Different Flavor, Same Price: The Puzzle of Uniform Pricing for Differentiated Products*

Rob McMillan†

April 2005

Abstract

Retailers typically sell many different products from the same manufacturer at the same price. I consider retailer-based explanations for this uniform pricing puzzle, using a structural model to estimate the counterfactual profits that would be lost by a retailer switching from a non-uniform to a uniform pricing regime in the carbonated soft drink category. Applying the demand model developed in McMillan (2005) to household-level panel data on purchases of carbonated soft drinks, I estimate that the retail store I observe earned an additional $36.56 (1992 dollars) in average weekly profits by charging non-uniform prices. This corresponds to roughly a 3% difference in profits, and suggests that when a retail store faces even a relatively small cost to determine the optimal set of non-uniform prices, it may be optimal to charge the same price for many products.

*I thank Peter Reiss, Frank Wolak, Pat Bajari, Michaela Draganska, Liran Einav, Cristobal Huneeus, Navin Kartik, Tom MaCurdy, Mikko Packalen, James Pearce, and many others for helpful discussions.

†Federal Trade Commission, Bureau of Economics. The views expressed herein are not purported to reflect those of the Federal Trade Commission, nor any of its Commissioners. Email: `rmcmillan@ftc.gov`.

1 Introduction

Retailers typically sell many different products from the same manufacturer at the same price. For example, in the yogurt category, all flavors of six ounce Dannon Fruit-on-the-Bottom yogurt are sold at one price, while all flavors of six ounce Yoplait Original yogurt are sold at a second (uniform) price. This practice is common in many product categories, including frozen dinners, ice cream and salsa. In other product categories, however (e.g., frozen juice), items are typically sold at different prices within manufacturer brands. While not true at all times, in all stores, and for all products, the extent of these uniform prices across different retailers, product categories, and time is stunning. Why is it optimal for the retailer to sell many different items at the same price? Consider the following anecdotal examples:

Tea and Juice Although many teas and juices are sold at uniform prices, there are notable exceptions. Frozen orange juice is almost always priced differently from other frozen juice. Within the premium juice category (e.g., brands such as Odwalla and Naked Juice), prices are frequently completely non-uniform. Similarly, although most teas are sold at uniform prices, some varieties of tea are frequently sold at a higher price. These non-uniform prices seem to correlate with marked differences in marginal costs. Although most tea leaves cost roughly the same, some cost more to produce. Similarly, differences in cost and juiciness across different fruits can lead to different marginal costs for the same volume of liquid. Furthermore, both tea and juice are products that are difficult for manufacturers to adjust the amount of input per unit of output. If manufacturers adjust the amount of real juice or tea leaves, consumers are likely to notice.

Wine Different varieties of wine from the same vineyard and vintage are typically sold at the same price when they are sold for less than $15 a bottle. For example, the much-reviewed Charles Shaw wines are $1.99, regardless of variety (e.g., Merlot, Cabernet Sauvignon, Shiraz, Gamay Beaujolais, Chardonnay, or Sauvignon Blanc). However, for more expensive wine, different varieties are generally priced non-uniformly. As prices increase, we see more and larger deviations from uniform prices.

Clothing Within a particular style, clothing is typically sold at the same price for different colors and sizes. There are however, exceptions to this rule: while S,M,L and XL sizes are typically the same price, many retailers charge more for XXXL and "tall" sizes. These sizes generally cost the retailer more, either because of the amount of fabric used, or because average costs are higher due to lower volumes. Also, although men's shirts are usually priced uniformly across colors, striped shirts are frequently priced differently. Retailers frequently claim that this is because striped shirts are a different style than solid colored shirts, with different demand. Finally, upscale clothing stores are more likely to charge different prices for different colors.

Books Books, even harlequin romance novels, are not sold at uniform prices. Even different books by a single author generally have different prices. At first this seems puzzling. But unlike many products, books are frequently sold with a suggested retail price stamped on their cover or dust jacket. Furthermore, this price is the same everywhere. While many retailers offer lower prices (e.g., discounts for New York Times Bestsellers), these are nearly always offered as a percentage difference from this suggested price. If one is willing to take as given the constraint that most books are sold at a single price (or at most 2-3 prices) nationally, it becomes clear that the demand for different books is almost certainly quite different.

While many products exhibit uniform prices, there seem to be clear patterns characterizing products that are not priced uniformly. Products with ostensibly different marginal costs, such as different flavors of tea, varieties of frozen juice, and odd sizes of clothing are frequently sold at different prices. Other products, with demand that clearly differs across varieties such as different colors of designer clothes, colors of cars, expensive varieties of wine also tend to be priced differently.

A suggestive pattern emerges: unless there are *clearly* additional profits from non-uniform pricing, there is a strong tendency towards uniform prices. If marginal costs are sufficiently different, we tend to see non-uniform prices. If demand for the products is sufficiently different, we tend to see non-uniform prices. If prices are sufficiently high, we tend to see non-uniform prices. This suggests that managerial menu costs on the part of the retailer may be able to explain the observed uniform pricing behavior. In determining what price to charge, the retailer incurs a cost, most obviously the opportunity cost of a price-setting executive's time. Given this cost, it may be optimal for the retailer to group products with similar costs and demands and sell them at a single price.

Given the pervasiveness of uniform pricing in retail environments, the dearth of papers on the subject is quite surprising. To my knowledge, only Orbach & Einav (2001) confront the uniform pricing puzzle directly (and then only in the movie industry). They observe that tickets for movies that are a priori expected to be blockbusters are sold at the same price as movies that are a priori expected to be box office bombs. Unfortunately, they are hamstrung by the fact that they never observe non-uniform pricing for movies, and are unable to find any convincing explanation.

The answer to the uniform pricing puzzle must revolve around the key question: How do retailers set prices? Menu costs are not the only potential explanation. Indeed, in addition to the menu cost explanation, there are a variety of explanations, which fall into two groups: demand-side (consumer-based) and supply-side (retailer-based) explanations.

Potential demand-side explanations involve explicit consumer preferences for uniform prices. Based on discussions with price-setters, Kashyap (1995) and Canetti, Blinder & Lebow (1998) find that many firms believe they face a kinked demand curve, containing so-called "price points" where marginal revenue is discontinuous. Two consumer preferences that might yield a "uniform price" price point are that more prices makes it harder to figure out what to buy, and that more prices make the consumer feel that the retailer is trying to take advantage of her. Shugan (1980) and Hauser & Wernerfelt (1990) develop theoretical models of costly optimization and Draganska & Jain (2001) find some evidence of this in yogurt. Kahneman, Knetsch & Thaler (1986) look at consumers' perceptions of fairness, and show that consumers may perceive unfairness in retailers pricing policies. Evidence that these perceptions of fairness can affect demand can be found in recent popular press surrounding actions by Coke (Hays 1997) and Amazon (Heun 2001).

In addition to the "null" hypothesis that the observed prices are actually optimal in a traditional supply and demand framework, the two principal supply-side explanations are: (1) that the observed uniform pricing behavior is either driven by menu costs or (2) that it stems from an attempt by retailers to soften price competition.

The former explanation is favored by Ball & Mankiw (2004), who use it to explain "sticky" prices. The puzzle of uniform pricing across differentiated products is closely related to the long-standing macro-economic issue of sticky prices. Sticky prices can be thought of as uniform prices for products that are differentiated by time – an example of *inter-temporal* price uniformity. Addressing this issue, Ball & Mankiw (2004) suggest that much of the observed inter-temporal price uniformity can be explained by the menu costs associated with “the time and attention required of managers to gather the relevant information and make... decisions.”(p.24-25).

Leslie (2004) also believes that menu costs play a role in pricing decisions. He finds that Broadway theaters can earn higher profits by charging prices that differ across seats. While he finds that charging different prices for different seating categories results in higher profits, he cannot explain why theaters use only two or three different categories. Because he does not observe seat-level price variation he cannot estimate the implied menu costs.

Although they do not describe it as such, Chintagunta, Dubé & Singh (2003) provide a measurement of implied menu costs. They find that multi-store retail chains can earn higher profits by charging different prices in different geographic areas. They predict that if retailers charged a different price in each store (rather than using only three or four different menus of prices for over 80 stores), the chain would have earned an additional $10,000 per week in the orange juice category alone! Such a large result seems unreasonable in light of the salary levels found in Levy, Bergen, Dutta & Venable (1997), and may be due to the potential demand-side explanations discussed above or the restrictive assumptions that they make about the nature of competition between retailers in order to identify demand.

Taking a different tack and inferring managerial menu costs from salary data, Levy et al. (1997) guesstimate that the annual price-setting managerial costs are $2.3-$2.9 million at the chain level, which translate to average annual per-store costs of roughly $7000.[1] These are average, not marginal costs, however.

As mentioned above, an alternative supply-side explanation for uniform pricing is that it leads to a softening of price competition. There is an extensive literature investigating the effects of multi-market interaction on firms' abilities to collude (for example Nevo (2001) looks at the case of breakfast cereal, while Carlton (1989) suggests this as an explanation for inter-temporal price uniformity). Most relevant to my analysis is Corts (1998) who shows that firms engaged in multi-market competition may prefer to commit themselves to charging the same price in both markets in order to soften price competition. Corts models the interaction between two firms which compete in a Bertrand setting in two markets. In his model, the two firms have identical costs, but these costs differ across markets. If they are able to charge different prices in each market, the firms drive the price down to marginal cost. However, if they are able to restrict themselves to charge the same price in both markets, then they are able to earn positive profits in expectation. Viewing two

[1] These figures are in 1992 dollars. One 1992 dollar is equivalent to $1.35 2004 dollars.

different flavors as different markets, his result suggests that retailers may benefit if they are able to tacitly agree to charge fewer prices than the number of distinct products that they sell. Alternatively, charging fewer prices may allow colluding firms to more easily detect cheating. However, both of these theories of collusion require retaliation by the cartel in the face of detected cheating. This does not coincide with what I see in my data: I see repeated non-uniform pricing by one retailer that is not met by non-uniform pricing by any other retailers in my sample. As we will see in more detail later (see Figures 1-4), I do not see the stores in my data retaliating through the use of non-uniform prices in the face of deviations.

Finally, there has also been a growing body of literature exploring the implications of line length – the dual of the uniform pricing puzzle (See, for example, Draganska & Jain (2001), Bayus & Putsis (1999), and Kadiyali, Vilcassim & Chintagunta (1999)). These papers model the retailer's decision to add products to a "line". This literature has taken for granted that all the products in a line have the same price. Indeed, many authors have defined a product line as the set of products from a single manufacturer sold by the retailer at a uniform price. Rather than examine the pricing decision, this literature has focused exclusively on the decision of whether to introduce additional products. Clearly in addition to facing the problem of maximizing product line length, manufacturers face the decision of when to price products differently, that is, when to *split* a line. To my knowledge, this question has not been directly addressed by the line length literature.

Although many of these explanations for uniform pricing are plausible (and may be the cause for some cases of uniform pricing), this paper examines the plausibility of the managerial menu cost explanation, which is essentially a story of bounded rationality on the part of the retailer. There are several reasons for this focus. The first reason for this focus is that the menu cost explanation uses standard assumptions concerning consumer choice behavior, and at a minimum will provide a useful benchmark for comparison when considering other explanations. Second, it is certainly plausible that for many goods, menu costs may be able to explain uniform pricing. Third, while there is reason to believe that demand-side explanations may play a role, I am limited by a lack of data. Exploring demand-side explanations in more detail cannot be done with the data presently available.[2] By contrast, if I assume that consumers do not have explicit preferences for uniform prices,

[2] Doing so would, at a minimum, require exogenous switching between uniform and non-uniform pricing strategies. In addition, Kahneman et al. (1986) suggest that consumer responses to alternative price behavior are heavily dependent on framing (e.g., explanations for alternative pricing behavior).

it is possible to examine the implications of uniform pricing for retailers.

This paper estimates the economic profits that retailers appear to lose by following optimal uniform pricing strategies. These lost profits place bounds on the retailer's costs of implementing these alternative pricing regimes. In the next section, I describe a framework for estimating the menu costs to the retailer from following non-uniform pricing strategies, and identify soft drinks as a product category for investigation. The remainder of the paper proceeds as follows: section 3 lays out the empirical demand model, and section 4 describes the data. Section 5 contains both the parameter estimates of the demand model and the results of the coutnerfactual experiment. Section 6 interprets these results in light of other available evidence on menu costs and section 7 concludes.

2 An Overview of Managerial Menu Costs for Soft Drinks

Assuming that uniform pricing is driven entirely by supply-side factors, then clearly the first-order question is: "How much profit is forgone?" To answer this question, I assume that retailers are able to choose among a variety of different pricing strategies of differing sophistication. These strategies are functions that map each period's state space to a vector of prices. Consider the array of potential pricing strategies that a retailer of carbonated soft drinks might choose from:

- Charge a constant percentage markup of 30% over wholesale price.
- Charge a markup of a constant amount of $0.25 over wholesale price.
- Charge a constant percentage markup of x% over wholesale price, where x is chosen optimally.
- Charge a single per unit price (e.g., $0.25 per 12 ounces) for all soft drinks, regardless of size or flavor.
- Charge a single per unit price for all soft drinks from the same manufacturer, regardless of size, but charge different prices across manufacturers.
- Charge a single per unit price for all soft drinks of the same size, regardless of manufacturer, but charge different prices across sizes.

- Charge a different price for each product (e.g., one price for a 2-Liter bottle of Diet Coke, a second price for a 2-Liter bottle of Coke Classic, a third price for a 2-Liter bottle of Diet Pepsi, etc.).

- Charge two different per-unit prices for all soft drinks, but determine the groupings of products and these two price levels optimally.

In this context, uniform pricing by manufacturer-brand-size and completely non-uniform prices are just two of many potential pricing strategies. The retailer's implementation costs for these pricing strategies clearly differ. For example, charging a constant markup of 30% on all products requires no knowledge on the part of the retailer about the residual demand curve that it faces. In fact, many books on applied pricing for small retailers (e.g., Burstiner (1997)) suggest that they simply charge a 100% markup on their entire inventory, a practice known as "keystone pricing". By contrast, charging a different (and profit maximizing price) for each product requires intimate knowledge on the part of the retailer of the residual demand curve that it faces. It must be cognizant not only of consumers' preferences, but also of the current state space (competitors' current prices, current advertising activity, current wholesale prices, holiday periods, etc.). Furthermore, mapping these state variables to optimal prices for each product involves solving a high dimensional optimization problem *every period*.

The following framework is useful for analyzing the retailer's decision process. Assume that each period, the retailer maximizes expected profits, less menu costs:

$$\begin{array}{c}\text{Expected}\\\text{Profit}\\\text{this Period}\end{array} = \begin{array}{c}\text{Expected}\\\text{Revenue}\\\text{from Sales}\end{array} - \begin{array}{c}\text{Expected}\\\text{Cost}\\\text{of Goods}\end{array} - \begin{array}{c}\text{Menu Costs}\\\text{this Period}\end{array} \tag{1}$$

The menu costs incurred by the retailer each period can be thought of as having two components:

$$\begin{array}{c}\text{Menu Costs}\\\text{this Period}\end{array} = \begin{array}{c}\text{Maintenance}\\\text{Costs}\\\text{(Recurring)}\end{array} + \begin{array}{c}\text{Upgrade}\\\text{Costs}\\\text{(if any, Non-Recurring)}\end{array} \tag{2}$$

The first component of the current period's menu cost is the recurring cost of maintaining the current pricing strategy. Obviously, this cost would vary depending on the pricing strategy chosen. A simple pricing strategy, such as "charge the same prices as last period" would incur zero maintenance costs. But in the case of a complex strategy, this cost could be quite high. Each period, the retailer would have to learn the current state space (entering wholesale prices into a computer pricing program, learning competitors' prices, etc.) and apply the pricing rule to determine the prices for that period. Even within the class of complicated pricing rules, costs might vary. Due to the difficulties inherent in numerical optimization, uni-dimensional pricing strategies (e.g., charging a single optimal markup) are much easier to implement (and hence require less managerial time) than high-dimensional pricing strategies.

The second component of the current period's menu cost is the cost of upgrading to a better pricing strategy. This fixed cost is non-recurring (or at least infrequently recurring). If the retailer decides to use the same pricing strategy as in the previous period, no upgrade costs would be incurred. But if, for example, a retailer decided to switch from a "charge the same prices as last period" strategy to a "charge the profit-maximizing price for each good" strategy, they would potentially have to incur several costs.

First, in order to learn their demand curve, the retailer may want to introduce exogenous price variation.[3] In addition to the opportunity cost of the time it takes for a manager to determine these prices, this experimentation involves forgone profits, because it explicitly requires charging prices that are believed to be non-optimal. Fortunately, this experimentation is required only infrequently – when it is believed that the structural parameters of the demand system have changed.

Second, the retailer must analyze the data to recover the structural parameters of the demand system. Here the retailer faces a choice regarding the level of sophistication used in estimating the demand curve. For example, the retailer could employ a homogenous logit model, a heterogenous logit model, or another model (such as that found in this paper) in estimating demand. In choosing among alternative models, the retailer must trade off the opportunity cost of a manager's time, or the cost of hiring consulting services against the expected cost of mis-specification. Like the costs associated with experimentation, this cost must be incurred only when it is believed that the structural parameters of the demand system have changed. Third, the retailer may potentially have to purchase software (such as optimization software) to allow it to convert the estimated demand system to a set of

[3]As discussed in more detail later, it is not possible to estimate the cross-price elasticities between two products if their relative prices are constant.

optimal prices.

Unlike the maintenance costs, which are incurred each period, these upgrade costs would only be incurred by the retailer infrequently, when the retailer perceives that the upgrade costs are less than the present discounted value of the additional profits gained from following the new pricing strategy.[4] This means that the upgrade costs must be incurred whenever there is reason to believe that the structural parameters of the demand system have changed. If the demand system changes substantially over time, or across distance, this will lead to additional upgrade costs. For example, if Coke is more popular in some areas, while Diet Coke is more popular in other areas, then demand must be estimated separately in each of these areas. This explains why even large chains might charge uniform prices - because demand may differ structurally across geographic areas, and hence it may need to be re-estimated for each area, eliminating returns to scale in upgrade costs. Similarly (though less likely), the retailer will have to re-estimate demand more frequently in areas where the distribution of consumers' preferences change frequently. We are less likely to see costly-to-implement pricing strategies when the retailer cannot expect to recoup the upgrade costs.

For tractability, this paper assumes a static model, and estimates the additional per-period maintenance costs for non-uniform pricing relative to uniform pricing. A dynamic structural model would have to include a model of the retailer's expectations about the demand curve it would find if it experimented, as well as the retailer's expectations how marginal costs and demand would change over time, and their competitor's actions.

My approach to measuring the menu costs that would rationalize the observed uniform pricing behavior is to look at the counterfactual expected profits earned by the retailer

[4] I choose to remain agnostic about the process that allows the retailer to form expectations about the additional profit to be gained from upgrading to a new pricing system without actually implementing the system.

under uniform and non-uniform pricing strategies. That is, for each week, I compute:

$$\begin{array}{c}\text{Minimum Menu Cost}\\\text{this Period}\\\text{that Rationalizes}\\\text{Uniform Prices}\end{array} = \left(\begin{array}{c}\text{Expected}\\\text{Revenue from Sales}\\\text{this Period}\\\text{with Non-Uniform}\\\text{Prices}\end{array} - \begin{array}{c}\text{Expected}\\\text{Cost of Goods}\\\text{this Period}\\\text{with Non-Uniform}\\\text{Prices}\end{array}\right) - \tag{3}$$

$$\left(\begin{array}{c}\text{Expected}\\\text{Revenue from Sales}\\\text{this Period}\\\text{with Uniform}\\\text{Prices}\end{array} - \begin{array}{c}\text{Expected}\\\text{Cost of Goods}\\\text{this Period}\\\text{with Uniform}\\\text{Prices}\end{array}\right) \tag{4}$$

The left-hand side of this equation corresponds to the amount of profit that the retailer actually earned by charging non-uniform prices instead of uniform prices. Therefore, the model predicts that if the retailer had faced menu costs higher than this amount, I would have observed that retailer charging uniform prices if the choice was made on a week-by-week basis. More likely, the uniform versus non-uniform pricing strategy decision is made infrequently. This would mean that the relevant profit difference to consider would be the discounted value of the sum of the profit differences over several weeks. The difference between the profits earned under uniform and non-uniform prices depends only on the demand function faced by the store, and it's marginal costs.

In order to make this comparison empirically, it is necessary to learn (a) the demand system faced by the retailer and (b) the cost structure faced by the retailer. Learning these two pieces in order to perform the counterfactual experiment posed above, requires developing a structural model of demand and supply. The remainder of this section discusses (a), deferring discussion of (b) to section 5.3.

In order to actually calculate the implied menu costs, I must learn both the marginal costs and the demand curves faced by the retailer. In practice, I only observe data on weekly prices and quantities purchased by households. Economic theory suggests I can recover the marginal costs if I know the residual demand function, but I must observe price and quantity data that includes variation in prices - I must observe a retailer charging non-uniform prices.[5] I solve this problem by considering a product that is frequently (but not

[5] Strictly speaking this is not quite true. For example, it is possible to empirically estimate a homogenous logit model and hence derive cross-price elasticities between two goods whose price ratio is constant as long as the their price levels are changing (i.e., get identification off sales). To see this, consider the case with two products and an outside good, with product dummies for characteristics. In this case, a homogenous logit model of household choice implies a system of 2 *linear* equations, where the dependent variable in equation j

always) priced uniformly: different flavors of carbonated soft drinks.

In addition to the necessary price variation, there are a number of features that make the carbonated soft drink category amenable to this investigation. First, the soft drink category has a large number of products, and a large number of different varieties of similar products. Second, contrary to many other product categories, these products could plausibly be grouped in a number of alternative ways. This stems from the fact that the product packaging for soft drinks is largely identical across brands and manufacturers. The label on a 2-Liter bottle of Coke may differ from its Pepsi counterpart, but the physical shape of the bottle is identical. This physical similarity is much different than other product categories, like yogurt, where consumers might be less likely to accept line pricing by flavors. Third, soft drinks are the most frequently purchased item in scanner data. According to scanner data, in most product categories, the median household makes a purchase only a couple of times per year. By contrast, the median household in the sample purchases soft drinks on eleven occasions over the sample period of two years. This gives us the hope of obtaining reasonably good estimates.

Unfortunately, there are several potential drawbacks to looking at carbonated soft drinks. The first potential drawback is that anecdotal evidence suggests that consumers may stockpile soft drinks. If true, this would greatly complicate the demand model. However, Hendel & Nevo (2002*b*) find no evidence of stockpiling of soft drinks. This is not definitive however, since they also find little evidence of stockpiling in detergents, while their structural paper on detergents (Hendel & Nevo 2002*a*) does find such effects.

A second potential problem with using soft drinks is that retailers may view the soft drink category as a "loss-leader" category, pricing it low in order to drive store traffic. Although it features in some theoretical work (see for example, Hess & Gerstner (1987)) there is little empirical evidence of loss leader behavior. To the extent that cross-category

is the log of the ratio of the market share of good j to the market share of the outside good: $ln(s_{jt}) - ln(s_{0t})$. The independent variables are the price of good j and indicator variables for each good. Because the logit assumes that the coefficients on these prices and dummy variables are equal in each equation, the system collapses to a single equation with three variables: price and indicator variables for the two goods. This equation is usually estimated by OLS. If the two goods are sold at the same price, say $1 in every period, clearly the price variable is co-linear with the product dummies, and the model cannot be estimated. If, however, the two goods are sold at different prices in each period, but at the same price relative to each other, then the model can be estimated. Unfortunately, the identification in this case is coming entirely from the functional form. The logit model assumes that households choose the alternative yielding the highest indirect utility. Because indirect utility functions are homogenous of degree zero – meaning that a change in the level of all prices is equivalent to a change in the level of income – the effect of this price variation is equivalent to variation in household income.

loss-leader behavior by the retailer does occur, it would only affect the assumptions I use to recover marginal costs. My method would interpret low prices as evidence of low marginal costs, when in fact prices would be low in order to drive store traffic. While my approach could be modified to account for cross-category loss-leader behavior, I lack strong candidates for an alternative to assuming the retailer maximizes profits on a category-by-category basis, and simultaneously estimating demand for several product categories is beyond the scope of this paper.

A third potential problem with the soft drink industry is the argument that the soft drink category is different because the two main manufacturers, Coke and Pepsi, have extraordinarily strong brands and, as a result, may exert pressure on retailers to price their products in a particular way. If Coke and Pepsi exerted influence over pricing, it does not present a problem for my demand estimation, which relies solely on the fact that I observe price variation. It could potentially affect my profit calculations. However, what I see in the data does not appear to be consistent with the Coke and Pepsi influencing pricing (at least not at the store that charged non-uniform prices). What I observe in the data is one store charging non-uniform prices, and all other stores charging uniform prices. It seems highly implausible that Coke and Pepsi would (or could) dictate to every other store in the area to charge uniform prices, but allow the retailer I observe to charge non-uniform prices.

Within the soft drink category, there is often a great deal of price variation, both for the same product over time and between products from different manufacturers. Indeed, Coke and Pepsi frequently alternate promotion weeks, with Coke on sale one week and Pepsi on sale the next. This behavior can be seen in Figures 1 and 2, which plot the weekly price (normalized to 12-ounce servings) of 2-Liter containers of Coke and Pepsi over a two year period at two different stores. In contrast to the price variation seen in these figures, soft drinks are typically sold at uniform prices by manufacturer-brand-size. Within a size, all flavors of Pepsi are typically sold at one price, and all flavors of Coke at another uniform price. Figure 3 illustrates this by plotting the ratios of the price of Diet Pepsi and Diet Caffeine Free Pepsi to Regular Pepsi. From the graph, it is easy to see that at this store (the same store as in Figure 1), the three Pepsi UPCs[6] were always sold at the same price. However, not all stores charged uniform prices during this time. Figure 4 shows the same

[6]The Store Keeping Unit (SKU), also sometimes known as a Universal Product Code (UPC) is a number that uniquely identifies each product/size. For example, a 2-Liter plastic bottle of Caffeine-Free Diet Coca-Cola has a different SKU than a 2-Liter plastic bottle of Diet Coca-Cola, or a 12oz can of Caffeine-Free Diet Coca-Cola.

price ratios as in Figure 3 but at another store (the same store as in Figure 2). The graph for this store clearly shows a great deal more variation in the prices of different flavors of Pepsi. This variation allows us to estimate demand separately for each Pepsi variety. Similar price variation among varieties of other manufacturer brands at this store allows us to estimate demand for many different items.[7]

[7] As an aside, note that Figures 3 and 4 (and the puzzle addressed in this paper) highlight two features of the data: (1) the prices of the three goods are moving in lock-step (i.e., Figure 3 is essentially two straight lines) and (2) they are all the same price (i.e., these lines are at 1). Although this paper considers (2), observing only (1) but not (2) would also be explained by menu costs. For example, if a retailer knew that Diet Coke was more popular than Coke (but not how this varied with their absolute prices or the prices of other products) then one unsophisticated pricing strategy that would result from this would be to always charge $0.10 more for Diet Coke than Coke.

Figure 1: Graph of Prices for Coke and Pepsi at Store B, 6/91-6/93

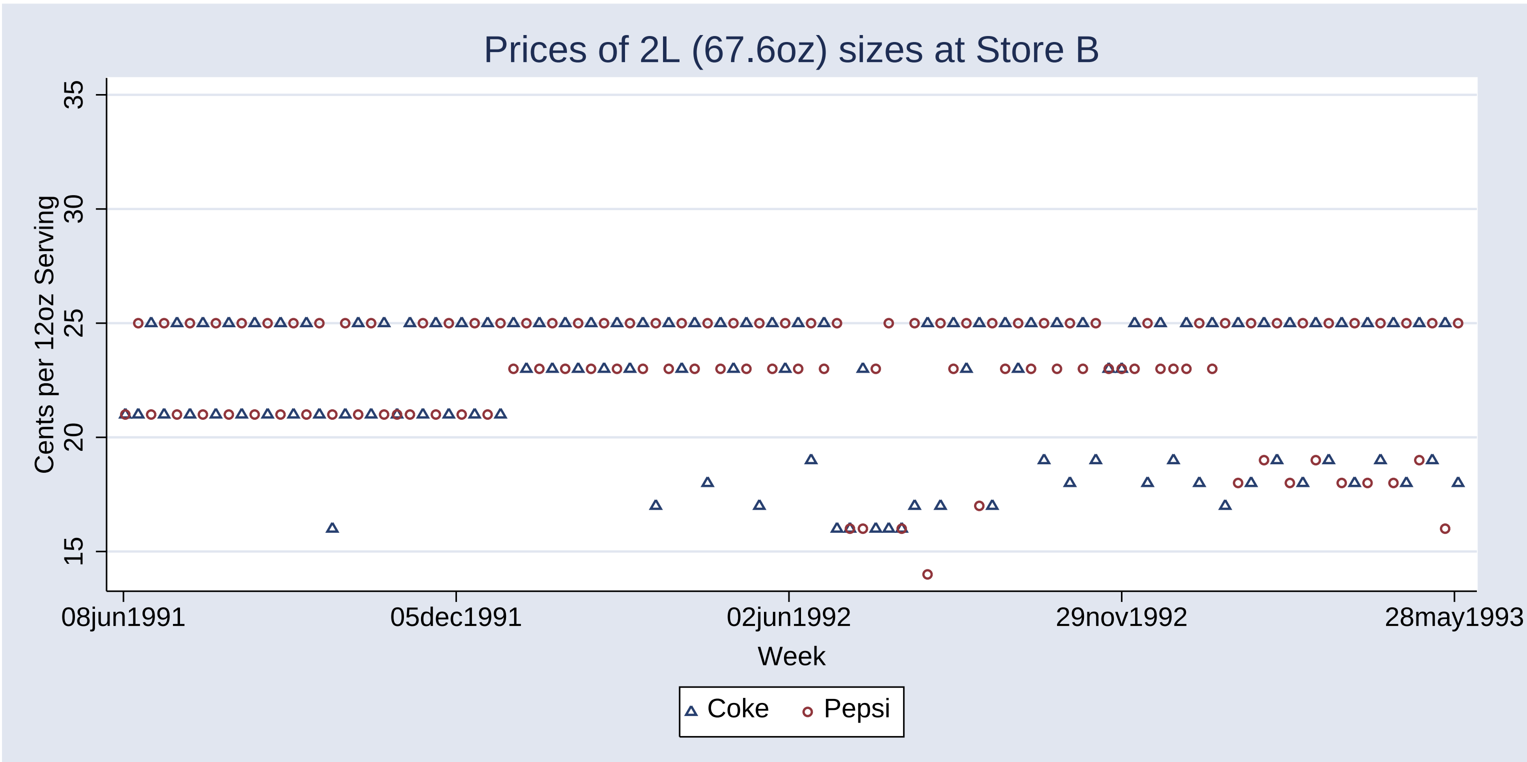


Figure 2: Graph of Prices for Coke and Pepsi at Store A, 6/91-6/93

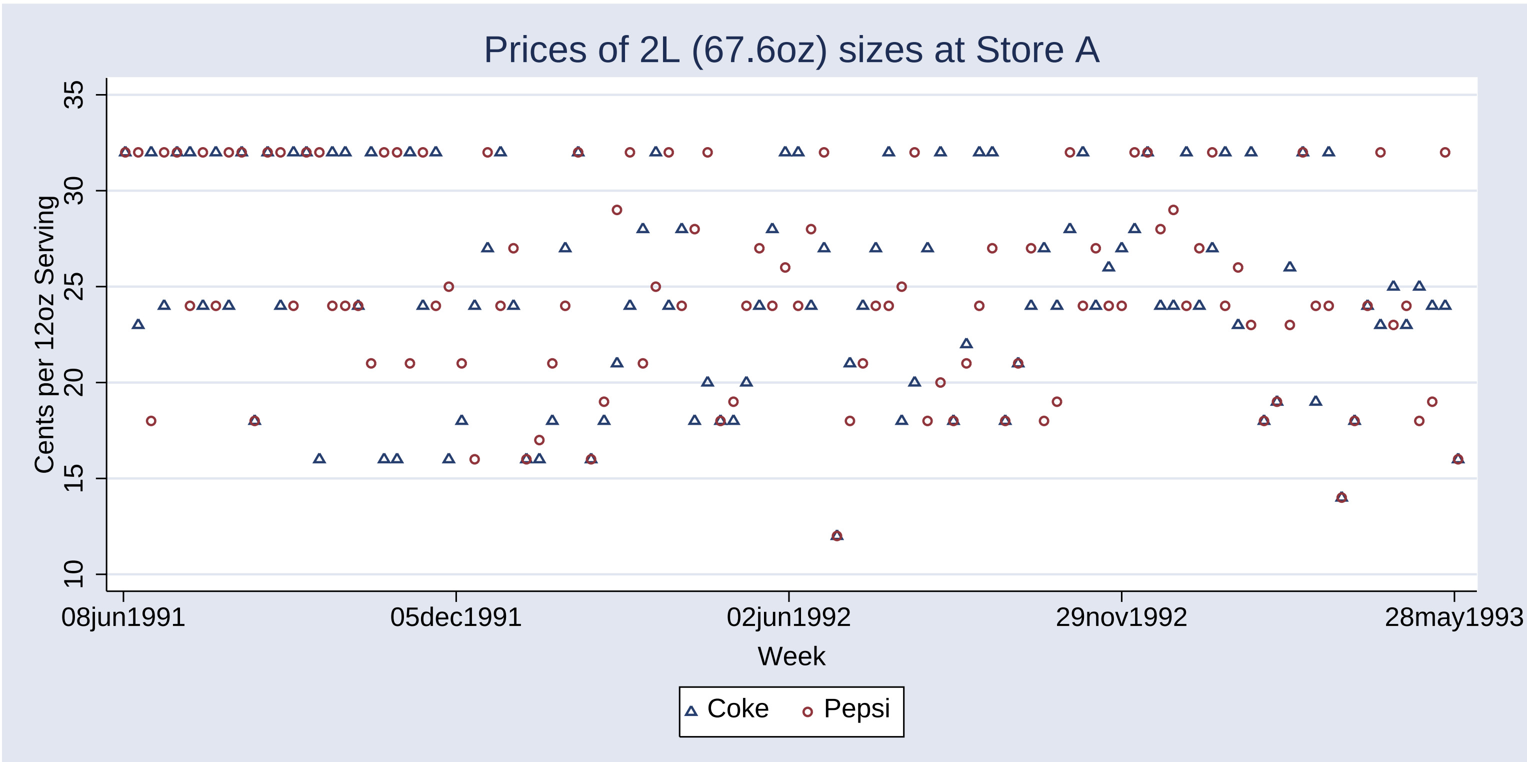


Figure 3: Graph of Prices Ratios of Pepsi Varieties to Regular Pepsi at Store B, 6/91-6/93

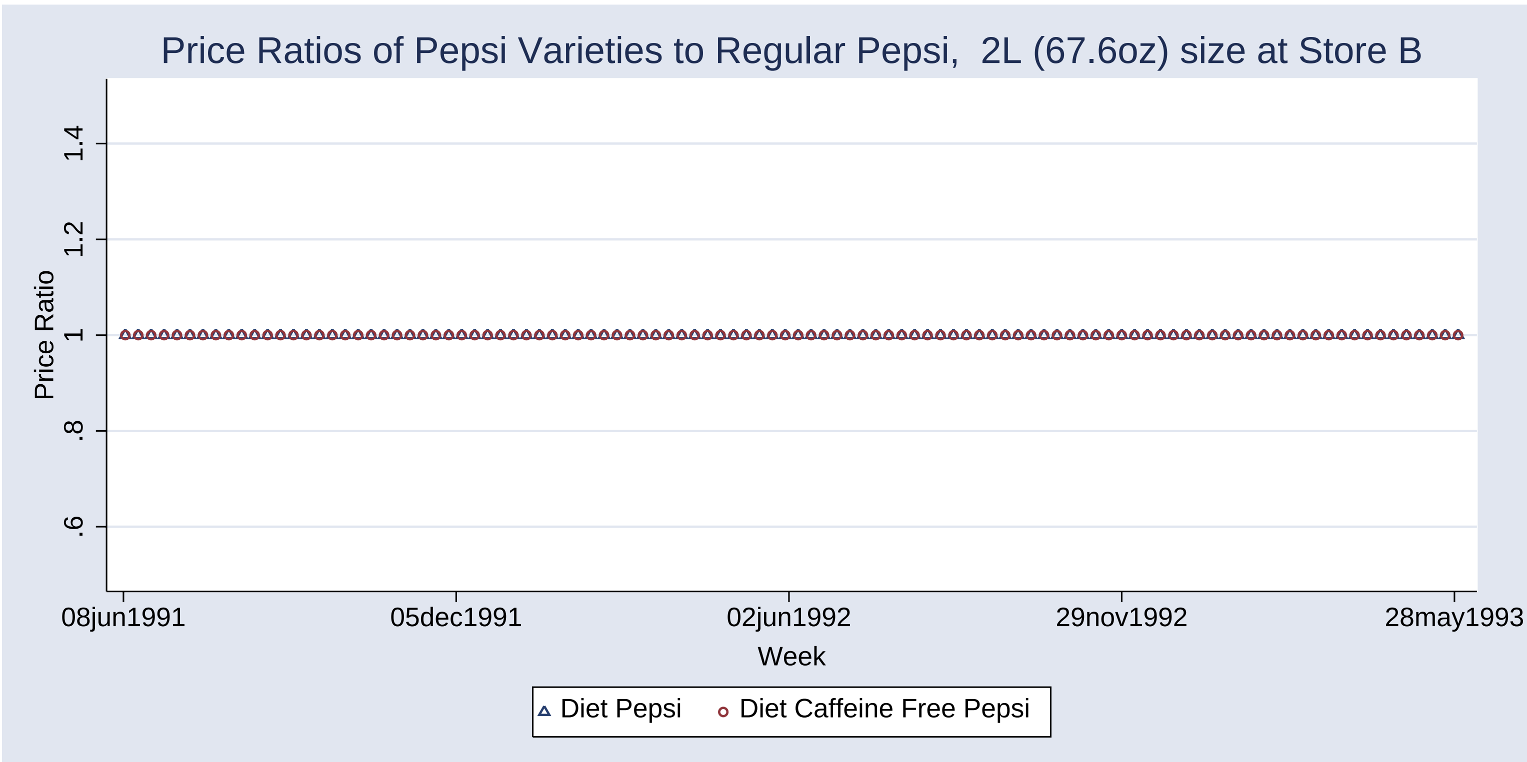
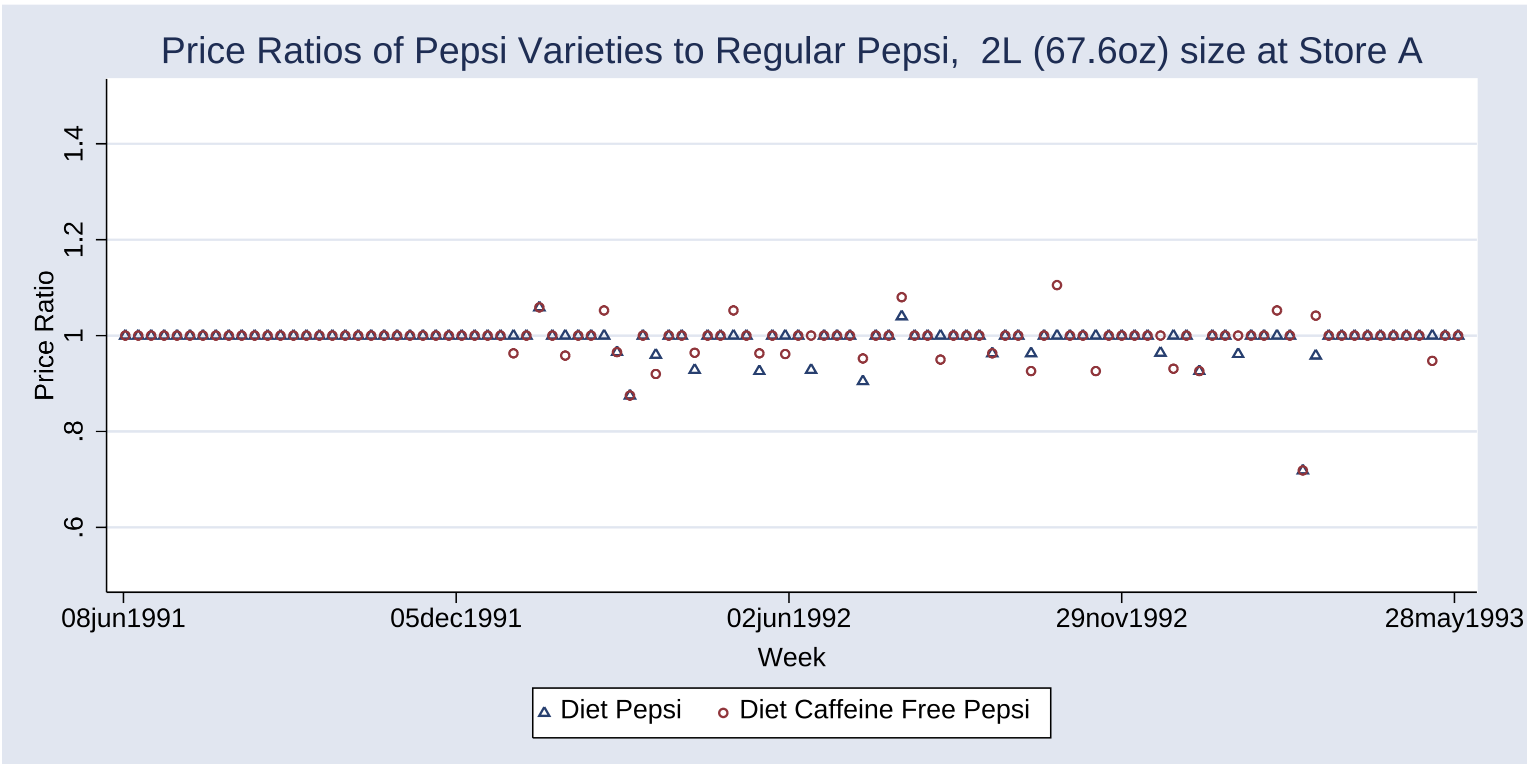


Figure 4: Graph of Prices Ratios of Pepsi Varieties to Regular Pepsi at Store A, 6/91-6/93



3 Empirical Demand Model

I need a model of consumer demand that will allow me to estimate the residual demand curve in period t for each product j: $Q_{jt}(\cdot)$. It is important to note that I need to estimate the *residual* demand function faced by a single store. Unless the retailer is a monopolist, this is the not the same as the market demand function faced by all stores. The residual demand function reflects the presence of other stores in the market. The difference is that the residual demand function accounts for the fact that the prices charged at other stores affect demand at store A. This means that if store B has a clearance sale, I should expect demand at store A to decline. Several previous empirical demand studies have ignored this aspect for the very good reason that in most datasets, this information is simply not available – prices for other stores are not observed. In my case, however, I observe the prices charged at four other competing stores. Furthermore, as explained later, the additional stores in the dataset were chosen precisely because they were the stores that shoppers were most likely to visit.[8]

I assume that the households' choice process is as follows. First, I assume that an exogenous process governs consumers' decision of whether to shop in a given week, as well as their total grocery expenditure in that week. Second, conditional on deciding to shop, I model households' store-choice decision using a conditional logit. Then, conditional on a household's choice of store, I assume that they optimally allocate expenditure between soft drinks and all other groceries. Hence, the residual demand faced by store A is equal to the sum over all households i (that went shopping in that week) of the probability that the household chose store A, multiplied by their expected purchases q_{it}, conditional on choosing store A. The resulting expected residual demand vector faced by store A in week t is:

$$E_t\left[Q(\mathbf{p})\right] = \sum_i E\left[i\text{'s purchases}|i \text{ goes to } A\right] \cdot P\left(i \text{ goes to } A \,\middle|\, \begin{array}{c} i\text{'s characteristics} \\ \text{and prices at all stores} \end{array}\right) \tag{5}$$

or, more formally:

$$E\left[Q(\mathbf{p}_t|\Omega_{obs,t}, \Omega_{unobs,t})\right] = \sum_i E\left[q_{it}|\text{parameters}, \Omega_{obs,t}\right] \cdot P\left(A|\text{parameters}, \Omega_{obs,t}\right) \tag{6}$$

where $\Omega_{obs,t}$ and $\Omega_{unobs,t}$ represent observed and unobserved variables specifying the state of

[8]Throughout this paper I assume that the retailer uses a best-response to other retailers, and does not account for the fact that deviations may lead to changes in rivals' pricing strategy. This is the same as the criterion for a Nash equilibrium.

the world in week t. For tractability, I assume that the store choice decision is conditionally independent of the households' soft drink purchase decision. Economically, this rules out, for example, going to store A because it is the only store that carries product j. More importantly, it also rules out going to a particular store based on a purchase-occasion-specific shock. This means that I am assuming that households' idiosyncratic preference shocks for varieties of soft drinks are independent of their idiosyncratic shock for their store-choice decision. Unfortunately, this means that the complete demand model is not consistent with utility maximization.[9] The next two subsections describe the specification and estimation procedure used for each of these two components of residual demand on more detail.

3.1 Product-Level Demand Model

The majority of extant demand models are unable to address simultaneously: complementarities between goods, continuous choice, and a large number of products. Given that households typically purchase several different soft drink products within the same purchase occasion, there is reason to believe that complementarities exist in my data. This behavior is summarized in table 1, which replicates a similar table in Dubé (2001). The same phenomenon is also evident, though to a lesser degree, when I consider only purchases among the top 25 SKUs by sales in Table 2.

The importance of addressing these issues, particularly in the current setting, is explored in McMillan (2005) through a series of Monte Carlo experiments. In light of this, I use the household-level demand model developed in that paper. This model is a hybrid between Kim, Allenby & Rossi (2002)'s and Chan (2002). It improves on Chan (2002) by explicitly solving for the household's budget constraint, by extending it to SKU-level demand. The non-linearity also allows a more flexible matrix of characteristics, with more characteristics than products. I also use physical characteristics, rather than brand-level (as in Chan) or product-level (as in Kim et al. (2002)) dummy variables.[10]

In order to reduce the dimensionality of the parameter space, I assume that, as in the logit model, households derive utility from product characteristics. Each product j (shown in Table 5) can then be expressed as a vector of C different characteristics (described in

[9]For example, the household's choice of store is assumed to be conditionally independent of the household's preferences for particular items. This is reflected in my use of a generic average price, rather than an average of the bundle of products the household would purchase.

[10]McMillan (2005) contains a detailed discussion of the properties and features of this utility function.

Table 1: Distribution of Purchase Occasions by Number of Items and Number of SKUs Purchased (all Carbonated Soft Drinks)

Total Number of Items Purchased	Total Number of SKUs Purchased					
	1	2	3	4	5+	Total
1	1,332	0	0	0	0	1,332
2	967	377	0	0	0	1,344
3	177	195	86	0	0	458
4	370	111	45	16	0	542
5	25	50	22	9	3	109
6	122	46	22	5	2	197
7	7	23	13	4	2	49
8	83	16	17	1	1	118
9	15	11	3	4	0	33
10+	169	39	11	5	4	228
Total	3,267	868	219	44	12	4,410

This table shows the distribution of household purchase occasions across multiple units and multiple products, replicating a similar table found in (Dubé 2001).

Table 2: Distribution of Purchase Occasions by Number of Items and Number of SKUs Purchased (Among Top 25 Carbonated Soft Drinks)

Total Number of Items Purchased	Total Number of SKUs Purchased					
	1	2	3	4	5	Total
1	1,163	0	0	0	0	1,163
2	274	108	0	0	0	382
3	86	38	6	0	0	130
4	49	23	3	1	0	76
5	17	23	5	2	0	47
6	25	15	6	0	0	46
7	5	6	2	0	0	13
8	8	3	3	0	0	14
9	3	2	2	0	0	7
10+	60	5	2	0	1	68
Total	1,690	223	29	3	1	1,946

This table shows the distribution of household purchase occasions across multiple units and multiple products, replicating a similar table found in (Dubé 2001).

section 4), and the menu faced by the household can be represented by a $J \times C$ matrix A where the rows of A are the products, and the columns are characteristics. Hence, A is just a stacked matrix consisting of the X_{jt}'s from the logit model.

To illustrate how the dataset fits into the model, consider an example with just two characteristics (Diet and Cola) and three available products (Coke, Diet Coke, and Diet 7Up). Because it is a non-diet cola, Coke has characteristic vector [0 1].[11] As a diet cola, Diet Coke has characteristic vector [1 1]. Finally, as a diet non-cola, Diet 7Up has characteristic vector [1 0]. Stacking these three products' characteristic vectors yields the following A matrix[12]:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Again following the logit, I assume that a household's utility function is additively separable in these characteristics. Unlike the logit, however, this model allows households the ability to consume multiple units of a single product, as well as consuming several different products. In particular, I assume that in week t, household i myopically maximizes the utility function[13]:

$$U_{it}(\mathbf{q}_t, z_t) = \sum_{c \in C} \beta_c (A'_{ct}\mathbf{q}_t)^{\rho_c} + \varepsilon'_{it}\mathbf{q}_t + z_t \tag{7}$$

with respect to the vector $\mathbf{q}_t$ and the scalar z_t. A_{ct} is the cth column of A in week t, $\mathbf{q}_t$ is a column vector of length J comprising the household's purchases of the J goods described by A, z_t is the amount of outside good consumed, and β_c and ρ_c are characteristic-specific scalar parameters. The J dimensional vector ε_{it} represents the household/shopping-occasion marginal utility shock, which is observed by the utility-maximizing household, but not by the econometrician.[14] The household maximizes this utility function subject to the budget

[11] Note that although in this case the product characteristics are indicator variables, in general they need only be non-negative. For example, in the estimated model one of the characteristics is the number of milligrams of caffeine per 12-ounce serving.

[12] Although the A matrix shown here is time invariant, the empirical A matrix will typically vary from week to week, because I include feature and display as time-varying characteristics.

[13] For simplicity, the model assumes households are homogenous in their preferences (β and ρ). Extending the model to account for heterogeneity (observed or unobserved) is straightforward, though computationally burdensome.

[14] I currently assume that ε_{it} is i.i.d. across products, time, and households, and negatively log-normally distributed on the interval $(-\infty, 0)$. It is necessary to bound ε_{it} from above in order to prevent unreasonable choice behavior. If, for example, the realization of ε_{itj} is greater than the price of good j, a household may never consume the outside good on that purchase occasion, regardless of the level of w_t.

constraint:

$$\sum_{j \in J} p_{jt} q_{jt} + z_t \leq w_{it} \tag{8}$$

where w_{it} is the household's total grocery expenditure in the store in week t. Returning to the three-good example above and substituting the A matrix into the utility function, we can see that the household maximizes [15]:

$$\begin{aligned} U_{it}(\mathbf{q}_t, z_t) \quad = \quad & \beta_{Diet} \left(q_{DietCoke,t} + q_{Diet7Up,t} \right)^{\rho_{Diet}} + \beta_{Cola} \left(q_{Coke,t} + q_{DietCoke,t} \right)^{\rho_{Cola}} + \\ & \varepsilon_{i,Coke,t} q_{Coke} + \varepsilon_{i,DietCoke,t} q_{DietCoke,t} + \varepsilon_{i,Diet7Up,t} q_{Diet7Up,t} + z_t \end{aligned}$$

with respect to $\mathbf{q}$ and z, subject to:

$$p_{Coke,t} q_{Coke,t} + p_{DietCoke,t} q_{DietCoke,t} + p_{Diet7Up,t} q_{Diet7Up,t} + z_t \leq w_{it}$$

The model is estimated using the Method of Simulated Moments (MSM), developed independently by McFadden (1989) and Pakes & Pollard (1989)[16]. This estimation method uses the fact that the expectation of the difference between the expected purchases and the actual, observed purchases, is zero at the true parameter values. More formally, I use the $(J+2) * (J+1)$ moment conditions:

$$H_{I,T,R}(\beta, \rho) = \frac{1}{IT} \sum_{i=1}^{I} \sum_{t=1}^{T} \left(\mathbf{q}_{it} - E[\mathbf{q}_{it} | \beta, \rho, \mathbf{p}_t, w_{it}] \right) x_{it} \tag{9}$$

where $\mathbf{q}_{it}$ is the household's vector of actual purchases. Each moment is an average across all purchase occasions of the difference between expected purchases and the actual, observed purchases, interacted with the instruments. The vector of instruments, x_{it} consists of all exogenous variables in the model (more on this in the following paragraphs), namely: the prices of each good, the household's budget, and a constant.

Because exact computation is infeasible, I simulate $E[\mathbf{q}_{it} | \beta, \rho, \mathbf{p}_t, w_{it}]$ by drawing $R = 30$

[15] In reality, households are forced to choose between the discrete sizes offered by the store. I make no attempt to model this feature of the data, as it introduces an extraordinary amount of computational cost with little clear return. In the estimation, I do not restrict the predicted purchases to the discrete purchases that the household could have actually made, instead allowing purchase quantities to vary continuously. Although Dubé (2001) suggests selecting the purchasable grid point adjacent to the unconstrained maximum, this is not numerically feasible in my case as it would require the examination of $2^{25} \simeq 3.3 \times 10^7$ points for each household maximization.

[16] Gouriéroux & Monfort (1996) contains the best summary and discussion of various simulation estimators and their properties that I have found.

sets of ε_{it}'s (which I hold constant as I search the parameter space). Hence, I use the fact that:

$$E[\mathbf{q}_{it}|\beta, \rho, w_{it}, \mathbf{p}_{it}] \approx \frac{1}{R} \sum_{r=1}^{R} \mathbf{q}(\beta, \rho, w_{it}, \mathbf{p}_t, \varepsilon_{it}^r) \tag{10}$$

Using these moments, I define my estimates as minimizing the distance function:

$$[\widehat{\rho}, \widehat{\beta}] = argmin \left\{ (H_{I,T,R}(\rho, \beta))' \, W \, (H_{I,T,R}(\rho, \beta)) \right\} \tag{11}$$

Ideally, I would implement this as a two-stage procedure. The first stage of this procedure would involve choosing the weighting matrix, and finding consistent estimates of the parameters. The second stage takes these estimates and uses them to calculate the optimal weighting matrix, and then re-estimates the parameters. In practice, however, estimation currently takes several weeks. Therefore, I report only the (inefficient) first-stage estimates.

Each of these two stages of estimation consists of iterating over several steps:

1. Choose starting values for the parameters: β and ρ.

2. Take the actual characteristics of the households in the sample that went shopping in that week. In this case, a household is completely characterized by it's budget (w_{it}). This amounts to assuming that the retailer knows the distribution of the households that would go shopping (not necessarily at its store) in each week.

3. Draw R sets of ε_{it}'s for each observed purchase occasion. I use $R = 30$. This means drawing $R*$(Number of Purchase Occasions)$*$(Number of Products)$= 30*16,008*25 = 12,006,000$ i.i.d. negative lognormal random numbers. These random numbers are held constant across iterations.

4. Using the expenditures from the actual purchase occasion as the budget constraint, and the actual menu of prices in the week of the purchase occasion, take the current parameters and solve explicitly (numerically) the household's utility maximization problem. This step is non-trivial and accounts for the bulk of the computational power involved in this estimation procedure. This means solving $R*$(Number of Purchase Occasions)$=30 * 16,008 = 480,240$ utility maximization problems *for each iteration*

of the parameter values.[17]I discuss this process, and suggest numerical algorithms at greater length in the Appendix.

5. For each purchase occasion, average over the R different purchase vectors to get the expected purchases for that purchase occasion at the current set of parameter values.

6. Using the difference between the actual vector of purchases on that purchase occasion and the expected purchases calculated in Step 4, calculate the current moment equations.

7. Interact these moment equations with the weighting matrix W to calculate the current distance function. The weighting matrix is of dimension $(J+2)*(J+1)$ by $(J+2)*(J+1)$.

8. Using a numerical minimization algorithm[18], choose a new set of parameter values (β and ρ) and repeat steps until a minimum is found.

In addition to the computational cost, this simulation method forces me to make assumptions about the distribution of the unobservables (ε). The assumption that I choose to make is that these unobservables are distributed independently of all observable variables. In particular, I assume that they are distributed independently of prices. The distributions of ε could be made dependent upon prices (or other observables). I do not do so here for two reasons. First, it seems at least plausible that brand, holiday, feature, and display variables account for much of the potential for unobserved correlation between price and these unobservables, but to the extent that the retailer observes time-varying changes in the household error terms, the unconditional distribution of the error term will differ significantly from its distribution conditional on prices. In this case my estimates may be both biased and inconsistent. Short of simulation, I cannot think of a way to "bound" the effects of violations of this assumption. The second, and principal justification for this assumption, is that it is crucial in making the estimation tractable. Even implementing a recursive routine to match predicted with observed market shares (as in Berry, Levinsohn & Pakes (1995)) would be prohibitively computationally expensive. Economically, with respect to prices, I am assuming that the retailer does not observe any (or at least does not adjust

[17]In solving the household utility maximization problem, I have had the most success using the gradient-based E04UGF routine available from the Numerical Algorithms Group (NAG).

[18]In searching for the global extremum of the distance function, I have had the most success using NAG's simplex-based E04CCF routine.

prices in response to) time-varying changes in the distribution of the idiosyncratic demand shocks. Given this assumption that the idiosyncratic shocks are distributed independently of prices, it is internally consistent to use prices as instruments (since they will be orthogonal to the difference between the actual and the expected demands).

3.2 Store Choice Model

As noted earlier, I need to estimate the store's residual demand, not market demand. The true process by which households choose where to shop is almost certainly related to their decisions about exactly what to purchase once they get there. However, I am unaware of any paper that simultaneously models the household's store choice and product-level purchasing decisions. A fully structural model would involve calculating the household's expected utility, net of travel costs, from visiting each of the stores in its choice set. Such a model would also involve consumers forming expectations of the menu of prices they would face at each store. Furthermore, the effects of these prices on store choice would almost certainly depend on the household's expected shopping basket on that purchase occasion. Given my estimation procedure, this approach is far too computationally burdensome. Instead, I approximate this choice, by assuming that households' choices among the stores in my sample follow a logit choice model. This model represents an approximation, of the true model and does not directly correspond to any model of utility maximization. I use it because it is computationally cheap, and because I believe the approximation is reasonable.

The model I estimate assumes that in week t, conditional on going shopping, household i derives indirect utility:

$$u_{ist} = D'_{ist}\delta^0 - \overline{\mathbf{p}}'_{st}\delta^1 + \eta_{ist} \tag{12}$$

from choosing store s at time t, where D_{ist} is a vector of household characterstics interacted with store indicator variables, $\overline{\mathbf{p}}$ is a vector of price indices for several product categories at store s, including soft drinks, and δ^0 and δ^1 are vectors of parameters.[19] I defer discussion of the exact specifications and discussion of the estimated coefficients to section 5.2.

This model implicitly assumes that households form expectations about current prices (see Ho, Tang & Bell (1998)). I estimate several specifications using current prices, implicitly assuming that the household is able to perfectly forecast (or learn) these prices. I also estimate several other specifications using prices from the previous two weeks as predictors of

[19] In principle, the elements of δ^0 and δ^1 could be allowed to vary across households, though I do not do so here.

store choice[20], though in general I do not find that prices substantially influence households' choice of store. These findings are consistent with those of Hoch, Dreze & Purk (1994) who also find that consumers are largely inelastic to short term price changes in their choice of store.

Although I am not aware of any papers that simultaneously address the household's decision of what to buy and where to shop, there is an extensive literature on store choice, which I will not attempt to summarize in detail here. Instead I focus on the portions of that literature that I have included in the specification of this model.

I follow Bell, Ho & Tang (1998) and Leszczyc, Sinha & Timmermans (2000) in incorporating household-level demographics and find that these are both statistically and economically significant in predicting store choice. While Rhee & Bell (2002), find that once unobserved heterogeneity is accounted for, shoppers' demographic characteristics are not statistically significant in predicting the probability of switching, they do not allow the effects of these characteristics to vary across stores. Although I do not control for unobserved heterogeneity, I find that the effect of household characteristics vary on a store-by-store basis.

After demographic variables, I find that one of the most significant predictors of store choice is whether the household visited the store in the previous two weeks. This is consistent with the finding by Rhee & Bell (2002), who find that households are highly path-dependent in their choice of store. However, this may simply be controlling for unobserved time-varying heterogeneity among consumers.

I also account for the possibility that, as suggested by Bell & Lattin (1998), households with higher expenditure levels tend to prefer stores with certain pricing formats. Specifically, they found that households with large average expenditure levels tend to prefer so-called Every Day Low Price (EDLP) stores to High-Low stores whose prices fluctuate more wildly from week-to-week. To account for this effect, I interacted the household's expenditure level for the purchase occasion with store indicator variables. This allows shoppers who expect to have high (or low) expenditure levels to seek out specific stores.

Finally, as mentioned earlier, I take the household's decision to go shopping to be governed by an exogenous process. This is consistent with evidence in Chiang, Chung & Cremers (2001) who find that consumers decision to shop is largely unaffected by marketing

[20] I also experimented with using longer lags, but found that they did not improve the predictive power of the model.

mix variables.

3.3 Residual Demand

I complete the construction of the residual demand system by bringing together the product choice and store choice models. The residual demand faced by store A is equal to the sum over all households (that went shopping in that week) of the probability that the household chose store A, multiplied by their expected purchases, conditional on choosing store A. Hence, the expected demand system faced by store A in week t is:

$$E_t\left[Q(\mathbf{p})\right] = \sum_i E\left[\mathbf{q}_{it}|\rho, \beta, \mathbf{p}_t, w_{it}\right] \cdot P\left(A|D_{ist}, \overline{p}_{At}, \overline{p}_{-At}\right) \tag{13}$$

$$\approx \sum_i \left[\frac{1}{R}\sum_{r=1}^{R} \mathbf{q}(\rho, \beta, \mathbf{p}_t, w_{it}, \varepsilon_{it}^r) \cdot \frac{exp\left(D'_{iAt}\delta^0 - \overline{p}_{At}\delta^1\right)}{\sum_{s\in\{A,B,C,D,E\}} exp\left(D'_{ist}\delta^0 - \overline{p}_{st}\delta^1\right)}\right] \tag{14}$$

In calculating this expected demand, I follow similar steps to those used in the estimation:

1. Take the estimated values of the parameters: β and ρ.

2. Draw R sets of ε_{it} for each observed purchase occasion. I use $R = 30$. These random numbers need not be the same as those used to estimate the parameters above. Note that in this case the number of purchase occasions is the total number of store trips (to any store) made in that week, not just the purchase occasions from store A.

3. Using the expenditures from the actual purchase occasions as the budget constraints, and the actual menu of prices in the week of the purchase occasion, take the current parameters and solve explicitly solve the household's utility maximization problem (using numerical methods discussed in the Appendix).

4. For each purchase occasion, average over the R different purchase vectors to get the expected purchases for that purchase occasion at the current set of parameter values.

5. For each purchase occasion, multiply these expected purchases by the probability of choosing store A in that week.

4 Data

Although a trip to nearly any store offers many cases of uniform pricing, I focus exclusively on grocery stores. There are a number of reasons for restricting my attention to grocery stores. First, they offer literally thousands of examples of uniform pricing. Second, given that grocery stores carry a large number of products,[21] one might expect them to use relatively sophisticated pricing techniques. Third, grocery stores are a significant portion of the economy. In 1997, U.S. grocery stores had sales in excess of $368 billion, with roughly 100,000 establishments (U.S. Economic Census, 1997). Fourth, with few exceptions, groceries are not characterized by consumer uncertainty. Consumers are presumably quite familiar with products' characteristics as well as their preferences over these characteristics. For example, there is not much uncertainty about what will be inside when you pop open a can of Diet Coke. Finally, grocery stores conveniently offer the availability of scanner panel data.

For a two-year period from 1991 to 1993, Information Resources Incorporated (IRI) collected a panel dataset in urban Chicago. This dataset has both aggregate and micro components. The aggregate component consists of weekly price and quantity[22] data at the store/SKU level for several different product categories at five geographically close stores. Throughout the paper, these stores are referred to as stores A through E. As mentioned earlier, one of these stores, which I will call store A, charged non-uniform prices for carbonated soft drinks during this period. The micro-level component of this dataset contains carbonated soft drink purchase histories for 548 households at these five grocery stores over the two-year period. The dataset also contains the households' total grocery expenditure on each purchase occasion. IRI paid these households to use a special electronic card that recorded their purchases when they shopped at these stores. For the majority of the analysis, I use only a subset of these households consisting of 262 households that visited store A (the store at which I estimate demand) at least once during the two year sample period. [23]

[21] A typical grocery store carries over 14,000 different products.

[22] Quantity sold includes sales to all customers, not just those in the panel.

[23] I supplement the IRI data with demographic and wholesale price data from Dominick's Finer Foods, a grocery chain located in the greater Chicago metropolitan area. From 1989 to 1997, through an arrangement with the University of Chicago Graduate School of Business, Dominick's kept track of store-level, weekly unit sales and price for every SKU for a number of product categories, including carbonated soft drinks. This dataset has weekly store sales totals in these product categories as well as store area demographics pulled from census data. In addition, the dataset contains the actual wholesale prices that Dominick's

According to the documentation provided with the data, these five stores and 548 panelists were selected by IRI using two criteria: First, although very little information is available on the actual sampling procedures used, IRI tried to create a stratified random sample of households, reflective of the population in the area. Second, in order to avoid the effects of unobserved market fluctuations, it was IRI's goal to, as much as possible, achieve a closed system. That is, IRI tried to include the stores that the households in the panel would be most likely to shop at, in order to observe as large a fraction of their grocery expenditure as possible. That IRI achieved this goal is supported by the fact for the vast majority of the households, grocery expenditure at stores within the sample universe appears to be fairly constant over time.

Tables 3 and 4 show the distribution of households' expenditure at different stores for all households, as well as for those who shopped at store A at least once. The mean weekly expenditure by a household shopping at Store A was \$22, while the median was \$15.[24] This is less than stores B and C, but similar to stores D and E. Even households that shopped at store A at least once tended to spend more at these other stores, although the majority (\$350,000) of their total expenditure of \$610,000 over the period was at store A.

---

paid for each good. For a more thorough description of the dataset, see Hoch et al. (1994). This dataset is publicly available from the Kilts Center at the University of Chicago Graduate School of Business web page: http://gsbwww.uchicago.edu/kilts/research/db/dominicks/ Finally, because this dataset covers the same time and geographic area as the IRI dataset, one can match many of the products across the two datasets, giving us a measure of the wholesale prices for these products. To the best of my knowledge this matching has not been done in previous work.

[24] For clarity, all dollar references in this section are nominal. Hence, prices in 1991 use 1991 dollars, etc. I use this approach because retailers and wholesale prices over the period do not appear to move with inflation. For reference, one 1991 dollar is equivalent to \$1.39 2004 dollars, a 1992 dollar is equivalent to \$1.35 2004 dollars, and a 1993 dollar is equivalent to \$1.31 2004 dollars.

Table 3: Distribution of Purchase Occasion Expenditure by Store, All Purchase Occasions

Store	Number of Purchase Occasions	Expenditure in Dollars							
		Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum	Total ($000's)
Store A	16,008	22	21	0.25	8	15	27	228	350
Store B	10,063	38	36	0.16	14	27	50	281	390
Store C	13,733	44	38	0.14	17	33	60	378	600
Store D	7,835	22	30	0.34	6	12	23	325	170
Store E	5,637	26	28	0.34	9	16	33	254	150
Total	53,516	31	32	0.14	10	20	40	378	1,700

Table 4: Distribution of Purchase Occasion Expenditure by Store, Purchases Made by Panelists Who Visited Store A at least Once.

Store	Number of Purchase Occasions	Expenditure in Dollars							
		Mean	Standard Deviation	Minimum	25th Percentile	Median	75th Percentile	Maximum	Total ($000's)
Store A	16,008	22	21	0.25	8	15	27	228	350
Store B	615	32	27	1.24	14	23	41	177	20
Store C	2,234	34	38	0.14	13	24	40	378	8
Store D	6,435	15	29	0.34	5	10	17	188	99
Store E	3,479	19	19	0.34	7	13	26	201	68
Total	28,782	21	23	0.14	8	14	26	378	610

Unlike many previous papers which have estimated brand-level demand, this paper estimates SKU-level demand. Over a two year period, a typical grocery store sells over 200 different items in the carbonated soft drink category. The vast majority of these products are offered only rarely, or quickly enter and exit. Because this paper uses panel purchases to estimate demand, and many of these products are only rarely (or never) purchased by the panel, it not practical to estimate the households' demand for them. Instead, I estimate the households' demand for the 25 products with the largest market share by volume. These products represent 71% of the Store A's carbonated soft drink sales by volume, and 69% of their total soft drink sales by dollar value.

The products included in the analysis are shown in Table 5. Of these, three varieties (8 items) were distributed by the Coca Cola Corp., two varieties (8 items) were distributed by Pepsi Co., two varieties (4 items) were distributed by Dr. Pepper/7Up, two varieties (4 items) were distributed by the Royal Crown Corp., and one variety (1 items) was distributed by an independent producer under a private label.

Table 5: Variety and Size Distribution of in the Dataset, grouped by Manufacturer

Manufacturer	Variety	Number of 12oz servings (Liters in parentheses) 1 (0.36)	5.63 (2.0)	6 (2.13)	8.45 (3.0)	12 (4.26)	24 (8.52)	Number of Sizes Avail.
Coca Cola	Coke		✓			✓	✓	3
	Diet Coke		✓			✓	✓	3
	Diet CF Coke					✓	✓	2
Pepsico	Pepsi		✓	✓		✓	✓	4
	Diet Pepsi		✓		✓	✓	✓	4
RC Corp.	RC		✓					1
	Diet Rite		✓		✓		✓	3
DP/7Up	7Up		✓	✓				2
	Diet 7Up		✓				✓	2
PL	Private Label	✓						1
Total Number of Items		1	8	2	2	5	7	25

Some descriptive statistics on the price and sales volume for these products are shown in Table 6. The price of a 12-ounce serving of carbonated soft drink varied from a high of $0.49 as part of a 12-pack of 12-ounce cans of Diet Coke, to a low of $0.12 for a single can of the Private Label cola. Most products appear to have had either an end-of-aisle display, or a mention in the store's circular in between one-third to one-half of the weeks. The notable exceptions to this were the 3L bottle of Pepsi, the 2L bottle of Diet 7up, and the store brand which received significantly less advertising (as measured by circular and display activity).

Table 6: Summary Statistics for Prices and Quantities Sold at Store A, grouped by Manufacturer

		Minimum Price ($)	Mean Price ($)	Maximum Price ($)	S.D. of Price	Mean Price at DFF	Mean DFF Wholesale Price	S.D. of DFF Wholesale Price	Mean Number of Servings Sold Per Week	S.D. of Number of Servings Sold Per Week	% of Weeks on End of Aisle Display	% of Weeks Featured in Circular	% Market Share by Volume	% Market Share by Sales
Coke	2L Bottle	0.12	0.22	0.32	0.06	0.26	0.23	0.04	30	26	47	61	3.93	3.63
	12-pack 12oz cans	0.17	0.31	0.47	0.09	0.34	0.29	0.06	29	30	44	51	3.22	3.77
	24-pack 12oz cans	0.17	0.27	0.34	0.06	0.24	0.18	0.07	17	27	38	46	2.67	2.61
Diet Coke	2L Bottle	0.12	0.23	0.32	0.06	0.26	0.23	0.05	21	27	47	56	2.65	2.26
	12-pack 12oz cans	0.17	0.31	0.49	0.10	0.34	0.31	0.04	28	31	33	49	3.19	3.78
	24-pack 12oz cans	0.15	0.27	0.34	0.06	0.25	0.19	0.08	28	41	38	45	3.95	4.16
Diet Caffeine Free Coke	12-pack 12oz cans	0.17	0.33	0.47	0.10	0.36	0.30	0.04	15	20	22	40	2.06	2.37
	24-pack 12oz cans	0.17	0.27	0.34	0.06	0.21	0.15	0.06	14	22	33	45	2.09	2.26
Pepsi	2L Bottle	0.12	0.22	0.32	0.06	0.25	0.22	0.04	32	23	46	58	5.03	4.68
	6-pack 12oz cans	0.24	0.39	0.48	0.10	0.22	0.17	0.09	13	15	17	33	1.60	2.33
	12-pack 12oz cans	0.16	0.31	0.47	0.10	0.33	0.28	0.08	25	26	38	50	2.61	2.91
	24-pack 12oz cans	0.17	0.26	0.34	0.06	0.20	0.15	0.06	50	84	42	52	5.23	4.96
Diet Pepsi	2L Bottle	0.12	0.23	0.32	0.06	0.25	0.22	0.05	14	16	40	56	2.02	1.86
	3L Bottle	0.20	0.20	0.20	0.00	0.04	0.03	0.05	7	7	2	1	1.41	1.25
	12-pack 12oz cans	0.10	0.31	0.47	0.10	0.34	0.30	0.05	21	24	46	50	2.06	2.30
	24-pack 12oz cans	0.17	0.26	0.34	0.06	0.24	0.18	0.07	23	35	42	47	3.19	2.99
RC	2L Bottle	0.12	0.21	0.32	0.06	0.25	0.22	0.04	20	16	57	54	2.74	2.37
Diet Rite	2L Bottle	0.17	0.37	0.48	0.11	0.39	0.30	0.10	15	22	28	32	1.67	2.23
	3L Bottle	0.12	0.21	0.32	0.06	0.24	0.21	0.05	12	16	55	53	1.49	1.17
	24-pack 12oz cans	0.17	0.26	0.34	0.06	0.19	0.14	0.05	14	24	43	44	2.03	1.92
7up	2L Bottle	0.12	0.17	0.32	0.02	0.19	0.15	0.05	30	17	31	35	5.88	4.51
	6-pack 12oz cans	0.12	0.21	0.32	0.05	0.24	0.19	0.04	13	14	37	24	1.68	1.41
Diet 7up	2L Bottle	0.11	0.18	0.20	0.03	0.19	0.15	0.00	10	20	11	9	1.43	1.00
	24-pack 12oz cans	0.17	0.26	0.32	0.05	0.21	0.18	0.02	19	34	37	31	2.67	2.56
Private Label	12oz Can	0.12	0.19	0.22	0.03	NA	NA	NA	20	13	2	25	4.13	3.39

All prices are in nominal Dollars per 12-ounce serving. Source: IRI and DFF Data.

Both the traditional logit model and the new model estimated here reduce the dimensionality of the demand system parameter space by assuming that households' preferences over products are driven by product characteristics. The store's residual demand $Q_{jt}(\cdot)$ for product j is denominated in twelve ounce servings of carbonated soft drink. The characteristics used in the analysis are: calories (per 12 ounce serving), milligrams of sodium (per 12 ounce serving), milligrams of caffeine (per 12 ounce serving), grams of sugar (per 12 ounce serving), as well as indicator variables for the presence of citric acid, phosphoric acid, and whether it is a diet drink. These physical characteristics were obtained by contacting the manufacturers of the products, and, to the best of my knowledge, represent the characteristics of the products during the relevant time period. I also include indicator variables for size, brand, and whether it was featured in store A's weekly circular, or an in-store display (in store A), as well as a constant common to all products. These characteristics were chosen based on earlier work by Dubé (2001). These characteristics are the elements of the A matrix, and are shown in Table 7. This table also shows the number of weeks that each product was available. For example, the 12-pack of 12-ounce cans of Diet Pepsi was unavailable for 16 of the 104 weeks, while the 24-pack of 12-ounce cans of Diet Caffeine Free Coke, and the 24-pack of 12-ounce cans of Diet Caffeine Free Coke were not available for 15 weeks.

Table 7: Characteristics of Products in the Dataset, grouped by Manufacturer

Manufacturer	Variety	Size	Weeks Sold	Calories	Sodium (mg)	Sugar (g)	Caffeine (mg)	Contains Phosphoric Acid	Contains Citric Acid	Diet	% of Weeks Featured in Circular	% of Weeks on End of Aisle Display
Coca Cola	Coke	2L Bottle	104	140	50	39	34	1	0	0	47	35
		12-pack 12oz cans	99	140	50	39	34	1	0	0	32	27
		24-pack 12oz cans	99	140	50	39	34	1	0	0	42	32
	Diet Coke	2L Bottle	104	0	40	0	45	1	1	1	46	36
		12-pack 12oz cans	99	0	40	0	45	1	1	1	32	23
		24-pack 12oz cans	99	0	40	0	45	1	1	1	41	32
	Diet Caffeine Free Coke	12-pack 12oz cans	100	0	40	0	0	1	1	1	32	19
		24-pack 12oz cans	89	0	40	0	0	1	1	1	44	31
Pepsico	Pepsi	2L Bottle	104	150	37.5	40.5	38	1	1	0	44	28
		6-pack 12oz cans	104	150	37.5	40.5	38	1	1	0	20	11
		12-pack 12oz cans	90	150	37.5	40.5	38	1	1	0	36	27
		24-pack 12oz cans	100	150	37.5	40.5	38	1	1	0	42	33
	Diet Pepsi	2L Bottle	104	0	37.5	0	36	1	1	1	44	27
		3L Bottle	94	0	37.5	0	36	1	1	1	1	2
		12-pack 12oz cans	88	0	37.5	0	36	1	1	1	33	30
		24-pack 12oz cans	101	0	37.5	0	36	1	1	1	42	37
Royal Crown	RC	2L Bottle	104	160	50	42	43	1	0	0	32	26
	Diet Rite	2L Bottle	104	0	45	0	48	1	1	1	20	29
		3L Bottle	104	0	45	0	48	1	1	1	5	6
		24-pack 12oz cans	98	0	45	0	48	1	1	1	23	34
DP/7up	7up	2L Bottle	104	140	75	39	0	0	1	0	45	42
		6-pack 12oz cans	104	140	75	39	0	0	1	0	22	21
	Diet 7up	2L Bottle	104	0	45	0	0	0	1	1	44	42
		24-pack 12oz cans	89	0	45	0	0	0	1	1	39	38
Private Label	Private Label	12oz Can	104	140	50	39	34	1	0	0	20	3

Characteristics are per 12oz serving. Source: Coca Cola Corp., Pepsico, Royal Crown Corp., and Cadbury Beverages.

5 Results

5.1 Structural Model

The parameter estimates and standard errors[25] from the structural model are presented in Table 8. The nonlinearity of the model makes these parameter estimates difficult to interpret directly, however we can make some inferences, particularly relative to each other.

The characteristic that appears to have the largest affect on households' soft drink purchasing decisions is the indicator variable for whether the product is sold on a holiday. This characteristic has both the largest maximal marginal effect[26] ($\beta_{holiday}\rho_{holiday}$) and a β that is statistically significantly greater than zero. The effect of the Holiday characteristic is offset somewhat by the fact that $\rho_{holiday}$ is quite close to zero. This means that although the marginal utility from soda on holidays starts off higher, its second derivative is more negative than for other characteristics. Taken together, these two facts imply that households are more likely to purchase soft drinks on holidays, but not likely to substantially increase the quantity that they purchase.

After the Holiday characteristic, the Coke and Pepsi characteristics have the largest effects, based on their high maximal marginal utility values. Interestingly, while β_{Pepsi} is statistically significantly different from zero, β_{Coke} is not. Given that both β_{Pepsi} and ρ_{Pepsi} are relatively large, consumers differentially prefer Pepsi (and Coke, to a lesser degree) to other soft drinks – they are both more likely to purchase these brands, and more likely to purchase more of them.

At the other end of the spectrum, due to the fact that both their β's and their ρ's are relatively small, Sodium and Caffeine do not appear to significantly affect consumers purchasing behavior (although their coefficients are imprecisely estimated).

More readily interpretable than the parameter estimates are the own and cross-price elasticities that they imply. In the logit model, the cross-price elasticities are infamously

[25] The standard errors shown in Table 8 are calculated using the fact that: $\sqrt{n}(\widehat{\theta}_n - \theta_0) \rightarrow_d N(0, V)$ as the number of observations n goes to infinity. $V = \left[G_0' W G_0\right]^{-1} G_0' W \left[\Omega_0 (1 + \frac{1}{R})\right] W G_0 \left[G_0' W G_0\right]^{-1}$, where W is the weighting matrix. I obtain a consistent estimator of V by using $\widehat{G}_{IT} = \frac{1}{IT} \sum_{i=1}^{I} \sum_{t=1}^{T} x_{it} \frac{\partial E[q_{it}|\theta, \mathbf{p}_t, w_{it}]}{\partial \theta}$ and $\widehat{\theta} = [\widehat{\beta}, \widehat{\rho}]$. I use the estimated parameters β and ρ to compute $\widehat{\Omega}_{IT} = \frac{1}{IT} \sum_{i=1}^{I} \sum_{t=1}^{T} ((q_{it} - E[q_{it}|\widehat{\beta}, \widehat{\rho}, \mathbf{p}_t, w_{it}])x_{it})((q_{it} - E[q_{it}|\widehat{\beta}, \widehat{\rho}, \mathbf{p}_t, w_{it}])x_{it})'$ which is a consistent estimator of Ω_0. For further discussion, see Gouriéroux & Monfort (1996). I use a diagonal weighting matrix (W), with the elements scaled by the sum of squares of each of the instruments.

[26] The term "maximal marginal effect" is somewhat misleading. It is actually the most positive marginal effect. When β_c is negative, ρ_c is necessarily greater than one, and hence, the marginal effect is increasing in magnitude as quantity increases.

Table 8: Parameter Estimates from Structural Model of Product Choice

Characteristic	Units	β coeff.	β s.e.	ρ coeff.	ρ s.e.	Maximal Marginal Effect ($\beta\rho$)
Constant		0.0001	(0.5435)	0.7471	(0.6382)	0.0001
Calories	per 12oz	-0.0001	(0.6231)	1.2581	(0.5957)	-0.0001
Sugar	g per 12oz	0.0054	(1.2754)	0.8134	(0.5749)	0.0044
Sodium	mg per 12 oz	0.0000	(0.8227)	0.1608	(0.2259)	0.0000
Caffeine	mg per 12oz	0.0000	(0.3318)	0.0957	(0.0737)	0.0000
Phosphoric Acid	Indicator Variable	0.5431	(0.1842)	0.3742	(0.2111)	0.2032
Citric Acid	”	0.1674	(0.0708)	0.6585	(0.2603)	0.1102
Cola	”	0.2588	(0.1187)	0.4188	(0.3062)	0.1084
Flavored	”	0.3671	(0.1610)	0.5139	(0.3254)	0.1886
Single Serving	”	0.3174	(0.0311)	0.5690	(0.0640)	0.1806
288oz	”	-0.0003	(0.0046)	1.6013	(0.4515)	-0.0005
Diet	”	-0.0076	(0.0059)	1.3875	(0.4267)	-0.0105
Coke	”	0.5836	(0.3054)	0.4090	(0.2517)	0.2387
Diet Coke	”	-0.0115	(0.0062)	1.6398	(0.0793)	-0.0189
Pepsi	”	0.4477	(0.0216)	0.4764	(0.0778)	0.2133
7up	”	0.2333	(0.1140)	0.5892	(0.2415)	0.1374
RC	”	0.2380	(0.0906)	0.4824	(0.1868)	0.1148
Jewel	”	0.4449	(0.3412)	0.3507	(0.2108)	0.1560
Holiday	”	5.3081	(1.8195)	0.0586	(0.0305)	0.3111
Feature	”	0.2232	(0.1304)	0.3683	(0.2217)	0.0822
Display	”	0.4830	(0.4114)	0.2712	(0.2434)	0.1310

This table shows parameter estimates from the structural model under two different specifications. The right-most column is the product of β and ρ.

dependent upon market shares. These restrictions are evident in Table 9, which shows a matrix of own and cross price elasticities for several sizes and varieties of Coke and Pepsi. The logit model predicts, for example, that the demand for 24-packs (288oz) of Diet Coke increases by 0.46% from either a 1% increase in the price of a 2-liter bottle of Diet Coke or a 1% increase in the price of a 2-liter bottle of Pepsi. This contrasts sharply with the price elasticities implied by my structural model, presented in Table 10. As clearly seen, my model allows for a rich pattern of cross-price elasticities. In addition to allowing cross-price elasticities to vary across products, my model also allows for negative off-diagonal price elasticities, suggesting that complementarities exist. For example, if households tend to purchase both 12-packs of Coke and 6-packs of 7Up, then an increase in the price of Coke may cause more people to buy Pepsi, but it will also lead people both to buy less Coke *and* to buy less of the Coke/7Up bundle. This flexibility is not possible using the traditional logit model.

Table 9: Selected Own and Cross Price Elasticities from Homogenous Logit Model of Product Choice

	Coke		Diet Coke		Pepsi		Diet Pepsi	
	288oz	2L	288oz	2L	288oz	2L	288oz	2L
Coke - 288oz	-2.398	0.015	0.046	0.014	0.073	0.019	0.051	0.013
Coke - 2L	0.052	-1.992	0.046	0.014	0.073	0.019	0.051	0.013
Diet Coke - 288oz	0.052	0.015	-2.404	0.014	0.073	0.019	0.051	0.013
Diet Coke - 2L	0.052	0.015	0.046	-2.001	0.073	0.019	0.051	0.013
Pepsi - 288oz	0.052	0.015	0.046	0.014	-2.203	0.019	0.051	0.013
Pepsi - 2L	0.052	0.015	0.046	0.014	0.073	-2.002	0.051	0.013
Diet Pepsi - 288oz	0.052	0.015	0.046	0.014	0.073	0.019	-2.392	0.013
Diet Pepsi - 2L	0.052	0.015	0.046	0.014	0.073	0.019	0.051	-2.000

Table 10: Matrix of Estimated Average Own and Cross-Price Elasticities from Structural Model of Product Choice

	7up, 2L Bottle	7up, 6-pack 12oz cans	Diet 7up, 24-pack 12oz cans	Diet 7up, 2L Bottle	Coke, 12-pack 12oz cans	Coke, 24-pack 12oz cans	Coke, 2L Bottle	Diet Coke, 12-pack 12oz cans	Diet Coke, 24-pack 12oz cans	Diet Coke, 2L Bottle	Diet Caffeine Free Coke, 12-pack 12oz cans	Diet Caffeine Free Coke, 24-pack 12oz cans	Private Label, 12oz Can	Pepsi, 12-pack 12oz cans	Pepsi, 24-pack 12oz cans	Pepsi, 2L Bottle	Pepsi, 6-pack 12oz cans	Diet Pepsi, 3L Bottle	Diet Pepsi, 12-pack 12oz cans	Diet Pepsi, 24-pack 12oz cans	Diet Pepsi, 2L Bottle	RC, 2L Bottle	RC, 3L Bottle	Diet Rite, 24-pack 12oz cans	Diet Rite, 2L Bottle
7up, 2L Bottle	-4.158	0.054	0.056	0.131	0.023	0.014	0.036	-0.015	-0.001	0.008	-0.005	0.000	0.013	0.030	0.020	0.035	0.044	0.001	-0.002	0.000	0.007	0.018	0.006	0.003	0.006
7up, 6-pack 12oz cans	0.081	-7.921	0.049	0.121	0.040	0.034	0.034	0.011	0.025	-0.003	0.024	0.013	0.026	0.056	0.024	0.028	0.061	-0.008	0.033	0.016	-0.001	0.012	0.000	0.026	-0.003
Diet 7up, 24-pack 12oz cans	0.082	-0.003	-3.409	0.109	0.034	0.019	-0.049	0.064	0.039	0.011	0.061	0.033	0.015	0.035	0.026	0.000	0.025	0.007	0.028	0.034	0.016	-0.020	0.023	0.057	0.023
Diet 7up, 2L Bottle	0.063	0.072	0.036	-6.704	-0.019	-0.010	-0.009	0.010	0.008	0.014	0.006	0.008	-0.010	0.005	0.002	0.009	0.001	0.010	0.012	0.009	0.016	-0.004	0.019	0.011	0.026
Coke, 12-pack 12oz cans	0.052	-0.131	0.020	-0.056	-4.584	0.071	0.088	0.084	0.068	0.090	0.100	0.064	0.063	0.109	0.049	0.065	0.078	0.028	0.083	0.045	0.056	0.042	0.060	0.054	0.076
Coke, 24-pack 12oz cans	0.048	0.055	0.028	-0.032	0.119	-2.971	0.134	0.093	0.108	0.119	0.109	0.082	0.070	0.106	0.059	0.070	0.127	0.033	0.083	0.056	0.073	0.059	0.071	0.076	0.065
Coke, 2L Bottle	0.027	0.028	-0.017	-0.022	0.094	0.053	-4.048	0.074	0.043	0.065	0.067	0.037	0.032	0.039	0.022	0.050	0.074	0.026	0.010	0.012	0.045	0.034	0.033	0.022	0.068
Diet Coke, 12-pack 12oz cans	0.020	-0.113	0.041	0.070	0.088	0.070	0.105	-4.757	0.075	0.134	0.188	0.067	0.044	0.081	0.045	0.052	0.054	0.054	0.065	0.053	0.066	0.021	0.090	0.072	0.093
Diet Coke, 24-pack 12oz cans	-0.003	0.155	0.057	0.068	0.111	0.121	0.143	0.113	-2.891	0.138	0.140	0.115	0.058	0.120	0.057	0.049	0.094	0.070	0.137	0.079	0.098	0.028	0.111	0.094	0.145
Diet Coke, 2L Bottle	0.012	0.029	0.006	0.027	0.081	0.044	0.072	0.103	0.055	-3.981	0.093	0.043	0.017	0.037	0.019	0.041	0.054	0.041	0.038	0.028	0.055	0.017	0.069	0.036	0.082
Diet Caffeine Free Coke, 12-pack 12oz cans	0.000	-0.057	0.040	0.030	0.099	0.069	0.110	0.180	0.083	0.117	-4.979	0.071	0.042	0.100	0.044	0.047	0.081	0.055	0.093	0.063	0.085	0.018	0.093	0.071	0.104
Diet Caffeine Free Coke, 24-pack 12oz cans	0.003	-0.013	0.048	0.044	0.123	0.089	0.124	0.099	0.115	0.133	0.113	-2.516	0.060	0.057	0.055	0.056	0.084	0.066	0.075	0.071	0.074	0.025	0.131	0.080	0.102
Private Label, 12oz Can	0.041	0.036	0.016	-0.039	0.088	0.057	0.083	0.066	0.044	0.049	0.059	0.044	-2.191	0.068	0.059	0.065	0.069	0.032	0.054	0.047	0.053	0.041	0.050	0.062	0.059
Pepsi, 12-pack 12oz cans	0.060	0.053	0.033	0.033	0.124	0.067	0.097	0.110	0.056	0.069	0.116	0.045	0.060	-4.085	0.060	0.075	0.074	0.055	0.101	0.051	0.042	0.046	0.068	0.066	0.092
Pepsi, 24-pack 12oz cans	0.105	-0.131	0.058	0.055	0.151	0.109	0.110	0.121	0.097	0.108	0.156	0.088	0.125	0.212	-2.591	0.133	0.121	0.150	0.199	0.121	0.137	0.113	0.138	0.124	0.146
Pepsi, 2L Bottle	0.080	-0.052	0.003	0.098	0.071	0.050	0.094	0.061	0.034	0.073	0.053	0.036	0.052	0.069	0.057	-3.286	0.096	0.075	0.089	0.059	0.096	0.063	0.105	0.049	0.087
Pepsi, 6-pack 12oz cans	0.076	0.037	0.033	0.012	0.061	0.064	0.124	0.087	0.050	0.067	0.095	0.045	0.051	0.127	0.047	0.088	-6.029	0.076	0.123	0.059	0.121	0.066	0.109	0.070	0.092
Diet Pepsi, 3L Bottle	-0.002	0.000	0.006	0.013	0.029	0.017	0.044	0.060	0.031	0.057	0.050	0.028	0.019	0.044	0.039	0.050	0.078	-2.972	0.040	0.038	0.096	0.039	0.075	0.042	0.058
Diet Pepsi, 12-pack 12oz cans	0.006	-0.008	0.041	0.040	0.072	0.054	0.039	0.068	0.068	0.077	0.100	0.053	0.051	0.114	0.063	0.094	0.090	0.088	-4.340	0.064	0.117	0.057	0.088	0.091	0.130
Diet Pepsi, 24-pack 12oz cans	0.006	-0.095	0.064	0.090	0.112	0.084	0.065	0.109	0.110	0.114	0.168	0.090	0.084	0.132	0.097	0.118	0.120	0.113	0.147	-3.122	0.126	0.085	0.134	0.112	0.126
Diet Pepsi, 2L Bottle	0.017	0.011	0.011	0.075	0.027	0.022	0.052	0.049	0.032	0.060	0.067	0.029	0.020	0.014	0.036	0.054	0.086	0.069	0.049	0.047	-4.250	0.041	0.069	0.054	0.063
RC, 2L Bottle	0.052	0.074	-0.027	-0.005	0.077	0.061	0.097	0.034	0.024	0.042	0.039	0.021	0.050	0.065	0.066	0.091	0.113	0.087	0.057	0.060	0.095	-1.825	0.094	0.060	0.089
RC, 3L Bottle	0.010	0.021	0.009	0.029	0.024	0.014	0.023	0.030	0.023	0.045	0.035	0.024	0.012	0.022	0.015	0.029	0.055	0.030	0.027	0.020	0.033	0.018	-4.157	0.026	0.068
Diet Rite, 24-pack 12oz cans	0.006	0.074	0.052	0.053	0.096	0.058	0.056	0.135	0.066	0.074	0.092	0.050	0.056	0.069	0.054	0.054	0.097	0.063	0.075	0.057	0.083	0.047	0.103	-3.693	0.128
Diet Rite, 2L Bottle	0.008	0.002	0.009	0.042	0.039	0.016	0.044	0.041	0.027	0.045	0.036	0.019	0.016	0.030	0.019	0.029	0.037	0.035	0.031	0.028	0.037	0.022	0.064	0.041	-4.980

5.2 Store Choice Model

As discussed in section 3.2, I estimate a conditional logit model of store choice, where the choice is conditional on shopping at one of the stores that we observe.[27] The results from eight different specifications of the store choice model are presented in Tables 11, 12, and 13. Specifications I-III have identical demographic variables, but involve progressively longer lagged price indices. Specifications IV-VI are nearly identical to I-III. They differ only in the fact that they include an additional variable that measures whether the household made a purchase at the store in the previous two weeks. Specifications VII and VIII simultaneously incorporate price indices from several periods for a subset of products, both with and without the lagged store choice variable.[28]

The two main results of my analysis of store choice are: (1) that observable demographics significantly affect households' choice of store, even after incorporating a measure of path dependence, and (2) for the product categories for which we have data, households are (at least in the short term) relatively inelastic with respect to store choice.

Table 11 presents the coefficients on the demographic variables in Specifications III and VI. These coefficients remain essentially unchanged with respect to different combinations of price index variables. The demographic variables include an indicator variable equal to one if the household made a purchase (of any kind) at the store in the previous two weeks. This lagged store choice variable accounts for two things: First, it acts like a household-level fixed effect, and second, it accounts for fact that it is easier to shop at a store when you know that store's layout.[29]

With respect to observable demographics, I find that people with lower incomes, were less likely to shop at stores A and B, and more likely to shop at stores C, D and E. Having an unemployed female in the household at the beginning of the sample period was a significant factor in store choice (although having an unemployed male was not), with unemployed female households much more likely to shop at store C. Non-white households were more likely to shop at store A, and households that subscribed to a newspaper were

[27] For two different sets of stores, there are a non-trivial number of households that shop at both stores. This is true for stores $A\&D$ and $B\&C$. I treat going to both stores as a separate alternative. In generating the price variables in this case, I use the lower of the two price indices of the stores in the bundle.

[28] Note: The coefficients are all measured with respect to Store C. If a household made multiple purchases at the same store in the same week, I collapsed these into a single purchase occasion.

[29] In the case of the "bundled" stores, the variable is created slightly differently. For example, if you went to stores A and D last two weeks, then this week the variable would be one for store A, store D, and the bundle of stores A and D. If you only went to Store A in the last two weeks, then this week the variable would be one for store A, zero for store D, and 0.5 for the bundle of stores A and D.

substantially more likely to shop at stores A, D, and to a lesser extent E. They are less likely to shop at store B. Furthermore, the coefficients on these demographic variables are largely invariant to the other aspects of the specification (depending primarily on whether lagged store choice is included). For this reason, I only present these coefficients for Specifications III and VI in Table 11.

Although the effects of expenditure levels on store choice are statistically significant, they are not economically so.[30]

As mentioned in section 3.2, I explored a variety of specifications for the store choice model, including lags of price indices, alternative measures of price, additional demographic variable, and additional index variables measuring the fraction of the category that was featured in the store circular or an end-of-aisle display.

The evidence from the effects of price on store choice were less encouraging, though, as noted earlier, they are in substantial agreement with the literature. The coefficients for price of Cookies and Detergent have the expected sign, and are statistically significant. Unfortunately, although some of the price index variables are significant, many are only significant at the five percent level. Given the number of coefficients, it is not surprising that a subset are statistically significant. Furthermore, many of the price index variables do not have the expected sign. Cat food, bar soap, and yogurt, for example, both have statistically significantly positive coefficients in several specifications. This suggests that these price indices may be capturing effects other than (i.e., that they are correlated with an omitted variable). This gives me less confidence in interpreting the coefficient on soft drinks, which (although the point estimates do not move too wildly) is only significant when I do not account for path dependence.

While I do not report their results here, I also estimated models using alternative price indices, including the Stone price index and a variety of indices measuring the extent of discounts offered. These alternative measures of price did not appear to have any effect on store choice.

[30] I explored using logged expenditure, as well as nonlinear effects from expenditure levels, but the effects were not substantially different.

Table 11: Coefficients on Demographic Variables from Specifications III and VI of the Conditional Logit Model of Store Choice

Specification III

	Store					
	A	A and D	B	B and C	D	E
Constant	4.633	4.595	2.38	1.18	-0.544	-4.904
	[0.190]	[0.235]	[0.203]	[0.279]	[0.276]	[0.296]
Expenditure ($)	-0.012	-0.010	-0.001	0.002	-0.014	-0.012
	[0.000]	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]
Log(Income)	-0.443	-0.507	-0.256	-0.246	0.030	0.452
	[0.019]	[0.024]	[0.020]	[0.028]	[0.027]	[0.028]
Unemployed	-0.768	-0.479	-0.070	0.209	-0.640	-1.110
Female	[0.038]	[0.049]	[0.036]	[0.049]	[0.055]	[0.060]
Unemployed	-0.053	-0.542	-0.427	-0.176	-2.183	0.127
Male	[0.069]	[0.097]	[0.078]	[0.105]	[0.256]	[0.106]
Non-white	0.921	0.243	0.271	-0.710	-0.907	-0.819
	[0.035]	[0.045]	[0.037]	[0.063]	[0.067]	[0.065]
Subscribes to	0.428	0.089	-0.463	-0.311	0.361	0.150
Newspaper	[0.034]	[0.045]	[0.038]	[0.052]	[0.044]	[0.043]
Household Has	-0.920	-16.726	-1.071	1.752	-16.315	-0.807
No Kids	[0.209]	[742.304]	[0.248]	[0.170]	[713.396]	[0.299]

Specification VI
Adds an Indicator Variable for Whether the
Household Visited that Store in the Last Two Weeks)

	Store					
	A	A and D	B	B and C	D	E
Constant	2.586	3.887	2.462	1.849	-0.679	-1.668
	[0.309]	[0.321]	[0.294]	[0.306]	[0.387]	[0.432]
Expenditure ($)	-0.011	-0.007	0.000	0.002	-0.013	-0.014
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Log(Income)	-0.263	-0.410	-0.261	-0.260	0.027	0.178
	[0.030]	[0.032]	[0.029]	[0.030]	[0.037]	[0.041]
Unemployed Female	-0.552	-0.306	-0.026	0.288	-0.586	-0.631
	[0.062]	[0.065]	[0.052]	[0.054]	[0.080]	[0.086]
Unemployed Male	0.331	-0.027	-0.278	-0.013	-1.42	0.091
	[0.120]	[0.131]	[0.112]	[0.113]	[0.280]	[0.165]
Non-white	0.551	0.093	0.090	-0.747	-0.642	-0.538
	[0.056]	[0.060]	[0.054]	[0.066]	[0.082]	[0.086]
Subscribes to	0.450	0.151	-0.400	-0.254	0.475	0.373
Newspaper	[0.055]	[0.059]	[0.053]	[0.056]	[0.065]	[0.067]
Does not Have Kids	1.088	-14.635	-0.368	1.773	-13.314	1.455
	[0.310]	[631.998]	[0.279]	[0.197]	[676.928]	[0.372]

Table 12: Coefficients on Price Index Variables for Specifications I-VI of Conditional Logit Model of Store Choice. Standard errors are in brackets.

	Specification					
	I	II	III	IV	V	VI
Product Category	Current Prices	Prices Lagged 1 Week	Prices Lagged 2 Weeks	Current Prices	Prices Lagged 1 Week	Prices Lagged 2 Weeks
Bacon	0.022	-0.005	-0.036	0.019	-0.042	-0.063
	[0.023]	[0.023]	[0.023]	[0.032]	[0.032]	[0.032]
BBQ Sauce	-0.121	-0.019	0.139	-0.101	0.087	0.180
	[0.093]	[0.094]	[0.095]	[0.128]	[0.130]	[0.132]
Butter	-0.065	-0.050	-0.023	-0.002	-0.069	-0.056
	[0.067]	[0.068]	[0.068]	[0.092]	[0.094]	[0.094]
Cat Food	1.217	0.993	1.070	0.286	0.181	0.718
	[0.284]	[0.288]	[0.283]	[0.412]	[0.422]	[0.417]
Cereal	-0.123	0.006	0.166	-0.139	0.159	0.240
	[0.051]	[0.052]	[0.052]	[0.074]	[0.074]	[0.075]
Cleansers	0.129	0.101	-0.041	0.039	-0.055	-0.217
	[0.061]	[0.062]	[0.062]	[0.088]	[0.091]	[0.091]
Coffee	0.094	0.060	0.010	0.123	-0.007	-0.001
	[0.035]	[0.036]	[0.036]	[0.049]	[0.050]	[0.049]
Cookies	-0.468	-0.479	-0.526	-0.437	-0.291	-0.410
	[0.071]	[0.071]	[0.071]	[0.101]	[0.103]	[0.102]
Crackers	-0.023	-0.01	0.038	-0.090	-0.001	0.065
	[0.043]	[0.043]	[0.043]	[0.062]	[0.062]	[0.062]
Detergents	-0.064	-0.076	-0.042	-0.095	-0.003	0.026
	[0.020]	[0.020]	[0.020]	[0.028]	[0.028]	[0.028]
Eggs	0.073	0.149	0.011	-0.101	0.051	-0.303
	[0.057]	[0.058]	[0.057]	[0.084]	[0.085]	[0.083]
Fabric Softener	-0.086	-0.086	-0.091	-0.131	-0.157	-0.065
	[0.054]	[0.055]	[0.054]	[0.076]	[0.078]	[0.078]
Frozen Pizza	-0.010	-0.028	-0.032	-0.031	-0.026	-0.042
	[0.031]	[0.033]	[0.032]	[0.042]	[0.046]	[0.045]
Hot Dogs	0.015	0.059	0.105	-0.028	0.043	0.122
	[0.032]	[0.032]	[0.032]	[0.044]	[0.045]	[0.045]
Ice Cream	-0.051	-0.048	-0.002	-0.015	-0.035	0.029
	[0.036]	[0.037]	[0.036]	[0.049]	[0.051]	[0.050]
Peanut Butter	-0.091	-0.093	0.003	-0.194	-0.072	0.092
	[0.045]	[0.046]	[0.044]	[0.064]	[0.065]	[0.061]
Snacks	-0.173	-0.247	-0.270	-0.160	-0.198	-0.237
	[0.061]	[0.061]	[0.061]	[0.085]	[0.085]	[0.085]
Bar Soap	0.304	0.252	0.288	0.465	0.324	0.470
	[0.064]	[0.065]	[0.064]	[0.089]	[0.089]	[0.088]
Soft Drinks	-0.041	-0.101	-0.082	-0.026	-0.037	-0.019
	[0.031]	[0.032]	[0.032]	[0.044]	[0.045]	[0.045]
Sugarless Gum	0.007	-0.012	0.058	0.026	0.058	0.209
	[0.069]	[0.069]	[0.069]	[0.094]	[0.096]	[0.096]
Toilet Tissue	-0.066	-0.003	-0.002	-0.042	0.055	-0.008
	[0.045]	[0.045]	[0.046]	[0.063]	[0.063]	[0.064]
Yogurt	0.432	0.195	0.681	0.149	-0.053	0.869
	[0.105]	[0.105]	[0.105]	[0.148]	[0.150]	[0.151]
Shopped at Store in Past 2 Weeks				4.372	4.376	4.386
				[0.029]	[0.029]	[0.029]
Number of Observations	277,011	274,337	273,489	277,011	274,337	273,489
Pseudo R2	0.124	0.1241	0.1246	0.5243	0.5282	0.5291
Log Likelihood	-67,458	-66,800	-66,586	-36,632	-35,982	-35,821
Effect on Market Share from a increase in Soft Drink Prices:						
Before:	.3405	.3407	.3414	.3405	.3407	.3412
After:	.3395	.3384	.3395	.3402	.3404	.3412

5.3 Counter-Factuals

5.3.1 Marginal Costs

In order to use the estimated demand system to recover estimates of the expected profits lost from uniform pricing, I need to make an assumption about the actual price-setting

Table 13: Coefficients on Price Indices for Specifications VII and VIII of Conditional Logit Model of Store Choice. Standard errors are in brackets.

Product Category	Specification VII Current Prices	Prices Lagged 1 week	Prices Lagged 2 weeks	Specification VIII Current Prices	Prices Lagged 1 week	Prices Lagged 2 weeks
Cat Food	0.657	0.469	0.715	-0.192	0.283	0.681
	[0.317]	[0.335]	[0.316]	[0.461]	[0.485]	[0.460]
Cereal	-0.129	-0.019	0.126	-0.153	0.121	0.247
	[0.054]	[0.055]	[0.055]	[0.077]	[0.077]	[0.077]
Coffee	0.091	0.009	0.002	0.162	-0.066	-0.003
	[0.036]	[0.036]	[0.036]	[0.050]	[0.049]	[0.050]
Cookies	-0.372	-0.189	-0.365	-0.391	-0.021	-0.248
	[0.078]	[0.081]	[0.077]	[0.112]	[0.117]	[0.112]
Detergents	-0.049	-0.036	-0.014	-0.076	0.042	0.035
	[0.021]	[0.021]	[0.021]	[0.029]	[0.029]	[0.029]
Eggs	0.023	0.148	-0.043	-0.183	0.112	-0.337
	[0.060]	[0.061]	[0.059]	[0.088]	[0.089]	[0.086]
Hot Dogs	-0.036	0.013	0.074	-0.060	0.005	0.105
	[0.032]	[0.032]	[0.033]	[0.045]	[0.045]	[0.047]
Peanut Butter	-0.034	-0.064	0.025	-0.131	-0.072	0.090
	[0.048]	[0.049]	[0.046]	[0.066]	[0.068]	[0.063]
Salty Snacks	-0.069	-0.220	-0.134	-0.152	-0.122	-0.135
	[0.069]	[0.072]	[0.069]	[0.095]	[0.096]	[0.094]
Bar Soap	0.241	0.035	0.198	0.362	-0.075	0.384
	[0.077]	[0.083]	[0.075]	[0.106]	[0.113]	[0.103]
Fabric Softener	-0.064	-0.081	-0.025	-0.126	-0.097	0.008
	[0.059]	[0.062]	[0.058]	[0.083]	[0.087]	[0.083]
Soft Drinks	-0.008	-0.101	-0.081	0.010	-0.070	-0.018
	[0.033]	[0.033]	[0.033]	[0.046]	[0.045]	[0.046]
Yogurt	0.327	0.037	0.586	0.075	-0.182	0.812
	[0.109]	[0.108]	[0.107]	[0.156]	[0.154]	[0.153]
Shopped at Store in Past 2 Weeks						4.389 [0.029]
Number of Observations		273,489			273,489	
Pseudo R2		0.1255			0.5296	
Log Likelihood		-66,524			-35,786	
Effect on Market Share from a increase in Soft Drink Prices:						
Before:		0.3414			0.3414	
After:		0.3370			0.3406	

behavior of the retailer during the sample period. The assumption I choose to make is that the retailer maximizes total weekly profit for the soft drink category, and charges the profit-maximizing price for each product in each week. This assumption implies the following J first-order conditions (one for each good j) for each week t:

$$\frac{\partial \Pi_{t,Non-Uniform}}{\partial p_{jt}} = E_t\left[Q_{jt}(\mathbf{p}_t)\right] + \sum_{k \in J} (p_{kt} - c_{kt}) \frac{\partial E_t\left[Q_{kt}(\mathbf{p}_t)\right]}{\partial p_{jt}} = 0 \qquad (15)$$

By taking these first-order conditions numerically[31], I am able to solve the system of J equations and J unknowns for each week – the c_{jt}'s – and recover the implied weekly marginal costs for each product. Note that in recovering these marginal costs, the level of Q_{jt} drops out. That is, the implied marginal cost is independent of the total number of households shopping in that week.

In recovering the marginal costs from the first-order conditions of the retailer, I am implicitly assuming that the demand system that I have estimated is the true demand system (and by association, that it is the demand system that store A used in setting its prices), and that in each week, the retailer knows the distribution of the budgets of the households. Solving this system of equations, gives me the implied marginal cost c_{jt} for each good, *at store A, during week t*.

Table 14 contains summary statistics for these implied marginal costs. In general, the estimated marginal costs are substantially lower than the wholesale prices reported in the Dominick's dataset (taken from a geographically proximate competing grocery retailer) and shown in Table 6. This discrepancy may be explained by the fact that store A is part of a large chain, and therefore may have received preferential wholesale prices. Additionally, these implied marginal costs may be capturing the effects of slotting allowances or non-linearities in wholesale prices (such as block discounts) not accounted for in the Dominick's data (see Israilevich (2004)).

If I am underestimating the true marginal costs, the likely source would be that suggest that either retailers are pricing non-optimally with respect to the soft drink category, that the estimated demand model is incorrect, or that the assumed supply model is incorrect (e.g., retailers may be engaging in cross-category subsidization).

[31] I do this by (1) choosing a fixed number of households, (2) simulating demand from these households at the observed prices at store A in week t, (3) numerically taking the derivatives of demand for each good with respect to all other goods.

Table 14: Summary Statistics for Marginal Costs (in Dollars per 12oz Serving) Implied by the Model

Product	Mean Marginal Cost	Std. Dev. of Marginal Cost	Number of Weeks w/ Negative Costs	Mean Markup ($)	Std. Dev. of Markup	Mean Margin (%)
7up, 2L Bottle	0.123	0.090	6	0.108	0.035	53.2
7up, 6-pack 12oz cans	0.310	0.204	2	0.081	0.135	27.7
Diet 7up, 24-pack 12oz cans	0.148	0.088	3	0.119	0.032	49.1
Diet 7up, 2L Bottle	0.148	0.096	3	0.083	0.046	41.7
Coke, 12-pack 12oz cans	0.160	0.120	2	0.179	0.048	58.5
Coke, 24-pack 12oz cans	0.093	0.079	12	0.181	0.026	71.1
Coke, 2L Bottle	0.099	0.075	10	0.150	0.027	65.3
Diet Coke, 12-pack 12oz cans	0.165	0.133	3	0.179	0.054	58.8
Diet Coke, 24-pack 12oz cans	0.085	0.084	19	0.189	0.031	74.8
Diet Coke, 2L Bottle	0.094	0.080	16	0.154	0.032	67.6
Diet Caffeine Free Coke, 12-pack 12oz cans	0.165	0.127	3	0.180	0.048	58.4
Diet Caffeine Free Coke, 24-pack 12oz cans	0.080	0.080	18	0.190	0.029	75.5
Private Label, 12oz Can	0.036	0.034	19	0.154	0.018	82.9
Pepsi, 12-pack 12oz cans	0.149	0.130	3	0.190	0.078	61.6
Pepsi, 24-pack 12oz cans	0.082	0.077	14	0.191	0.024	75.0
Pepsi, 2L Bottle	0.087	0.068	14	0.161	0.023	69.3
Pepsi, 6-pack 12oz cans	0.249	0.136	1	0.164	0.080	43.7
Diet Pepsi, 3L Bottle	0.077	0.014	0	0.123	0.014	61.6
Diet Pepsi, 12-pack 12oz cans	0.170	0.137	4	0.171	0.088	55.6
Diet Pepsi, 24-pack 12oz cans	0.109	0.074	9	0.167	0.023	64.7
Diet Pepsi, 2L Bottle	0.111	0.067	3	0.134	0.023	58.5
RC, 2L Bottle	0.014	0.035	29	0.165	0.020	93.8
RC, 3L Bottle	0.084	0.033	6	0.108	0.017	58.1
Diet Rite, 24-pack 12oz cans	0.116	0.062	1	0.147	0.023	58.7
Diet Rite, 2L Bottle	0.105	0.068	3	0.113	0.025	56.0

As is frequently the case (see Villas-Boas (2002)) my estimates imply that in some weeks, for some products, marginal costs are negative. Although this seems economically bizarre, in theory this result could be explained by slotting allowances. In practice, (see Israilevich (2004)) the arrangements between retailers and manufacturers regularly involves nonlinear contracting schemes such as block discounts (which imply negative marginal costs over some regions). To the extent that my assumption of constant marginal costs is violated, I may be picking up some of these nolinearities.

Figure 5 shows a typical path of prices and implied marginal costs over time, The key feature to notice here is that intertemporal marginal cost (e.g., wholesale price) variation is responsible for nearly all of the inter-temporal price variation. This feature is mirrored in the Dominick's data. Figure 6 plots the average markup in cents per 12oz serving over the sample period implied by the derived marginal costs. The average average markup across products appears to be roughly fourteen cetns per 12oz serving, with occasional spikes upwards (and one large spike downwards). Together with the low standard deviations on margins shown in Table 14, this also agrees with what is observed in the Dominick's data.

5.3.2 "Optimal" Uniform Prices

In order to calculate the profits the firm would have earned by following a uniform pricing strategy, I must first solve for the "optimal" uniform prices. I do this by restricting the prices each week to be uniform by manufacturer-brand-size[32]. Then for each week, I numerically solve for the set of prices that maximizes expected profits, subject to this restriction.

Table 15 presents summary statistics on the differences between these "optimal" uniform prices and the non-uniform prices actually charged by the retailer. Contrary to (my) expectations, the majority of the differences were not uniformly positive or negative. That is, in some weeks the non-uniform price was higher than the optimal uniform price, while in other weeks it was lower. In hindsight, this is actually suggested by the variation in the price ordering of the varieties in Figure 4. Furthermore, for all but two products (2L containers of RC and Diet Rite), the average difference between the non-uniform and the optimal uniform prices was less than one cent per 12oz serving.

[32]For example, I restrict 12-packs of 12oz cans of Coke, Diet Coke, and Diet Caffeine Free Coke to all sell at the same price each week, although I allow this price to vary across weeks.

Figure 5: Graph of the Price and Implied Marginal Cost (in cents per 12oz serving) for a 2L Bottle of Regular Pepsi, 6/91-6/93

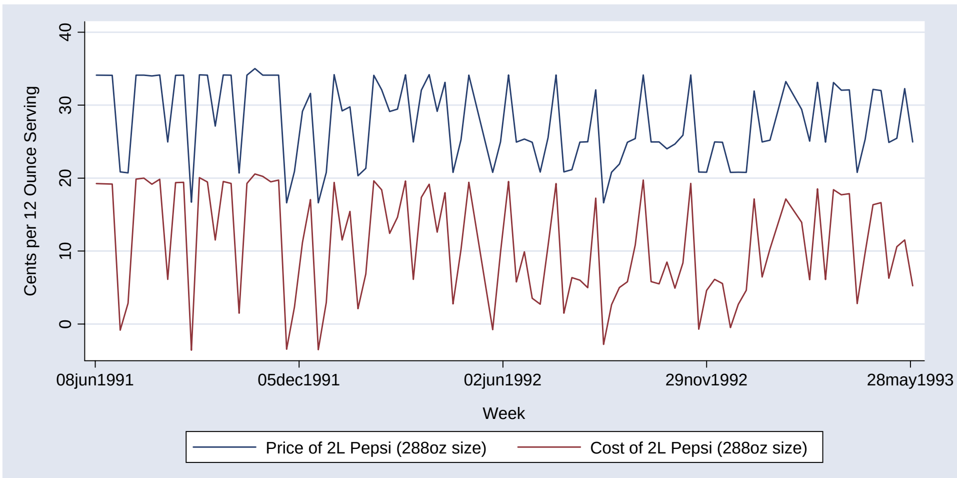


Figure 6: Graph of the Average Markup (in cents per 12oz serving) Across Products, 6/91-6/93

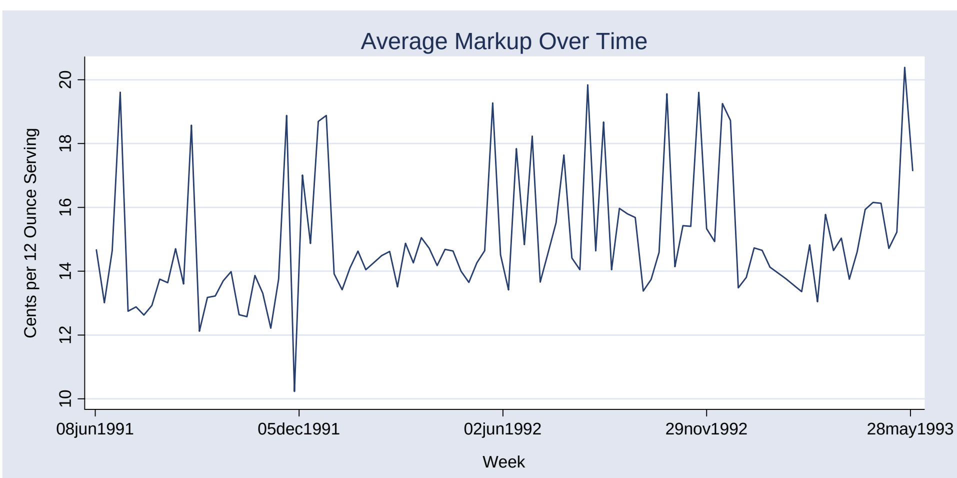


Figure 7: Graph of the Maximum Difference Across Products (in cents per 12oz serving) Between a Product's Uniform and Non-Uniform Prices, 6/91-6/93

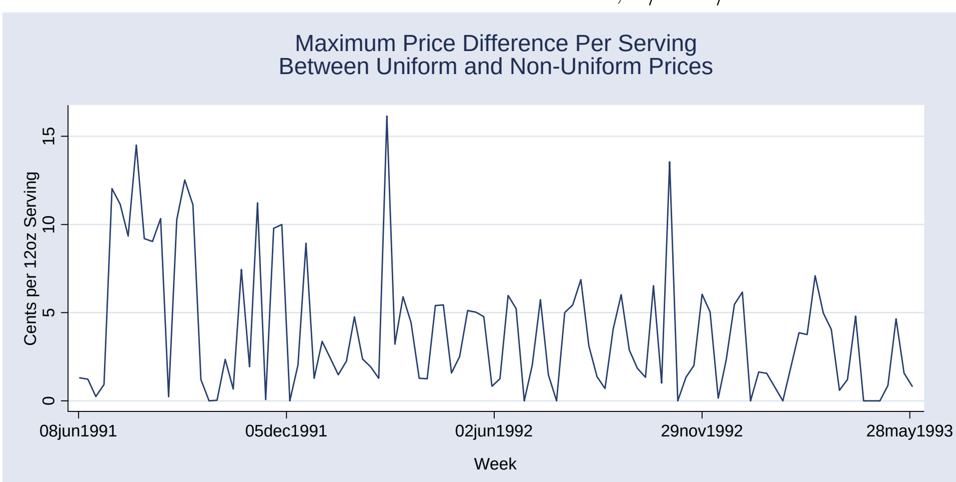


Table 15: Summary Statistics on the Differences Between Observed Non-Uniform Prices and "Optimal" Uniform Prices (in Dollars per 12oz Serving)

Product	Mean Difference	Std. Dev. of Difference	Greatest Increase from Uniform	Greatest Decrease from Uniform
7up, 2L Bottle	-0.001	0.015	0.145	0.028
Diet 7up, 2L Bottle	-0.001	0.016	0.152	0.023
Coke, 12-pack 12oz cans	0.004	0.024	0.053	0.161
Coke, 24-pack 12oz cans	0.000	0.018	0.111	0.135
Coke, 2L Bottle	0.001	0.005	0.021	0.023
Diet Coke, 12-pack 12oz cans	-0.002	0.012	0.054	0.035
Diet Coke, 24-pack 12oz cans	0.001	0.011	0.046	0.084
Diet Coke, 2L Bottle	0.001	0.006	0.021	0.018
Diet Caffeine Free Coke, 12-pack 12oz cans	-0.001	0.010	0.051	0.030
Diet Caffeine Free Coke, 24-pack 12oz cans	-0.001	0.007	0.046	0.016
Pepsi, 12-pack 12oz cans	-0.005	0.031	0.232	0.031
Pepsi, 24-pack 12oz cans	0.000	0.004	0.017	0.019
Pepsi, 2L Bottle	-0.001	0.005	0.031	0.011
Diet Pepsi, 12-pack 12oz cans	-0.001	0.022	0.184	0.046
Diet Pepsi, 24-pack 12oz cans	-0.002	0.007	0.039	0.013
Diet Pepsi, 2L Bottle	0.002	0.009	0.019	0.065
RC, 2L Bottle	0.011	0.015	0.014	0.056
Diet Rite, 2L Bottle	-0.028	0.037	0.125	0.023

5.3.3 Profit Differences

Using my marginal cost estimates and the actual quantities sold, I can estimate the profits that the store A *actually earned* in each week. In addition, by simulating expected demand, I can also calculate the profits that the firm *expected to earn* each week. The difference between these two numbers is that the former is scaled by the number of shoppers who actually went shopping in that week. These give me a measure of the profits earned by the firm under the non-uniform pricing regime. Comparing the these expected profit figures yields a weekly estimate of the percentage profit decrease that store A would have experienced if it had charged uniform prices.

Table 16 shows the detailed results of these calculations for a typical week of the sample: the week beginning July 7, 1991. Several features are apparent. The first is that demand is strongly skewed towards the lowest priced products. The ten products priced at $0.21 cents per 12oz serving or lower sell by far the lasrgest share of the quantity. The second feature is that many of the prices are the same or nearly the same under both uniform and non-uniform pricing policies. In this week, store A actually charged the same price for 24-packs of 12oz cans of both Coke and Diet Coke. Because I assume (in order to identify the marginal costs) that the retailer charged the optimal prices in each week, the results are skewed towards finding a smaller estimate of the profit difference. Third, much of the increase in profits comes from a significant *decrease* in the price of a single product: 2L bottles of Royal Crown (RC) cola. Finally, I note in passing that demand for some goods increased, in spite of an increase in the price going from the uniform to the non-uniform. This can be attributed to the effects of cross-price elasticities – the prices of many other goods also moved.

Table 16: Uniform and Non-Uniform Prices, Quantities and Estimated Profits for the Week of July 7, 1991. Prices and marginal costs reported are in cents per 12oz Serving. Quantity is measured in 12oz servings. Profits are measured in dollars. All prices and profits are in nominal terms. The total difference in profits for the week from the two pricing strategies is $61.52. The 3L size of Pepsi was not offered in this week.

Product	$p_{Uniform}$	$Q_{Uniform}$	$p_{Non-Uniform}$	$Q_{Non-Uniform}$	Marginal Cost	$\Pi_{Uniform}$ ($)	$\Pi_{Non-Uniform}$ ($)
7up, 2L Bottle	31.9	58	31.8	57	24.9	4.05	3.94
7up, 6-pack 12oz cans	50.9	12	48.0	24	58.5	-0.90	-2.55
Diet 7up, 24-pack 12oz cans	20.8	5381	20.8	5406	6.6	761.30	764.38
Diet 7up, 2L Bottle	31.9	8	31.8	10	29.1	0.23	0.25
Coke, 12-pack 12oz cans	47.1	57	47.0	58	33.5	7.80	7.76
Coke, 24-pack 12oz cans	20.8	5609	20.8	5538	0.7	1126.52	1115.18
Coke, 2L Bottle	32.0	83	31.8	85	18.8	10.88	10.98
Diet Coke, 12-pack 12oz cans	47.1	81	47.0	81	32.8	11.58	11.48
Diet Coke, 24-pack 12oz cans	20.8	5788	21.4	5463	-0.7	1242.81	1209.34
Diet Coke, 2L Bottle	32.0	83	31.9	86	20.4	9.58	9.82
Diet Caffeine Free Coke, 12-pack 12oz cans	47.1	70	47.0	71	32.9	10.01	9.99
Diet Caffeine Free Coke, 24-pack 12oz cans	20.8	5944	20.8	5896	-0.7	1272.49	1265.51
Private Label, 12oz Can	17.0	3050	17.0	3016	1.1	484.78	479.22
Pepsi, 12-pack 12oz cans	48.8	68	47.0	87	31.0	12.04	13.88
Pepsi, 24-pack 12oz cans	20.8	9099	20.8	8964	0.0	1887.23	1862.89
Pepsi, 2L Bottle	31.8	284	31.8	274	19.0	36.41	35.04
Pepsi, 6-pack 12oz cans	50.4	108	48.0	151	40.8	10.27	10.80
Diet Pepsi, 12-pack 12oz cans	48.8	51	47.0	69	31.6	8.79	10.69
Diet Pepsi, 24-pack 12oz cans	20.8	8337	20.7	8283	2.9	1492.67	1479.67
Diet Pepsi, 2L Bottle	31.8	94	31.8	90	20.5	10.75	10.20
RC, 2L Bottle	19.7	2703	14.1	4495	-1.7	578.59	708.66
RC, 3L Bottle	20.1	298	20.0	296	9.7	30.78	30.44
Diet Rite, 24-pack 12oz cans	32.5	285	31.2	348	18.9	38.73	42.75
Diet Rite, 2L Bottle	19.7	367	31.7	16	24.5	-17.43	1.16

The weekly expected differences in profits are presented in Figures 8 and 9. I estimate a distinct mass point at zero lost profits. As noted above, this is largely due to my assumption that, in each week, the prices charged by store A were optimal. This assumption implies that for weeks in which store A actually charged uniform prices, it could not have expected to lose any profits doing so.

My estimates imply that by charging uniform, rather than non-uniform prices, the retailer would have lost \$36.56 per week in profit, or a total of \$3,803 over the two year period of my sample. The prospect of earning an additional \$3,803 in profits (roughly \$5,135 2004 dollars) over a two-year period for the soft drink category may seem small, but this is only a single store in a much larger chain of more than 100 stores. If the chain were able to realize similar profit increases at other stores in the chain, a rough estimate of the profit increase would be over \$250,000 dollars per year in 2004 dollars. This would presumably be more than enough to hire an empirical economist to determine the optimal prices for each product in each store in each week. Furthermore, this estimate is solely for the soft drink category. While it is not clear what the results would be for other categories, similar profit increases may be possible.

6 Interpreting "Lost" Profits

6.1 Menu Costs

As mentioned in the introduction, if we set aside demand-side explanations, the two reasons for retailers to charge uniform prices are: to reduce menu costs and to soften price competition with other retailers. In the event that the difference between uniform and non-uniform prices is close to zero, this would suggest that retailers do not expect to lose much (if any) profit by charging uniform prices. On the other hand, if the predicted profit differences are positive, we must try to differentiate between these (and potentially other) explanations for the hypothetical "lost profits". When talking to store managers, the most frequently offered explanation for the observed price uniformity is some form of menu costs. When pressed, Safeway store managers respond that the reason for uniform pricing is that it is "too much trouble" to price every good separately. In understanding what is meant by "too much trouble" it is important to distinguish between two different kinds of menu costs: physical menu costs and managerial menu costs.

One type of menu cost comes from the costs associated with *physically* changing prices.

According to Tony Mather, Director Business Systems, Safeway (U.K.): "Pricing at the moment is very labor-intensive. Shelf-edge labels are batch printed, manually sorted and changed by hand while customers are out of the store."[33] Levy et al. (1997) estimate the average menu costs to a large chain-owned grocery store for physically changing a single price tag to be $0.52. To put this in perspective, a typical large grocery store usually changes the price tags on about 4,000 items each week, changing as many as 14,000 tags in some weeks. Although their study was conducted on behalf of a company selling electronic price display tags, their stated aim was to put a lower bound on menu costs and they report that grocery store executives generally agreed with their findings.

One might think physical menu costs promote uniform pricing – that stores reduce their physical menu costs by charging uniform prices. However, two pieces of evidence suggest that physical menu costs do not explain uniform prices. First, grocery stores typically post prices for every SKU *even when they are uniformly priced*. Hence, the physical menu costs are the same, regardless of whether the prices are priced uniformly or non-uniformly. Second, in cases where the physical menu cost is presumably small or insignificant, we still observe uniform prices. Even grocery stores that have implemented electronic display tags and that can change prices throughout the store at the touch of a button from the store's central computer continue to charge uniform prices. Furthermore, online grocery stores – who presumably have nearly zero physical menu costs – also sell at uniform prices.

A second kind of menu cost, and one that has not typically been discussed in the literature is the *managerial cost* associated with figuring out what price to charge for that product. While academic papers generally assume that retailers learn optimal prices costlessly, this is an abstraction from reality. In order to learn its demand function, a retailer must experiment by charging a variety of prices – introducing exogenous price variation – and this experimentation can be costly. In addition, the retailer may have to hire personnel or consulting services to determine "optimal" prices. These costs may not be insubstantial. A recent article in Business Week (Keenan 2003) suggests that implementing the advanced techniques offered by pricing consultants typically requires a "12-month average installation" time and a price that "start[s] at around $3 million." If the additional expected profit to be gained from charging different prices for two products is less than the cost of figuring out what those prices should be, then we will see uniform prices.

[33] http://www.symbol.com/uk/Solutions/case_study_safeway.html

This suggests that the store's choice of whether to follow a uniform or non-uniform pricing strategy is more likely a long-term decision rather than a week-by-week decision. In this case, the relevant cost to consider is the present discounted value of the sum of the lost expected profits across weeks and represents the one-time or infrequent cost either of experimentation or consulting services.

Managerial menu costs also suggest a reason that pricing strategies may vary across stores – leading some stores to charge uniform prices while others charge non-uniform prices. Pricing decisions for most large grocery chains are made at the chain level. Store managers at these chains typically receive the week's prices electronically from company headquarters, and are only responsible for making sure that price labels are printed and placed on shelves. This centralization allows large chains to spread out these managerial costs across many stores. However, evidence suggests that even large chains may be influenced by managerial menu costs. Chintagunta et al. (2003) document the fact that Dominick's Finer Foods grouped its stores into three different categories based on the levels of competition the stores faced, with each of roughly one hundred stores charging one of three menus of prices. Such pricing heuristics presumably lower managerial costs by reducing the dimensionality of the optimal pricing problem, but at the cost of non-optimal prices.

Other pricing heuristics seem to be in widespread use. Both small retailers and large grocery stores[34] frequently use constant-markup pricing heuristics, such as pricing all goods at wholesale cost plus a fixed percentage or amount. The apparent widespread use of these pricing heuristics may explain why soft drink prices tend to vary dramatically over time, but not cross-sectionally – while wholesale prices move a good deal over time, wholesale prices are typically uniform within manufacturer-brand. Unfortunately, this raises the question of why manufacturers would choose to price their products uniformly.

[34]Data suggests that Dominick's Finer Foods (described in section 5) frequently followed a constant-markup pricing strategy.

Figure 8: Graph of the Difference Between Profits from Uniform and Non-Uniform Price Strategies, as a Percent of the Profits Earned at Non-Uniform Prices, 6/91-6/93

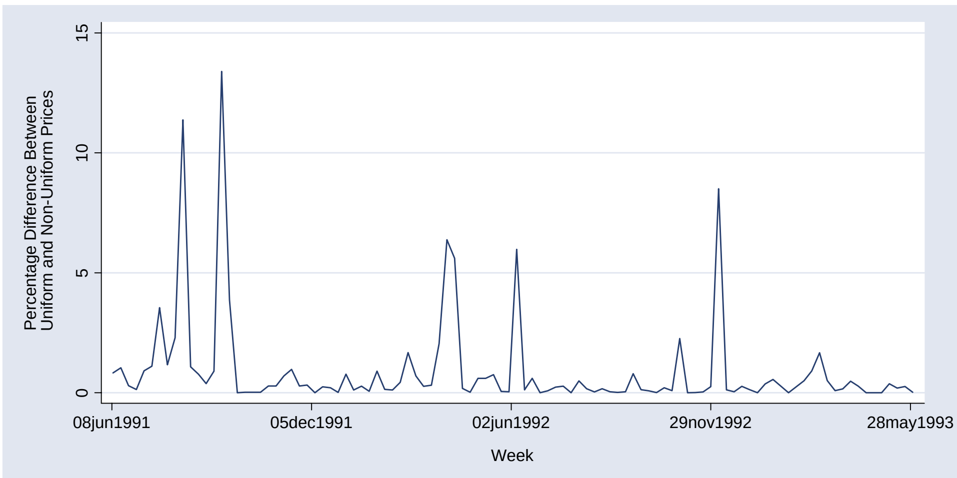
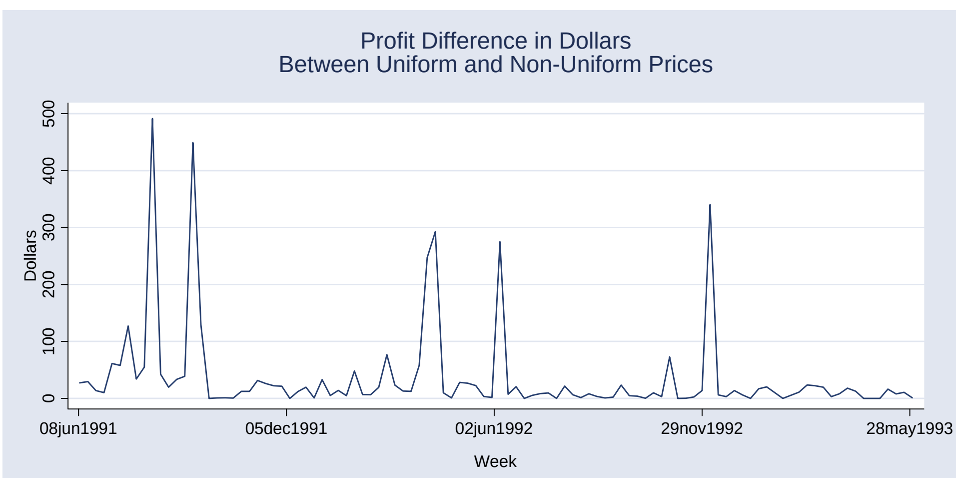


Figure 9: Graph of the Counterfactual Dollars Lost from Charging Uniform Prices, 6/91-6/93



7 Conclusion

In retail environments, many differentiated products are sold at uniform prices. Explanations for this behavior can be grouped into demand-side and supply-side explanations. Lacking the necessary data to investigate demand-side explanations, I look at supply-side explanations. Using grocery store scanner panel data and household grocery purchase histories, I examine the market for carbonated soft drinks – a product that is frequently, but not always, sold at uniform prices – and evaluate the validity of several supply-side explanations.

Using the demand system developed in McMillan (2005), I conduct the counter-factual experiment of forcing the prices of a particular store to be uniform, and comparing the resulting profits to the non-uniform case. This comparison of the expected profit earned by a single retailer at the weekly prices actually charged to the expected profit that same retailer would have earned, had it charged uniform prices that were optimal in each week (subject only to the restriction that they be uniform by manufacurer-brand-size), I am able to infer that the retailer would have experienced a profit loss of roughly \$36.56 (in 1992 dollars) per week if it had charged uniform, rather than non-uniform prices. This suggests that uniform pricing would have led to a total profit loss for the retailer over the two year sample period, of roughly \$5,135 in 2004 dollars.

This result suggests that there are additional profits to be earned from non-uniform pricing, under the assumption that the retailer charged optimal prices. Clearly, however, it may not be profitable for single-store retailers to take advantage of this opportunity. Without the benefits of multiple stores over which to spread the managerial costs of determining optimal prices, single store retailers may find it optimal to charge uniform prices. Unfortunately, this "scale" explanation cannot be the whole story. Anecdotal evidence suggests that Walmart charges uniform prices for many products, even though that company has almost certainly realized most returns to scale with respect to managerial menu costs. Although additional research is necessary regarding demand-side reactions to non-uniform pricing, these results suggest that pricing managers, particularly those at large retail chains, should be aware of potential additional profits available from non-uniform pricing. Moreover, they suggest that for single-store retailers, relatively small managerial menu costs are able to generate the behavior observed at many retailers – that of uniform prices.

References

Ball, L. & Mankiw, N. G. (2004), ‘A sticky-price manifesto’, *NBER Working Paper* (4677).

Bayus, B. L. & Putsis, W. P. (1999), ‘Product proliferation: An empirical analysis of product line determinants and market outcomes’, *Marketing Science* **18**(2), 137–153.

Bell, D., Ho, T.-H. & Tang, C. (1998), ‘Determining where to shop: Fixed and variable costs of shopping’, *Journal of Marketing Research* **35**(3), 352–369.

Bell, D. R. & Lattin, J. M. (1998), ‘Shopping behavior and consumer preference for store price format: Why “large basket” shoppers prefer edlp’, *Marketing Science* **78**, 66–88.

Berry, S., Levinsohn, J. & Pakes, A. (1995), ‘Automobile prices in market equilibrium’, *Econometrica* pp. 841–890.

Burstiner, I. (1997), *The Small Business Handbook: a Comprehensive Guide to Starting and Running Your Own Business*, third edn, Simon and Schuster.

Canetti, E., Blinder, A. & Lebow, D. (1998), *Asking About Prices: A New Approach to Understanding Price Stickiness*, Russell Sage Foundation Publications.

Carlton, D. W. (1989), *The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?*, Vol. 1 of *The Handbook of Industrial Organization*, Elsevier Science Publishers, chapter 15, pp. 909–946.

Chan, T. Y. (2002), ‘Estimating a continuous hedonic choice model with an application to demand for soft drinks’. Mimeo, Washington University, St. Louis, Olin School of Business.

Chiang, J., Chung, C.-F. & Cremers, E. T. (2001), ‘Promotions and the pattern of grocery shopping time’, *Journal of Applied Statistics* **28**(7), 801–819.

Chintagunta, P., Dubé, J.-P. & Singh, V. (2003), ‘Balancing profitability and customer welfare in a supermarket chain’, *Quantitative Marketing and Economics* **1**, 111–147.

Corts, K. (1998), ‘Third degree price discrimination in oligopoly: All-out competition and strategic commitment’, *RAND Journal of Economics* **29**(2), 306–323.

Draganska, M. & Jain, D. (2001), Product line length decisions in a competitive environment. Mimeo, Stanford University, Graduate School of Business.

Dubé, J.-P. (2001), Multiple discreteness and product differentiation: Strategy and demand for carbonated soft drinks. Mimeo, University of Chicago, Graduate School of Business.

Gouriéroux, C. & Monfort, A. (1996), *Simulation-Based Econometric Methods*, Oxford University Press.

Hauser, J. R. & Wernerfelt, B. (1990), 'An evaluation cost model of consideration sets', *The Journal of Consumer Research* **16**(4), 393–408.

Hays, C. L. (1997), 'Variable-price coke machine being tested', *The New York Times* p. C1. October 28.

Hendel, I. & Nevo, A. (2002*a*), Measuring the implications of sales and consumer stockpiling behavior. Mimeo, University of Wisconsin, Madison and University of California, Berkeley.

Hendel, I. & Nevo, A. (2002*b*), 'Sales and consumer inventory', *NBER Working Paper* (9048).

Hess, J. D. & Gerstner, E. (1987), 'Loss leader pricing and rain check policy', *Marketing Science* **6**(4), 358–374.

Heun, C. T. (2001), 'Dynamic pricing boosts bottom line', *Information Week* (861), 59.

Ho, T.-H., Tang, C. S. & Bell, D. R. (1998), 'Rational shopping behavior and the option value of variable pricing', *Marketing Science* **44**(12), 115–160.

Hoch, S. J., Dreze, X. & Purk, M. (1994), 'EDLP, hi-lo, and margin arithmetic', *Journal of Retailing* **58**, 16–27.

Israilevich, G. (2004), 'Assessing product-line decisions with supermarket scanner data', *Quantitative Marketing and Economics* **2**(2), 141–167.

Kadiyali, V., Vilcassim, N. & Chintagunta, P. (1999), 'Product line extensions and competitive market interactions: An empirical analysis', *Journal of Econometrics* **89**, 339–363.

Kahneman, D., Knetsch, J. L. & Thaler, R. (1986), 'Fairness as a constraint on profit seeking: Entitlements in the market', *American Economic Review* **76**(4), 728–741.

Kashyap, A. K. (1995), 'Sticky prices: New evidence from retail catalogs', *Quarterly Journal of Economics* **110**(1), 245–274.

Keenan, F. (2003), 'The price is really right', *Business Week* (3826), 62.

Kim, J., Allenby, G. & Rossi, P. (2002), 'Modeling consumer demand for variety', *Marketing Science* **21**, 229–250.

Leslie, P. (2004), 'Price discrimination in broadway theatre', *RAND Journal of Economics* **35**(3).

Leszczyc, P. P., Sinha, A. & Timmermans, H. (2000), 'Consumer store choice dynamics: An analysis of the competitive market structure for grocery stores', *Journal of Retailing* **76**, 323–345.

Levy, D., Bergen, M., Dutta, S. & Venable, R. (1997), 'The magnitude of menu costs: Direct evidence from large u.s. supermarket chains', *Quarterly Journal of Economics* **112**(3), 792–825.

McFadden, D. (1989), 'A method of simulated moments for estimation of discrete response models without numerical integration', *Econometrica* **57**(5), 995–1026.

McMillan, R. S. (2005), Estimating Demand for Differentiated Products with Continuous Choice and Variety-Seeking: An Application to the Puzzle of Uniform Pricing, PhD thesis, Stanford University.

Nevo, A. (2001), 'Measuring market power in the ready-to-eat cereal industry', *Econometrica* **69**(2), 307–342.

Orbach, B. Y. & Einav, L. (2001), 'Uniform prices for differentiated goods: The case of the movie-theater industry', *Harvard John M. Olin Discussion Paper Series* (337).

Pakes, A. & Pollard, D. (1989), 'Simulation and the asymptotics of optimization estimators', *Econometrica* **57**(5), 1027–1057.

Rhee, H. & Bell, D. R. (2002), 'The inter-store mobility of supermarket shoppers', *Journal of Retailing* (1), 225–237.

Shugan, S. M. (1980), 'The cost of thinking', *The Journal of Consumer Research* **7**(2), 99–111.

Villas-Boas, S. B. (2002), Vertical contracts between manufacturers and retailers: An empirical analysis.