D. [1973]. "Conditional Logit Analysis of Discrete Choice Behavior," in P. Zarembka, ed., *Frontiers of Econometrics*, New York, Academic Press.

[31] McFadden, D. [1981]. "Econometic Models of Probabilistic Choice." in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. Manski, and D. McFadden. MIT Press, Cambridge, MA.

[32] McFadden, D. [1989]. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration ." *Econometrica, 57, 995-1026.*

[33] Muris, T.J., D.T. Scheffman and P.T. Spiller [1992a]. "Strategy and Transaction Costs: The Organization of Distribution in the Carbonated Soft Drink Industry." *Journal of Economics and Management Strategy*, 1, 83-128.

[34] Muris, T.J., D.T. Scheffman and P.T. Spiller [1992b]. "Strategy and Transaction Costs: The Organization of Distribution in the Carbonated Soft Drink Industry" in *Case Studies in contracting and organization, ed. Scott Masten, Oxford University Press.*

[35] Nevo, A. [2000]. "Measuring Market Power in the Ready-To-Eat Cereal Industry." forthcoming *Econometrica.*

[36] Nevo, A. [1999]. "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry." U.C. Berkeley, *Working Paper.*

[37] Newey, W.K. and K.D. West [1987]. "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55, 703-708.

[38] Pakes, A. and D. Pollard [1989]. "Simulation and the Asymptotics of Optimization Estimators." *Econometrica* , 57, 1027-57.

[39] Petrin, A. [1999]. "Quantifying the Benefits of New Products: The Case of the Minivan." University of Chicago, *WorkingPaper*.

[40] Seetharaman, P.B. [1999], "Modeling Serial Correlation in Household Brand Choices using a Multinomial Probit Model of Brand Choice." Washington University, *Working Paper.*

[41] Slade, M.E. [1995]. "Product Market Rivalry with Multiple Strategic Weapons: an analysis of price and advertising competition." *Journal of Economics and Management Strategy*, 4, 445-476.

[42] M. Vilas-Boas and R. Winer [1999]. "Endogeneity in Brand Choice Models." forthcoming *Management Science.*

[43] Walsh, J.W. [1995], "Flexibility in Consumer Purchasing for Uncertain Future Tastes." *Marketing Science*, 14, 148-165.

[44] Werden G.J. and L.M. Froeb [1994]. "The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy." *Journal of Law, Economics and Organization*, 10, 407-26.

[45] White, L.J [1989]. "The Proposed Merger of Coca-Cola and Dr Pepper." In *The antitrust revolution*, ed. Kwoka, J.E. and L.J. White (Glenview, Ill: Scott, Foresman), 76-95.