Retail Market Power in a Shopping Basket Model of Supermarket Competition

Timothy J. Richards, Stephen F. Hamilton, and Koichi Yonezawa*

Paper presented at AAEA Annual Meetings, Boston, MA, July 31 - August 2, 2016

May 25, 2016

Abstract

Supermarket consumers typically purchase more than one item at a time. Retail prices, in turn, are likely to depend on demand relationships between multiple categories of goods in consumers' shopping baskets. In this paper, we develop a model of retail price competition that explicitly models the effect of complementary demand relationships between products that appear in consumer shopping baskets. We derive inferences for retail market power when shopping baskets contain products from complementary categories and compare outcomes with the predictions derived from conventional models that assume consumers make discrete choices among independent product categories. Our findings reveal that cross-category product complementarity in consumer shopping baskets facilitates substantially greater retail market power relative to the benchmark case of discrete choice over independent goods.

Keywords: complementarity, retailing, pricing, supermarkets, oligopolies.
JEL Codes: D43, L13, L81, M31.

*Richards is Professor and Morrison Chair of Agribusiness, Morrison School of Agribusiness, W. P. Carey School of Business, Arizona State University; Hamilton is Professor, Department of Economics, Orfalea College of Business, California Polytechnic State University San Luis Obispo; Yonezawa is Post-doctoral researcher, Technical University of Munich, Munich, Germany. Contact author, Richards: Ph. 480-727-1488, email: trichards@asu.edu. Support from the Agricultural and Food Research Initiative of the National Food and Agriculture Institute, USDA (grant #2016-67023-24807), is gratefully acknowledged.

1 Introduction

Consumers typically purchase more than one item at a time from supermarkets. This behavior suggests that demand relationships among products that typically appear in consumers' shopping carts play an essential role in determining how retailers set prices. Despite this fact, empirical models that examine equilibrium pricing behavior by food retailers frequently rely on data from only one product category, or else implicitly assume independence across product categories in consumers' shopping baskets by appling an identical, single-category demand model across several categories at once. In this paper, we develop and test a model of retail market power that allows us to estimate equilibrium price-setting behavior for multiple categories of products with inter-related demands.

We examine demand relationships among typical items in consumers' shopping cart using a structural model of equilibrium retail prices, conditioned on a basket-level model of demand. We adopt a flexible demand structure that encompasses the entire range of substitution effects between products from perfect substitutes to perfect complements, which nests shopping baskets comprised of independent products as a special case. Our nested empirical structure allows us to develop counterfactual experiments on the impact of substitution effects among items in consumers' shopping baskets relative to the limiting case of "category independence" commonly specified in models of supermarket pricing.

Empirical models of competition among retailers typically assume consumers make discrete choices among products within a category, among categories within a store, and among different retail stores (Bell and Lattin 1998). Here, we maintain the assumption that consumers make discrete choices among retail stores, which Smith and Thomassen (2012) show is approximately true, but recognize that consumers shopping within a given store typically purchase groceries with a shopping basket comprised of items from multiple product categories at a time, rather than purchasing a single item –or more precisely purchasing multiple items with independent demands– on each shopping occasion (Ainslie and Rossi 1998; Manchanda, Ansari, and Gupta 1999; Russell and Petersen 2000; Chib, Seetharaman, and Strijnev 2002; Kwak, Duvvuri, and Russell 2015).

It is well-recognized that demand relationships between items in a shopping basket is essential to understanding how retailers compete for store traffic. Indeed, the presumption of such relationships undelies the notion of "loss-leader" products, for instance turkey at Thanksgiving or tuna at Lent (Lal and Matutes 1994; Chevalier, Kashyap and Rossi 2003), in which price decreases on heavily-advertised products are offset by commensurate price increases on complementary goods. Our analysis explicitly models the effect of complementarity among items in a shopping basket on basket-level retail prices, which allows us to measure retail market power at the basket level rather than at the individual product level.

We base our observations on a structural model of consumer demand for items in a representative shopping basket at two competing stores. Our empirical model consists of a nested multi-variate logit (MVL) specification that encompasses a rich set of cross relationships between shopping basket items and between stores (Russell and Petersen 2000; Niraj, et al. 2008; Moon and Russell 2008; Kwak, Duvvuri, and Russell 2015). Conditioned on this MVL demand structure, we analytically derive the Bertrand-Nash solution to a retail pricing game over multiple products. Given that retailer rents are determined by store-level sales rather than brand-level sales, we condition equilibrium pricing decisions among competing supermarkets according to how consumers respond to price changes.

Understanding how consumers respond to retail price changes of individual products requires internalizing demand relationships between the various goods in a shopping basket. Because a typical shopping basket consists of many different categories of products, the net effect of internalizing substitute and complement relationships between products on retail prices is largely an empirical question. We derive implications for the role of cross-product demand effects on supermarket pricing behavior by relying on the structure of the nested MVL model to compare the equilibrium prices generated by our unrestricted MVL model to those generated by traditional multinomial logit (MNL) models that restrict the demand structure between goods.

We examine the effect of product complementarity on retail prices using panel data from supermarket retailers in the Eau Claire, Wisconsin market. Specifically,w we construct shopping baskets for households in our sample and examine the effect of shopping basket composition on retail prices for four categories of goods: Milk, breakfast cereal, soft drinks, and snacks. Our empirical results provide strong support for the hypothesis that selling complementary goods softens retail price competition. That is, relative to the case of independent goods we find evidence of higher overall retail prices for shopping baskets composed of complementary products.

Our findings indicate that price competition between oligopoly supermarkets is significantly less intense when retailers sell complementary products. We provide intuition for this outcome by developing a model of supermarket behavior that provides a clear decomposition between the *intra*-retailer margin and *inter*-retailer margin of supermarket behavior. On the intra-retailer margin, supermarkets act as monopolists for consumers that enter the store, fully internalizing cross-effects in demand when setting prices for items in the shopping basket. Lower prices facilitate complementary purchases within the store, providing retailers with an incentive on the intra-retailer margin to set lower prices when supermarket products are complementary goods. On the inter-retailer margin, supermarkets compete with rivals to acquire store traffic. Oligopoly retailers ignore the effect of price changes on the profit of rival retailers, so that introducing demand relationships between products in consumers' shopping baskets conveys an additional externality to the inter-retailer margin through product composition effects.

For complementary products, a selective price discount in one product category raises cross-category sales of complementary products, thereby increasing the value of a typical shopping basket. Because retailers internalize the effect of lower prices on facilitating complementary purchases only on the intra-retailer margin, retailers have an incentive to raise prices on the inter-retailer margin when shopping baskets are composed of complementary products, tempering the business-stealing effect of a selective price decrease. Put differently, providing complementary product categories softens price competition between retailers.

Our analysis reveals that the effect of product complementarity on retail market power depends on the intensity of inter-retailer competition. In markets with relatively weak competition between retailers, for instance when transportation costs between retailers is "high", retailers set lower prices for shopping baskets containing complementary products than in the case of independent goods; however, stocking complementary product categories also softens retail price competition, resulting in higher retail prices when competition between retailers is relatively intense (e.g., "low" transportation cost).[1] Thus, our findings suggest a somewhat counter-intuitive result that the effect of product complementarity on retail market power is likely to be accentuated as retail markets become more saturated.

Our findings are critically important to our understanding of the inherent competitiveness of not only grocery retailers, but the retailing function more generally. Many retailers sell shopping baskets of items that consist of a mix of items from substitute and complementary categories. Home improvement stores such as Home Depot and Lowes, drug stores (CVS and Walgreens), and even online retailers (Amazon and eBay) each price in order to manage pricing on the intra- and inter-retailer margin of rivalry. When complementarity is more important, our findings show that the adverse welfare effects on equilibrium pricing can be substantial.

Our paper contributes to the literatures on retail pricing, and demand modeling more generally. While others have used the MVL model to examine shopping-basket demand (Russell and Petersen 2000; Niraj, et al. 2008; Moon and Russell 2008; Kwak, Duvvuri, and Russell 2015), the link between shopping basket composition, store choice, and retail pricing between competing stores has remained unexplored. We show that accounting for the mix of complementary and substitute relationships among product categories has essential

[1] Our argument is a variation on the harvest-invest story of Dube, Hitsch, and Rossi (2009) and Pavlidis and Ellickson (2012) in that retailers compete relatively more intensively in order to earn high-margin customers. On the other hand, when categories are complements within the store – when retailers sell only store-brands, which are complementary due to "umbrella branding" (Erdem and Chang 2012), for example – then price competition is less intense, and market conduct is less competitive.

effects on retail market power. Unlike recent theoretical models that suggest accounting for the "incidental complementarity" associated with shopping-basket purchases has a pro-competitive effect on retail pricing (Rhodes 2015), we show that *product* complementarity has anti-competitive effects on shopping basket prices.[2]

In the next section, we derive a theoretical model of retail pricing under shopping-basket purchasing and show that equilibrium prices can rise under pure complementarity. We test this theory using the empirical model of store-chioce and shopping-basket demand described in Section 3. Section 4 describes our data source and provide some stylized facts that support the use of a MVL model to estimate demand inter-relationships among grocery categories. Section 5 summarizes our empirical findings and offers some implications for the conduct of retailing more generally, while we conclude and offer some suggestions for future research in Section 6.

2 Basket Composition and Pricing

In this section we present a simple theoretical model that isolates retail pricing incentives on the *intra*-retailer margin and *inter*-retailer margin of supermarkets. Consumers in the model engage in one-stop-shopping, acquiring a basket of goods comprised of multiple products on each shopping occasion. Retail pricing at the basket level, in turn, is driven by two opposing incentives: (*i*) on the inter-retailer margin, retailers wish to lower retail prices on all retail products to steal business from rivals; whereas (*ii*) on the intra-retailer margin, retailers maintain an optimal mix of prices, fully internalizing externalities between goods in the representative shopping basket by setting "Ramsey" prices. Customers that visit a given supermarket by low prices on items in the desired shopping basket purchase multiple products on a single shopping trip, which stimulates retailers to internalize complementary brand relationships in the multi-product demand system.

[2] The difference is due to the fact that "incidental complementarity" is driven by economies of scope in consumer transportation cost, rather than by explicit demand relationships among products in the shopping basket.

Consider duopoly supermarkets that stock products in multiple categories. The retailers differ in their spatial proximity to consumers in the Hotelling (1929) sense and stock product categories that contain complementary goods. Our focus is on how equilibrium prices change with the degree of complementarity between products, and we accordingly simplify the model by considering a fixed number of products.[3]

Each retailer is located at the end of a unit line segment and consumers are distributed uniformly along the line segment so that no one retail location is inherently superior to any other retail location. Consumers incur increasing transportation costs of τ per unit of distance to visit retailers.[4] The decision to shop with a given retailer consequently depends on the transportation cost required to visit the retailer relative to the consumption opportunity afforded by that retailer's product assortment and prices.

Consumer preferences over retail products are represented by the utility derived from one-stop shopping. Specifically, given her choice of retailer and consumption bundle $(x_1, x_2, ..., x_n)$ utility of the representative consumer is

$$u(x_1, x_2, ..., x_n) - \sum_i p_{j,i} x_i.$$

Solving this problem for the optimal consumption bundle selected at retailer j yields the indirect utility function

$$v^*(\mathbf{p}_j) = \max_{x_1, x_2, ..., x_n} u(x_1, x_2, ..., x_n) - \sum_i p_{j,i} x_i,$$

where $\mathbf{p}_j = (p_{j,1}, p_{j,2}, ..., p_{j,n})$ is the vector of prices selected by retailer j.

Aggregate demand facing each retailer depends on the decisions made by consumers at all points on the line segment regarding where to shop. Given consumer transportation costs of τ per unit distance, a consumer at a distance of $\theta \in (0, 1)$ from retailer j could achieve surplus of $v^*(\mathbf{p}_j) - \theta\tau$ by purchasing from that retailer. Letting θ^* denote the location of

[3] See Anderson and dePalma (1992, 2006) and Hamilton and Richards (2009) for analysis of product variety choices among multi-product retailers.

[4] Transportation costs for visiting retailers is asumed to be sufficiently high that consumers purchase multiple products on each shopping occasion and compare between supermarkets at the basket level rather than the individual product level.

the consumer who is indifferent between the alternative of shopping with either retailer, θ^* solves $v^*(\mathbf{p}_1) - \theta\tau = v^*(\mathbf{p}_2) - \tau(1-\theta)$, which yields

$$\theta^*(\mathbf{p}_1; \mathbf{p}_2) = \frac{1}{2} + \frac{1}{2\tau}\left[v(\mathbf{p}_1) - v(\mathbf{p}_2)\right]. \tag{1}$$

All consumers located at a distance of $\theta \leq \theta^*$ prefer to shop with retailer 1 and all consumers located at a distance of $\theta^* \leq \theta$ prefer to shop with retailer 2. The demand for retail product i, at retailer 1 accordingly, is $X_i(\mathbf{p}_1; \mathbf{p}_2) = \theta^*(\mathbf{p}_1; \mathbf{p}_2)x_i(\mathbf{p}_1)$ and total store demand for retailer 1 is defined accordingly by aggregating products in the representative consumer's basket: $X(\mathbf{p}_1) = \theta^*(\mathbf{p}_1; \mathbf{p}_2)\sum_i x_i(\mathbf{p}_1)$.

Now consider the problem of retailer 1. Suppose each retailer pays a fixed set-up cost, F, and a constant unit cost of c to stock an individual product. Denoting *per-customer* profit for retailer 1 as

$$\pi(\mathbf{p}_1) = \sum\nolimits_i (p_{1,i} - c)x_i(\mathbf{p}_1), \tag{2}$$

total retailer profit for retailer 1 is given by

$$\Pi(\mathbf{p}_1; \mathbf{p}_2) = \theta^*(\mathbf{p}_1; \mathbf{p}_2)\pi(\mathbf{p}_1) - F. \tag{3}$$

Differentiating (3) with respect to $p_{1,i}$ gives the first-order necessary condition

$$\theta^*(\mathbf{p}_1; \mathbf{p}_2)\frac{\partial \pi(\mathbf{p}_1)}{\partial p_{1,i}} + \frac{\pi(\mathbf{p}_1)}{2\tau}\frac{\partial v(\mathbf{p}_1)}{\partial p_{1,i}} = 0, \qquad i = 1, 2, ..., n, \tag{4}$$

where $\partial v(\mathbf{p}_1)/\partial p_{1,i} = -x_i(\mathbf{p}_1) < 0$ holds by Roy's identity. Notice that condition (4) decomposes the effect of a price change into an *inter*-retailer margin and an *intra*-retailer margin of profit. The first term on the left-hand side of equation (4) is the effect of a price change on the *intra*-retailer margin. For a given amount of store traffic (θ^* fixed), the retailer sets relative prices like a monopolist, selecting "Ramsey" prices that fully internalize demand relationships products. The second term on the left-hand side of equation (4) defines the effect of a price change on the *inter*-retailer margin. A small decrease in price of $dp_{1,i}$

units shifts $(x_i/\tau)dp_{1,i}$ customers towards retailer i and away from his rival through the so-called *business-stealing effect*. Because each customer purchases multiple products from the retailer, a unit increase in custom results in retail profit of $\pi(\mathbf{p}_1)$, resulting in rents of $(x_i/\tau)\pi(\mathbf{p}_1)dp_{1,i}$ from a unilateral decrease in the price of good i by $dp_{1,i}$ units.

Condition (4) implies oligopoly prices are proportionately lower than monopoly prices. A monopoly retailer not disciplined by competition on the inter-retailer margin would choose category prices such that $\partial\pi(\mathbf{p}_1)/\partial p_i = 0$; however, an oligopoly retailer selects prices below the monopoly price level, that is $\partial\pi(\mathbf{p}_1)/\partial p_i > 0$ in equation (4), because the business-stealing effect of a price increase (the second term on the left-hand side of (4)) is negative. Oligopoly prices are lower than monopoly prices, because retailers fail to internalize the positive externality of raising prices on the profits of rivals.

When retailers set prices simultaneously for multiple products, the business-stealing effect involves shifting the entire shopping basket of the marginal consumer. The composition of products in consumers' shopping baskets introduces a second externality on the inter-retailer margin. Because retailers fail to internalize product composition effects in the shopping baskets of consumers switching to rival retailers, retailers have insufficient incentive to lower retail prices when consumer shopping baskets are comprised of complementary goods. Relative to the case of independent goods, retail price competition softens on the inter-retailer margin when selling complementary products.

The retail pricing implications of a wide variety of product configurations is completely characterized by equation (4). To understand how demand complementarity affects supermarket prices, it is helpful to examine how demand complementarity alters pricing incentives on the intra- and inter-retailer margins. To simplify this comparison, consider the retail market equilibrium with n symmetric product categories, each of which contains one good. Dropping arguments for notational convenience, condition (4) can be written in the symmetric case as

$$\frac{p-c}{p} = \frac{1}{\varepsilon_{ii} - (n-1)\varepsilon_{ij} + S/\tau}, \tag{5}$$

where $\varepsilon_{ii} = -\frac{\partial x_i}{\partial p_i}\frac{p_i}{x_i} > 0$ is the own-price elasticity of demand, $\varepsilon_{ij} = \frac{\partial x_j}{\partial p_i}\frac{p_i}{x_j}$ is the cross-price elasticity of demand between product categories, and $S = npx$ is total retail sales. The first two terms in the denominator determine the intra-retailer margin. Absent competition from rival retailers, retail margins are given by $\frac{p-c}{p} = \frac{1}{\varepsilon_{ii}-(n-1)\varepsilon_{ij}}$, which is the Ramsey pricing condition for a multi-product monopolist.[5] On the intra-retailer margin, complementarity between retail categories, $\varepsilon_{ij} < 0$, is associated with narrower margins than in the case of independent goods, as retailers internalize the cross-product effect of a price discount on facilitating sales of complementary goods among consumers entering the store.

The third term in the denominator on the right-hand side of equation (5) captures the retail pricing incentive on the inter-retailer margin. This term is larger when consumer transportation costs are small; thus, the intensity of retail competition is essential for determining the effect of basket composition on retail market power. When unit transportation cost (τ) is "large", equilibrium retail prices are driven predominantly by incentives on the intra-retailer margin, resulting in decreased retail market power for categories with more complementary demand; however, when τ is "small", highly complementary categories sold by supermarkets result in greater retail market power relative to the case of independent goods.

To confirm this intuition, we simulate price changes in response to changes in demand complementarity using two demand structures: (i) quadratic utility; and (ii) constant elasticity of substitution (CES). The outcome is qualitatively similar in each case, with a range of τ emerging in which retail provision of complementary product categories results in higher retail prices than in the case of independent (or substitute) goods.

Consider the quadratic utility structure

$$u(x_1, x_2, ..., x_n) = \alpha \sum\nolimits_i x_i - \frac{1}{2}\beta \sum\nolimits_i x_i^2 - \lambda \sum\nolimits_i \sum\nolimits_{j \in n} x_i x_j,$$

where $\alpha, \beta, \lambda > 0$ and $\beta > \lambda$. This utility structure leads to demands of the form

$$x_i = a - bp_i + \delta \sum\nolimits_{j \neq i} p_j,$$

[5] This equation reduces to the usual monopoly markup in the case of independent goods, $\varepsilon_{ij} = 0$.

where $a = \alpha/(\beta + \lambda)$, $b = \beta/(\beta + \lambda)^2$, and $\delta = \lambda/(\beta + \lambda)^2$. Products are complements when $\delta < 0$.

Table 1 shows the results of numerical simulation for the case where $a = 1$, $b = 1$, $c = \frac{1}{2}$ and $n = 2$ for variations in transportation cost (τ) and demand complementarity (δ). Notice that retail prices decrease monotonically as the categories become more complementary when transportation costs are "high" ($\tau = 1$), decreasing from $p =$ \$0.73 for the case of independent categories ($\delta = 0$) to $p =$ \$0.53 for the case of "strong" complements ($\delta = -0.8$). The reason is that pricing outcomes are predominantly driven by the intra-retailer margin when consumers face high transportation costs.

[table 1 in here]

For the case when transportation costs are "low" ($\tau = 0.01$), retail prices follow a non-monotonic pattern with changes in demand complementarity. The retail price peaks at $\delta = -0.7$, rising with the degree of complementarity between product categories for $\delta \in (-0.7, 0)$, and then falling at still higher degrees of complementarity. The reason for this is that retail prices converge towards marginal cost on the intra-retailer margin for high levels of product complementarity, dampening business-stealing incentives and the impact of ignoring product composition externalities on the inter-retailer margin. In all cases, retail prices are higher when retailers stock complementary product categories than in the case of independent product categories when $\tau = 0.01$. Because transportation costs are unobservable, it is unclear *a priori* whether the effect of complementarity in the first column or the second column is a better description of retail pricing in the real world. Therefore, in the next section we describe an empirical approach designed to test the effect of complementarity on equilibrium retail pricing.

3 Empirical Model of Retail Pricing and Demand Complementarity

Overview

We estimate a structural model of retail demand, and retailer pricing. Our primary objective is to determine the effect of demand structure in consumers' shopping baskets on retail market power. We begin by specifying a hierarchical empirical model in which consumers first choose stores, and choose baskets conditional on store choice, and then derive expressions for retail prices that are consistent with equilibrium in a Bertrand-Nash environment, conditional on the structure of demand between stores and within the shopping basket. We conclude this section by describing how both the demand and pricing elements of the model are estimated.

Model of Retail Demand

Consumers $h = 1, 2, 3, ..., H$ in our model select items from among $i = 1, 2, 3, ..., N$ categories, c_{iht}, in assembling a shopping basket, $\mathbf{b}_{ht} = (c_{1ht}, c_{2ht}, c_{3ht}, ..., c_{Nht})$ on each trip, t, conditional on their choice of store, r. Define the set of all possible baskets in r as $\mathbf{b}^r_{ht} \in \mathbf{B}^r$. Our focus is on purchase incidence, which is the probability of choosing an item from a particular category on a given shopping occasion, and we model demand at the category level by assuming consumers purchase one item per category across multiple categories.

Consumers choose among categories to maximize utility, U^r_{ht}, and we follow Song and Chintagunta (2006) in writing utility in terms of a discrete, second-order Taylor series approximation:

$$
\begin{aligned}
U^r_{ht}(\mathbf{b}^r_{ht}|r) &= V^r_{ht}(\mathbf{b}^r_{ht}|r) + \varepsilon^r_{ht} \\
&= \sum_{i=1}^{N} \pi^r_{iht} c^r_{iht} + \sum_{i=1}^{N} \sum_{j \neq i}^{N} \theta^r_{ijh} c^r_{iht} c^r_{jht} + \varepsilon^r_{ht},
\end{aligned}
\tag{6}
$$

where π^r_{iht} is the baseline utility for category i earned by household h on shopping trip t in store r, c^r_{iht} is a discrete indicator that equals 1 when category i is purchased in store r, and 0 otherwise, ε^r_{ht} is a Gumbel-distributed error term that is iid across households and shopping trips, and θ^r_{ijh} is a household-specific parameter that captures the degree of interdependence in demand between categories i and j in store r. Specifically, $\theta^r_{ijh} < 0$ if the categories are substitutes, $\theta^r_{ijh} > 0$ if the categories are complements, and $\theta^r_{ijh} = 0$ if the categories are independent in demand. To ensure identification, we restrict all $\theta^r_{ii} = 0$ and impose symmetry on the matrix of cross-purchase effects, $\theta^r_{ijh} = \theta^r_{jih}, \forall i, j \in r$ (Besag 1974, Cressie 1993, Russell and Petersen 2000).

The probability that a household purchases a product from a given category on a given shopping occasion depends on both perceived need, and marketing activities from the brands in the category (Bucklin and Lattin 1992, Russell and Petersen 2000). Therefore, we write baseline utility for each category as dependent on a set of category ($\mathbf{X}_i$) and household ($\mathbf{Z}_h$) specific factors such that:

$$\pi^r_{iht} = \alpha^r_{ih} + \beta^r_{ih}\mathbf{X}^r_i + \gamma^r_{ih}\mathbf{Z}_h, \tag{7}$$

where perceived need, in turn, is affected by the rate at which a household consumes products in the category, and the frequency that they tend to purchase in the category, which we combine to form a measure of the amount of inventory on hand (INV_h).[6] Need is also determined by more fundamental household factors such as the size of the household (HH_h), the age distribution of family members (AGE_h), and state dependence that arises from either loyalty, habituation, or some other source of intertermporal correlation in purchase incidence (LOY_h).[7] Marketing mix elements at the category level include a price index of the individual items in each category (PR^r_i), the proportion of items featured during the purchase occasion (FT^r_i), the percentage on display (DP^r_i), and the share on temporary price reduction (TPR^r_i). While there are likely other factors that influence category choice, this set covers those typically used in the category-choice literature (Manchanda, Ansari, and Gupta 1999) and exhaust those available in our data.

Each of the variables entering (7) represent sources of observed heterogeneity, whether

[6] We infer household inventory using methods that are standard in this literature (Bucklin and Lattin 1992).

[7] Each of these demographic variables was tested in the empirical model, and found to be not significant, so were excluded from the results reported below.

at the category ($\mathbf{X}_i^r$) or household ($\mathbf{Z}_h$) levels. However, there is also likely to be substantial unobserved heterogeneity in household preferences and in attributes of the category that may affect incidence. Therefore, we specify each of the estimated parameters as randomly distributed in order to capture unobserved heterogeneity in category preference, marketing mix responsiveness, and the marginal effect of demographic attributes, respectively. In the most general form of the model, therefore, we estimate:

$$\begin{aligned}
\alpha_{ih}^r &= \alpha_{i0}^r + \alpha_{i1}^r v_1,\ v_1 \sim N(0, \sigma_1), &(8)\\
\beta_{ikh}^r &= \beta_{ik0}^r + \beta_{ik1}^r \nu_2,\ v_2 \sim N(0, \sigma_2), &(9)\\
\gamma_{ilh}^r &= \gamma_{il0}^r + \gamma_{il1}^r \nu_3,\ v_3 \sim N(0, \sigma_3),
\end{aligned}$$

for each k element of the marketing-mix matrix, and l element of the matrix of household attributes.

With the error assumption in equation (6), the conditional probability of purchasing in each category assumes a relatively simple logit form. Following Kwak, Duvvuri, and Russell (2015), we simplify the expression for the conditional incidence probability by writing the cross-category purchase effect in matrix form, suppressing the store index on the individual elements, where: $\Theta_h^r = [\Theta_{1h}, \Theta_{2h}, ..., \Theta_{Nh}]$ and each Θ_{ih} represents a column vector of the NxN cross-effect Θ_h^r matrix which is defined as:

$$\Theta_h^r = \begin{bmatrix} 0 & \theta_{12h} & \theta_{13h} & ... & \theta_{1Nh} \\ \theta_{21h} & 0 & \theta_{23h} & ... & \theta_{2Nh} \\ \theta_{31h} & \theta_{32h} & 0 & ... & \theta_{3Nh} \\ . & . & . & ... & . \\ . & . & . & ... & . \\ \theta_{N1h} & \theta_{N2h} & \theta_{N3h} & ... & 0 \end{bmatrix}, \tag{10}$$

so that the conditional utility of purchasing in category i is written as:

$$U_{ht}^r(c_{iht}^r | c_{jht}^r) = \pi_{ht}^{r\prime}\mathbf{b}_{ht}^r + \Theta_{ih}^{r\prime}\mathbf{b}_{ht}^r + \varepsilon_{ht}, \tag{11}$$

for the items in the basket vector $\mathbf{b}_{ht}^r$. Conditional utility functions of this type potentially convey important information, and are more empirically tractable that the full probability distribution of all potential assortments (Moon and Russell 2008), but are limited in that they cannot describe the entire matrix of substitute relationships in a consistent way, and are not econometrically efficient in that they fail to exploit the cross-equation relationships implied by the utility maximization problem. To see this more clearly, we derive the estimating equation implied by the Gumbel error-distribution assumption, conditional on the purchases made in all other categories, c^r_{jht}. With this conditional assumption, the probability of purchasing in category $i = 1$ is written as:

$$\Pr(c^r_{1ht} = 1 | c^r_{jht}, r) = \frac{[\exp(\pi^r_{1ht} + \Theta^{r\prime}_{1h}\mathbf{b}^r_{ht})]^{c_{1ht}}}{1 + \exp(\pi^r_{1ht} + \Theta^{r\prime}_{1h}\mathbf{b}^r_{ht})}, \tag{12}$$

and $\mathbf{b}^r_{ht}$ represents the basket vector. Estimating all N of these equations together in a system is one option, or Besag (1974) describes how the full distribution of $\mathbf{b}^r_{ht}$ choices are estimated together.

Assuming the Θ^r_h matrix is fully symmetric, and the main diagonal consists entirely of zeros, then Besag (1974) shows that the probability of choosing the entire vector $\mathbf{b}^r_{ht}$ is written as:

$$\Pr(\mathbf{b}^r_{ht} | r) = \frac{\exp(\boldsymbol{\pi}^{r\prime}_{ht}\mathbf{b}^r_{ht} + \frac{1}{2}\mathbf{b}^{r\prime}_{nt}\Theta^r_h\mathbf{b}^r_{ht})}{\sum\limits_{\mathbf{b}^r_{ht} \in \mathbf{B}^r} [\exp(\boldsymbol{\pi}^{r\prime}_{ht}\mathbf{b}^r_{ht} + \frac{1}{2}\mathbf{b}^{r\prime}_{ht}\Theta^r_h\mathbf{b}^r_{ht})]}, \tag{13}$$

where $\Pr(\mathbf{b}^r_{ht})$ is interpreted as the joint probability of choosing the observed combination of categories from among the 2^N potentially available from N categories, still conditional on the choice of store r.[8] Assuming the elements of the main diagonal of Θ^r is necessary for identification, while the symmetry assumption is required to ensure that (13) truly represents

[8] The practical limitations of describing 2^N choices are somewhat obvious. Recently, others have developed ways to either reduce the dimensionality of the $\mathbf{b}_{ht}$ vector, or of estimating it more efficiently. Kwak, Duvvuri, and Russell (2015) focus on "clusters" of items within conventional category definitions, while Moon and Russel (2008) project the $\mathbf{b}_{ht}$ vector into household-attribute space, so only 2 parameters are estimated. Kamakura and Kwak (2012) use the random-sampling approach of McFadden (1978) to reduce the estimation burden while leaving the size of the problem intact. Because our problem is well-described with only a small number of categories (5), we estimate the MVL model in its native form.

a joint distribution, a multi-variate logistic (MVL, Cox 1972) distribution, of the category-purchase events. Essentially, the model in (13) represents the probability of observing the simultaneous occurrence of N discrete events – a shopping basket – at one point in time. Due to the iid assumption of the logit errors associated with each basket choice, the model in (13) implicitly assumes that the baskets are subject to the independence of irrelevant alternatives (IAA), but the categories within the basket are allowed to assume a more general correlation structure (Kwak, Duvvuri, and Russell 2015).

To this point, the model describes basket choice, conditional on the choice of store. Because we are interested in the unconditional probability of choosing a particular basket of groceries, we next describe a model of store choice that allows us to derive the joint probability of choosing a basket and a store such that: $\Pr(\mathbf{b}_{ht}, r) = \Pr(\mathbf{b}_{ht}^{r}|r)\Pr(r)$. Estimating the model as a nested variant of the MVL also allows us to test the hypotheses that follow from our theoretical model as complementarity will have implications for equilibrium retail prices on both the inter- and intra-retailer margin. Although consumers do not necessarily choose one store exclusively in practice, we model their choice of *primary* store, which typically accounts for over 80% of their grocery expenditure (Smith and Thomassen 2012). With this assumption, we can describe store choice within a tractable, nested-logit framework that is well-accepted for this purpose in the literature (Bell and Lattin 1998).

Store choice depends upon attributes of the store, and both observable and unobservable attributes of the consumer that may mean that one store is more desirable than another. Bell, Ho, and Tang (1998) differentiate between fixed costs of visiting a particular store – those that do not vary with the amount purchased – and the variable costs, or those that depend on how much the consumer buys. Fixed costs include the distance between the household and each store, as well as measures of state dependence that capture opportunity costs of departing from habits or familiar patterns of behavior. Similar measures are also used by Bell and Lattin (1998) and Briesch, Chintagunta, and Fox (2009) to describe store choice. Variable costs include measures of the relative cost of shopping at each store, typically captured by a price index (Bell, Ho, and Tang 1998), the assortment available at each store (Briesch, Chintagunta, and Fox 2009), and the relative attractiveness of purchasing a basket of items from each store (Bell and Lattin 1998).

We capture each of these elements of store choice in writing the indirect utility from choosing store r as:

$$V_{ht}(r_{ht}) = \delta_k \mathbf{Y}_{hr} + \lambda IV_{hr} + \nu_{rht} = \delta_{0r} + \delta_1 DIST_{hr} + \delta_2 LOY_{hr} + \delta_3 VAR_r + \delta_4 PRI_r + \lambda IV_{hr} + \nu_{rht}, \tag{14}$$

for a set of household-and-store variables $\mathbf{Y}_{hr}$,where $DIST_{hr}$ is the distance, measured in Euclidean terms, between household h and store r, LOY_{hr} is a measure of state-dependence, defined as loyalty to a particular store, which we operationalize as the percentage of visits to store r in the 6 months prior to the estimation period (Bell and Lattin 1998), VAR_r is a measure of the depth of assortment across all categories in store, calculated as the total number of universal product codes (UPCs) offered each week in store r, PRI_r is a price index calculated as a sales-weighted average across all categories in the model, and the inclusive value (IV_{hr}) for store r and household h, which captures the relative attractiveness of the store as measured by the expected utility from a basket of groceries estimated in the conditional basket choice model, written as: $IV_{hr} = \log \left(\sum_{\mathbf{b}_{ht}^r \in \mathbf{B}^r} [\exp(\boldsymbol{\pi}_{ht}^{r\prime} \mathbf{b}_{ht}^r + \frac{1}{2} \mathbf{b}_{ht}^{r\prime} \Theta_h^r \mathbf{b}_{ht}^r)] \right)$. The nesting parameter, λ, also referred to as the substitution parameter, measures the extent to which groups, stores in this case, substitute for each other. If $\lambda = 1$, our nested MVL model collapses to a MVL / simple logit model of choice among baskets and stores.

To complete the nested MVL model structure, we assume the error ν_{rht} is distributed Type I Extreme Value, which means that the choice among stores takes a familiar logit form:

$$\Pr(r) = \frac{\exp(\delta_k \mathbf{Y}_{hr} + \lambda IV_{hr})}{\sum_{s \in R} \exp(\delta_k \mathbf{Y}_{hs} + \lambda IV_{hs})}, \tag{15}$$

with the same hierarchical structure as a more traditional nested logit, for each household,

h. With estimates of the store choice model, we can then calculate the joint probability of each basket-and-store choice, and derive the effect of a price change within one basket on equilibrium prices within each store on the intra-retailer margin, and across rival stores on the inter-retailer margin.

Demand Elasticities

Intra- and inter-retailer price effects are measured by price elasticities of demand. We derive the individual-item elasticities implied by (13) and (15) in this section, and apply them to the analysis of competitive price-and-assortment response in the subsequent section. To foreshadow our results, we find that the elasticity expressions look very similar to the usual logit price-elasticity expressions, but with one critical difference: Because each category can appear in several bundles, the derivative must sum over the marginal effect of a change in price on the probability of observing each bundle that contains that category. Formally, the joint probability of observing c^r_{jht} in store r is given by:

$$\Pr(c^r_{jht}, r) = \Pr(c^r_{jht}|r)\Pr(r) = \sum_{c^r_{jht}\in\mathbf{b}^r_{ht}} \left(\frac{\exp(\boldsymbol{\pi}^{r\prime}_{ht}\mathbf{b}^r_{ht} + \frac{1}{2}\mathbf{b}^{r\prime}_{nt}\Theta^r_h\mathbf{b}^r_{ht})}{\sum\limits_{\mathbf{b}^r_{ht}\in\mathbf{B}^r}[\exp(\boldsymbol{\pi}^{r\prime}_{ht}\mathbf{b}^r_{ht} + \frac{1}{2}\mathbf{b}^{r\prime}_{ht}\Theta^r_h\mathbf{b}^r_{ht})]} \right) \left(\frac{\exp(\delta_k\mathbf{Y}_{hr} + \lambda IV_{hr})}{\sum\limits_{s\in R}\exp(\delta_k\mathbf{Y}_{hr} + \lambda IV_{hr})} \right), \tag{16}$$

so that the household-level marginal effect of a change in a same-category, same-store price is:

$$\frac{\partial \Pr(c^r_{jht}, r)}{\partial PR_j} = \frac{\beta_{ph}\Pr(c^r_{jht}, r)}{\lambda}(1 - (1-\lambda)\Pr(c^r_{jht}|r) - \lambda\Pr(c^r_{jht}, r)), \tag{17}$$

where β_{ph} is the household-specific marginal utility of income, and $\Pr(c_{jht}, r)$ includes all baskets that contain the category j in store r. Similarly, the marginal effect of a change in the price index for a different category (i) in the same store on the probability of purchasing category j, when the categories are in the same baskets is given by:

$$\frac{\partial \Pr(c^r_{jht}, r)}{\partial PR_i} = \frac{-\beta_{ph} \Pr(c^r_{jht}, r)}{\lambda}((1-\lambda) \Pr(c^r_{iht}|r) + \lambda \Pr(c^r_{iht}, r)), \tag{18}$$

and the marginal effect of change in the price of a category that is not in the same bundle as j is given by:

$$\frac{\partial \Pr(c^r_{jht}, r)}{\partial PR_k} = \frac{-\beta_{ph}(\Delta \Pr)}{\lambda}((1-\lambda)\Delta \Pr + \lambda\Delta \Pr), \tag{19}$$

where the argument $\Delta \Pr = \Pr((c^r_{jht}, r), (c^r_{kht}, r)) - \Pr(c^r_{jht}, r) \Pr(c^r_{kht}, r)$ is interpreted as the difference between the joint probability of observing categories j and k purchased together in the same basket and store, less the product of the marginal probabilities of observing each category purchase, again in the same store.[9] Finally, we derive the cross-price elasticity for a pair of products that are not in the same bundle, and are not purchased in the same store. Logically, the demand relationship between pairs of products purchased in different stores is only manifest through the store-choice component of the model, so the cross-price elasticity becomes simply:

$$\frac{\partial \Pr(c^r_{jht}, r)}{\partial PR_l} = -\beta_{ph}\lambda \Pr(c^r_{jht}, r), \tag{20}$$

for all products l not in the same store as j.

With these expressions, we can estimate an entire matrix of price responses, for all categories with respect to all other categories, accounting for the fact that they may or may not be purchased in the same shopping basket.

Price Response

In this section, we complete the structural model of price response by deriving the optimal retailer response to choices made by rivals. Retailers maximize profit by choosing category-prices in Bertrand-Nash rivalry, conditioned on market demand aggregated over the household behavior described in (13). Retailers are assumed to purchase items sold in

[9] Note that these probabilities are marginal with respect to categories, but still joint with respect to stores.

each category from manufacturers, and pass through input cost increases to consumers.[10] Dropping time subscripts for clarity, the profit equation for retailer r is written as

$$\tau_r = M \sum_{i \in I_r} s_{r_i}(p_{r_i} - c_{r_i}) - F_r, \tag{21}$$

where M is the size of the aggregate market for all products, I is the set of all categories, and F_r reflects the retailer's fixed cost of selling items in all categories.

Retailing costs, which include the wholesale price charged by manufacturers in each category, are specified as a linear function of input prices. This results in the following expression for retailing costs:

$$c_{r_i}(\mathbf{v}_r) = \sum_{l \in L} \eta_{wl} v_{rl} + \epsilon_{ijr}, \tag{22}$$

where $\mathbf{v}_r$ is a vector of L input prices, η_{wl} are estimates of the contribution of each input price to unit costs, and ϵ_{ijr} is an iid error term. Input prices include a category-specific primary ingredient price (fluid milk for the milk category; wheat, rice, and oats for the cereal category; sugar for soft drinks; and wheat and salt for the snack category, each from the Bureau of Labor Statistics (BLS)), indices of wages earned by workers in the food retailing and food manufacturing industries (BLS), and producer price indices for utilities, energy, packaging, advertising, and other business services (BLS). Retailing costs are estimated after substituting equation (22) into the first-order conditions derived below. Conditional on the structure of demand, retailer r's first order condition for the price of category i is given by

$$\frac{\partial \tau_r}{\partial p_{ir}} = M s_{ir} + M \sum_{i \in I} (p_{ir} - c_{ir}) \frac{\partial s_{ir}}{\partial p_{jr}} = 0, \ \forall i \in I, \ r \in R, \tag{23}$$

where $\partial s_{ir} / \partial p_{jr}$ is one element of a matrix of share derivatives with respect to price for all categories, i and j. Notice that equation (23) implies that each retailer internalizes all

[10] To the extent that retailers have some purchasing power over manufacturers, this assumption may be an oversimplification. Our assumption simply implies that what we refer to as retailer margins below, may in fact be a combination of manufacturer and retailer margins. However, if complementarity does not influence upstream market power, which it should not, our qualitative conclusions will not be affected by this assumption.

cross-sectional pricing externalities across categories in the store, but does not take into account the effect of his category pricing on the sales of other retailers. Stacking the first-order conditions across retailers we define the ownership matrix as Ω, which has element $\omega_{ir} = 1$ if category i is sold by retailer r (and zero otherwise). Making use of this notation, we write the first-order condition as:

$$\mathbf{p} = \mathbf{c} - \boldsymbol{\phi}(\boldsymbol{\Omega}\mathbf{S}_p)^{-1}\mathbf{s} + \boldsymbol{\epsilon}, \tag{24}$$

where bold notation indicates a vector (or matrix), and $\mathbf{S}_p$ is the matrix of share-derivatives with element $\partial s_{ir}/\partial p_{jr}$.[11] We include the parameter ϕ in this model, the "conduct parameter" in order to measure the extent of deviation from the maintained form of the pricing game. If the estimate of $\phi = 1$, then retailers do indeed compete as Bertrand-Nash rivals and any markup is due entirely to the extent of product differentiation (category differentiation) reflected in the matrix of share derivatives. If, however, the estimate of $\phi = 0$, then margins are zero and retail prices are consistent with perfect competition. How complementarity affects margins, and pricing power, therefore is reflected in estimates of ϕ and the implied equilibrium prices that result.

We estimate equation (24) using GMM to recover the parameters of the retail cost function using information from the demand side and the structure of the game. Based on these estimates, we next conduct a set of counter-factual simulations to investigate the importance of complementarity on equilibrium retail prices. In another estimation-and-simulation exercise, we compare fitted prices from our maintained model to alternative estimates that assume prices are conditioned on a more usual, logit model of retail demand. In this way, we achieve our dual objectives of both describing the "reality" of retail pricing, and testing theoretical models of how complementarity affects retail prices.

Estimation Methods

In the absence of unobserved heterogeneity, the MVL model is estimated using maxi-

[11] The specific form of these derivatives for the random-coefficient MVL model are provided in the technical appendix.

mum likelihood in a relatively standard way. However, because we allow a range of parameters to vary across panel observations, the likelihood function no longer has a closed form. Therefore, we estimate the model using simulated maximum likelihood (Train 2003), using $r = 1, 2, 3....R$ simulations. For clarity of the likelihood function, we index the possible baskets ($15 = 2^N - 1$, excluding the null basket) by k, and define a set of indicator variables z_k that assume a value of 1 if basket k is chosen and 0 otherwise. We then estimate the likelihood function as a panel over h cross-sections and t shopping occasions per household so that the simulated likelihood function is written (Kwak, Duvvuri, and Russel 2015):

$$\mathcal{L}(\mathbf{b}_{ht}) = \frac{1}{R}\sum_{r=1}^{R}\prod_{h}\prod_{k}(\Pr(\mathbf{b}_{ht} = \mathbf{b}_{ht}^{k})^{z_k}, \tag{25}$$

where the joint distribution function for all possible baskets is given in (13). To increase the efficiency of the SML routine, the simulated draws follow a Halton sequence with 50 draws, as suggested by Bhat (2003). We experimented with a range of Halton draws, and our results did not change substantially from one trial to the next, so we conclude that our estimates are relatively stable. In the Results and Discussion section that follows, we present results from a number of alternative specifications in order to establish the validity of our maintained model.

We estimate the structural model, which consists of the demand model in (25) and the supply model in (24), sequentially, first estimating the demand model and then the supply, or pricing, model conditional on the demand estimates. In the pricing model (24), the markup term on the right-side is clearly endogenous. Consequently, we estimate the pricing model using generalized method of moments (GMM). Correcting for endogeneity using GMM requires a set of instruments that are likely to be correlated with retail margins, but mean-independent of the error term. Intuitively, identifying retailer pricing conduct requires instruments that shift the demand curve facing retailers. For this purpose, we use marketing mix variables for the other store in the market, average values across the sample for each demographic variable, and a set of category dummy variables. A first-stage regression of this set of instruments on the retail margin yields an F-value of 168.724, so our set of instruments cannot be described as weak in the sense of Staiger and Stock (1997). Our identification strategy is well-accepted in the literature (Villas-Boas 2007; Richards and Hamilton 2015) so should at least yield results that are comparable to others.

4 Data Summary and Stylized Facts

Our sample is comprised of weekly scanner data from the IRI Academic Data Set (Bronnenberg, Kruger, and Mela 2008) in the Eau Claire, WI market over the period 2009 - 2011. For each household in our sample, we construct a shopping basket consisting of milk, cereal, carbonated soft drinks, and salty snack purchases.[12] Generalizing the analysis of consumer purchasing behavior to consider demand relationships among these four categories allows us to examine the impact of expected demand complements (milk and cereal; soft drinks and snacks) as well as anticipated substitutes (milk and soft drinks) on supermarket pricing of shopping baskets.

We chose these four categories because they are representative of a typical shopping cart, include products from only a few large manufacturers, and are well-described by the IRI data. In fact, all four were in the top eight grocery categories in the US in 2013 (#1 soda, #2 milk, #4 snacks, #8 cereal (Supermarket News 2014)), while the others in the top eight were not chosen because they are either dominated by in-store products (#3 bread), or highly fragmented, speciality categories (#5 beer, #6 cheese, and #7 wine). Each of our selected categories has a penetration rate above 70% in the IRI data, resulting in a large number of transactions for all households across the four categories. The high purchase frequency among categories in our sample allows us to capture other forms of complementarity besides use-complementarity, such as when retailers market umbrella brands (Erdem 1998), or when frequently-purchased items happen to follow similar purchase cycles. Moreover, store brands are well-represented in the milk and soft drinks categories, and are purchased in approxi-

[12]While typical shopping baskets for households in our sample contain more than four items, the MVL model quickly becomes intractable for unrestricted choice sets (Kamakura and Kwak 2012).

mately the same frequency, leading to a large number of observations in which products from these categories are purchased together in a household shopping basket.

To ensure a rich, within-subject data set, we only retain households with at least 50 purchase occasions over the 3-year sample period. Focusing on households with a large number of repeat purchases allows us to control for state-dependence in demand using household-varying inventory variables for each cross-sectional observation.

In panel data, it is necessary to have data on prices for not only the product that was purchased, but those that were not purchased as well. For this purpose, we merged the household- and store-level data sets by store, week, and UPC. By combining the household and store-level data, we observe the complete set of prices, and other marketing mix variables, for all UPCs available on a given purchase trip.

Despite the relatively small nature of the Eau Claire retail market, the retail supermarket industry appears to be highly fragmented, with two dominant stores, and a number of smaller, "fringe" competitors. In order to minimize the confounding effect of store-specific loyalties to lessor players in the market, we include only households who purchase cereal from one of the two most popular stores in the data set (IRI keys 257871 and 1085053), which together account for well over 60% of weekly sales for each category in our sample. By focusing on the two most important stores in the market, our estimates are most likely to capture strategic pricing behavior, rather than decisions taken in isolation by a niche competitor with a highly loyal, price-insensitive, customer base.

We choose to examine shopping basket composition effects on supermarket pricing in Eau Claire, because Eau Claire is a relatively small (approximately 65,000 population), homogeneous city, with two, competing supermarket chains and a relatively small number of alternatives. Nevertheless, there is considerable variation in household demographics across stores in our sample. Table 2 provides a summary of household demographics in our sample, indexed according to store of purchase. Notice that households shopping at store 1 have annual income nearly $10,000 (20.7%) higher than households shopping at store 2. This difference in average income suggests that consumers shopping at store 1 may have less price sensitivity for products and greater demand for high-quality items relative to consumers shopping at store 2. Shoppers at store 1 are similar in age, but slightly more educated, and have fewer children than shoppers at store 2. Milk and soft drinks are important drivers of basket volume, suggesting that the difference in family size may have important implications for how each supermarket sets shopping basket prices. Across all households in our sample, 73.4% switch among stores, while 19.9% visit store 1 exclusively and 6.7% visit store 2 exclusively. Among households switching among stores, 58.5% of the time these customers visit store 1 and 41.5% of the time they visit store 2. Because store 1 appears to be inherently more attractive to households than store 2, which is potentially due to other factors besides basket-pricing, we include a store fixed effect in the empirical model described below.

[table 2 in here]

Variation in shopping-basket composition is necessary to identify cross-category demand relationships in our data. Ideally, observed shopping baskets in our sample would span the entire set of combinations among the four categories, with some households purchasing one of each item, some purchasing two items together, some three, and some all four items together in the shopping basket. Table 3 provides a summary of the shopping behavior of all households across both stores, and the amount spent on each type of shopping basket. As expected, the most common shopping basket item across the four categories we examine is milk, with 28.1% of all shopping occasions representing a single-purchase "milk run." Two-item shopping baskets are purchased on 29.4% of all shopping occasions, with the most common basket consisting of milk and soft drinks, followed closely by milk and snacks. The popularity of two-item shopping baskets containing both milk and soft drinks baskets is not surprising given the frequent incidence of both milk and soft drink purchases.[13] Three-item shopping baskets are purchased on 10.2% of all shopping occasions, with the most common combination being milk, soft drinks and snacks. Customer shopping baskets contain

[13]On each visit to the store, households purchase in anticipation of several consumption occasions, and for several different family members, likely with heterogeneous tastes (Dube 2004).

items from all four categories on 1.8% of all shopping occasions. Overall, the distribution of shopping basket purchases in our sample involves sufficient representation from each category-combination to identify parameters of the MVL model.

[table 3 in here]

The summary information in Table 3 reveals several interesting patterns for shopping basket expenditure. First, notice that consumer expenditure is higher for shopping baskets that contain soft drinks than for any other basket with a comparable number of items. This anecdotal evidence suggests that retailers have an incentive to set a mix of prices that promotes soft drink purchases in as many baskets as possible, which may explain why soft drinks are promoted and displayed more often than any other category in our sample (see Table 4). Second, as expected, larger shopping baskets generally entail greater expenditure, even on a per item basis. Third, expenditure on larger shopping baskets is more stable and, hence, more predictable for retailers. The coefficient of variation of shopping basket expenditure among single-item purchases is over 75%, whereas it is only 49% for three-item purchases and 43% for four-item purchases. Larger and more stable patterns of basket expenditure are inherently more attractive for retailers, although whether this leads to more pricing power is an empirical question.

[table 4 in here]

Examining marketing mix data by store provides some summary insights in this regard. The entries in Table 4 provide a summary of retail prices, marketing mix activity and inventory holding behavior for each category and each store. Notice that there are marked differences in category prices charged by each store, particularly in the milk and soft drink categories. Given that the stores are of the same general format (traditional supermarkets containing approximately 35,000 SKUs each), marketing many of the same national brands, and they are only 1.5 miles apart, this pricing evidence suggests that the stores do indeed compete using different multi-category pricing strategies. In light of the importance of soft drinks in generating larger shopping baskets, for example, the greater frequency in which store 1 displays and promotes soft drinks can partially explain the fact that store 1 attracted over 50% more customers in the Eau Claire market than store 2. Given the lower average household income of customers at store 2, it is surprising that milk and soft drink prices are set higher at store 2, particularly given the greater frequency of milk and soft drinks in larger, multi-item shopping baskets.

In the next section, we present the results of our structural pricing model on the link between shopping basket composition and market power.

5 Results and Discussion

Overview

We begin our presentation of the results with the MVL demand estimates, comparing fixed-coefficient versions with a random-coefficient version in order to examine the importance of unobserved heterogeneity. We then compare the MVL estimates with a logical alternative – a binary logit category-choice model that assumes independence among categories. After establishing the preferred demand model, we than present estimates of the supply side, or pricing conduct model, comparing equilibrium prices under the maintained MVL model and the logit alternative. Because the MVL model consists of a mix of complementary and substitute categories, however, and the point of our research is to examine the effect of complementary and substitute relationships among categories on equilibrium prices, we present results from a cleaner simulation experiment in which we compare equilibrium prices under varying levels of full-complementarity and full-substitute relationships. This counterfactual experiment clearly demonstrates the impact of consumers' shopping-basket purchase behavior on equilibrium prices.

Demand Estimates

Estimates from two versions of the general MVL model for category choice in each store are shown in table 5. In this table, Model 1 does not allow for unobserved heterogeneity in the response to any variable as all coefficients are assumed to be fixed. We relax this assumption in Model 2 by allowing for random category preferences, as well as price-response substitution coefficients.[14] Because Model 1 is nested within the more general version, likelihood ratio (LR) tests are appropriate tests of model mis-specification. The LR test statistic used to compare Model 1 to Model 2 is Chi-square distributed with 14 degrees of freedom (there are 14 restrictions involved in nesting Model 1 in Model 2), which implies a critical Chi-square value of 23.685. The calculated LR statistic for Store 1 is $1,730.91$, and is 862.75 for Store 2, so we easily reject Model 1 in favor of Model 2 in both cases, and conclude that unobserved heterogeneity is an important factor driving category choice. Consequently, we interpret the results from Model 2 for category choice in both stores.

[table 5 in here]

Among the parameter estimates in Table 5, each of the own-price coefficients is less than zero and strongly significant.[15] With respect to the marketing mix variables, we find that promotion activity is statistically significant for most category and store combinations, and apparently highly effective, but display and feature are only significant in a handful of cases, and only with with respect to soft drinks and snacks. This finding is perhaps due to the fact that soft drinks and snacks are more impulse buys than milk and cereal, so are more susceptible to in-store marketing methods such as feature and display. Inventory is only significant for cereal and snacks – and only in Store 1 for the former – where the effect is negative as expected. Despite the logic of including inventory in a category-choice model, few studies report statistically significant inventory effects. Overall, the model provides a relatively good fit to the data.

Our primary interest in the MVL model is our θ_{ij} estimates, which define the extent and significance of substitution effects among categories in consumer shopping baskets. These estimates are shown in the second part of table 5. In this table, we impose symmetry as

[14] The general version of the model described above allowed for demographic effects in addition to marketing-mix and category-specific variables. However, none of the demographic variables proved to be statistically significant, so were excluded from the model.

[15] Because these estimates are scaled by different units of measure, the magnitudes of each of these coefficients is not interpretable *per se*.

a condition for identification, resulting in six estimated parameters for each store. Among these, we find a significant positive demand relationship in Store 1 between all pairs of items, except for milk and soft drinks. Interestingly, the model that does not account for unobserved heterogeneity mistakenly attributes a substitute relationship, albeit a not statistically significant one, also to cereal and soft drinks, and cereal and snacks. In general, our estimates for Store 1 suggest that complementary demand relationships among shopping-cart items are both more frequent and stronger than previously believed. For Store 2, the complementary relationships found for Store 1 again appear, although less statistically significant for milk and cereal, while the cereal / snack and cereal / soft drink substitute relationships reappear. Finding that cereal and snacks are net substitutes is perhaps due to shifting use patterns for cereal, as more consumers choose cereal in the evening and less at breakfast time. Overall, categories in consumer shopping baskets are more likely to be complements than they are substitutes, even after controlling for the traditional price-driven notion of complementarity. Therefore, complementarity cannot be ignored in models of equilibrium pricing.

To this point, however, our within-store estimates are only able to capture the effect of complementarity on the intra-store margin, or how a retailer internalizes demand relationships. With estimates of a store-substitution model, taking the demand relationships within each store into account, we are also able to determine the effect of complementarity on the inter-retailer, or store-choice, margin. Estimates of the store choice model are shown in table 6 below. Opposite to the category-choice results presented above, the random-coefficient version of the store choice model is, in fact, not preferred to the fixed coefficient variant. Although our model does include several elements that are likely to explain store choice, the obvious implication that we have captured all heterogeneity in store choice is likely not the case, but rather there simply is not enough heterogeneity in choosing stores to identify the random coefficients. More important from this table, however, is the estimate of store substitution, λ. Because this estimate is significantly below 1, our estimates imply that consumers do not readily substitute among stores, but are still sufficiently willing to visit another store that substitution remains a possibility. Whether the estimated substitution parameter is sufficient to affect equilibrium prices, however, remains an empirical issue that we address below.

[table 6 in here]

The first step in this comparison is determining whether a nested, multi-product, multi-store model is preferred to one in which all choices are binary. Clearly, the nature of the demand model used to condition the equilibrium pricing estimates depends on the nested MVL model as a viable description of demand at the shopping basket level. Conventional models of purchase incidence, or category choice, assume categories are purchased according to a discrete, typically binary logit process (Bucklin and Lattin 1991). If consumers make unwavering purchases from a "shopping list" irrespective of posted prices, then there is no clear multinomial choice among categories on a given shopping occasion, but rather a binary decision as to whether or not to make a purchase from each category. For this reason, the logical alternative to our MVL model is a binary logit model applied to each category; however, the error assumption in the logit model implies that each category decision is independent of any other category decision, which implicitly assumes a shopping basket demand structure with independent goods.[16] The MVL model nests the binary logit category choice model as a special case with all θ_{ij} parameters jointly restricted to zero, which allows us to compare these specifications using LR tests.

To remain consistent with the MVL estimation strategy, we estimate several versions of the binary-logit category-choice model, accounting for differing characterizations of unobserved heterogeneity. Models 1 - 3 in table 7 represent fixed-coefficient, random category-preference, and fully random-parameter versions of the binary logit category choice model, respectively. As in the case of our MVL model, the preferred logit model is the most comprehensive version (model 3). Comparing this specification with the most general MVL model,

[16] Alternatively, if we were to apply a multinomial logit model to category choice, the implication would be that categories are strict substitutes.

the LR test statistic with six degrees of freedom (recall we are restricting all of the θ_{ij} parameters to equal zero) has a critical Chi-square value of 12.59, whereas the estimated test statistic value is $37,456.82$. The LR test clearly rejects the binary logit model in favor of the maintained MVL alternative.

[table 7 in here]

Given widespread use of the logit model in the purchase incidence literature, it is worthwhile to compare the parameter estimates from the binary logit model to those from the MVL. This comparison is marked by several notable features. First, with respect to the estimated price coefficients, there is a clear pattern on the direction of mis-specification bias in the logit model as the price parameter is over-estimated by the logit model in each case (less negative). Second, a similar pattern emerges for the category-preference pattern estimates, as each is significantly under-estimated by the logit model. This outcome is intuitive when product categories are complementary, for instance purchasing cereal with milk increases the value of milk in a shopping basket, additional value that is unaccounted for in the binary logit model under the restriction of independent categories. Whether these biases have any implications for equilibrium prices, however, is addressed by comparing equilibrium prices between the MVL and binary logit models of category-and-store choice.

Pricing Model Estimates

Our structural pricing model implies retailers set prices in oligopoly equilibrium, conditional on estimates from a model of consumer demand. How category demand changes with variation in category prices, therefore, is key to deriving the resulting equilibrium prices. We first describe the likely state of retail equilibrium between our two sample stores using our preferred MVL model, and then compare these results to counterfactual experiments that show what equilibrium prices would be like in alternative environments with independent product categories, perfect complements, and perfect substitutes. Our MVL estimates necessarily contain a mix of complementary and substitute relationships among product categories, and for this reason, we motivate our examination of the effect of shopping basket composition on retail pricing behavior by focusing on extreme cases.

Table 8 shows our estimates from the MVL pricing model. Model 1 presents OLS estimates that do not correct for the endogeneity of retail margins, whereas Model 2 presents our GMM estimates. Correcting for endogeneity is necessary in both cases, as a Hausman (1978) test rejects the notion that margins are exogenous.[17] Based on these estimates, the primary input is the most important input cost, as expected, but retailing and food manufacturing wages also have a substantial impact on equilibrium prices. Most importantly, the estimated conduct parameter, ϕ, in the MVL model reveals that our retailers are more competitive than the maintained Bertrand-Nash assumption. Specifically, the estimated value of $\phi = 0.16$ implies that retailers are more competitive than under Bertrand-Nash behavior, but less than perfectly competitive, as to be expected when retailers are differentiated both by space, the variety of items on offer, and other unobserved store-quality factors.

[table 8 in here]

Table 9 shows how estimated retail conduct changes when the model is restricted to impose demand independence among categories in consumer shopping baskets. In this case, there are no pricing externalities of the type described by Smith and Thomassen (2012) that can be internalized either to increase margins (if substitutes) or to enhance cross-category sales (if complements). Based on our GMM estimates, notice that estimated market conduct is now $\phi = 1.16$, nearly seven times greater than for the MVL pricing model. This outcome suggests that retailers would behave in a substantially less competitive fashion when setting category-level prices under independent category demand relative to setting prices that internalize demand relationships between categories in the shopping basket. Indeed, the average price across all four categories implied by the binary logit conduct parameter is approximately \$3.30, compared with an implied equilibrium basket price in the MVL case of only \$3.16.[18]

[17] The Hausman (1978) specification test compares one estimator that is consistent under the null hypothesis (exogeneity) with one that is efficient under the null, but inconsistent under the alternative hypothesis (GMM). The resulting test statistic is Chi-square distributed with one degree of freedom.

[18] In grocery retailing, a 5% difference in prices is indeed substantial when margins average less than 2%.

Our comparison so far considers retailer pricing behavior in shopping baskets with a mix of substitute and complementary categories to the outcome under independent goods. Our finding of higher market power in the case of independent goods may be an artifact of the particular product categories we consider, rather than a general outcome for equilibrium prices when retailers stock categories comprised of complementary goods. For this purpose, we extend our analysis to numerically consider the extreme cases of perfect substitutes and perfect complements.

[table 9 in here]

Simulation Exercise

Recall that the logit model alternative describes category purchases as completely independent. While the nested structure of the MVL model is instructive in revealing the specification bias in models of discrete category choice, it does not provide insight on how substitute or complementary relationships among categories affect equilibrium retail prices. For this purpose, we conduct a simple simulation exercise in which we fix the θ_{ij} parameters in the MVL model that govern substitute and complement relationships to values that represent varying levels of product relationships from perfect substitutes to perfect complements. The result provides a clean test of how different relationships among categories affect equilibrium prices within the context of our maintained model, resulting in a clear *reducto ab absurdum* demonstration of the importance of shopping basket composition effects on equilibrium prices when retailers provide complementary categories of goods.

Table 10 shows the results of various simulation exercises. In this table, we allow the θ_{ij} parameters to take one of four levels, ranging from high substitutability ($\theta_{ij} = -2$) to high complementarity ($\theta_{ij} = 2$). Under high substitutability, the estimated conduct parameter of 0.16 implies an equilibrium average price of approximately \$3.12. Prices are more competitive than the MVL model estimates based on a mix of demand relationships, as the equilibrium price is slightly lower. Under strong complementarity, equilibrium prices rise to roughly \$3.31 on average, indicating far less competitive behavior and correspondingly higher equilibrium prices than in the case of mixed category relationships.

[table 10 in here]

Our findings provide counterpoint to the recent theoretical literature on how shopping-basket demand relationships affect retail prices (Smith and Thomassen 2012; Rhodes 2015). In these models, the authors focus on the intra-retailer margin, which suggests that internalizing demand complementarity between categories in a shopping basket reduces retail prices; however, these analyses ignore the counteracting incentives to raise prices of complementary goods on the inter-retailer margin. When the interaction of oligopoly retailers is included, we find that complementarity between goods in consumers' shopping baskets increases equilibrium retail prices.

6 Conclusions and Implications

In this study, we investigate the role of category-level complementarity on equilibrium retail prices. An emerging theoretical literature argues that the inherent complementarity associated with purchasing groceries by the shopping-basket leads to more competitive pricing than would otherwise be the case. However, current empirical models are insufficient to test this theory as they implicitly assume categories are independent in demand. We derive an empirical model that is able to accommodate a full range of complementarity and substitute relationships at the category level, and use this model to test whether complementary goods in consumer shopping baskets results in increased retail market power.

Our findings have important implications for our understanding of retail pricing, and the competitiveness of supermarket retailers. Take the growth of store brands, for example. Perhaps the dominant retail trend of the last 20 years, the rise of store brands has been attributed to a host of causes, from increasing bargaining power over manufacturers (Mills 1995) to building brand loyalty for the store (Corstjens and Lal 2000). Our findings suggest another, fundamental reason for introducing store brands. If a retailer can market private labels under one umbrella brand (Erdem and Sun 2002; Erdem and Chang 2012), then the resulting complementarity created within the store can reduce the intensity of price competition on the inter-retailer margin. Retail strategies designed to induce cross-category purchases, such as erecting potato displays in the meat aisle, or offering salad dressings in the produce aisle, are easily explained in terms of our findings as retailers have a clear motive to create as many opportunities for complementarity purchases in a shopping basket as possible.

From a broad, policy perspective, the food retailing sector is generally regarded as being highly competitive. However, the growth of super-center retailing through the likes of Walmart and Target, and the emergence of online retail giants (e.g., Amazon) are based on cross-selling over multiple categories. To the extent that these firms have been, and will likely continue to be, successful in expanding the scope of their customers' shopping baskets, our findings suggest that supermarket retailing may become decidedly less competitive.

The MVL model used at the core of this study is currently the state-of-the-art for analyzing shopping basket demand. However, it is not inherently scalable to the level required to understand an entire shopping basket. Because the MVL becomes intractable for any more than four or five items, a more general model is necessary. We leave this for future research.

References

[1] Ainslie, A., and P. E. Rossi. 1998. Similarities in choice behavior across product categories. Marketing Science 17: 91-106.

[2] Bell, D. R., and J. M. Lattin. 1998. Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. Marketing Science 17: 66-88.

[3] Bell, D. R., T. H. Ho, and C. S. Tang. 1998. Determining where to shop: Fixed and variable costs of shopping. Journal of Marketing Research 35: 352-369.

[4] Bhat, C. R. 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transportation Research Part B: Methodological 37: 837-855.

[5] Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society. Series B (Methodological), 36: 192-236.

[6] Briesch, R. A., P. K. Chintagunta, and E. J. Fox. 2009. How does assortment affect grocery store choice? Journal of Marketing Research 46: 176-189.

[7] Bronnenberg, B. J., M. W. Kruger, and C. F. Mela. 2008. Database paper-The IRI marketing data set. Marketing Science 27: 745-748.

[8] Bucklin, R. E., and J. M. Lattin. 1991. A two-state model of purchase incidence and brand choice. Marketing Science 10: 24-39.

[9] Bucklin, R. E., and J. M. Lattin. 1992. A model of product category competition among grocery retailers. Journal of Retailing 68: 271.

[10] Chevalier, J. A., A. K. Kashyap, and P. E. Rossi. 2003. "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data." American Economic Review 93: 15-37.

[11] Chib, S., P. B. Seetharaman, A. Strijnev. 2002. Analysis of multicategory purchase incidence decisions using IRI market basket data. Advances in Econometrics 16: 55–90.

[12] Corstjens, M., R. Lal. 2000. Building store loyalty through store brands. Journal of Marketing Research 37: 281-291.

[13] Cox, D. R. 1972. The analysis of multivariate binary data. Journal of the Royal Statistical Society Series C 21: 113–120.

[14] Cressie, N. A.C. 1993. Statistics for spatial data. New York: John Wiley and Sons.

[15] Erdem, T. 1998. An empirical analysis of umbrella branding. Journal of Marketing Research 35: 339–351.

[16] Erdem, T., and B. Sun. 2002. An empirical investigation of the spillover effects of advertising and sales promotions in umbrella branding. Journal of Marketing Research 39: 1–16.

[17] Erdem, T., S. R. Chang. 2012. A cross-category and cross-country analysis of umbrella branding for national and store brands. Journal of the Academy of Marketing Science 40: 86-101.

[18] Hamilton, S. F. and T. Richards. 2009. Variety competition in retail markets. Management Science 55: 1368-76.

[19] Hausman, J. A. 1978. Specification tests in econometrics. Econometrica 46: 1251-1271.

[20] Hosken, D. S., L. Olson, and L. Smith. 2012. Do retail mergers affect competition? Evidence from grocery retailing. Federal Trade Commission, Bureau of Economics, Working paper no. 313.

[21] Hotelling, H. 1929. Stability in competition. The Economic Journal 37: 41-57.

[22] Kamakura, W., and K. Kwak. 2012. Menu-choice modeling. Working paper, Rice University, Department of Marketing.

[23] Kwak, K., S. D. Duvvuri, and G. J. Russell. 2015. An analysis of assortment choice in grocery retailing. Journal of Retailing 91: 19-33.

[24] Lal, R. and Matutes, C., 1994. Retail pricing and advertising strategies. Journal of Business: 345-370.

[25] Manchanda, P., A. Ansari, S. Gupta. 1999. The shopping basket: A model for multicategory purchase incidence decisions. Marketing Science 18: 95–114.

[26] McFadden, D. 1978. Modelling the choice of residential location. Institute of Transportation Studies, University of California. pp. 75-96.

[27] Mills, D. E. 1995. Why retailers sell private labels. Journal of Economics & Management Strategy 4: 509-528.

[28] Moon, S. and G. J. Russell. 2008. Predicting product purchase from inferred customer similarity: an autologistic model approach. Management Science 54: 71–82.

[29] Niraj, R., V. Padmanabhan and P.B. Seetharaman. 2008. A cross-category model of households' incidence and quantity decisions. Marketing Science 27: 225–35.

[30] Pavlidis, P., and P. B. Ellickson. 2012. Switching costs and market power under umbrella branding. Working paper, University of Rochester, Department of Economics, Rochester, NY.

[31] Rhodes, A. 2015. Multiproduct retailing. The Review of Economic Studies 82: 360-390.

[32] Richards, T. J., and S. F. Hamilton. 2006. Rivalry in price and variety among supermarket retailers. American Journal of Agricultural Economics 88: 710-726.

[33] Richards, T. J. and S. F. Hamilton. 2015. Variety pass-through: An examination of the ready-to-eat cereal market. forthcoming in the Review of Economics and Statistics.

[34] Russell, G. J. and A. Petersen. 2000. Analysis of cross category dependence in market basket selection. Journal of Retailing 76: 367–92.

[35] Smith, H. 2004. Supermarket choice and supermarket competition in market equilibrium. The Review of Economic Studies 71: 235-263.

[36] Smith, H., and D. Hay. 2005. Streets, malls, and supermarkets. Journal of Economics and Management Strategy 14: 29-59.

[37] Smith, H., O. Thomassen. 2012. Multi-category demand and supermarket pricing. International Journal of Industrial Organization 30: 309-314.

[38] Song, I. and P. K. Chintagunta. 2006. Measuring cross-category price effects with aggregate store data. Management Science 52: 1594–609.

[39] Staiger, D., and J. H. Stock. 1997. Instrumental variables with weak instruments. Econometrica 65: 557-586.

[40] Sudhir, K. 2001. Structural analysis of manufacturer pricing in the presence of a strategic retailer. Marketing Science 20: 244-264.

[41] Supermarket News. 2014. Data table: Supermarket categories by dollar, unit sales. (http://supermarketnews.com/%5Bprimary-term%5D/data-table-supermarket-categories-dollar-unit-sales). Accessed on Nov. 16, 2015.

[42] Train, K. 2003. Discrete choice methods with simulation. Cambridge, U.K.: Cambridge University Press.

[43] United States Department of Labor. Bureau of Labor Statistics. Inflation and prices. (http://www.bls.gov/data/) Accessed on July 31, 2015.

[44] Villas-Boas, S. B. 2007. Vertical relationships between manufacturers and retailers: Inference with limited data. Review of Economic Studies 74: 625-652.

Table 1. Retail Prices with Complementarity

	Retail Price Level	
Complementarity	$t = 1$	$t = 0.01$
$\delta = -0.8$	\$0.528	\$0.523
$\delta = -0.6$	\$0.562	\$0.524
$\delta = -0.4$	\$0.605	\$0.517
$\delta = -0.2$	\$0.661	\$0.512
$\delta = 0$	\$0.733	\$0.509

Table 2. Data Summary: Sample Buyers

		Store 1			Store 2		
		Mean	Std. Dev.	N	Mean	Std. Dev.	N
Income	$,000	57.610	21.116	85062	47.728	30.793	51560
Family Size	#	2.458	0.028	85062	2.319	1.100	51560
Age	Years	58.224	0.616	85062	58.579	9.986	51560
Education	Years	12.305	0.030	85062	11.845	2.566	51560
Number of Children	#	2.049	0.971	85062	2.207	0.854	51560
Trips	#	149.845	75.716	85062	158.138	90.429	51560
Single Store	%	19.948	39.981	85062	6.702	25.081	51560
Multi-Store	%	58.459	34.607		41.541	34.607	

Note: Single Store refers to households that choose one store only, while multi-store refers to the market share among households that visit both stores.

Table 3. Data Summary: Shopping Basket Composition and Prices

	Share %	Std. Dev.	Expenditure	Std. Dev.
Milk Only	0.281	0.449	3.478	1.979
Cereal Only	0.052	0.222	5.344	3.951
Soft Drinks Only	0.139	0.346	8.158	7.607
Snacks Only	0.116	0.320	4.664	3.591
Milk, Cereal	0.048	0.215	9.945	5.694
Milk, Soft Drinks	0.080	0.272	11.023	6.387
Milk, Snacks	0.078	0.267	8.871	4.580
Cereal, Soft Drinks	0.014	0.117	13.369	7.568
Cereal, Snacks	0.018	0.132	10.814	6.183
Soft Drinks, Snacks	0.055	0.229	14.334	8.771
Milk, Cereal, Soft Drinks	0.019	0.135	17.139	8.323
Milk, Cereal, Snacks	0.024	0.152	15.861	7.477
Milk, Soft Drinks, Snacks	0.049	0.216	18.407	9.161
Cereal, Soft Drinks, Snacks	0.010	0.100	20.411	10.229
Milk, Cereal, Soft Drinks, Snacks	0.018	0.134	24.699	10.504

Note: Data are pooled over both stores in the data set.

Table 4. Data Summary: Category Marketing Mix by Store

		Store 1			Store 2		
	Units	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Milk Price	\$ / gallon	3.072	1.244	85062	3.653	0.932	51560
Cereal Price	\$ / 16 oz box	3.052	0.571	85062	3.075	0.603	51560
Soft Drink Price	\$ / case	4.775	1.784	85062	5.093	1.609	51560
Snack Price	\$ / 16 oz unit	3.760	0.955	85062	3.749	0.934	51560
Milk Feature	%	0.046	0.197	85062	0.054	0.179	51560
Cereal Feature	%	0.188	0.255	85062	0.270	0.283	51560
Soft Drink Feature	%	0.239	0.307	85062	0.346	0.330	51560
Snack Feature	%	0.193	0.274	85062	0.195	0.271	51560
Milk Display	%	0.071	0.218	85062	0.103	0.230	51560
Cereal Display	%	0.278	0.273	85062	0.374	0.267	51560
Soft Drink Display	%	0.556	0.312	85062	0.446	0.331	51560
Snack Display	%	0.596	0.303	85062	0.498	0.328	51560
Milk Promotion	%	0.393	0.426	85062	0.227	0.323	51560
Cereal Promotion	%	0.424	0.282	85062	0.588	0.254	51560
Soft Drink Promotion	%	0.705	0.288	85062	0.567	0.324	51560
Snack Promotion	%	0.682	0.282	85062	0.624	0.296	51560
Milk Inventory	oz	36.687	139.600	85062	44.740	138.157	51560
Cereal Inventory	oz	3.002	10.224	85062	4.134	13.742	51560
Soft Drink Inventory	oz	5.047	30.638	85062	15.480	111.814	51560
Snack Inventory	oz	3.833	12.238	85062	4.366	12.984	51560

Table 5. MVL Model of Multi-Category, Multi-Store Demand

	Store 1				Store 2			
	Model 1		Model 2		Model 1		Model 2	
	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio
Milk	12.1046*	15.0422	18.3686*	13.4786	11.3163*	12.3082	14.3290*	7.3797
Milk (σ)			1.4732*	5.4483			-1.0068*	-2.5760
Price	-7.0103*	-22.5891	-8.7592*	-17.0160	-8.4098*	-21.0969	-6.7089*	-16.8612
Price (σ)			-0.3270*	-2.7713			-0.5251*	-3.0329
Feature	-0.0457	-0.0477	-0.2907	-0.2305	0.0995	0.1823	0.2486	0.2544
Display	0.5164	0.5224	0.2380	0.1743	1.2369*	2.5197	0.0912	0.1046
Promotion	4.4391*	11.3927	1.9469*	3.7882	2.7757*	6.1200	-0.0453	-0.0891
Inventory	-1.0188	-0.6111	-1.6277	-0.5530	-4.3854*	-10.6605	-0.1932	-0.2747
Cereal	8.7656*	15.3344	7.6561*	20.7291	8.7481*	8.0958	13.1149*	3.3835
Cereal (σ)			0.0062	0.0869			0.1105	0.1857
Price	-1.3222*	-20.9479	-1.6086*	-31.3378	-1.2886*	-10.1593	-1.3096*	-3.0512
Price (σ)			-0.1973*	-13.6323			-0.0146	-0.1564
Feature	0.3235	0.7021	0.6371*	2.8273	0.0349	0.0547	0.8322	0.5633
Display	0.4449	1.0399	0.8440*	3.7111	0.1965	0.3879	0.9886	0.6180
Promotion	1.1991*	3.6694	1.6560*	8.3196	2.0798*	3.5125	3.1410	1.4853
Inventory	-0.7226	-0.6610	-2.4873*	-7.1274	-0.6979	-0.7105	0.4535	0.1563
Soft Drinks	14.4464*	7.5102	20.1588*	16.6782	13.9115*	8.3108	13.4149*	7.9896
Soft Drinks (σ)			0.0022	0.0137			-0.0471	-0.1330
Price	-0.6533*	-7.9134	-0.6567*	-19.4569	-0.6824*	-9.0667	-0.4435*	-7.8320
Price (σ)			0.3113*	13.7696			-0.0495*	-2.0024
Feature	-0.3287	-0.3551	-0.1030	-0.1473	0.7579	1.2045	0.4547	0.5187
Display	0.2800	0.5002	1.0130*	2.7668	-0.6661	-1.2127	1.0837	1.4469
Promotion	2.1956*	3.1785	8.7373*	15.7332	3.8631*	6.6824	1.9725*	2.6090
Inventory	0.0861*	21.6407	0.1051*	21.0581	0.0745*	19.1542	-0.0080	-0.5787
Snacks	9.2274*	7.9075	11.7258*	5.5514	9.3144*	6.6513	11.7982*	14.7533
Snacks (σ)			0.2510	0.9298			-0.2192*	-2.6926
Price	-0.7055*	-7.9421	-0.8645*	-7.2633	-0.7734*	-7.8032	-0.9909*	-15.6052
Price (σ)			0.0247	0.9463			0.0128	0.6982
Feature	0.1178	0.1049	-0.1529	-0.1835	0.4117	0.6125	0.4577*	2.0016
Display	1.0131	1.6019	1.4348*	2.0717	0.8784	1.4477	0.5221*	2.8932
Promotion	2.4482*	3.7841	3.3948*	3.6058	3.1019*	3.4486	1.6551*	7.2059
Inventory	0.1808	1.2787	-0.6784*	-5.5508	-0.2967*	-3.7248	-1.1160*	-22.8735
LLF	-3047.573		-2,182.12		-2525.1680		-1,750.74	
AIC	0.072		0.0521		0.1082		0.606	

Note: A single asterisk indicates significance at a 5% level.

Table 5 (con't). MVL Interaction Parameters

	Store 1				Store 2			
	Model 1		Model 2		Model 1		Model 2	
	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio
θ_{mc}	2.4374*	8.5632	2.1185*	11.4372	2.6727*	5.6804	1.0866	1.3189
$\theta_{mc}(\sigma)$			0.1343	1.1430			0.1197	0.1434
θ_{ms}	-1.2219*	-3.6901	-2.4622*	-5.4319	-0.3363	-0.8844	-3.7839*	-2.7763
$\theta_{ms}(\sigma)$			0.7692*	3.6420			0.0998	0.3360
θ_{mk}	1.7597*	5.8743	3.4071*	6.3082	3.1418*	7.6064	6.2546*	10.1128
$\theta_{mk}(\sigma)$			0.2890	0.5544			1.3787*	2.7741
θ_{cs}	-1.2903*	-4.1471	0.7789*	4.4442	0.5483	1.1205	-0.7963	-1.1997
$\theta_{cs}(\sigma)$			0.0216	0.2101			0.1751	0.3423
θ_{ck}	-0.1989	-0.8019	1.8649*	10.0422	-0.2914	-0.6101	-0.5671*	-1.9815
$\theta_{ck}(\sigma)$			0.1716	1.1772			0.4157	1.8034
θ_{sk}	9.8468	0.0136	3.3516*	6.9557	3.3509*	4.3560	0.5862*	2.1764
$\theta_{sk}(\sigma)$			0.1770	0.3229			0.1124	1.0371

Note: A single asterisk indicates significance at a 5% level. Subscripts are: m=milk, c=cereal, s=soda, and k=snacks.

Table 6. Models of Store Choice

	Model 1 Fixed Coefficient		Model 2 Random Coefficient	
	Estimate	t-ratio	Estimate	t-ratio
Loyalty: Store 1	3.5220*	61.3050	3.5226*	61.3051
			0.0030	0.1953
Price: Store 1	-0.1329*	-5.4603	-0.1332*	-5.4727
			0.0022	0.7817
Variety: Store 1	-0.1834*	-6.7112	-0.1834*	-6.7134
			0.0026	0.6014
Distance: Store 1	-0.3864*	-27.1758	-0.3864*	-27.1518
			0.0020	0.8919
Loyalty: Store 2	1.1756*	20.1444	1.1754*	20.1336
			0.0214	1.1318
Price: Store 2	-0.1714*	-5.6645	-0.1714*	-5.6639
			0.0007	0.2347
Variety: Store 2	0.6235*	11.8790	0.6231*	11.8669
			0.0055	0.9048
Distance: Store 2	-0.3602*	-25.2560	-0.3600*	-25.2461
			0.0012	0.4491
Lambda	0.0713*	51.6667	0.0713*	51.6667
			0.0003	0.3140
LLF	-44,498.6		-44,496.5	
AIC/N	0.6520		0.6520	

Note: A single asterisk indicates significance at a 5% level.

Table 7. Independence Model of Multi-Category Demand

	Model 1		Model 2		Model 3	
	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio
Milk	5.9040*	34.5304	5.6435*	37.8427	6.4187*	34.5036
Milk (s)			0.1274*	2.7293	0.3667*	6.0146
Price	-4.5445*	-48.4849	-4.6356*	-57.8656	-5.0295*	-52.3194
Price (s)					0.1020*	2.5417
Feature	0.3797*	2.0962	0.1934	1.1797	0.4138	1.8360
Display	0.1982	1.1523	-0.0863	-0.5736	-0.0790	-0.3476
Promotion	1.4473*	12.1008	1.8720*	17.5820	1.5418*	12.1647
Store 2	-0.2733*	-3.9351	-0.1996*	-2.8402	-0.2147*	-2.4431
Inventory	-1.6766*	-12.9070	-0.6157*	-2.6928	-0.4089	-1.0807
Cereal	6.2055*	21.9081	6.6360*	23.0954	6.5087*	21.1574
Cereal (s)			0.1222*	2.1657	0.0441	0.5745
Price	-0.7604*	-21.0105	-0.9819*	-25.2801	-0.9326*	-22.5048
Price (s)					0.0351*	2.6546
Feature	0.1594	1.0642	0.2522	1.8654	0.3689*	2.0133
Display	0.1916	1.2773	0.2544	1.8990	0.3984*	2.1155
Promotion	1.7898*	9.5387	1.7664*	11.1040	1.2274*	6.8687
Store 2	-0.2812*	-2.7384	0.0516	0.4792	0.2441	1.9425
Inventory	-0.9277*	-3.2418	-0.3656	-0.9992	0.1449	0.2716
Soft Drinks	6.1217*	52.7002	6.2300*	36.6426	6.8830*	23.0455
Soft Drinks (s)			0.4160*	6.6355	0.1636	1.6951
Price	-0.3795*	-47.7997	-0.3266*	-32.6264	-0.3182*	-20.9074
Price (s)					-0.0123	-1.8363
Feature	0.4043*	3.6608	0.2049	1.5605	0.1121	0.5031
Display	0.4226*	4.2849	0.3014*	2.6448	0.5613*	3.0634
Promotion	2.0888*	22.8035	2.2722*	19.3408	2.0322*	11.5250
Store 2	0.4164*	6.3566	0.1085	1.3617	-0.3072*	-2.3735
Inventory	0.3675*	331.0811	0.2129*	183.5431	0.0232*	10.7721
Snacks	5.7296*	27.0353	5.9175*	28.9929	6.6734*	21.4089
Snacks (s)			0.0439	1.0844	0.1681*	2.4123
Price	-0.5451*	-28.4324	-0.5973*	-32.2011	-0.5252*	-21.0501
Price (s)					-0.0161	-1.6016
Feature	0.1422	1.2175	0.1631	1.4607	0.2561	1.4682
Display	0.5923*	5.5301	0.5831*	5.9169	0.6582*	4.0599
Promotion	1.3806*	11.5636	1.4184*	12.8946	1.5728*	8.5900
Store 2	-0.0949	-1.1938	-0.2661*	-3.4305	-0.2937*	-2.4390
Inventory	0.2766*	16.2700	0.1924*	7.4213	-0.1252*	-4.0475
LLF	-30,724.2371		-27,227.9084		-20,103.8512	
AIC	0.4502		0.3991		0.2950	

Note: A single asterisk indicates significance at a 5% level.

Table 8. Equilibrium Pricing: MVL Demand Model

	Model 1: OLS		Model 2: GMM	
	Estimate	t-ratio	Estimate	t-ratio
Primary Input	38.1003*	60.4488	46.9283*	69.5749
Retail Wage	0.7294	1.8795	5.5260*	2.7584
Food Mfg Wage	1.1155*	2.8445	-1.4505*	-2.1192
Energy	-1.3991*	-2.9753	-9.5150*	-5.4911
Business Services	-7.9635*	-2.7497	-16.1443	-1.5003
Packaging	-1.0403	-1.4966	6.1348*	1.9888
Margin	0.1202*	28.4834	0.1608*	37.2199
LLF / GMM	-2313.4360		464.1010	
Chi-Square	459.2150		232.0510	

Note: A single asterisk indicates significance at a 5% level.

Table 9. Equilibrium Pricing: Independent Demand Model

	Model 1: OLS		Model 2: GMM	
	Estimate	t-ratio	Estimate	t-ratio
Primary Input	0.3555*	57.4313	0.4558*	67.7325
Retail Wage	0.0203*	4.9754	0.1443*	3.5153
Food Mfg Wage	0.0152*	3.7230	0.0113	0.5516
Energy	-0.0149*	-3.0553	0.0830	1.8495
Business Services	-0.1588*	-5.2316	0.1602	0.7203
Packaging	-0.0102	-1.4202	-0.0569	-1.0901
Margin	0.6584*	25.1586	1.1610*	35.7876
Store 1	-0.0976*	-1.9689	-0.8393*	-3.1399
LLF / GMM	-2,374.5783		244.1453	
Chi-Square	264.4269		122.0726	
	Mean	Std. Dev.	Min.	Max.
Price	3.3006	1.8374	0.0025	13.8877

Note: A single asterisk indicates significance at a 5% level. Price is the fitted value of the average retail price over all categories.

Table 10. Equilibrium Prices with Simulated Demand

	C_{ij} = -2	C_{ij} = -1	C_{ij}=Base	C_{ij} = +1	C_{ij} = +2
Mean	3.1235	3.1592	3.1584	3.2932	3.3131
Standard Deviation	1.8488	1.8475	1.7833	1.7897	1.7903

Note: Base case calculated at estimated parameters in Table 7.