A Supermarket-Level Analysis of Demand for Breakfast Cereals: A Random Coefficients Approach

Benaissa Chidmi, Rigoberto A. Lopez, and Ronald W. Cotterill

Department of Agricultural and Resource Economics
The University of Connecticut

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, July 24-27, 2005

A Supermarket-Level Analysis of Demand for Breakfast Cereals: A Random Coefficients Approach

1. Introduction

Demand analysis constitutes a cornerstone in the analysis of the industry and consumer behavior. Unlike the demand for homogeneous products, the demand for differentiated products presents two challenges: the dimensionality problem and the consumer heterogeneity issue. Due to differentiation, the number of brands is too large (for instance, in ready-to-eat cereal industry there 200 different brands) making the classical methods of demand such as the Linear Expenditure (Stone, 1954), the Rotterdam (Thail, 1965; and Barten 1966), the Translog (Christensen, Jorgensen and Lau, 1975), and the Almost Ideal Demand System models (Deaton and Muellbauer, 1980) practically impossible to implement without assuming a restricted pattern of substitution. Furthermore, these methods ignore the consumer heterogeneity issue, which offers the basis for marketing segmentation, targeting and positioning, as well as micro marketing (Kamakura et al., 1996).

To solve the dimensionality problem, Spence (1976) and Dixit and Stiglitz (1977) proposed the constant elasticity of substitution (CES) utility function where the dimensionality problem is solved by imposing symmetry between different brands, which implies the estimation of a single parameter regardless of the number of brands. This strong restriction implies that the cross-price elasticities are equal, thus leading to inappropriate conclusions. Another approach to solve the dimensionality problem is to divide the brands into smaller categories and use a flexible functional form to estimate the demand within each category, thus reducing the number of parameters to be estimated (Hausman, Leonard, and Zona (1994), Cotterill (1994), Hausman (1996), and Ma (1997))[1]. The problem with this approach resides in the difficulty of an arbitrary division across categories. For instance, Hausman (1996) uses only three categories (adult, kid, and family) while Cotterill (1994) and Ma (1997) use four categories (all family, taste enhanced wholesome, simple health, and kid cereal).

For the consumer heterogeneity issue, most recent advances have been in the area of marketing (For a summary of a selected number of studies including consumer heterogeneity, see Leszczyc and Bass (1998)). Ignoring consumer heterogeneity in studying consumer brand processes may lead to biased results and hence inaccurate inferences regarding marketing strategies (Segmentation, positioning, targeting and micro marketing).

This paper applies the Berry, Levinsohn and Pakes (BLP, 1995) discrete choice procedure developed by McFadden (1973, 1981, and 1984) to estimate the demand for 37 brands of ready-to-eat cereal (RTEC) at the supermarket chain level in Boston area. The dimensionality problem is solved by projecting the consumer choices onto a set of brand characteristics giving smaller dimension than the number of brands. The product and consumer heterogeneity are taken into account by the use of random coefficients discrete choice models (Cardell, 1989; Berry, Levinsohn, and Pakes, 1995, Nevo 2000), which offer the advantage to solve the problems of substitution pattern implied, by the use of classical discrete choice models such as the Logit, the Probit, and the Nested Logit. The

[1] Hausman, Leonard, and Zona (1994) and Hausman (1996) divide the demand estimation problem in three stages: the first stage concerns the demand for the product (RTEC), the second stage corresponds to different categories (kids, adults, family cereals) and the last stage analyzes the demand for different brands within each category.

main contribution of this paper is that it is the first study to estimate a discrete choice andom coefficients demand system for branded products at the chain as opposed to market level[2]. Unlike the previous works, the level of desegregation is such that one is able to estimate the demand for RTEC brands at the supermarkets level, hence combining the brands and supermarkets choices, and approximating the relevant levels of competition in the real world. Nevo (2000), for instance, used brand at the city level, sidestepping the importance of supermarkets.

The level of desegregation is such that one is able to estimate the demand for RTEC brands at the supermarkets level, hence combining the brands and supermarkets choices. It also study uses four-week data while prior brand demand analysis uses quarterly (Nevo (2001); Hausman et al. (1994)) or weekly (Kadiyali et al. (1999); Cotterill (1994)) observations[3].

2. A Model with Product and Consumer Heterogeneity

This section presents a general model of demand that takes into account the product differentiation and the consumer heterogeneity. The general intuition is that the consumer chooses the brand that maximizes its utility. The choice is driven by the brand characteristics and the consumer characteristics. While some of the consumer and brand characteristics are observable (to the researcher), other characteristics are not. Consumers with different observed and unobserved characteristics make different choices. The indirect utility[4] of consumer *i* from buying the brand *j* is given by

[2] Cotterill and Dhar (2002) is the only prior chain level demand study and it uses a nested logit model.
[3] There is no consensus concerning which time unit is desirable. Quarterly may be too aggregate, while weekly may be too disaggregate to measure strategic pricing moves in a static equilibrium model.
[4] The indirect utility comes from a quasi-linear utility function.

$$U_{ij} = \beta_j + x_j\beta_i - \alpha_i p_j + \zeta_j + \varepsilon_{ij} \ , \qquad i = 1,...n \qquad j = 1,..., J \tag{1}$$

where β_j represents the store/brand fixed effects, x_j are the observed product characteristics of brand *j*, p_j is the price of the brand j, ζ_j are the unobserved (by the econometrician) product characteristics, and ε_{ij} represents the distribution of consumer preferences about the unobserved product characteristics, with a density $f(\varepsilon)$. The parameters to be estimated are α_i and β_i. Note that those parameters are allowed to vary across consumers, therefore taking into account the heterogeneity taste of consumer.

These coefficients can be decomposed into a fixed component and a variable component (changing with consumers' observed and unobserved characteristics). This decomposition can be expressed as:

$$\alpha_i = \alpha + \lambda D_i + \gamma v_i \tag{2}$$

$$\beta_i = \beta + \varphi D_i + \rho v_i \tag{3}$$

where the D_i represents the consumers' observed characteristics such as demographics variables (e.g., income and age), and v_i denotes the unobserved consumers' characteristics.

Substituting (2) and (3) in (1) yields

$$U_{ij} = \beta_j + x_j\beta + \lambda D_i x_j + \gamma v_i x_j - \alpha p_j - \lambda D_i p_j - \gamma v_i p_j + \zeta_j + \varepsilon_{ij} \tag{4}$$

Unobserved consumer characteristics v_i are assumed to be normally distributed $N(0, I)$, where *I* is the identity matrix; and the observed consumer characteristics D_i have an empirical distribution $h(D)$, not necessarily a normal distribution, from the demographic data.

The indirect utility given in equation (4) can be decomposed into two parts: a mean utility given by $\delta_j = \beta_j + \beta x_j - \alpha p_j + \zeta_j$ and a deviation from that mean, which is a function of the interaction between the observed and unobserved consumer's characteristics and the price and observed brand characteristics, given by

$$\mu_{ij} = \lambda D_i x_j - \lambda D_i p_j + \gamma v_i x_j - \gamma v_i p_j + \varepsilon_{ij} . \tag{5}$$

To complete the model, an outside good is included to give the consumer the possibility not to buy any one of the J brands included in the choice set[5]. The utility of the outside good is normalized to be constant over time and equal zero. For the case at hand, the outside good can include all other brands, or the residual brands not included in the study.

Given the observed and unobserved consumer characteristics define the set of choice by

$$S_j(x_j, p_j, \zeta_j; \theta) = \{(D_i, v_i, \varepsilon_{ij}) : U_{ijt} \geq U_{ik} \forall k = 0,1,...N\} , \tag{6}$$

where θ is a vector that includes all the parameters of the model.

The consumer purchases the brand that gives the highest utility. The market share of the jth brand corresponds to the probability the jth brand is chosen. That is,

$$s_j = \int I\{(D_i, v_i, \varepsilon_{ij}) : U_{ij} \geq U_{ik} \forall k = 0,1,...N\} dH(D) dG(v) dF(\varepsilon). \tag{7}$$

Depending on the assumptions regarding D, v , and ε , the integral in (7) can have or not a closed formula. In a general setting, the integral in (7) does not have a closed formula and should be solved numerically (BLP, 1995; Nevo, 2000).

Using (7), the price elasticities of the market shares are

[5] The inclusion of the outside good is necessary in order to accomplish with the exhaustiveness of alternatives of the discrete choice model. For a detailed discussion, see Train (2002).

$$\eta_j = \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = \begin{cases} \dfrac{p_j}{s_j}\displaystyle\int \alpha_i s_{ij}(1-s_{ij})dH(D)dG(\nu), & \text{for } j=k, \\ -\dfrac{p_k}{s_j}\displaystyle\int \alpha_i s_{ij} s_{ik} dH(D)dG(\nu), & \text{otherwise.} \end{cases} \tag{8}$$

In this setting, each consumer can have different price elasticity for each individual brand. The substitution patterns are not constrained by a priori segmentation of the market. If, for example, the price of a brand increases consumers are more likely to switch to brands with similar characteristics, rather than to the most popular brand. Also individual with similar characteristics will tend to have similar purchasing patterns.

3. Model Estimation

The estimation proceeds by computing the integral in (7), either analytically or numerically. Different models can be implemented depending on the assumptions on the distribution of the unobserved consumer characteristics. When consumer heterogeneity is integrated in the random shock ε_{ij}, the integral in (7) can be solved analytically. In the case ε_{ij} is distributed i.i.d. with a Type I extreme value distribution, the solution of (7) yields the Multinomial Logit model, henceforth the Logit model.

3.1. The Multinomial Logit Model

The simplest case is to assume that consumer heterogeneity enters only through the error term ε_{ij} In this case, there are not random coefficients since the parameters α_i and β_i are not varying across consumers. Under this assumption equation (4) becomes

$$U_{ij} = \beta_j + x_j\beta - \alpha p_j + \zeta_j + \varepsilon_{ij}. \tag{9}$$

Further assume that ε_{ij} is distributed i.i.d. with a type I extreme value distribution, i.e., $f(\varepsilon) = e^{-e^{-\varepsilon}}$. Then the integral in equation (7) can be solved analytically and the market shares for the *jth* brand (corresponding to the probability that the *jth* is chosen) is given by the following equation:

$$s_{jt} = \frac{\exp(x_j \beta - \alpha p_j + \zeta_j)}{1 + \sum_{k=1}^{J} \exp(x_k \beta - \alpha p_k + \zeta_k)}. \qquad (10)$$

Equation (10) corresponds to the Logit model. This model presents the advantage to be simple to implement. The estimation of its parameters is based on the inversion proposed by Berry, Levisohn and Pakes (BLP, 1995) given by

$$\delta_j = \ln(s_j) - \ln(s_0) = \beta_j + \beta x_j - \alpha p_j + \zeta_j\,, \qquad (11)$$

where s_0 is the market share of the outside good, obtained by subtracting the sum of observed market shares of all the inside brands from 1. Note that the logit model is transformed to a simple linear regression where the natural logarithm of the ratio between the observed market shares inside good with respect to outside good is regressed on product characteristics and the price variables.

The price elasticities of the market shares given by equation (8) reduce to

$$\eta_{jk} = \frac{\partial s_j}{\partial p_k} \frac{p_k}{s_j} = \begin{cases} \alpha p_j (1 - s_j), & if \;\; j = k, \\ -\alpha p_k s_k\,, & otherwise. \end{cases} \qquad (12)$$

Equation (12) poses two problems. The first problem is what is called independent from irrelevant alternatives (IIA) (McFadden, 1981; Train (2000)). Notice that the ratio of the logit market shares for any two brands *j* and *k* does not depend on any brands other than *j* and *k*. That is, the relative odds of market shares of brand *j* over brand *k* are the same no matter what other brands are available or what the characteristics of the other brands are.

While the IIA property is appropriate in some cases, it is very restrictive in many cases. For instance consider the case of two RTEC: a kids' cereal (KA) and a family's cereal (F) that have similar market shares. If the price of another kids' cereal (KB) increases, the substitution from KB to F and from KB to KA will be the same. However, in real life, one would expect the substitution to take place across kids' cereals only, i.e. from KB to KA.

The other problem is that the own-price elasticities are related to own prices. A lower price implies a lower elasticity, which may imply inaccurate measure of price-cost margins. The simplicity of the Logit model gives the researcher an opportunity to test how well the data is behaving. However, due to its limitations, the Logit model should not be used to infer any type of conclusion regarding market structure or market power (Nevo, 2000). To do so, one needs a more elaborate model that circumvents the Logit model shortcomings. The random coefficients model constitutes a good candidate that, at least, provides a more flexible substitution pattern and own price elasticities.

3.2. The Random Coefficients Model

This model is much more complicated than the Logit model since it allows for consumer heterogeneity α_i and β_i as described in (2) and (3), that is each consumer is different from another consumer in their response to price and brand characteristics. The Random Coefficients model (henceforth RCM) poses two challenges. First, the integral in equation (7) has no closed formula and should be solved numerically[6]. Second, information on the distribution of demographics is needed to compute the individual market shares.

[6] The integral in (7) is solved using the simulation technique proposed by Pakes (1986).

The solution of the integral in (7) is based on the choice of the parameters that minimize the distance between the predicted market shares given by equation (7) and the observed market shares. That is,

$$\underset{\theta}{Min}\left\| s(x, p, \delta; \theta) - S \right\|, \tag{13}$$

where $s(.)$ represents the market shares given by equation (7) and S are the observed market shares. However, this approach implies a costly non-linear minimization procedure because all the parameters enter (12) in a non-linear manner.

To avoid this difficulty, Berry (1994) suggests inverting the market share function giving the mean utility valuation δ that equates the predicted market shares with observed market shares $s_{.}(\delta_{.};\theta_2) = S_{.}$, where θ_2 (the notation is borrowed from Nevo (2000)) is a vector of parameters that enter the indirect utility function non-linearly.

Once the mean utility valuation δ is obtained the next step is to define the error term as the deviation from that mean. That is,

$$\omega_j = \delta_j(S_{.};\theta_2) - (x_j\beta + \alpha p_j). \tag{14}$$

Note that it is the observed and not the predicted market shares that enter the error term. The error term is then interacted with instruments to form an objective function to be minimized using the generalized methods of moments (GMM) estimation.

4. Data Sources and Management

The data used in the above analysis consists of two kinds of variables: retail sales variables and demographic variables.

The sales data were obtained from the Information Resource, Inc. (IRI) Infoscan database at the Food Policy Marketing Center of University of Connecticut. It covers RTEC sales for 37 brands at the four leading supermarkets in Boston (Stop & Shop, Shaw's, DeMoulas and Star Market) for four-weekly periods between April 1995 and December 1997. One important feature of this period is that it covers significant price drops in the 1990s when the RTEC industry was being questioned on market power (Cotterill, 1999, and Connor, 1999). The sales data collected consists of the following variables: dollar sales, volume (in pounds) sales, and the percent volume sold with any feature.

From the RTEC sales data, the market shares and the retail prices were computed for each brand and supermarket. Market shares are obtained by converting volume sales into number of servings sold and dividing by the potential market size. This is done by using the serving weight found on the box of cereals. The potential market size is assumed to be one serving per capita and per day as in Nevo (2001). The real retail prices were computed by dividing the dollar sales of each brand by the number of servings sold and then deflated using the urban consumers CPI for Boston (with CPI=100 for 1981).

The analysis is conducted using a set of 37 RTEC brands produced by six manufacturers (Kellogg's, General Mills, Post-Kraft, Quaker, Ralston and Nabisco) sold in four supermarket channels (Stop & Shop, Shaw's, DeMoulas, Star market and all other chains) in Boston market from April 1995 to December 1997 for 6475 observations.

Primary data on product characteristics were collected by examining the cereal boxes. The variables collected were the content of calories, sugar, proteins, vitamins, minerals, sodium, potassium, fiber and total fat. The characteristics also included dummy variables for corn, oat, rice and fruits and a dummy variable for kids' cereals. It is assumed that those characteristics did not change since between 1995 and 1997. Besides the sales data, the analysis uses the demographic data to take account of the heterogeneity of consumer taste. This paper uses two demographic variables: the natural logarithm of age and income. Further it is assumed that those variables are jointly normally distributed with mean given by the grocery data and variance-covariance matrix given by the CPS data at Boston level.[7]

The demand model presented above implies endogeneity of RTEC prices, and, hence, can lead to biased parameter estimates[8]. This implies that prices are correlated with product characteristics. This study uses a set of instrumental variables to control for retail price endogeneity in a particular supermarket. The set has two subcomponents. The first one consists of the interaction between input prices and brand dummy variables, where input prices (wages in the Boston area and the price of gas, the price of industrial and commercial electricity at the location of manufacturers, the Federal Funds Effective interest rate, and the 3-month Commercial Paper interest rate) were interacted with brand dummy variables. The second subcomponent consists of time dummy variables describing the jawboning campaign events that induced price drops (change in conduct) by RTEC manufacturers, as described by Cotterill (1999) and Connor (1999).

All the price instruments mentioned above were interacted with the error terms when applying the GMM estimation procedure. The use of GMM technique implies the need for an optimal weighting matrix. This paper follows Hansen (1982) who shows that setting the weighting matrix equal the inverse of an asymptotic covariance matrix is optimal in the sense that it gives parameter estimates with the smallest asymptotic variance.

[7] Romeo (2005) shows that knowing the joint distribution for demographics at the city level is sufficient to infer the distribution at the county or zip code levels.
[8] This endogeneity comes from the fact that retail prices depend on observed and unobserved product characteristics. Any variation in those characteristics induces a variation in retail prices.

5. Empirical Results

5.1. The Logit Demand

Table 1 presents the results of the regression of equation (11) ($\ln(s_{jt}) - \ln(s_{0t})$) on prices, promotion and product characteristics. The characteristics included are the content of calories, sugar, fiber and a dummy variable for kids' cereals.

The second column of Table 1 presents the results of the ordinary least squares (OLS) regression, while column 3 presents the instrumental variables (IV) results of a two-stage least squares regression. The OLS and IV results are mixed. Hence, as one would expect, the parameter estimates of the price, calories and sugar are negative, though the sugar parameter estimate is not significant. In the other hand, the promotion, fiber and kid dummy variable are not of the expected sign. For example, one would expect a positive effect of the fiber content and promotion on the market share of the RTEC brand. For the kid dummy variables, a possible explanation of the negative sign would be that given the high level of retail prices, household with kids tend to opt for other cheaper breakfast alternatives.

Table 2 presents the own-price elasticities estimated from the Logit model as given in equation (11). As expected, all the own-price elasticities are negative with a magnitude greater than one in absolute value. This implies that at the supermarket level the demand for differentiated brands is elastic. The elasticities range from –7.6819 for Ralston Cookie Crisp in Star Market to –2.6032 for Kellogg's Corn Flakes in Stop & Shop. Figure 1 shows the box plots for the own price elasticities for each supermarket. Note that the outlier in each supermarket is the Ralston Cookie Crisp. Figure 1 also shows that in average the elasticities are low in absolute value in Shaw's compared to the other supermarkets, the Star Market being the most elastic. In the other hand, the analysis of the variance (Table 3) shows that the difference between own-price elasticities across supermarkets is statistically significant.

Turning now to the cross-price elasticities, equation (12) shows that the cross-price elasticity of the brand *j* with respect to brand *k* does not depend on brand *j* price or market shares, it only depends on the share and the price of brand *k*. This implies the same cross-price elasticity for all *k*. This means that a price increase for brand *j* will increase the market share of the other brands by the same percentage. This issue constitutes the biggest disadvantage of the logit model.

5.2. The Random Coefficients Model Results

The advantage of the RCM is allowing for heterogeneity of tastes across consumers. The approach taken above will allow having a sort of measurement of the heterogeneity, which includes the influence of demographic variables (such as income and age) and also the effect of the unobserved consumer characteristics. The RCM model also remedies to the cross-price substitution problem posed by logit model.

The results of the estimation are presented in Table 3. The first set of parameter estimates (Means (β's)) gives the parameter that enter the indirect utility function linearly or the mean valuation utility. The second set gives the interaction between the unobserved consumer characteristics (random draws from a multivariate normal distribution) and the brand characteristics. The third and fourth set gives the interaction of the brand characteristics with the age and income variables, respectively.

The parameter estimates of the means of the distribution of the marginal utility (β's) or the mean valuation of utility are all significant and have expected signs except few. For the average consumer, the price has negative marginal utility, as one would expect. Similarly, the calories content has a negative effect on the mean valuation of the utility. The fiber coefficient is negative and significant, this result is consistent with Stanley Tschirthart (1991) but different from the finding of Nevo (2001). Stanley and Tschirthart attribute their finding to the taste component of fiber dominating the nutrition component, while Nevo attributes his finding to the nutrition component that dominates the taste component. For the average consumer, sugar has positive marginal utility; this is consistent with the finding of Stanley and Tschirthart (1991) and Nevo (2001). Notice also that the promotion coefficient is positive and significant, as one would expect that merchandising would increase the marginal utility of the brand.

The estimates of the interaction between the taste parameters and the unobserved consumer characteristics are mainly not significant except for sugar and the calories contents. It seems that the unobserved consumer characteristics react negatively with calories content while consumers approve, for some reason, the sugar content. A surprising and totally different from Nevo's findings (2001) is that the interaction of the product attributes with age are not significant, making one to conclude that the age of the consumer is not an important when consumers make their decisions to by RTEC. The interactions with income are all significant in other hand.

The distribution of the individual price sensitivity is given in figure 3. The figure shows that the individual price coefficient is normally distributed.

The implied own-price RCM elasticities are given in table 4. Compared to the logit-implied elasticities, the RCM elasticities are lower in absolute value; this pattern conforms to Nevo (2001) and Villas-Boas (2002). Regarding the magnitude of the elasticities, the elasticities of this paper are greater (in absolute value) to those found by Nevo (2001). However, the latter were at the city level while the present elasticities are at the supermarket level. Figure 4 gives the box plots of the own-price elasticities for each supermarket. The box plots show that in Shaw's supermarkets RTEC brands show more inelastic behavior that in the other supermarket chains, with Star Market being the most elastic. The own-price elasticities in Shaw's varies between –6.0686 for Ralston Cookie Crisp to –2.5697 for Kellogg's Corn Flakes suggesting a margin varying between 16% and 39 %. For Stop & Shop, the own price elasticities vary between –6.8381 for Ralston Cookie Crisp to –2.4439 for Kellogg's Corn Flakes. Notice also that in the four supermarket chains, Kellogg's Corn Flakes, is the most inelastic, while the Ralston Cookie Crisp is the most elastic. Analysis of variance was conducted to see the effect of the retailers and the manufacturers on the level of elasticities. The results show that linear effects of those two factors are significant, while the interaction term is not significant, indicating a little evidence that the origin of the brand (manufacturer) has an effect on where the brand is sold.

For the cross-price elasticities, due to the size of the data set it is impossible to present all the cross-price elasticities. In general, all the cross-price elasticities are positive and vary across brands as oppose to those implied by the logit model. To better interpret these elasticities, the brands are grouped in segments as in Cotterill (1994) and Ma (1997). Table 5 presents a sample of own- and cross-price elasticities in Stop $ Shop. For example, Kellogg's Corn Flakes is more sensitive to a change in the price of GM Cheerios, also an all family brand, than to GM Lucky Charms, a kids cereal. Another pattern that is obvious from the table is that the Kellogg's and GM brands are less sensitive to the brands from other manufacturers. The latter seem to be very sensitive to a change in the price of the leading brands, mainly GM Cheerios.

Table 6 presents the implied cross-price elasticities for three manufacturers brands across supermarkets. The results show that for stop & shop, the leading supermarket chain, the brands are more sensitive to the change in price of the brands sold in their supermarkets. Interestingly, in the rest of the supermarket chains, the brands are mostly sensitive to the change in the price of the brands sold in Stop & Shop. For example, Kellogg's Corn Flakes sold in Shaw's is more sensitive to a change in the price of Kellogg's Corn Flakes sold in Stop & Shop than the change in the price of GM Cheerios sold in Shaw's. These substitution patterns are consistent across the table and show that for a given manufacturer brand, the Stop & Shop supermarket chain is the price leader for the RTECs in Boston.

6. Conclusion

This paper used a random coefficients discrete choice approach to estimate the demand for RTECs at the top four supermarkets in Boston area. The empirical results provide a wealth of consumer behavior information, including own- and cross-price elasticites for 37 brands of RTECs at the four leading supermarkets in Boston. Consumers respond positively and strongly to promotion, negatively and strongly to price, calories and fiber, and weakly to sugar content. Income has a strong interactive effect with product characteristics and thus is a useful variable for market segmentation. In comparison, the results with the more commonly used Logit model indicate significantly lower price elasticities, provide a limited window on consumer behavior, and yield predicted brand and supermarket market shares that are quite divergent from observed values. Since the BLP approach, by construction, predicts market shares calibrated to observed ones, it also provides a preferred benchmark for consumer behavior and marketing decisions.

The demand for RTECs is generally price elastic (ranging between -3 and -8). The substitution patterns given by the implied cross-price elasticities show that brands of RTEC are more sensitive to the brands that are in same segment/category than to the brands in other categories. At the manufacturer level, Post, Quaker, Nabisco and Ralston brands seem to be more sensitive to the leading manufacturers brands, mainly GM Cheerios. At the retail level, the results show that for a given manufacturer brand, Stop &Shop is the price leader for this particular product.

References

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

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

Cotterill, R. W. 1999. "Jawboning Cereal: The Campaign to Lower Cereal Prices." *Agribusiness: An International Journal*, 15, No. 2, pp. 197-205.

Cotterill, Ronald W. and Haller, Lawrence E. 1997. "An Economic Analysis of the Demand for RTE Cereal: Product Market Definition and Unilateral Market Power Effects." University of Connecticut Food Marketing Policy Center Research Report No. 35 (September).

Connor, J.M., 1999. "Breakfast Cereals: The Extreme Food Industry," *Agribusiness: An International Journal*, 15, No. 2, pp. 257-259.

Hausman, J., G. Leonard, and J.D. Zona, 1994. "Competitive Analysis with Differentiated Products." *Annales D'Economie et de Statistique*, 34: 159-180.

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

Kamakura, Wagner A., B.D. Kim and J. Lee, 1996. "Modeling Preference and Structural Heterogeneity in Consumer Choice." *Marketing Science*, 15, No. 2, pp. 152-172.

Leszczyc, Popkowski P.T.L., and F.M. Bass, 1998. "Determining the Effects of Unobserved and Observed Heterogeneity on Consumer brand Choice." *Applied Stochastic Models and Data Analysis*, 14, pp. 95-115.

Ma, L.Y., 1997. "An Econometric Analysis of Competition in a Differentiated Product Industry: The U.S. Ready-to-Eat Cereal Industry," Ph.D. Dissertation. Department Of Agricultural & Resource Economics, University of Connecticut.

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

McFadden, D. and K. Train, 2002. "Mixed MNL Models of Discrete Response," *Journal of Applied Econometrics*, 15, No. 5, pp.447-470.

Nevo, A., 2000. “A Practitioner’s Guide to Estimation of random Coefficients Logit Models of demand,” *Journal of Economics and management Strategy,* 9, No. 4, pp.513-548.

Nevo, A., 2001. “Measuring Market Power in the Ready-To-Eat Cereal Industry,” *Econometrica,* 69, No. 2, pp.307.342.

Romeo, C., 2005. “Estimating discrete joint probability distributions for demographic characteristics at the store level given store level marginal distributions and a city-wide joint distribution.” *forthcoming Quantitative Marketing and Economics*.

Scherer, F.M., 1979. “The Welfare Economics of Product Variety, An Application to the Ready-To-Eat Cereals Industry,” *Journal of Industrial Economics,* 28 (December), pp. 113-134.

Stanley, Linda R., and John Tschirhart (1991), “Hedonic Prices for a Nondurable Good: The Case of Breakfast Cereals”, *Review of Economics and Statistics*, 73, No. 3, pp.537-541.

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

Villas-Boas, B.S., 2002. “Vertical Contracts Between manufacturers and Retailers: An Empirical Analysis,” *Working Paper*, University of California Berkeley.

Table 1 RTEC Demand in Boston: Logit Parameter Estimates.

Variables	OLS Estimates	t-Statistic	IV Estimates	t-Statistic
Price	-27.6796***	-91.1868	-29.3303***	-79.2719
Promotion	-2.1818***	-48.1693	-1.9268***	-17.2523
Calories	-1.0058***	-19.7234	-0.8489***	-17.9813
Fiber	-0.0786***	-7.7998	-0.1003***	-7.6590
Sugar	-0.0047	-0.1550	-0.0154	-0.1541
Kid Dummy	-0.0805***	-2.6299	-0.0478***	-2.5793
Goodness of Fit[1]	R-Squared	0.9826	Over Identification Test	433
			Critical value (95%)	79.49

Table 2. Logit Own-Price Elasticity Estimates for RTECs in Boston Supermarkets

RTEC Brand	Stop & Shop	Shaw's	DeMoulas	Star Market	Simple Average
KApple Jacks	-5.3171	-4.9800	-4.8116	-5.5592	-5.1670
KComplete Bran	-4.2344	-4.1182	-4.1363	-4.5140	-4.2507
KCorn Flakes	-2.6032	-2.7239	-2.9059	-3.0153	-2.8121
KCorn Pops	-4.4817	-4.3886	-4.9545	-5.0138	-4.7097
Kcrispix	-4.8514	-4.6789	-5.3242	-5.5049	-5.0899
Kfroot Loops	-4.8821	-5.1089	-4.5405	-4.9259	-4.8644
Kfrosted Flakes	-3.6847	-3.8024	-3.5929	-3.9250	-3.7513
Kfrosted Mini Wheats	-3.6570	-3.5713	-3.3485	-3.9077	-3.6211
Kraisin Bran	-3.2185	-3.4598	-3.1845	-3.6343	-3.3743
Krice Krispies	-4.5093	-4.4151	-4.5518	-4.7120	-4.5471
Kspecial K	-5.1552	-5.1877	-5.5871	-5.5188	-5.3622
GMcheerios	-3.9962	-3.6751	-4.4856	-4.5667	-4.1809
GMCinammon Crunch	-4.9142	-4.5534	-5.1227	-5.1516	-4.9355
GMCoco Puffs	-4.7944	-4.6863	-5.1181	-5.0469	-4.9114
GMGolden Grahams	-5.2746	-4.8709	-5.2885	-5.8284	-5.3156
GMHoney Nut Cheerios	-4.4274	-4.3088	-4.2448	-4.5235	-4.3761
GMKix	-5.1054	-4.8513	-5.9457	-5.9259	-5.4571
GMLucky Charms	-4.9513	-4.5748	-5.2368	-5.0749	-4.9595
GMMulti Gain Cheerios	-5.4907	-5.4033	-5.2242	-6.0752	-5.5484
GMTotal	-5.2002	-4.7145	-5.4030	-5.6860	-5.2509
GMTotal Raisin Bran	-4.1328	-4.2822	-4.1110	-4.1125	-4.1596
GMWheaties	-3.5949	-3.5154	-3.8788	-4.2747	-3.8160
GMApple Cinnamon	-4.4796	-4.5976	-4.5272	-4.9038	-4.6271
Pbanana Nut Crunch	-4.2908	-3.7730	-4.5177	-4.6385	-4.3050
Pcocoa Pebbles	-5.0948	-4.5375	-4.9543	-5.4942	-5.0202
Pfruit Pebbles	-4.9605	-4.3430	-4.8640	-5.4885	-4.9140
Pgrape Nuts	-2.9363	-2.9337	-3.1192	-3.0590	-3.0120
Phoney Comb	-4.6623	-4.2232	-4.6402	-5.0609	-4.6466
Praisin Bran	-3.3051	-3.0566	-3.1255	-3.5552	-3.2606
Qcap N Crunch	-4.2815	-4.0460	-4.2214	-5.0515	-4.4001
Qoat	-4.4348	-4.0197	-3.4754	-4.8470	-4.1942
QToasted	-4.7299	-4.4149	-4.6436	-5.4524	-4.8102
N Frosted Wheat Bites	-4.2244	-3.9999	-4.1750	-4.3994	-4.1997
N Spoon Size	-4.0467	-3.6852	-3.9742	-4.1195	-3.9564
R Cookie Crisp	-7.1637	-6.3598	-6.0344	-7.8619	-6.8549
R Corn Chex	-5.2256	-4.6543	-5.4042	-5.8800	-5.2910
R Rice Chex	-5.2039	-4.6667	-5.4641	-5.8481	-5.2957

Table 3. RTEC Demand: RCM Parameter Estimates.

	Variable	Estimate	t-Statistic
Means (β 's)	Price	-27.7599***	13.4325
	Promotion	0.7422***	5.2475
	Calories	-2.8994***	10.2532
	Fiber	-0.1226***	3.1106
	Sugar	0.2137*	1.6691
	Kid Dummy	-0.2875**	2.5394
Standard Deviation			
	Price	-0.8678	0.3944
	Constant	0.2424	0.4348
	Calories	-1.6272***	9.2714
	Fiber	0.0343	0.5693
	Sugar	0.5864***	4.9153
	Kid Dummy	0.2938	0.6008
Interaction with Age			
	Price	2.8931	0.1309
	Constant	-6.7248*	1.5894
	Calories	0.9916	0.3066
	Fiber	0.5260	1.0648
	Sugar	0.3469	0.2187
	Kid Dummy	1.7769	1.3489
Interaction with Income			
	Price	-32.8321***	10.9825
	Constant	3.7235***	3.9783
	Calories	-2.4213***	5.1761
	Fiber	-0.3812***	4.1623
	Sugar	1.2334***	4.7853
	Kid Dummy	-1.1670***	4.9963

Table 4. RCM Own-Price Elasticity Estimates for RTECs in Boston Supermarkets

RTEC Brand	Stop & Shop	Shaw's	DeMoulas	Star Market	Simple Average
KApple Jacks	-5.0538	-4.7361	-4.5712	-5.2891	-4.9126
KComplete Bran	-4.0203	-3.9114	-3.9273	-4.2890	-4.0370
KCorn Flakes	-2.4439	-2.5697	-2.7390	-2.8427	-2.6488
KCorn Pops	-4.2322	-4.1493	-4.6857	-4.7486	-4.4540
Kcrispix	-4.6042	-4.4406	-5.0598	-5.2341	-4.8347
Kfroot Loops	-4.6290	-4.8549	-4.3061	-4.6756	-4.6164
Kfrosted Flakes	-3.4815	-3.6043	-3.4005	-3.7183	-3.5512
Kfrosted Mini Wheats	-3.4461	-3.3784	-3.1623	-3.6964	-3.4208
Kraisin Bran	-3.0110	-3.2719	-2.9969	-3.4274	-3.1768
Krice Krispies	-4.2751	-4.1947	-4.3234	-4.4774	-4.3176
Kspecial K	-4.8778	-4.9197	-5.3034	-5.2355	-5.0841
GMcheerios	-3.7578	-3.4781	-4.2353	-4.3027	-3.9435
GMCinammon Crunch	-4.6603	-4.3195	-4.8653	-4.8935	-4.6846
GMCoco Puffs	-4.5527	-4.4510	-4.8665	-4.7949	-4.6663
GMGolden Grahams	-5.0053	-4.6197	-5.0213	-5.5405	-5.0467
GMHoney Nut Cheerios	-4.1925	-4.0908	-4.0275	-4.2923	-4.1508
GMKix	-4.8463	-4.6135	-5.6565	-5.6408	-5.1893
GMLucky Charms	-4.6996	-4.3449	-4.9809	-4.8259	-4.7128
GMMulti Gain Cheerios	-5.2208	-5.1478	-4.9787	-5.7828	-5.2825
GMTotal	-4.9320	-4.4749	-5.1367	-5.4025	-4.9865
GMTotal Raisin Bran	-3.9104	-4.0669	-3.9033	-3.8982	-3.9447
GMWheaties	-3.4049	-3.3307	-3.6783	-4.0563	-3.6176
GMApple Cinnamon	-4.2531	-4.3708	-4.3033	-4.6634	-4.3976
Pbanana Nut Crunch	-4.0538	-3.5724	-4.2966	-4.3949	-4.0794
Pcocoa Pebbles	-4.8463	-4.3115	-4.7108	-5.2312	-4.7750
Pfruit Pebbles	-4.7125	-4.1204	-4.6203	-5.2237	-4.6692
Pgrape Nuts	-2.7828	-2.7841	-2.9619	-2.9013	-2.8575
Phoney Comb	-4.4319	-4.0115	-4.4126	-4.8175	-4.4184
Praisin Bran	-3.1093	-2.8918	-2.9604	-3.3596	-3.0803
Qcap N Crunch	-4.0652	-3.8474	-4.0157	-4.8111	-4.1848
Qoat	-4.2113	-3.8091	-3.2921	-4.5981	-3.9776
QToasted	-4.4876	-4.1944	-4.4128	-5.1835	-4.5696
N Frosted Wheat Bites	-4.0142	-3.8000	-3.9721	-4.1836	-3.9925
N Spoon Size	-3.8200	-3.5007	-3.7822	-3.9049	-3.7519
R Cookie Crisp	-6.8381	-6.0686	-5.7516	-7.5201	-6.5446
R Corn Chex	-4.9638	-4.4173	-5.1379	-5.5944	-5.0283
R Rice Chex	-4.9429	-4.4288	-5.1956	-5.5639	-5.0328

Table 5 A Sample of Implied Cross-Price Elasticities in Stop $ Shop

RTEC Brand	*K Corn flakes*	*GM Cheerios*	*R Corn Chex*	*K Raisin Bran*	*P raisin Bran*	*GM Total Raisn Bran*	*K Special K*	*P Grape Nuts*	*GM Total*	*K Froot Loops*	*Q CapN Crunch*	*GM Lucky Charms*
K Corn Flakes	-2.4439	0.0265	0.0061	0.0160	0.0097	0.0084	0.0274	0.0108	0.0129	0.0132	0.0089	0.0120
K Rice Krispies	0.0189	0.0588	0.0088	0.0220	0.0154	0.0107	0.0378	0.0155	0.0233	0.0191	0.0190	0.0196
K Crispix	0.0211	0.0490	0.0090	0.0201	0.0138	0.0101	0.0379	0.0140	0.0223	0.0188	0.0165	0.0176
K Raisin Bran	0.0174	0.0376	0.0067	-3.0110	0.0443	0.0285	0.0297	0.0139	0.0184	0.0142	0.0112	0.0231
K Frosted Mini Wheats	0.0190	0.0507	0.0089	0.0470	0.0331	0.0205	0.0362	0.0156	0.0228	0.0180	0.0157	0.0220
K Special K	0.0201	0.0505	0.0089	0.0206	0.0142	0.0105	-4.8778	0.0140	0.0230	0.0187	0.0168	0.0181
K Frosted Flakes	0.0205	0.0354	0.0062	0.0286	0.0189	0.0147	0.0291	0.0133	0.0165	0.0142	0.0108	0.0198
K Froot Loops	0.0206	0.0510	0.0092	0.0201	0.0139	0.0100	0.0382	0.0140	0.0229	-4.6290	0.0188	0.0184
GM Cheerios	0.0149	-3.7578	0.0091	0.0198	0.0145	0.0093	0.0384	0.0158	0.0269	0.0194	0.0230	0.0204
GM Wheaties	0.0225	0.0400	0.0076	0.0214	0.0143	0.0107	0.0334	0.0132	0.0186	0.0161	0.0127	0.0159
GM Total Raisin Bran	0.0185	0.0351	0.0065	0.0549	0.0378	-3.9104	0.0303	0.0133	0.0183	0.0141	0.0102	0.0238
GM Total	0.0168	0.0599	0.0093	0.0221	0.0159	0.0111	0.0398	0.0145	-4.9320	0.0194	0.0190	0.0207
GM Multi Grain Cheerios	0.0115	0.0835	0.0093	0.0215	0.0166	0.0106	0.0404	0.0157	0.0319	0.0204	0.0253	0.0248
GM Honey Nut Cheerios	0.0166	0.0493	0.0073	0.0284	0.0200	0.0143	0.0329	0.0145	0.0214	0.0156	0.0141	0.0226
GM Lucky Charms	0.0161	0.0490	0.0075	0.0292	0.0207	0.0150	0.0337	0.0144	0.0222	0.0170	0.0153	-4.6996
GM Kix	0.0145	0.0779	0.0101	0.0219	0.0161	0.0105	0.0414	0.0161	0.0300	0.0215	0.0249	0.0234
Post Garpe Nuts	0.0188	0.0472	0.0073	0.0222	0.0152	0.0107	0.0320	-2.7828	0.0192	0.0157	0.0149	0.0178
Post Raisin Bran	0.0157	0.0393	0.0066	0.0652	-3.1093	0.0293	0.0295	0.0140	0.0191	0.0140	0.0114	0.0244
N Frosted Wheat Bites	0.0164	0.0587	0.0098	0.0458	0.0337	0.0196	0.0384	0.0157	0.0263	0.0197	0.0186	0.0248
Q Oat	0.0195	0.0465	0.0087	0.0524	0.0365	0.0231	0.0358	0.0151	0.0222	0.0176	0.0144	0.0232
Q CapN Crunch	0.0165	0.0683	0.0097	0.0186	0.0134	0.0089	0.0399	0.0159	0.0261	0.0207	-4.0652	0.0201
Ralston Corn Chex	0.0217	0.0539	-4.9638	0.0212	0.0148	0.0105	0.0408	0.0147	0.0244	0.0205	0.0187	0.0185

Table 6 A Sample of Cross-Price Elasticities across Supermarkets

	Stop & Shop			Shaw's			DeMoulas			Star market		
	K Corn Flakes	GM Cheerios	P Grape Nuts	K Corn Flakes	GM Cheerios	P Grape Nuts	K Corn Flakes	GM Cheerios	P Grape Nuts	K Corn Flakes	GM Cheerios	P Grape Nuts
Stop & Shop												
K Corn Flakes	-2.4439	0.026536	0.010794	0.010114	0.015487	0.005306	0.017078	0.015112	0.012506	0.018851	0.016511	0.008716
GM Cheerios	0.014869	-3.7578	0.015832	0.00569	0.036441	0.007907	0.010103	0.049313	0.006965	0.011731	0.058155	0.013437
P Grape Nuts	0.018806	0.047169	-2.7828	0.007083	0.025677	0.007253	0.012228	0.029967	0.006234	0.013867	0.033608	0.01212
Shaw's												
K Corn Flakes	0.027939	0.027087	0.010986	-2.5697	0.015866	0.005425	0.01705	0.015547	0.004563	0.018695	0.016978	0.008886
GM Cheerios	0.015855	0.067083	0.015617	0.006137	-3.4781	0.00778	0.010766	0.045034	0.006847	0.012395	0.052632	0.013161
P Grape Nuts	0.018843	0.046955	0.014583	0.007104	0.025599	-2.7841	0.012261	0.029781	0.006232	0.013854	0.033547	0.012112
DeMoulas												
K Corn Flakes	0.027345	0.028677	0.011126	0.009914	0.016601	0.005512	-2.739	0.016658	0.004657	0.018522	0.018306	0.009104
GM Cheerios	0.013713	0.078844	0.015989	0.005292	0.038986	0.007962	0.009382	-4.2353	0.007108	0.010933	0.066833	0.013811
P Grape Nuts	0.018347	0.04903	0.014728	0.006922	0.026612	0.007332	0.011938	0.031819	-2.9619	0.013716	0.035765	0.012281
Star Market												
K Corn Flakes	0.027095	0.029538	0.011315	0.009825	0.017031	0.005572	0.016602	0.017252	0.004735	-2.8427	0.019019	0.009225
GM Cheerios	0.013374	0.082102	0.015976	0.005169	0.040347	0.00796	0.009226	0.0587	0.007116	0.010828	-4.3027	0.013784
P Grape Nuts	0.018573	0.048593	0.014656	0.006982	0.026148	0.007277	0.0121	0.03108	0.006229	0.013623	0.034967	-2.9013

Figure 1: Frequency Distribution of the Own-Price Elasticities for Logit Model

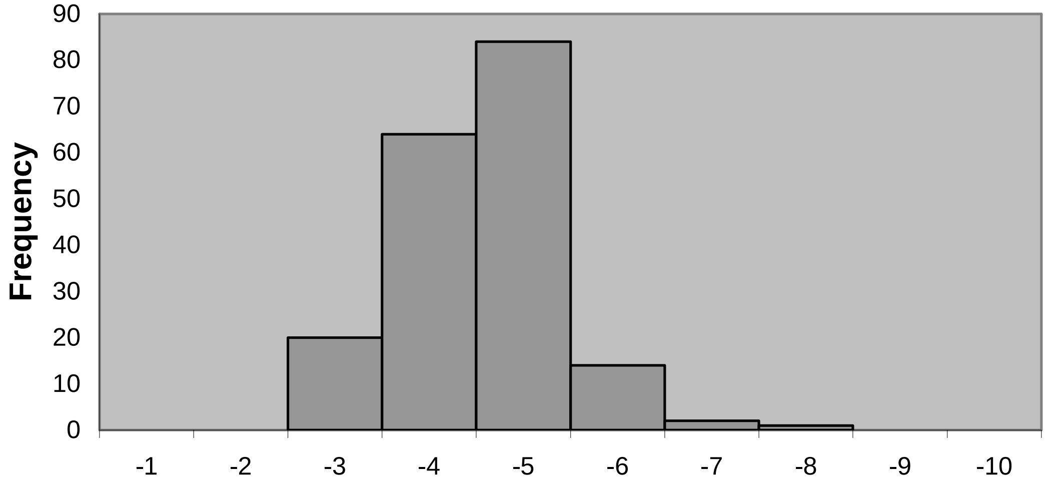
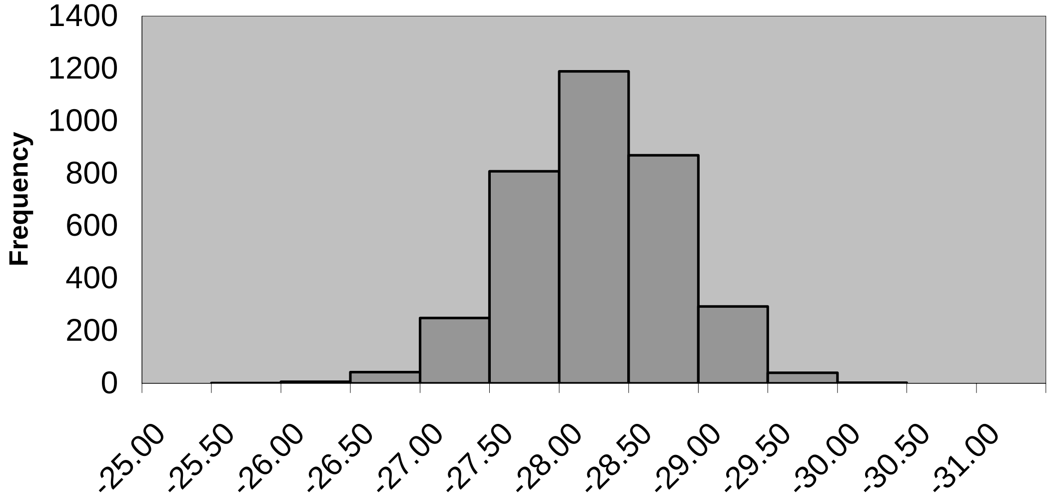


Figure 2 Frequency Distribution of Price Coefficient



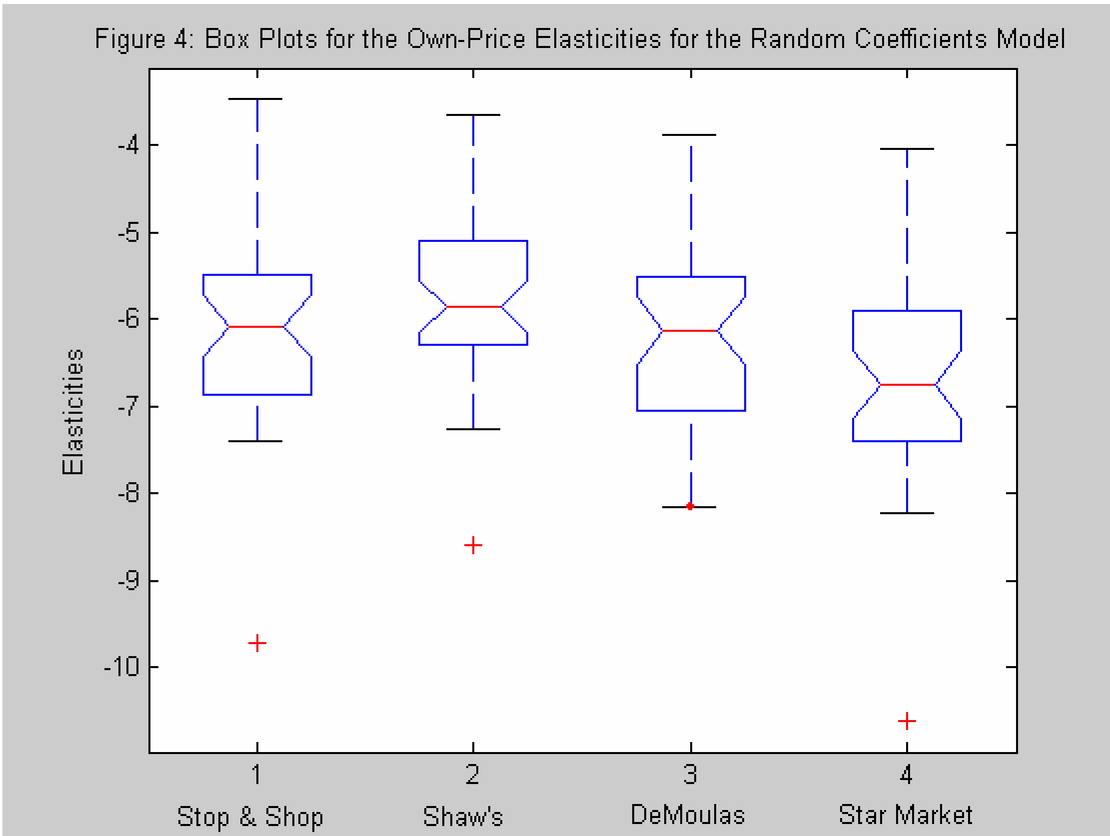


Figure 4: Box Plots for the Own-Price Elasticities for the Random Coefficients Model
Elasticities
-4
-5
-6
-7
-8
-9
-10
1
2
3
4
Stop & Shop
Shaw's
DeMoulas
Star Market

Figure 5 : Frequency Distribution of the Own-Price Elasticities for RCM

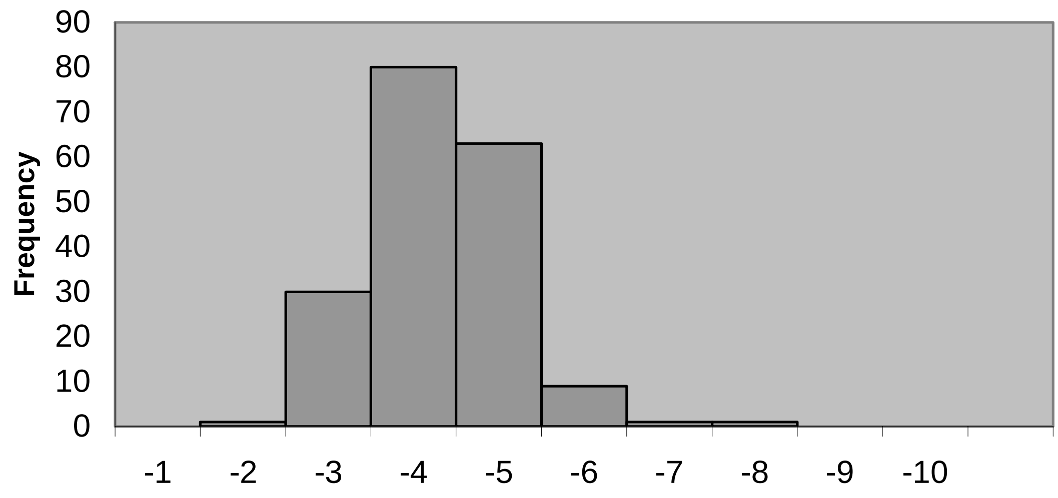
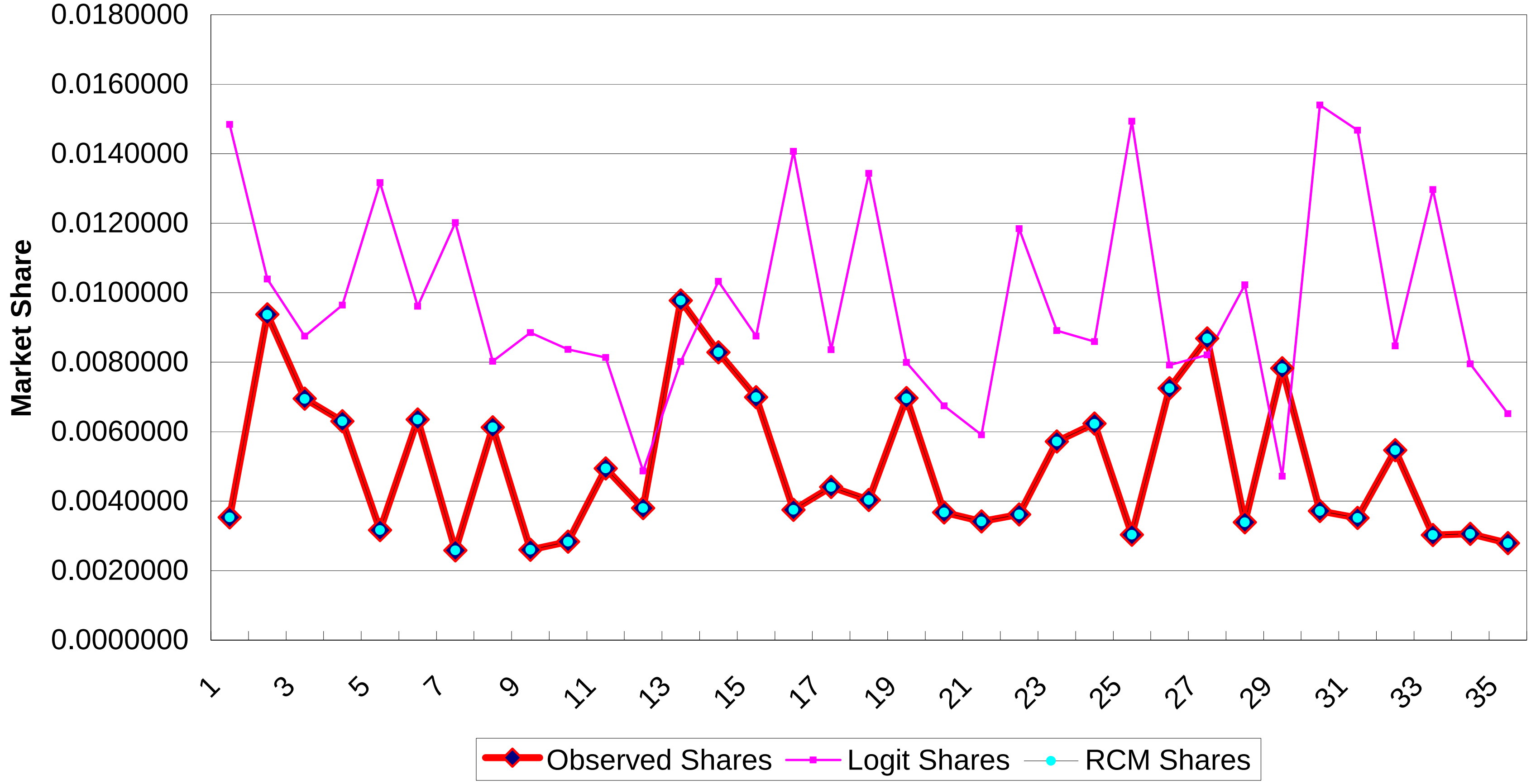


Figure 6: Observed and Predicted Market Shares for Logit and RCM for Kellogg's Corn Flakes in Stop & Shop Chain



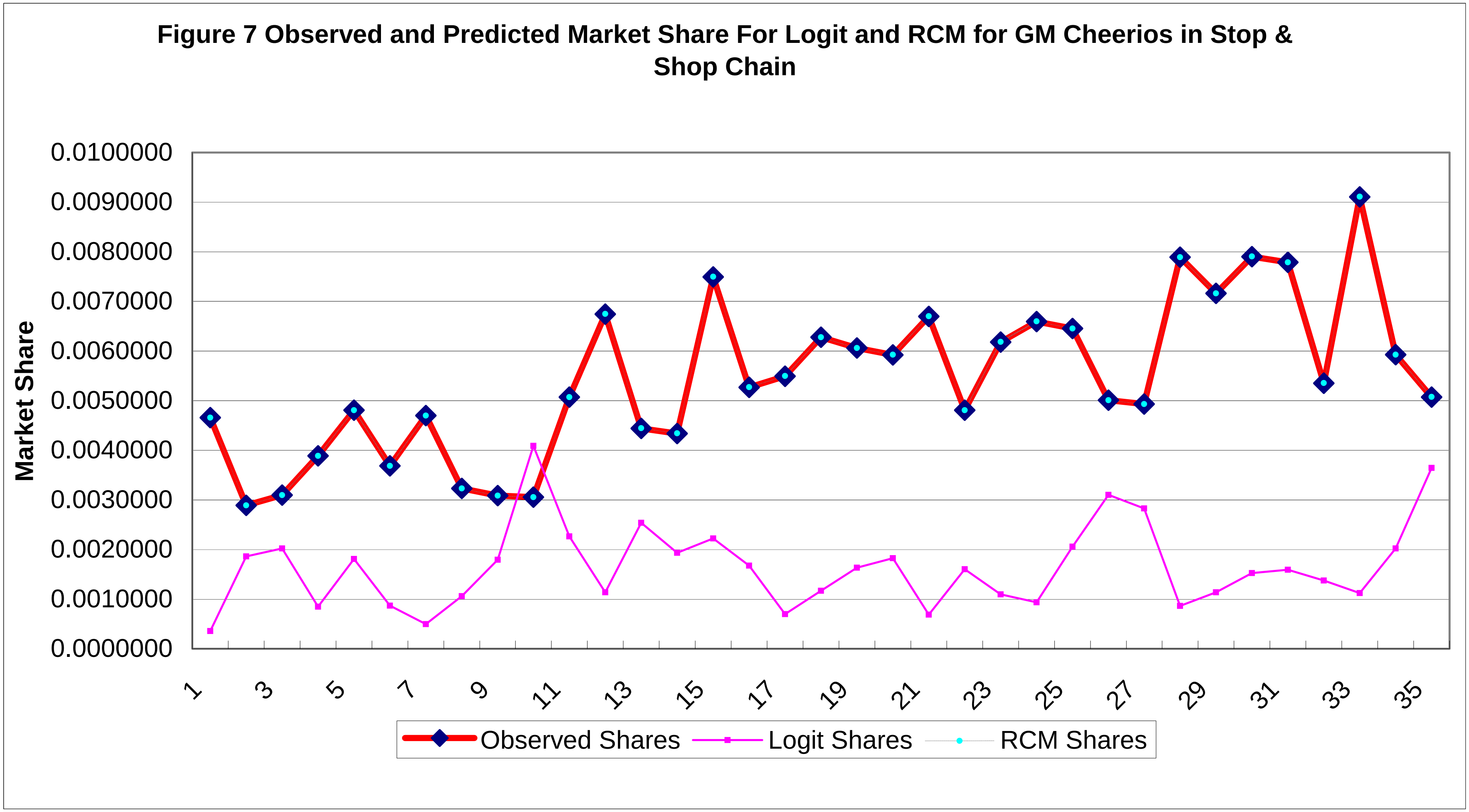


Figure 7 Observed and Predicted Market Share For Logit and RCM for GM Cheerios in Stop & Shop Chain
Market Share
0.0100000
0.0090000
0.0080000
0.0070000
0.0060000
0.0050000
0.0040000
0.0030000
0.0020000
0.0010000
0.0000000
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
Observed Shares
Logit Shares
RCM Shares