Product Differentiation and Mergers in the Carbonated Soft Drink Industry

Jean-Pierre Dubé
University of Chicago
Graduate School of Business
Chicago, IL 60637
jdube@gsb.uchicago.edu

I simulate the competitive impact of several soft drink mergers from the 1980s on equilibrium prices and quantities. An unusual feature of soft drink demand is that, at the individual purchase level, households regularly select a variety of soft drink products. Specifically, on a given trip households may select multiple soft drink products and multiple units of each. A concern is that using a standard discrete choice model that assumes single unit purchases may understate the price elasticity of demand. To model the sophisticated choice behavior generating this multiple discreteness, *I use a household-level scanner data set. Market demand is then computed by aggregating the household estimates. Combining the aggregate demand estimates with a model of static oligopoly, I then run the merger simulations. Despite moderate price increases, I find substantial welfare losses from the proposed merger between Coca-Cola and Dr. Pepper. I also find large price increases and corresponding welfare losses from the proposed merger between Pepsi and 7 UP and, more notably, between Coca-Cola and Pepsi.*

1. Introduction

With the advent of aggregate brand-level data collected at supermarket checkout scanners, researchers have begun to use structural econometric models for policy analysis. The rich content of scanner data enables the estimation of demand systems and their corresponding cross-price elasticities. The areas of merger and antitrust policy have been strong beneficiaries of these improved data. Recent advances in structural approaches to empirical merger analysis consist of combining

I am very grateful to my thesis committee: David Besanko, Tim Conley, Sachin Gupta, and Rob Porter. I also thank Arie Beresteanu, Judy Chevalier, Anne Gron, Gunter Hitsch, Amil Petrin, and John Woodruff for their comments. I also benefitted from the detailed comments of an area editor and an anonymous reviewer, as well as seminar participants at the University of Chicago, 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.

© 2005 Blackwell Publishing, 350 Main Street, Malden, MA 02148, USA, and 9600 Garsington Road, Oxford OX4 2DQ, UK.
Journal of Economics & Management Strategy, Volume 14, Number 4, Winter 2005, 879–904

the estimated demand system with a game-theoretic model of the competitive industry structure to simulate the impact of a merger on equilibrium prices (e.g., Baker and Bresnahan, 1985; Berry and Pakes, 1993; Hausman et al., 1994; Werden and Froeb, 1994; Nevo, 2000). The use of aggregate data (metropolitan level or national level) generally requires strong assumptions in order to build a demand system from primitives on a model of consumer choice while accommodating a large number of product alternatives. A concern for policy applications is whether these modeling assumptions could have adverse effects on the estimated cross-price elasticities and hence the implications for merger-related gains in market power. The increasing availability of more micro consumer scanner data is a useful starting point for estimating consumer demand in product categories that do not satisfy typical modeling assumptions, such as discrete choice purchase behavior.

I investigate the economic impact of several previously challenged mergers in the carbonated soft drinks (CSD) industry. The CSD industry presents an interesting opportunity for research. For the past two decades, CSD manufacturers have been under heavy scrutiny following aggressive attempts by the major players, Coca-Cola Co. and PepsiCo., to increase their market shares through acquisitions.[1] In a landmark case against Coca-Cola Co., the Federal Trade Commission (FTC) successfully blocked the proposed acquisition of Dr. Pepper. Just prior to the trial, Pepsi called off its proposed merger with 7 Up. Despite an unprecedented use of economics during the trial, the court resorted to a traditional market share concentration-based argument (White, 1989).[2] In ruling against Coca-Cola, the Federal District Court found insufficient empirical evidence to assess the economic claims. Several years later, in 1995, the FTC and Coke reached an agreement that Coca-Cola would not acquire the rights to Dr. Pepper. Furthermore, Coke would seek FTC approval for the acquisition of any CSD manufacturer with sales exceeding 75 million gallons for each of the three prior years (i.e., the seven largest CSD firms behind Coke) until 2004.[3]

To assess the economic impact of these mergers, I estimate demand for CSDs using household-level scanner panel data. An interesting feature of the observed household purchase behavior is the regular purchase of assortments of CSDs across households and shopping trips. Shopping baskets often consist of several different CSD products and multiple units of each. This behavior rules out popular *discrete choice* modeling approaches for consumer demand, such as the multinomial

1. Coca-Cola has also been challenged internationally for its moves to acquire Cadbury Schweppes brands in Europe.
2. *F.T.C.* v. *Coca-Cola Co.*, 641 F. Supp. 1128 (1986).
3. See the FTC press release at http://www.ftc.gov/opa/1995/05/coke7.htm.

logit or probit. These models would be convenient in the CSD industry as they accommodate a large number of differentiated products with a parsimonious set of parameters. Instead, I use Hendel's (1999) *multiple discreteness* model that accommodates assortment decisions. Recognizing the difference between the time of purchase and the time of consumption, consumers are assumed to make multiple decisions in anticipation of a stream of future consumption occasions. At the time of a shopping trip, the consumer makes several discrete choices—one for each anticipated consumption occasion.

Using the aggregate predicted individual demands, I then compute the manufacturer margins and marginal costs that would prevail in a Bertrand–Nash equilibrium. The combination of demand and marginal costs provides the basis for simulating the effects of several hypothetical CSD mergers. The simulations provide evidence supporting Coke's claim that the merger with Dr. Pepper would not lead to large price increases. However, the merger nevertheless generates substantial welfare losses to the Denver economy. The results for both pricing and welfare losses clearly support the FTC's opposition to the merger between Pepsi and 7 UP. Finally, the hypothetical merger between Coke and Pepsi results in very large price increases and welfare losses, as expected.

The remainder of the paper is organized as follows. Section 2 describes the CSD market structure at the time of the sample. Section 3 derives the demand and supply-side models used for analysis. Section 4 outlines the estimation procedure. Section 5 describes the data. Section 6 presents the empirical results from the demand estimation and the merger simulations. Finally, Section 7 concludes and outlines possibilities for further research.

2. Mergers in the CSD Industry

Pepsi's William C. Munro once noted, "The soft drink is not a serious thing. No one needs it."[4] However, a recent study by the national soft drink association (NSDA) estimates that the industry currently employs over 175,000 people, generating $8 billion per year in wages and salaries.[5] 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 per year. In contrast, the second largest beverage, beer, accounted for only 19.4%, roughly 22.1 gallons per capita

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

5. *Economic Impact Of The Soft Drink Industry*, The national softdrink association, www.nsda.org.

in 1998.[6] AC Nielsen estimates that the CSD category is the largest in the Dry Grocery department at US food stores, accounting for roughly one-tenth of such departments' national sales revenue. Finally, the Coca-Cola brand name, widely regarded as the world's most valuable brand, has an estimated value of $70 billion.[7]

As it became mature in the 1980s, a wave of consolidations swept the CSD industry. By 1989, Cadbury Schweppes had acquired Canada Dry, Hires Root Beer, and Crush; and Hicks and Haas had acquired 7 UP, 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 1 firm) announced plans to acquire Dr. Pepper (the number 3 firm) and Pepsi (the number 2 firm) announced plans to acquire 7 UP (the number 4 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 7 UP immediately canceled their merger plans. However, Coke persisted, bringing the case to the Federal District Court.

Although 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).[8] In an unusual departure from traditional market shared-based arguments, the FTC and Coca-Cola also presented several economic arguments. The FTC argued that CSD profits were the result of tacit collusion and that a merger would exacerbate this problem. In contrast, Coke argued that product differentiation complicated coordination so it would be virtually impossible even with a merger.[9] Coke further claimed that intense competition from Pepsi would keep prices low regardless of the merger. Coke also predicted that Dr. Pepper would benefit from more efficient production and, thus, would lower its price. Finally, Coke argued that only the merger between Coke and Pepsi would lead to an objectionable decrease in competition. Lacking sufficient empirical evidence supporting the economic arguments, the court was unable to take the economic arguments into consideration in its final decision.

6. *Beverage World*; East Stroudsburg; May 15, 1999; Greg W Prince.
7. Gerry Khermouch "The Best Global Brands," *Business Week*, August 5, 2002, p. 92.
8. The court rejected Coke's claim that the market was the national beverage industry (e.g., milk, juice, coffee), which generated a postmerger HHI of only 739.
9. The FTC reported a 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 (White, 1989).

Table I.
Distribution of Total CSD Products and Total Units Purchased on a Given Shopping Trip (Conditional on a CSD Purchase)

Prods/Units	1	2	3	4	5	6	7	8	9	10+	Total
1	20,652	11,238	1,447	2,454	245	454	33	282	19	215	37,039
2	0	6,928	2,215	1,817	436	464	146	259	45	166	12,476
3	0	0	1,322	768	302	247	114	109	45	130	3,037
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	20,652	18,166	4,984	5,374	1,199	1,350	403	757	165	654	53,704

3. The Model

3.1 Multiple-Unit Purchases

An interesting feature of the CSD purchase data is the frequent incidence of multiple-item purchases. In contrast to the behavior implied by the typical discrete choice models (e.g., logit or probit), households do not always select a single unit of a single CSD product on a given shopping trip. Table I breaks down the distribution of shopping trips during which a CSD was purchased by the total number of different CSD products purchased and the total number of units on a given trip.[10] In fact, only 39% of the trips result in a single-unit purchase. Past research has shown that ignoring quantity choices (e.g., looking only at product choices) can lead to underestimated price elasticities of demand (Chintagunta, 1993; Bell et al., 1999). In the current context, the problem is exacerbated by the fact that households may pick several different products and some quantity of each.

To capture the assortment decisions of households, I use the Hendel (1999) model, which combines utility-maximization problem and household-specific variables reflecting purchase history. One explanation for why households purchase assortments arises due to the separation between the time of purchase and the time of consumption. Typically, a consumer makes shopping decisions in anticipation of a stream of future consumption occasions before the next shopping trip.

10. In the current context, an alternative refers to a specific product as defined by its UPC and a unit refers to the number of units purchased of a given UPC.

A basket of goods is selected to satisfy each of these anticipated needs. Differences in tastes across consumption occasions leads to the purchase of an assortment. 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).

3.2 Demand

In this section I derive consumer demand for CSDs at the time of the shopping trip. The decision of when to visit a store and where to shop is treated as exogenous.[11] 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 (time subscripts are dropped for simplification): $Q^h = \sum_{j=1}^{J^h} Q_j^h$. The actual number J^h is not observed by the econometrician. I assume that J^h derives from a distribution characterized by household demographics and purchase history (inventory) summarized in a $(d \times 1)$ vector of household characteristics, D^h. Because the number of choices a consumer makes is an integer, I assume J^h is distributed Poisson with mean, λ, depending on the household characteristics, D^h, as in Hendel (1999),

$$\begin{aligned} J^h &\sim P\left(\lambda^h\right), \\ \lambda_h &= D^{h\prime}\delta, \end{aligned} \tag{1}$$

where $P(\cdot)$ is the CDF of a Poisson.

Each household has quasi-linear preferences for its vector of purchases of the I soft drink products available, Q^h, 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

$$U^h(z, Q^h) = \sum_{j=1}^{J^h} u_j^h\left(\sum_{i=1}^{I} \Psi_{ij}^h Q_{ij}^h\right) + \alpha^h z, \tag{2}$$

where Q_{ij}^h is the quantity purchased of product i and Ψ_{ij}^h captures the household's perceived quality of product i for consumption occasion j. The parameter α^h captures the marginal utility of income spent on

11. This is a common assumption for scanner data models. It is unlikely that prices in a single product category will influence a consumer's store choice. Berto Villas-Boas (2001), finds cross-store price elasticities to be extremely small in the yogurt category. Similarly, Slade (1995) conducted a survey of retailers that revealed consumers do not tend to search on specific item prices across stores each week. Nevertheless, if consumers search more for low prices when their demand for CSD is high, the correlation between prices and the household's error term could generate endogeneity bias.

other product categories during a store trip.[12] The additive separability of the J^h subutility functions enables solving of the decision for each consumption occasion independently. For a given consumption occasion, the perfect substitutes structure combined with curvature assumptions for $u_h^j(\cdot)$ ensure that households select a nonnegative quantity of each product. Because the perceived product qualities, Ψ_{ij}^h, vary across the J^h consumption occasions, households may purchase a variety of products on a given trip.

The household faces an expenditure constraint, $\sum_{j=1}^{J^h} \sum_{i=1}^{I} p_i Q_{ij}^h + z \leq y^h$, where p_i is the price of product i and y^h is the household's total shopping budget. Substituting the expenditure equation, which I assume is binding, into the original utility function gives

$$U^h(Q^h) = \sum_{j=1}^{J^h} u_j^h \left(\sum_{i=1}^{I} \Psi_{ij}^h Q_{ij}^h \right) + \alpha^h \left(y^h - \sum_{j=1}^{J^h} \sum_{i=1}^{I} p_i Q_{ij}^h \right). \tag{3}$$

Assuming a specific functional form for the subutility functions in (3), the household decision is broken into J^h separate problems, with a subutility for each expected consumption occasion j,

$$\begin{aligned} u_j^h(Q_j^h) &= \left(\sum_{i=1}^{I} \Psi_{ij}^h Q_{ij}^h \right)^{\gamma} - \alpha^h \sum_{i=1}^{I} p_i Q_{ij}^h \\ \Psi_{ij}^h &= \max\left(0, X_i \beta_j^h + \xi_i\right)^{m(D^h)}, \end{aligned} \tag{4}$$

where X_i is a $(1 \times k)$ vector of product i's observable attributes, β_j^h is a $(k \times 1)$ vector of random tastes for attributes during consumption occasion j, and ξ_i captures the influence of any remaining unmeasured characteristics of product i. If these attributes are observed by CSD manufacturers and incorporated into their pricing decisions, then they will be correlated with prices (Berry, 1994). This endogeneity of prices could bias the parameter estimates. To resolve this problem, I estimate the full set of product intercepts, ξ_i, as fixed effects.[13] Because $m^h = m(D^h)$ does not vary across products, it captures a household's taste for quality as a function of demographics.[14] Hence, a household with a higher

12. Because Hendel (1999) models a profit function, he normalized this term to 1.

13. It is possible that residual weekly variation in unobserved product characteristics could shift demand and prices (e.g., Villas-Boas and Winer, 1999). I assume any remaining sources of measurement error will be offset by the fact that households shop on different dates and in different stores, generating heterogeneity in the prices faced by households (e.g., Shum, 2000).

14. As discussed in Hendel (1999), this term enables some households to purchase relatively more expensive items regardless of their tastes for attributes. This will introduce a verical component to product differentiation.

$m(D_h)$ will put more weight on differences in products' qualities, Ψ_{ij}^h. Because the parameter $m(D_h)$ does not vary across choice occasions, it also adds another layer of heterogeneity across households (versus across choice occasions). The parameter γ determines the curvature of the utility function which, for simplicity, I assume is the same for all households. So long as the estimated value of γ lies between 0 and 1, the model maintains the concavity property needed for an interior solution.

For a given expected consumption occasion j, the household faces a set of latent utilities, $u_j^* = (u_{j1}^*, \ldots, u_{jI}^*)$, where $u_{ji}^* = \max_Q u_j^h(Q_j)$. The household selects product i if $u_{ji}^* = \max(u_{j1}^*, \ldots, u_{jI}^*)$. The optimal quantity of product i for occasion j, Q_{ij}^{h*}, satisfies the first-order condition:

$$\gamma \left(\Psi_{ij}^h\right)^{\gamma} \left(Q_{ij}^h\right)^{\gamma - 1} - \alpha^h p_i = 0. \tag{5}$$

Solving for the optimal Q_{ij}^h gives

$$Q_{ij}^{h*} = \left(\frac{\gamma \left(\Psi_{ij}^h\right)^{\gamma}}{\alpha^h p_i}\right)^{\frac{1}{1-\gamma}}. \tag{6}$$

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*}. Because Ψ may take on negative values, this specification also allows for zero demand (no purchase) for a given product and consumption occasion.

In the data, one does not observe the individual choices for each consumption occasion. The data contain the quantity of each product purchased on each trip, where such quantities are the sum of optimal decisions across each of the consumption occasions. The derived expected aggregate purchase vector has the following form:

$$\begin{aligned} E\,Q^h(D^h, X, \Theta) = \sum_{J^h=1}^{\infty} \sum_{j=1}^{J^h} \int_{-\infty}^{\infty} & Q_j^{h*}\left(D^h, X, \beta_j^h, \Theta\right) \\ & \times f(\beta | D^h, \Theta) p(J | D^h, \Theta) \partial\beta \partial J, \end{aligned} \tag{7}$$

where $f(\beta | D^h, \Theta)$ is the normal pdf associated with the taste vector conditional on household characteristics and model parameters, and $p\left(J | D^h, \Theta\right)$ is the Poisson pdf of the number of expected consumption occasions conditional on household characteristics and model parameters. Two advantages of this modeling approach are that it can accommodate many products, by the characteristics approach, and it can also accommodate large CSD assortments since the number of products purchased on a trip is driven by λ whose dimension is invariant to the size of the assortment.[15]

3.3 Controlling for Household-Specific Behavior

I now discuss how the model controls for household-specific effects such as heterogeneity in tastes as well as dynamics in purchase behavior. In addition to their effect on the expected number of consumption occasions, λ^h (mean of the Poisson), household characteristics also affect the marginal utility of income, a^h, and the taste for quality, m^h,

$$\begin{aligned} \alpha^h &= D^{h\prime}\phi, \\ m^h &= 1 + D^{h\prime}\kappa. \end{aligned} \tag{8}$$

These parameters control for observed heterogeneity across households.[16] Unobserved heterogeneity across consumption occasions enters the perceived quality function, Ψ_{ij}^h in (4), in the form of random tastes for product attributes,

$$\beta_j^h = \tilde{\beta} + D^{h\prime}\mu + \Omega\nu_j^h, \tag{9}$$

where $\tilde{\beta}$ captures the component of tastes for attributes that is common to all households and consumption occasions. The $(k \times d)$ matrix of coefficients, μ, captures the interaction of demographics and tastes. The matrix Ω is a diagonal matrix whose elements are standard deviations and ν_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 $\tilde{\beta} + D^{h\prime}\mu$ and variance $\Omega\Omega'$.

In the long run, households' price sensitivity for CSDs is captured by their marginal utility of income, α^h, which is the relevant metric for assessing the impact of mergers on prices. However, previous research has found that a household's short-run price sensitivity may vary due to in-store promotional activity, inventories, and brand loyalty, each of which could bias the long-run price sensitivity if ignored. I control for in-store promotions by including weekly feature ad and display variables. To control for inventories, I include the time since last CSD purchase in the total number of expected consumption occasions,

15. Kim et al. (2002) also propose a model for households purchasing assortments. However, their approach would not handle the large number of CSD alternatives considered in the current context. Hausman et al. (1995) also propose a model for the special case in which the number of consumption occasions is observed in the data.

16. The intercept of m_h is normalized to 1 because it enters the model as an exponent.

λ_t^h. Modeling consumer stock piling, price expectations, and forward-looking consumer behavior is beyond the scope of this paper, but see Erdem (1996) and Erdem et al. (2003). To control for loyalty, I include two indicator variables in the perceived quality equation, Ψ_{ij}^h (Erdem, 1996; Keane, 1997). The first captures whether a given brand was purchased on the previous trip. The second indicates whether a given product was purchased on the previous trip (i.e., a specific package size of a brand).[17]

3.4 Measuring Consumer Welfare

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 for such a change in consumer welfare is the Hicksian compensating variation, which captures the amount by which consumer incomes must be compensated to equalize the pre- and postmerger levels of utility. I find Δy^h such that optimal true and counterfactual utilities are equal,

$$U^{h*}(\mathbf{p}^0, y^h) = U^{h*}(\mathbf{p}^1, y^h + \Delta y^h),$$

where $U^{h*}(\cdot,\cdot)$ denotes the maximal level of utility attainable by household h at the given prices, and $\mathbf{p}^0$ and $\mathbf{p}^1$ are the pre- and post-merger prices, respectively. Given the form of the utility function in (3), the compensating variation for each household shopping trip can be written as

$$\Delta y^h = \frac{U^{h*}(\mathbf{p}^0, y^h) - U^{h*}(\mathbf{p}^1, y^h)}{\alpha^h}.$$

By dividing through by the marginal utility of income, α^h, this expression yields a money-metric assessment of the welfare change.

3.5 Supply

The soft drink industry is best described as an oligopoly with multi-product firms. Previous empirical research in the CSD industry finds no evidence of collusive pricing (Gasmi et al., 1992). Rather, the evidence suggests that product differentiation generates market power in itself (Langan and Cotterill, 1994; Cotterill et al., 1996). I use a short-run model in which firms choose profit-maximizing prices in each quarter, conditional on the product attributes. Given that a single distributor

17. I find that results do not change substantively when I use a more sophisticated exponentially smoothed loyalty history variable (as in Guadagni and Little, 1983) instead of indicator variables. See http://gsbwww.uchicago.edu/fac/jean-pierre.dube/research/ for these results.

typically bottles and distributes the entire product line under a given brand, 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 et al., 1992), I do not model the distribution channel structure. Retail pricing decisions are treated as exogenous because retail margins are typically very low for CSDs.[18] The retailer's timing of advertising and display decisions are also taken as exogenous. The decision of which items to promote in the weekly newspaper flyer is a store-wide decision reflecting retail competition rather than an intra-category management decision.

Each of the F firms are assumed to produce some subset, B_f, of the $i = 1, \dots, I$ CSD products, making quantity and price decisions at a quarterly frequency based on expected demand. Each firm f sets prices to maximize its expected profits,

$$\pi_f = \sum_{i \in B_f} (p_i - c_i)\, Q_i(p) - C_f,$$

where p_i is the price firm f charges for product i, Q_i are the total sales of product i, c_i is the marginal cost of producing product[19] i and C_f are firm f's fixed production costs. 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,

$$Q_i(p) + \sum_{k \in B_f} (p_k - c_k) \frac{\partial Q_k(p)}{\partial p_i} = 0, \quad i \in B_f, \quad f = 1, \dots, F. \tag{10}$$

Define the $(J \times J)$ matrix $\mathbf{\Delta}$ with entries

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

Stacking the prices, marginal costs, and expected quantities into $(J \times 1)$ vectors, $\mathbf{Q}$, $\mathbf{p}$, and $\mathbf{c}$, respectively, the system of first-order conditions can be rewritten in matrix form as mark-ups:

$$\mathbf{p} - \mathbf{c} = \Delta^{-1}\mathbf{Q}. \tag{11}$$

18. In an AC Nielsen study of national supermarket chains, soft drinks are found to have margins very close to zero and 33% lower than the average margin across all product categories.

19. I do not take into account retailer costs in this model. Carbonated soft drinks is a "direct-store-delivery" product category, meaning the CSD bottler delivers and stocks the product for the retailer. Hence, the wholesale price is the primary source of marginal costs to the retailer for CSDs.

As is typical in the literature, I estimate these mark-ups directly from the estimated demand parameters, without using information on costs (Bresnahan, 1989). Because retail margins on CSDs are typically close to zero, I assume that manufacturer price is simply the quarterly average retail price for a product. This approach is consistent with previous merger analyses using aggregate scanner data averaged across weeks and stores to model manufacturer competition (Baker and Bresnahan, 1985; Hausman et al., 1995; Hausman, 1996; Nevo, 2000). I recover c by solving (11) above. I also assume manufacturers are myopic in that they do not account for the long-run effects of loyalty when setting their prices.

Given the complexity of evaluating comparative statics analytically, mergers are evaluated by simulating the postmerger prices numerically (as in Nevo, 2000). In other words, I solve the equation

$$\mathbf{p}^* = \mathbf{c} + \boldsymbol{\Delta}(\mathbf{p}^*)^{-1}\mathbf{Q}(\mathbf{p}^*) \tag{12}$$

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

4. Model Estimation

In this section, I briefly describe the method of simulated moments (MSM) estimation procedure used to search for the demand parameters. More technical details about the estimation of the model are provided in Dubé (2004). Using the expected purchase vector for a household trip, equation (7), I define the prediction error,

$$\varepsilon_{ht}\left(D_t^h, \Theta\right) = E\,Q_t^h\left(D_t^h, \Theta\right) - q_t^h, \tag{13}$$

where q_t^h is the vector of observed purchases of each product by household h at time t. Orthogonality conditions are of the form

$$g(\mathbf{D}, \Theta) = \frac{1}{T}\sum_{h=1}^{H}\sum_{t=1}^{T_h} D_t^h * \varepsilon_t^h\left(D_t^h, \Theta\right), \tag{14}$$

where T is the total sample size (total shopping trips) and $\mathbf{D}$ is the $(T \times d)$ matrix of all household characteristics. Evaluating (14) is complicated by the multivariate integral in the expected demand function, (7). I compute the integral using direct Monte Carlo simulation, to obtain the simulated moments $\tilde{g}(\mathbf{D}, \Theta)$. The parameters are then estimated using MSM (McFadden, 1989; Pakes and Pollard, 1989), that is by minimizing

$$J_{HT}(\Theta) = [\tilde{g}(\mathbf{D}, \Theta)]'\, W[\tilde{g}(\mathbf{D}, \Theta)], \tag{15}$$

where I specify W as the estimated asymptotic variance of g for efficiency (Hansen, 1982).[20] I use 30 simulation draws, which was sufficient to eliminate any "simulation noise." The MSM estimate, Θ_{MSM}, is consistent and has asymptotic variance $\Xi = (\frac{dg^s(\Theta_0)}{d\Theta}' W \frac{dg^s(\Theta_0)}{d\Theta})^{-1}$.

To control for potential price endogeneity, recall that product fixed effects, ξ_i, are included in the model. Including these fixed effects prevents separately estimating parameters capturing the mean tastes for product characteristics that do not vary over time. Similar to Nevo (2000), I recover these latter parameters using a minimum distance procedure that projects the estimated fixed effects, $\hat{\xi}_i$, onto the non-time-varying product attributes, using the estimated covariance matrix of the fixed effects as a weight matrix.

5. Data

The data used for demand estimation consist of individual household purchase histories in supermarkets, weekly store-level prices and promotional activity, household characteristics, and physical product characteristics. The household and store-level data were provided by AC Nielsen, covering the Denver area between January of 1993 and March of 1995. Product characteristics were collected from the nutritional information printed directly on the packaging.

In the current analysis, a product is defined as a UPC (Universal Product Code) to distinguish among different package sizes of a given brand. Different package sizes of a brand are treated as different products due to the significant differences in storability of, for instance, small aluminum cans and plastic bottles. This definition also distinguishes between diet versus regular (e.g., diet Coke is a different product than Coke Classic), and caffeine-free versus regular (e.g., Caffeine-Free Coke is different than Coke Classic). The definition of a brand differs from that of a product. A given brand, such as Coke Classic, is available in three different pack sizes: 12-pack of cans, 6-pack of cans, and a 2-liter bottle. In the CSD category, we also see brand extensions such as Diet Coke and Caffeine-free Diet Coke. Even though these bear the Coke name, they are priced and promoted differently.[21] Hence, for the analysis below, I treat these brand extensions as separate brands.

20. The weight matrix is corrected for serial dependence across weeks for each household, analogous to Newey and West (1987). That is, I assume the vector of prediction errors is correlated across time. This correction increases some of the standard errors considerably. No evidence of dependence across households was detected.

21. Discussions with Craig Stacey, director of marketing at Coca-Cola, revealed that Coke and Diet Coke are treated as separate brands targeted at different consumer segments. Historically, they have also been advertised and promoted separately. Nevertheless, Coca-Cola has noticed strong halo effects from Coke Classic advertising on Diet Coke, which has led them to reduce the amount of recent Diet Coke advertising.

The household panel consists of 2,108 households' shopping histories (including trips during which no CSDs were purchased) in the 58 largest supermarkets in the Denver Scantrac, each with over $2 million in annual "all commodity volume." Each household also has a corresponding set of reported demographic variables that are used to control for heterogeneity in tastes. For instance, AC Nielsen tracks the age and education level of the female head-of-household as previous research has found the characteristics of the female head-of-household highly correlated with household-specific preferences. In this study, I use an indicator for whether the female head-of-household is under 35 years of age. On average, households make 88.43 shopping trips within the sample period. In the previous section, I discussed the break-down of the typical shopping basket and the incidence of variety purchases as described in Table I. In addition, the average household purchases 3.3 different brands during the sample period and 5.1 different UPC-denoted products. The latter statistic indicates households frequently purchase different package sizes of the same brand, which is captured in the model with a brand loyalty measure (in addition to product-specific loyalty). The average household also purchases 2.2 different package sizes during the sample period. At the same time, conditional on purchasing at least two distinct products on a trip, an average shopping trip only consists of 1.2 different package sizes. These findings suggest that households tend to switch among products of the same package size. A related finding is documented in Guadagni and Little (1983), who observe households switching across products of the same package size over time.

For each of the supermarkets, the data contain the weekly prices, newspaper feature ad, and in-aisle display activity for each of the CSD products carried in the store. To simplify the analysis, only products with at least 1% of the aggregate sales volume share (in ounces) are included, yielding 26 diet and regular products with a combined share of 51% of the category. Below I discuss the sensitivity of my results to the scope of products included. Summary statistics of the Nielsen data appear in Table II.

The product characteristics that are assumed to influence perceived quality include total calories, total carbohydrates, sodium content (in mg), all of which and a set of dummy variables that indicate the presence of caffeine, phosphoric acid, citric acid, caramel color, and no color. I also define four other 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 complete set of products and their characteristics are reported in Table III.

TABLE II.
DESCRIPTIVE STATISTICS (AVERAGED ACROSS TRIPS)

Variables	Mean	Std. Dev.
Kids	0.3865	0.4870
Family size	2.6976	1.4034
Income bracket	4.2470	1.9616
Female head under 35 years	0.1964	0.3973
Time since last trip	6.8498	5.9508
Time since last purchase	39.7690	74.9517
Max. temperature (°F)	64.6149	19.8264
Holiday	0.1747	0.3797
Shelf price ($)	2.1515	0.3782
Feature ad	0.3203	0.0579
Display	0.4174	0.0503
Total units purchased (per trip)	2.359	2.0644
Total brands purchase (per trip)	1.4179	0.7359
Average number of distinct brands purchased per household	3.3484	2.185
Average number of distinct package sizes purchased per household	2.2365	1.0767
Total shopping trips	169,788	
Total households	1,920	

Additional covariates are included in the Ψ_{ij}^{h} equation to explain some of the variation in shopping behavior across time that is not captured by the marketing variables such as feature ads and in-aisle displays. Two variables, the maximum daily temperature (degrees Fahrenheit) and a holiday dummy variable indicating weeks with a national holiday control for seasonality in demand.[22] Loyalty variables constructed using the household shopping histories are also included in the Ψ equations. Finally, proxies for inventory accumulation are assumed to shift the number of anticipated consumption occasions for which consumers shop each period, λ_{ht}, in response to unobserved (to the researcher) stocks of CSDs. These proxies include both the time since the last CSD purchase (in days) and the time since the last shopping trip (in days).[23]

22. These variables were found to explain more of the sales variation than seasonal dummy variables.

23. I also experimented with an inventory measure that depreciates purchases according to an exogenous household-specific consumption rate. Consistent with findings reported in the marketing literature, the effect of this variable was insignificant.

Table III.
List of Products and their Characteristics (Ordered by Market Shares)

Product	Sodium	Carbs.	Phos.	Citric	Caramel	Clear
PEPSI CAN 12P	35	41	1	1	1	0
COKE CAN 12P	50	39	1	0	1	0
PEPSI CAN 6P	35	41	1	1	1	0
COKE DIET CAN 12P	40	0	1	1	1	0
PEPSI BOTTLE 67.6 oz	35	41	1	1	1	0
PEPSI DIET CAN 12P	35	0	1	1	1	0
COKE CAN 6P	50	39	1	0	1	0
PEPSI DIET CAN 6P	35	0	1	1	1	0
COKE BOTTLE 67.6 oz	50	39	1	0	1	0
PEPSI DIET BOTTLE 67.6 oz	35	0	1	1	1	0
COKE DIET CAN 6P	40	0	1	1	1	0
DR PEPPER CAN 12P	55	40	1	0	1	0
MOUNTAIN DEW CAN 12P	70	46	0	1	0	0
DR PEPPER CAN 6P	55	40	1	0	1	0
7 UP CAFF-FREE BOTTLE 67.6 oz	75	39	0	1	0	1
COKE DIET CAFF-FREE CAN 12P	40	0	1	1	1	0
COKE DIET BOTTLE 67.6 oz	40	0	1	1	1	0
7 UP DIET CAFF-FREE BOTTLE 67.6 oz	35	0	0	1	0	1
MOUNTAIN DEW CAN 6P	70	46	0	1	0	0
SPRITE CAFF-FREE CAN 12P	70	38	0	1	0	1
PEPSI DIET CAFF-FREE CAN 12P	35	0	1	1	1	0
DR PEPPER BOTTLE 67.6 oz	55	40	1	0	1	0
MOUNTAIN DEW BOTTLE 67.6 oz	70	46	0	1	0	0
PEPSI BOTTLE 16 oz	35	41	1	1	1	0
PEPSI DIET CAFF-FREE CAN 6P	35	0	1	1	1	0
A & W CAFF-FREE CAN 6P	45	46	0	0	1	0

6. Results

6.1 Parameter Estimates

Estimation results are shown in Table IV, starting with the coefficients of the quality function, Ψ, in the first three columns. The abbreviation "s.d." in a variable's name indicates this is a standard deviation for a random coefficient.[24]

The results imply that marketing variables such as feature ads and displays have a strong positive influence on purchasing behavior. However, there is substantial heterogeneity in these responses as indicated by the significant s.d. coefficients. Loyalty to the brand and to the specific

24. All standard errors have been corrected for potential serial dependence for up to 15 days, which nearly doubles several of the standard errors.

Table IV.
Results of Demand Estimation

Perceived Quality Ψ			Nonlinear Terms, $\lambda, \alpha, m, \gamma$		
Variables	Param	SE	Variables	Param	SE
Feature ad	0.796	0.025	λ: kids	0.0813	0.029
S.d. feature ad	0.058	0.014	λ: family size	0.0418	0.0158
Display	1.588	0.048	λ: time since last CSD	0.0003	0.0001
S.d. display	0.204	0.027	λ: time since last trip	−0.0047	0.0018
Prod. loyalty	0.062	0.013	λ: temperature	0.0063	0.002
Brand loyalty	0.015	0.220	λ: holiday	−0.0015	0.0024
Intercept	−0.277	0.133	α: constant	5.0315	0.1556
S.d. intercept	2.281	0.023	α: family size	0.6292	0.0334
Diet	−0.085	0.005	α: time since last CSD	0.0098	0.003
S.d. diet	0.278	0.019	α: time since last trip	0.0145	0.0049
Sodium	−0.036	0.001	m: income	2.4484	0.0952
Carbs	0.336	0.017	γ	0.0596	0.0021
Caffeine	1.961	0.047			
Phos.	−1.289	0.080			
Citric	0.070	0.027			
S.d. citric	0.240	0.031			
Caramel	1.539	0.030			
S.d. caramel	1.193	0.102			
Nocolor	2.202	0.062			
Cans ∗ 6	2.082	0.034			
S.d. cans ∗ 6	1.137	0.050			
Can ∗ 12	1.571	0.030			
S.d. cans ∗ 12	0.144	0.016			
Bottles ∗ 16	0.683	0.156			
S.d. bottles ∗ 16	2.049	0.224			
Kids ∗ caffeine	0.542	0.045			
(Family size) ∗ (pack size)	0.012	0.001			
(Female < 35) ∗ diet	0.329	0.033			
Hansen's *J* statistic			184.48		
No. observations (trips)			169,788		

product (UPC) chosen on the previous trip explain only a small portion of the perceived quality. The results suggest that loyalty to a specific brand is stronger than loyalty to a given UPC.

As for demographic characteristics, I find that, 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. Larger households place slightly more weight on products with more servings, such as the 12-pack of cans. Households with kids place a higher weight on products with caffeine than without. This finding could be due to the fact that the caffeine-free products, such as 7 UP and Sprite, tend to appeal more to adults.

The fourth to sixth columns of Table IV contain the estimated coefficients for the mean of the Poisson, λ^h, the marginal utility of income, α^h, the vertical dimension of taste, m^h, and the curvature of the utility function, γ. Beginning with λ^h, the expected number of decisions a household makes on a given trip depends primarily on the presence of kids and on family size. Temperature and holidays also increase the number of decisions. Proxies for inventory, time since last shopping trip and time since last CSD purchase, have small effects.

The marginal utility of income, α^h, also increases with the number of people in the household. Once again, the effects of time since last trip and time since last CSD purchase are very small. The vertical component increases with income, so that households with higher income perceive more distance between products. Finally, the estimated values of γ are positive and below 1, which is consistent with the notion that utility is concave.[25]

6.2 Price Elasticities, Margins, and Marginal costs

Table V reports the aggregate price elasticities, marginal costs, and price-cost margins (PCM), $\frac{p-c}{p}$, corresponding to the aggregated quarterly consumer demand estimates. Marginal costs and PCMs are computed by assuming Nash equilibrium in manufacturer prices each quarter. The table reports the median level and standard error across the nine quarters in the sample.

Pepsi clearly has a pricing advantage over its rivals, setting margins of about 50–60%. In comparison, Coke's margins are around 40%. This fact is not surprising since Pepsi is the market leader in the Denver market. However, the relative pricing advantage of Pepsi in this market will likely have a downward bias on the welfare implications of the Coke and Dr. Pepper mergers. 7 UP margins are also close to 40%, whereas Dr. Pepper's are closer to 30%. The marginal cost of a 67.6 ounce bottle is markedly lower than that of a 6-pack of cans, possibly reflecting the relative cost advantage of a plastic bottle versus an aluminum can. The fact that 12-packs of cans have slightly higher marginal costs than 6-packs may reflect the more sophisticated cardboard box used to bundle 12 cans, as opposed to the simple plastic ring used to bind the 6-packs.

25. I also calculate Hansen's statistic to test our set of over-identifying restrictions. I obtain a statistic of 184.48. Since we have 54 parameters and 160 moment conditions, this statistic is asymptotically distributed χ^2 with 106 degrees of freedom. The critical value is roughly $\chi^2_{.05}(106) = 124$, and the model is rejected. In practice, this test is routinely rejected with large data sets and, hence, in our context it could be inconclusive.

Table V.
Own-Price Elasticities, Predicted Mark-Ups and Marginal Costs

Product	Price ($/12 oz) Median	Own-Price Elasticity Median	Own-Price Elasticity SE	MC ($/12 oz) Median	MC ($/12 oz) SE	PCM (%) Median	PCM (%) SE
PEPSI 12P	0.30	−3.36	0.15	0.14	0.03	51.5	1.66
PEPSI 6P	0.27	−3.07	0.17	0.11	0.02	55.7	1.37
PEPSI 67.6oz	0.18	−3.26	0.09	0.07	0.04	60.4	2.75
PEPSI DT 12P	0.30	−3.06	0.29	0.11	0.03	63.9	2.15
PEPSI DT 6P	0.26	−3.99	0.18	0.13	0.01	50.0	1.47
PEPSI DT 67.6oz	0.18	−3.99	0.28	0.08	0.05	58.1	3.31
PEPSI DT CF 12P	0.30	−5.00	0.62	0.15	0.04	47.5	2.77
PEPSI DT CF 6P	0.26	−4.98	0.27	0.17	0.03	37.5	1.64
PEPSI 16oz	0.33	−3.90	0.56	0.15	0.03	50.1	3.03
MT DW 12P	0.30	−4.37	0.33	0.19	0.03	37.4	3.11
MT DW 6P	0.27	−4.55	0.29	0.12	0.03	55.5	1.88
MT DW 67.6oz	0.18	−3.79	0.19	0.09	0.04	50.3	2.30
COKE CLS 12P	0.29	−3.64	0.18	0.17	0.06	43.3	3.66
COKE CLS 6P	0.27	−4.05	0.25	0.18	0.02	33.9	1.27
COKE CLS 67.6oz	0.19	−3.98	0.21	0.11	0.01	42.4	1.23
COKE DT 12P	0.29	−3.67	0.21	0.15	0.15	47.5	10.30
COKE DT 6P	0.27	−4.37	0.22	0.17	0.01	34.7	1.27
COKE DT CF 12P	0.29	−5.62	1.17	0.19	0.01	32.2	1.21
COKE DT 67.6oz	0.19	−4.04	0.18	0.11	0.06	42.3	4.48
SP CF 12P	0.29	−3.86	0.23	0.17	0.04	44.6	2.35
7 UP R CF 67.6oz	0.18	−4.39	0.19	0.13	0.12	30.0	7.84
7 UP DT CF 67.6oz	0.18	−3.41	0.13	0.11	0.01	39.1	1.11
DR PR 12P	0.31	−4.06	0.34	0.23	0.02	28.0	2.44
DR PR 6P	0.28	−6.03	0.39	0.21	0.14	19.9	8.81
DR PR 67.6oz	0.19	−3.89	0.23	0.13	0.11	30.3	8.54
A and W CF 6P	0.28	−3.85	0.31	0.20	0.04	31.4	2.16

Standard errors were computed using a parametric bootstrap. 150 draws were taken from the asymptotic distribution of the parameters and used to compute a sample of MC, PCM and elasticities.

I find that Dr. Pepper has noticeably higher marginal costs than other products. This result may reflect the differences in distribution costs. For instance, 7 UP 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 the colas.

In Table VI, I report own and cross-price elasticities for a sample of products (medians across the nine quarters).[26] Interestingly, the elasticities demonstrate that consumers tend to substitute primarily between products of the same size in response to price changes. For

26. The complete set of elasticities is available in a separate appendix.

Table VI.
A Sample of Aggregate Own and Cross-Price Elasticities (Medians Across the 9 Quarters)

Product	PEPSI 12P	COKE 12P	MT DW 12P	DR PR 6P	7 UP 67.6oz	PEPSI DT 12P	COKE DT 12P	PEPSI 6P	COKE 67.6oz
PEPSI 12P	−3.10	0.39	0.77	0.18	0.31	0.48	0.41	0.13	0.46
PEPSI 6P	0.15	0.12	0.15	0.80	0.42	0.05	0.16	−3.25	0.36
PEPSI 67.6oz	0.21	0.14	0.21	0.24	0.44	0.17	0.13	0.15	0.53
PEPSI DT 12P	0.53	0.30	0.45	0.20	0.24	−3.43	0.21	0.07	0.11
PEPSI DT 6P	0.08	0.15	0.01	0.61	0.15	0.07	0.15	0.55	0.07
PEPSI DT 67.6oz	0.06	0.05	0.05	0.05	0.24	0.11	0.05	0.05	0.20
PEPSI DT CF 12P	0.03	0.03	0.00	0.00	0.03	0.10	0.01	0.00	0.06
PEPSI DT CF 6P	0.02	0.01	0.04	0.63	0.02	0.02	0.02	0.02	0.04
PEPSI 16oz	0.07	0.05	0.12	0.00	0.16	0.04	0.11	0.08	0.11
MT DW 12P	0.08	0.14	−4.42	0.00	0.05	0.08	0.01	0.02	0.04
MT DW 6P	0.00	0.03	0.07	0.08	0.11	0.00	0.12	0.26	0.01
MT DW 67.6oz	0.05	0.04	0.08	0.12	0.17	0.02	0.03	0.04	0.12
COKE 12P	0.34	−3.52	0.50	0.24	0.25	0.45	1.15	0.16	0.18
COKE 6P	0.12	0.11	0.23	0.86	0.20	0.20	0.08	0.46	0.16
COKE 67.6oz	0.06	0.06	0.03	0.22	0.30	0.08	0.21	0.09	−3.89
COKE DT 12P	0.11	0.55	0.21	0.10	0.16	0.29	−3.96	0.03	0.36
COKE DT 6P	0.11	0.03	0.02	0.27	0.11	0.05	0.05	0.28	0.10
COKE DT CF 12P	0.02	0.06	0.07	0.06	0.02	0.08	0.08	0.00	0.03
COKE DT 67.6oz	0.05	0.11	0.04	0.02	0.27	0.08	0.05	0.05	0.81
SP CF 12P	0.07	0.17	0.18	0.09	0.11	0.07	0.09	0.06	0.05
7 UP CF 67.6oz	0.08	0.06	0.10	0.01	−4.25	0.08	0.07	0.07	0.12
7 UP DT CF 67.6oz	0.05	0.08	0.17	0.12	0.95	0.01	0.08	0.05	0.15
DR PR 12P	0.26	0.29	0.17	0.11	0.20	0.32	0.10	0.04	0.09
DR PR 6P	0.01	0.01	0.02	−5.76	0.02	0.01	0.04	0.09	0.09
DR PR 67.6oz	0.06	0.05	0.06	0.09	0.16	0.04	0.04	0.10	0.26
A & W CF 6P	0.05	0.02	0.02	0.23	0.07	0.04	0.04	0.13	0.03

instance, demand for 6-packs of Diet Pepsi are quite sensitive to the price of 6-packs of Pepsi, and demand for 67.6 oz bottles of 7 UP are sensitive to the price of 67.6 oz bottles of Diet 7 UP. This finding is consistent with our discussion of typical shopping baskets, in the data section above. In addition, Coke and Pepsi are clearly the primary substitutes of almost every brand, but the reverse is not true. Although not reported, I find that a random coefficients logit model estimated using household data provides substantially lower own and cross-price elasticities. The logit demand system is a more traditional approach to analyzing mergers that ignores multiple discreteness.

In a separate appendix, I conduct several sensitivity checks of the results to alternative controls for consumer heterogeneity. The findings indicate that fewer controls for unobserved heterogeneity in tastes leads to more inelastic demand estimates. I also find that the use of alternative specifications to control for state dependence in consumer tastes (i.e., loyalty) does not seem to have much impact on the price elasticities of demand. I also find my results to be fairly robust to the scope of products used. Increasing the product set to include the top 30 UPCs does not change the elasticity estimates, while decreasing the set to the top 20 does lead to more inelastic demand estimates.[27]

6.3 Mergers

In this section, I use my results to simulate the equilibrium prices and quantities for the proposed 1986 mergers between Coke and Dr. Pepper, and Pepsi and 7 UP as well as the hypothetical merger between Coke and Pepsi.[28] In evaluating the 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. These sunk costs consist of advertising outlays for launching new brands as well as slotting fees to retailers to obtain shelf space even for new package sizes of existing brands (Israilevich, 2003). In addition, I assume that a merger only affects the pricing decisions of firms (i.e., internalizes pricing of both merged firms' product lines) and, hence, consumer demand is only affected insofar as consumers face different prices. Hence, the definition of loyalty and other purchase-related variables do not change post-merger.

Table VII reports the median predicted quarterly price changes, in percent, for each merger across the nine quarters in the sample.[29] The first column reports the changes in prices associated with breaking apart Dr. Pepper and 7 UP, to replicate the 1986 market structure. This break-up appears to have had 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 VIII reports the corresponding changes in producer and consumer surplus for each merger. Producer surplus corresponds to variable profits and consumer surplus corresponds to the Hicksian compensating variation. The sample predictions are aggregated to the market level using a projection factor reported by AC Nielsen. Hence,

27. The results from these analyses and the appendix can be found on my website at: http://gsbwww.uchicago.edu/fac/jean-pierre.dube/research/.

28. Although the relevant set of brands has not changed since the time of the trial, in 1989, Dr. Pepper and 7 UP were both acquired by Hicks and Haas.

29. The approximation of Hausman et al. (1994), holding elasticities constant, tends to overstate the price increases compared to those computed numerically.

Table VII.
Median Simulated Percent Change in Quarterly Price from Mergers

Product	1986	Coke/Dr. Pepper	Pepsi/7 UP	Coke/Pepsi
PEPSI 12P	0.03	0.10	0.82	12.25
PEPSI 6P	0.01	0.11	0.55	19.19
PEPSI 67.6oz	−0.11	0.32	2.18	13.10
PEPSI DT 12P	0.09	−0.05	0.99	23.75
PEPSI DT 6P	0.04	−0.27	0.22	14.28
PEPSI DT 67.6oz	−0.04	0.19	1.53	16.84
PEPSI DT CF 12P	0.83	−1.12	−1.66	21.01
PEPSI DT CF 6P	1.27	−1.41	−1.37	19.25
PEPSI 16oz	0.06	−0.40	0.19	12.32
MT DW 12P	0.24	−0.90	0.08	18.49
MT DW 6P	0.33	−0.57	0.17	14.81
MT DW 67.6oz	0.03	−0.16	2.16	12.67
COKE CLS 12P	0.04	1.03	0.05	16.00
COKE CLS 6P	0.04	1.65	−0.14	20.95
COKE CLS 67.6oz	−0.09	2.16	0.46	24.94
COKE DT 12P	0.14	1.06	−0.05	21.42
COKE DT 6P	0.08	0.62	−0.48	21.99
COKE DT CF 12P	1.25	0.24	−1.93	15.33
COKE DT 67.6oz	0.09	0.68	0.12	20.93
SP CF 12P	0.24	−0.74	−4.89	7.82
7 UP R CF 67.6oz	−1.70	0.99	16.56	10.47
7 UP DT CF 67.6oz	−1.54	−0.12	14.15	11.44
DR PR 12P	−0.78	5.75	−0.30	7.07
DR PR 6P	−0.84	3.60	0.34	5.88
DR PR 67.6oz	−1.10	4.64	−0.02	9.39
A and W CF 6P	−1.26	−1.14	0.49	9.20

the merger simulations should be interpreted as applying to the entire Denver scantrac.

The first merger, between Coke and Dr. Pepper, did not seem to have a large effect on prices. Coke prices never rise by more than about 2% and the prices of Dr. Pepper increase by between 4% and 6%. Despite the seemingly moderate price increases, the merger had substantial implications for welfare. Interestingly, the rise in prices appeared to have a large benefit for Pepsi and 7 UP, whose quarterly variable profits rose by \$261,000 and \$9,000, respectively (about 3% in both cases). The total joint quarterly profitability of Coke and Dr. Pepper increased only slightly, just under 1% on average. The increased profitability of Coke products was generated by the less profitable pricing of Dr. Pepper products. The reduced profitability of Dr. Pepper products could be understated because the analysis does not take into account the potential cost advantages from leveraging Coca-Cola distribution.

TABLE VIII.
MEDIAN CHANGE IN QUARTERLY WELFARE (THOUSANDS OF DOLLARS PER QUARTER)

	Coke/Dr. Pepper	Pepsi/7 UP	Coke/Pepsi
Pepsi	261.41	168.62	567.02
Coke	60.65	203.01	133.17
Dr. Pepper	−41.88	36.31	179.21
7 UP	9.28	−44.39	100.35
Consumers	−1291.42	−1901.36	−29707.28
Total	−998.24	−1533.51	−28727.53

Quarterly consumer surplus fell, on average, by $1.2 million. Despite the gains in profits, aggregate welfare in the Denver economy fell overall by almost $1 million. Aggregating up to the entire US economy, (approximated by the 50 largest Nielsen scantracs), implied welfare losses of at least $50 million per quarter. In addition, because Denver is a mid-sized Nielsen scantrac and because Denver is a Pepsi-dominated market, the total US welfare losses could be even larger. These substantial losses warranted the efforts of the FTC, not only to block the merger in 1986, but to limit Coca-Cola's ability to acquire large competitors for the decade after 1994.

For the merger between Pepsi and 7 UP, cola prices did not rise by much more than 2%. However, the price of 7 UP rises between 14% and 16%. The joint profits of 7 UP and Pepsi rose by about 1.5%, mainly because increasing 7 UP product prices improved the profitability of Pepsi products. Profits also rise at both Coke and Dr. Pepper by 6% and 9%, respectively. At the same time, consumer surplus fell by almost $2 million. Overall, the merger led to roughly $1.5 million in lost surplus for the Denver market in each quarter. These much more dramatic price increases and welfare losses, at the expense of consumer well-being, likely explain why Pepsi did not fight the FTC's decision to contest the merger.

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 increased by well over 10%. In fact, many rose by more than 20%, especially the diet colas. This drastic reduction in cola competition had the indirect effect of allowing 7 UP, Dr. Pepper, and A & W Rootbeer each to raise their prices substantially. The large increase in prices resulted in about a $700,000 average increase in the quarterly joint variable profits of Coke and Pepsi, roughly an 8% improvement. These gains came at a huge cost to consumers, who lost an average of $30 million in quarterly consumer surplus. Overall, the Denver economy suffered an almost $29 million loss in aggregate surplus, on average.

This analysis is limited by the static nature of the model of producers. Hence, the analysis does not consider the possibility that postmerger profit gains could stimulate entry of new firms into the market. Moreover, the analysis does not account for potential efficiency gains from joint production of merged firms. During the 1986 trial, Coke argued that Dr. Pepper would benefit from increased production efficiency and scale economies in distribution, both of which would lower production costs substantially. However, little evidence was provided to quantify these efficiency gains. In the case of Pepsi and 7 UP, 7 UP already piggybacks off Pepsi's bottling and distribution network in most markets. Thus, any potential efficiency gains for 7 UP would need to reflect the production of concentrated syrup.

7. Conclusions

Recent advances in the collection of checkout scanner data have helped bring newer and more sophisticated structural econometric tools into merger analysis. Existing methods have relied on simplifying assumptions, such as discrete choice purchase behavior, to help build parsimonious demand systems for these analyses. However, in industries such as CSDs, the single unit purchase assumption is inappropriate. I resolve this problem by using disaggregate household-level point-of-purchase data to estimate demand. These microdata provides information on the variation in assortments purchased across households and shopping trips. The Hendel (1999) specification is used to model demand for these assortments. These demand estimates are then aggregated to simulate the impact of several hypothetical CSD mergers. Lacking comparable data and modeling techniques, the FTC was unable to conduct these types of merger simulations to evaluate the proposed acquisition of Dr. Pepper by Coca-Cola Co back in 1986. Despite the moderate impact on prices of the merger between Coca-Cola Co. and Dr. Pepper, I find evidence of fairly large welfare losses. The price implications and welfare losses are even more substantial for the merger of PepsiCo. and 7 UP and, especially, for the hypothetical merger of Coca-Cola Co. and PepsiCo.

The current work has some limitations. The treatment of unobserved heterogeneity pertains to the randomness of tastes across the latent consumption occasions. The random coefficients do not capture persistent household-specific unobserved tastes. To capture persistent sources of household-specific heterogeneity, I use a rich set of demographic information, which I interact with several aspects of the model. Nonetheless, I find some evidence of serial dependence in the unexplained portion of the model. Although this dependence could simply reflect sampling problems, it could also come from unmeasured heterogeneity. Unmeasured heterogeneity could bias the estimates of loyalty. To check this problem, I reestimate the model without the loyalty parameters and find little effect on the estimated elasticities, increasing my confidence that the current results are not biased. I also find my results to be robust across several different specifications of loyalty.

The current work is also based entirely on static modeling assumptions. Potential sources for future research include studying various forms of consumer dynamics insofar as they influence assortment decisions. For instance, expectations of short-run price fluctuations, such as future sales, could cause consumers to defer purchases until the time of the sale. One might expect consumers to "stock-up" during sale periods, purchasing large assortments of items while they are perceived to be a good deal (Erdem et al., 2003). Alternatively, as consumers become experienced with certain goods, they may begin to vary their purchase assortment to seek variety. On the supply side, the study of equilibrium pricing could be extended to include the timing of short-run price cuts and, subsequently, the impact of mergers on these types of consumer-driven promotions.

References

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.

Bell, D.R., J. Chiang, and V. Padmanabhan, 1999, "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18, 504–526.

Berry, S., 1994, "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25, 242–262.

—— and A. Pakes, 1993, "Some Applications and Limitations of Recent Advances in Empirical Industrial Organization: Merger Analysis," *American Economic Review*, 83, 247–252.

Berto Villas–Boas, S., 2001, "Vertical Contracts Between Manufacturers and Retailers: Inference with Limited Data," Working Paper, University of California, Berkeley.

Bresnahan, T.F., 1989, "Empirical Methods for Industries with Market Power," in R. Schmalensee and R. Willig, eds. *Handbook of Industrial Organization*, Amsterdam, North Holland: Elsevier Science.

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

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.

Dubé, J.P., 2004, "Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks," *Marketing Science*, 23(1).

Erdem, T., 1996, "A Dynamic Analysis of Market Structure Based on Panel Data," *Marketing Science*, 15, 359–378.

——, S. Imai, and M.P. Keane, 2003, "A Model of Consumer Brand and Quantity Choice Dynamics Under Price Uncertainty," *Quantitative Marketing and Economics*, 1(1), 5–64.

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.

Guadagni, P.M. and J.D.C. Little, 1983, "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2, 203–238.

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

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

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.

——, ——, and J.D. Zona, 1994, "Competitive Analysis with Differenciated Products," *Annales d'Economie et de Statistique*, 34, 159–180.

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

Israilevich, G., 2004, "Assessing Supermarket Product-Line Decisions: The Impact of Slotting Fees," *Quantitative Marketing and Economics*, 2, 141–167.

Keane, M.P., 1997, "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *Journal of Business and Economic Statistics*, 15, 310–327.

Kim, J., G.M. Allenby, and P.E. Rossi, 2002, "Modeling Consumer Demand for Variety," *Marketing Science*, 21, 229–250.

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.

McFadden, D., 1989, "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995–1026.

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

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

Nevo, A., 2000, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *The Rand Journal of Economics*, 31, 395–421.

Pakes, A. and D. Pollard, 1989, "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057.

Shum, M., 2000, "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast Cereals Market," Johns Hopkins University, Working Paper.

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.

Villas-Boas, M. and R. Winer, 1999, "Endogeneity in Brand Choice Models," *Management Science*, 45(10), 1324–1338.

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

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, 407–426.

White, L.J., 1989, "The Proposed Merger of Coca-Cola and Dr Pepper," in J.E. Kwoka and L.J. White, eds., *The Antitrust Revolution*, Glenview, Ill: Scott, Foresman, pp 76–95.