Zwick Center for Food and Resource Policy

Working Papers Series

No. 15

An Empirical Analysis of Equilibrium Pricing and Advertising in the Ready-To-Eat Cereal Market

Michael Cohen* and Adam Rabinowitz**

November 8, 2012

Department of Agricultural and Resource Economics
College of Agriculture and Natural Resources
1376 Storrs Road, Unit 4021
Storrs, CT 06269-4021
Phone: (860) 486-1927
Contact: ZwickCenter@uconn.edu

University of Connecticut

* Stern School of Business, New York University
** Zwick Center for Food and Resource Policy, University of Connecticut

An Empirical Analysis of Equilibrium Pricing and Advertising in the Ready-To-Eat Cereal Market †

Michael A. Cohen
New York University
Stern School of Business
40 West Forth Street, Suite 914
New York, NY 10012
(212) 992-6873
mcohen@stern.nyu.edu

Adam N. Rabinowitz
University of Connecticut
Zwick Center
1376 Storrs Road, Unit 4021
Storrs, CT 06269
(860) 486-2826
adam.rabinowitz@uconn.edu

†We thank Panle Jia Barwick, Josh Berning, Bryan Bollinger, Richard Briesch, Ron Cotterill, Rui Huang, Jozsef Molnar, Lars Nesheim, and Aviv Nevo for helpful discussions. The paper has benefited from comments by seminar participants at: Massachusetts Institute of Technology, The 2010 London Conference on the Econometric Analysis of Scanner Data, and University of Connecticut. We thank The Food Marketing Policy Center at University of Connecticut for financial support and Chris Jeffords and Elena Castellari for research assistance.

An Empirical Analysis of Equilibrium Pricing and Advertising in the Ready-To-Eat Cereal Market

Abstract

We introduce a model of dynamic price and advertising competition and use the model to investigate a popular segment of the Ready-To-Eat Cereal Market. It is well understood that advertising is a key non-price strategic demand determinant in differentiated product markets. Two popular and compatible models for the role of advertising have emerged. One model specifies advertising as a determinant of a consumer's level of product awareness (informative) and the other model specifies advertising as a complimentary product attribute (persuasive). Without specific information on a consumer's level of awareness the two roles cannot be simultaneously identified within a demand model. One must then decide how advertising enters the demand model. To explore the implications of this specification decision on the assessment of price and advertising competition we: Estimate heterogeneous consumer logit demand models for each advertising specification; define a model of price and advertising competition for multi-product firms; and compute the distribution of equilibrium price and advertising strategies to asses the fit against the observed price and advertising strategies. To illustrate how the market model may be applied we investigate a segment of the ready-to-eat cereal market and evaluate a hypothetical brand acquisition. The analysis provides three key empirical findings: Upward pricing pressure, due to the acquisition, is larger when advertising is incorporated into the demand model as a result of estimation bias; the informative advertising model yields downward pressure on price while the persuasive advertising model yields upward pressure on price; and the information theoretic model fits the data better.

Keywords: Pricing, Informative and Persuasive Advertising, Demand Estimation, Competition, Advertising Dynamics

1 Introduction

Economic theory recognizes and empirical studies document the crucial role advertising plays in determining prices and firm profitability. The classic article by Dorfman and Steiner (1954) demonstrates that in a static duopoly, a firm's brand advertising, as well as its pricing, influences its performance. Advertising can shift and rotate the demand curve, reducing a brands' own price elasticity and producing higher price-cost margins in equilibrium. These higher margins can also increase the optimal amount of advertising. Friedman (1983) generalizes the Dorfman and Steiner model to a dynamic setting that recognizes how current period advertising contributes to a stock that depreciates over time. In the Friedman model, a firm's brand advertising influences demand in future periods as well as in the current period. Structural empirical industrial organization models of differentiated product markets have focused on pricing or media advertising in isolation but not together. This paper introduces an empirical modeling framework inclusive of advertising stock dynamics to analyze pricing and advertising equilibria in differentiated product markets.

The stock effects of advertising on consumer demand have been recognized since Nerlove and Arrow (1962) who model the cumulative "goodwill" effect of advertising for a single firm. Until recently the empirical literature investigating the impact of advertising on demand has limited advertising's role to preference-enhancer. Dubé, Hitsch, and Manchanda (2005) incorporate the "goodwill" effect and offer a dynamic advertising model that identifies when pulsing, which is heavy ad placement in some periods and none at other times, or continuous advertising is optimal. Other empirical work by Shum (2004) measures how brand advertising expenditure, as a proxy for exposure, induces consumers to switch from other brands. Rather than television and other measured media advertisements, Slade (1995) examines retail advertising in newspapers produced within the context of a manufacturer's trade promotion. In "retailer push" promotion programs, manufacturers offer the retailer lower priced products in return for retail price reductions, aisle end display, or shelf signage, as well as a "price reduction" advertisement in their weekly promotion circular (Gerstner & Hess, 1991). An earlier structural econometric study by Gasmi, Laffont, and Vuong (1992) presents a static duopoly model to test for collusive advertising and pricing between Pepsi and Coca-Cola in the carbonated soft drink industry. Recently, Goeree (2008) investigates the personal computer market and develops a model that allows for consumer information heterogeneity within a Berry, Levinsohn, and Pakes (1995) (BLP) random coefficients logit demand model by incorporating consumer awareness as a function of current period advertising. In Goeree's model, product demand depends on the probability that consumers are "aware" of that product as well as competing products. Related research by Draganska and Klapper (2012) offers a similar conceptual demand model and exploits consumer survey data on brand recall to better isolate the awareness effects of advertising from the preference augmenting impact of advertising.

In the absence of data that reasonably allows one to separately identify awareness from preference effects, and without evidence that the two are independent, the model specification dictates how advertising influences demand and consequently industry price and advertising dynamics. For example Doraszelski and Markovich (2007) provide a theoretical investigation of advertising competition that analyzes the impact "goodwill" versus "awareness" advertising has on firm entry or exit from a market and find different industry dynamics for each role of advertising. Because we do not observe information about consumer awareness and because it is not part of market sales data, our empirical investigation explores the impact this modeling assumption has on implied firm pricing and advertising equilibria.

We consider two alternative demand models: One incorporates advertising as an instrument to augment the level of product awareness (Grossman & Shapiro, 1984), the other incorporates advertising as a preference shifter (Becker & Murphy, 1993). Our limited-awareness demand model extends Goeree's conceptual model by specifying awareness as a function of advertising stock to capture the dynamics of product salience in the marketplace. This view is consistent with a continuous level of awareness rather than a discrete one, and by allowing product set awareness to be a function of awareness levels for competing products, we are able to model a consumer's cognitive limitations in maintaining the salience of every product within a product category. If we believe that once a product has a consumer's attention and that they evaluate it independent of advertising, then this is an appealing model setup. However, if we believe all products are equally salient to the consumer and advertising is a complimentary attribute of the product, then the full awareness BLP model inclusive of an advertising goodwill component is appropriate. It is probable that reality lies somewhere between the two models yet such a model is not empirically identifiable without knowledge of consumer awareness levels.

To estimate both demand models, we use a mathematical programming equilibrium condition (MPEC) estimation approach (Dubé, Fox, & Su, 2009). Aside from the MPEC estimation method, our estimation of the limited awareness demand model differs from Goeree's because we do not simulate latent choice sets; rather we directly integrate over all possible latent choice sets, thereby eliminating error caused by choice set simulation.[1] In addition to the Berry et al. (1995) macromoments (market level), we use supplemental micromoments in the spirit of Petrin (2002) to aid identification of the heterogeneity parameters. The micromoments we specify are consistent with the approach outlined by Chintagunta and Dube (2005), who estimate heterogeneity parameters using observed consumer purchases conditional on mean market preference. However we incorporate them into the MPEC estimation framework, by specifying the score condition for the density of consumer purchases, with respect to heterogeneity parameters, as additional equilibrium constraints.

Our empirical analysis investigates the taste enhanced wholesome (TEW) segment of the U.S. ready to eat (RTE) breakfast cereal market. Our market data are from daily A.C. Nielsen Homescan purchases for 8 brands in 8 television markets and are aggregated at the biweekly market level. This aggregation allows us to match our Homescan data with market level advertising data. We also reap the benefits of the Homescan data by using observed consumer product choices to identify the distribution of consumer tastes. Using data at both levels has the additional benefit that we can instrument for endogenous regressors using IV estimation methods in the spirit of BLP. The Nielsen Media Research data provide brand level measures of advertising penetration (Gross Rating Points), advertising frequency, and expenditures on advertising. These data permit measurement of the price per unit of advertising penetration and thus estimation of the impact of both advertising demand and cost on optimal price and advertising strategies. To model oligopolistic competition, we use demand estimates in the firm's problem and solve for the optimal price and advertising levels. In our setting the advertising game is dynamic; thus we solve for the set of Markov perfect equilibrium pricing and advertising strategies. This approach parallels Dubé et al. (2005) who solve for the Markov perfect advertising strategy equilibria and Benkard (2004) who studies pricing dynamics. In contrast, our analysis solves pricing *and* advertising strategies for a set of multi-product firms. We apply the Pakes and McGuire (1994) general approach for solving games with strategic interactions.

To illustrate how the model may be applied, we generalize a unilateral effects merger analysis to incorporate advertising impacts and simulate a counterfactual merger. The evaluation of horizontal merger

[1] We are able to avoid choice set simulation for two reasons. One, our empirical analysis considers a manageable number of product choices. Two, we exploit the sparsity of choice sets with relatively few options to improve computational efficiency.

analysis in differentiated product industries has relied upon estimation of brand level demand systems and focused upon the price impacts that arise from the internalization of own price elasticities, while ignoring the internalization due to the impact of advertising (Hausman, Leonard, & Zona, 1994; Nevo, 2000). Recently, Tenn, Froeb, and Tschantz (2010) have expanded the theory of unilateral analysis to include promotions as well as price. To the best of our knowledge the literature has not incorporated advertising dynamics to analyze price impacts due to a merger in a differentiated product industry. We simulate and compare pre- and post- merger equilibrium implied by our model to that implied by the extant price only Nevo (2000) approach.

We demonstrate that a brand's advertising has two impacts on the merger analysis. First is the impact on demand for that and other differentiated products and hence an impact on the price after the merger. Generalizing the Bertrand oligopoly model to include demand side advertising effects destroys the unambiguous prediction wherein the internalization of positive cross price elasticities produces elevated prices for all products postmerger (Deneckere & Davidson, 1985; Hausman et al., 1994; Martin, 2002). Second advertising is viewed as a strategic, i.e. variable input that influences marginal costs. Thus one might observe "efficiency" gains that offset any estimated demand side price increases so that the prices are lower post-merger.

After an introduction to the taste enhanced wholesome segment of the ready to eat cereal market and our data, we introduce the demand model, the oligopolistic model of pricing and advertising, and the Markov perfect equilibrium concept we apply. Next, we discuss estimation identification details. Then we outline the demand estimation approach and present our demand estimation results. Before concluding, we use the demand estimates and our model of firm conduct to asses the impact of a merger on price and advertising levels. We compare the results implied by our model of competition to a price only model and investigate the implications of demand model specification.

2 Data

The market data we use is from the 2006 to 2008 A.C. Nielsen Homescan database for ready to eat breakfast cereals. It records daily purchasing from about 2,700 households in a variety of store types across 8 designated marketing areas (DMA) (Atlanta, Boston, Chicago, Houston, Los Angeles, New York, Philadelphia, and Seattle-Tacoma) over 152 consecutive weeks. In addition, we use data from Nielsen Media Research that also covers the 2006 to 2008 period and is provided on a weekly basis across multiple television outlets (cable, network, syndicated, and spot). To combine these data we aggregate household transactions to market level prices and quantities over 76 biweekly periods, creating a balanced panel to be used for analysis. This market level data is supplemented with the household level choice data to identify market idiosyncracies.

2.1 The Taste Enhanced Wholesome Cereal Segment

The RTE cereal market contains hundreds of products that are often classified in four major categories - All Family Basic, Simple Health Nutrition, Kids, and Taste Enhanced Wholesome. The taste enhanced wholesome segment makes up over 20% of the U.S. RTE cereal market[2], and due to the nature of competition, the TEW segment is identified as a segment appropriate for isolation in research. For example, in a letter to Kraft General Foods proposing to evaluate the impact of couponing and trade deals on brand volume, Nielsen Marketing Research executives state:

> The scope of this analysis will be Taste Enhanced Wholesome (TEW) cereals. ... We are choosing to evaluate this part of the entire RTE cereal category for the following reasons:
>
> - The TEW segment consists of brands that possess strong interactions with each other.
> - Including other category segments into the evaluation may "dilute" the switching patterns observed in the data. For example, a household may always purchase a Traditional Kid cereal and a TEW cereal for different members of their household. However, the data does not reflect the intended user of each brand, and in effect may falsely indicate a switch form [sic] a Traditional Kid brand to TEW.
>
> [State of New York v. Kraft General Foods et al. 93 Civ 0811 (KMW) @ KGF048666, PX853]

The reasoning precisely supports our focus on TEW brands in this paper.

Table 1 provides key descriptive statistics for the 8 brand, 8 DMA panel data set used in this study. The data set contains 2 Post, 1 Kashi, 4 Kellogg's, and 1 Private Label brand.[3] Demand shares are computed based on consumption weight of cereal. Post Honey Bunches of Oats has the largest market share, 4.38%. The next highest market share is Kellogg's Raisin Bran with 2.53%. Other brands from major RTE firms are Post Raisin Bran (1.44%), Kellogg's Special K Red Berries (1.42%), Kellogg's Raisin

[2]Author's calculations from IRI Marketing Fact Book, 2007

[3]Fixed product characteristics are provided in the appendix.

Bran Crunch (1.06%), and Kellogg's Smart Start (0.94%). Kashi GoLean Crunch! brand, a leading fringe firm brand in this industry, has a very respectable 1.21% market share and Private Label Raisin Bran has 1.24%.[4] These 8 brands account for 14.21%, on average, of the potential RTE cereal market. Average price per pound across the branded cereals ranges from \$3.17 for Kellogg's Special K Red Berries to \$1.71 for Post Raisin Bran. Private label Raisin Bran is 5 cents higher at \$1.76, thus Post Raisin Bran has virtually no brand equity.

2.2 Advertising in the Taste Enhanced Wholesome Cereal Segment

Gross Rating Points (GRPs) are a measure of the advertising penetration as composed of its reach (the size of the audience viewing the ad) and frequency (the number of times the audience views the ad). GRPs provide a measurement of the total exposure a typical household has to a particular advertisement in a given market period. The extent of advertising, as measured by biweekly GRPs in Table 1, varies greatly across the 8 brands. The Private Label Raisin Bran and Post Raisin Bran have no advertising. Post advertises its other TEW product, Honey Bunches of Oats, which has the highest advertising penetration in this segment (GRP=354.54). The four Kellogg's brands have advertising ranging from 0.01 GRP for Raisin Bran to a GRP of 235.33 for Raisin Bran Crunch. Kashi GoLean Crunch! has an advertising GRP of 91.00.

The advertising price per GRP in Table 1 ranges from \$259.38 for Kellogg's Raisin Bran whose advertises very little (GRP=0.01) to a low value of \$93.71 for Kashi GoLean Crunch! whose GRP are 91.00. The advertising price across the four Kellogg's TEW brands ranges from \$121.62 to \$259.38, averaging \$180.89, almost double that of Post Honey Bunches of Oats and Kashi GoLean Crunch!.

With respect to advertising conduct, past research formalizes pulsing of advertising as Effective Frequency Planning (EFP) (Naples, 1979), which justifies pulsing as a response to the existence of a critical level of goodwill stock and diminishing marginal returns to advertising. Ultimately pulsing is recommended as a means to generate a sufficient level of goodwill stock on a limited advertising budget. Dubé et al. (2005) note that pulsing is justified as technically optimal under general conditions, and when advertising stock slowly decays over time. Table 2 identifies brands that pulse advertising. It gives the

[4] The RTE cereal industry routinely recognizes adult cereals with more than 1% of actual sales as brands above minimum efficient scale. Brands with lower shares risk discontinuation over time.

percent advertising frequency which is the percent of biweek periods where one or more advertisements for a particular brand appeared. Kellogg's Smart Start, for example, advertised in only 10.5% of the possible biweek periods. Therefore, it is pulsing its advertisements into the market in a very sporadic fashion. Kashi GoLean Crunch! brand is also pulsing its ads, but ads are appearing in 38.2% of the bi week periods. Post Honey Bunches of Oats in comparison does very little pulsing with ads appearing in 93.3% of the possible bi week periods. No TEW cereal advertises in every bi week period in all 8 markets.

2.3 Reporting Errors in Homescan Data

Given that our research uses the Nielsen Homescan consumer panel data to determine market prices, we have considered and addressed issues of the accuracy and reporting problems with self-reported data. Einav, Leibtag, and Nevo (2010) perform a validation study on 2004 Homescan data in two unidentified metropolitan markets. From this study the authors document recording discrepancies and formulate a method using their validation sample to control for recording errors in Homescan data. In correcting for measurement errors one needs to assume that the distribution of errors in the validation sample are the same as the primary data. To evaluate if this is a reasonable assumption, we consider the variables observed in our data.

One variable of particular importance that we have, and is not available to Einav et al. (2010), is an indicator of the source of price for each observation. From the products we use in this study over 50% are store reported prices, and thus not subject to self-reporting errors. Store reported prices, however, may be subject to imputation errors in the way Nielsen determines prices. The distribution of errors between these two different reporting problems may be different, yet the validation sample is not capable of making this distinction. This is one reason that we do not believe the validation approach is plausible for our particular study.

An additional difference between Nielsen store imputed prices and actual purchase price for a household comes from the timing of prices used in imputation as opposed to actual date of purchase. The process of imputing a store level price for a given week requires aggregating across a 7 day period. While it is unknown what weekly period Nielsen uses to impute prices, it is known that different retail chains follow different pricing weeks. It is highly likely that imputed prices of a given week may incorporate two different prices for the same item when in fact the household actually faced only one of those prices. Our data aggregation method aggregates daily purchases to biweekly weighted average purchases, resulting in an average price for the period. We believe this reduces errors that may exist as a result of purchase timing and imputation process, and thus do not correct our data for differences in reported prices due to purchase and reporting variations.

3 Models

This section introduces our models of consumer and firm behavior and the market equilibrium concept we employ. It begins by specifying the limited awareness demand model and the full awareness demand model with advertising goodwill specified as a complementary consumption attribute. Next we describe how advertising produces goodwill stock which is specified to influence demand. Then it details firm profit and behavioral assumptions. Finally it establishes the equilibrium concept we apply.

3.1 Demand Models

We begin by specifying a discrete choice demand model with random utility. Consumers, denoted by i, enjoy utility from consuming product j in market t according to:

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt}. \tag{1}$$

Here $\delta_{jt} = x_j'\beta - \alpha^p p_{jt} + \xi_{jt}$ is the market mean utility a consumer derives from product j, β represents the mean marginal utility of product characteristics x_j; and α^p measures the mean marginal utility of income. ξ_{jt} is an *i.i.d.* product specific demand shock partially observed by firms and unobserved by the econometrician. The random shock $\mu_{ijt} + \epsilon_{ijt}$, captures consumer taste heterogeneity. ϵ_{ijt} is an independently and identically distributed (i.i.d) type I extreme value model error. The μ_{ijt} term specifies consumer specific deviations from the mean utility as a function of observed consumer attributes, D_{it}; unobserved to the econometrician consumer taste shocks, υ_i; and product characteristics and prices. That is,

$$\mu_{ijt} = [x_j' \; p_{jt}](\Omega D_{it} + \upsilon_i \Sigma^{1/2}), \quad \upsilon_i \sim N(0, I_k). \tag{2}$$

where Ω translates how tastes vary with D_{ij}. The model assumes that tastes are normally distributed with variance parameter, Σ. Consumers can choose an outside option that may include no purchase or purchase of alternatives outside of the set of brands considered for analysis. The outside option is the zero utility option, specified as:

$$u_{i0t} = \xi_{0t} + \epsilon_{i0t}. \tag{3}$$

We fix ξ_{0t} at zero to achieve identification of the inside option utility levels relative to the utility of the outside option.

Our demand model, following Goeree (2008), relaxes the classical assumption that consumers are fully informed about market offerings at the point of sale. In particular, consumer demand share is a function random latent choice sets where the probability that consumer i purchases product j, is a function of the probability that the consumer is aware of product j; the probability they are aware of other products competing with j; and the probability they would purchase j conditional on their choice set, $\mathcal{S}$. Recognizing that consumers can always choose to purchase an outside option or not to purchase at all, consumer i's demand share for product j in market t is:

$$s_{ijt} = \sum_{\mathcal{S} \in C_j} \prod_{l \in \mathcal{S}} \phi_{ilt} \prod_{r \notin \mathcal{S}} (1 - \phi_{irt}) \frac{exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k \in \mathcal{S}} exp(\delta_{kt} + \mu_{ikt})}, \tag{4}$$

where $\mathcal{C}_j$ is the set of all choice sets that include product j and ϕ_{ijt} is the probability that the consumer is informed and aware of products j. This model includes the standard full awareness logit share expression as a limiting case where the consumer is aware of all the choices in $\mathcal{C}_j$. The specific form of the awareness probability, after integrating out the type I extreme value consumer information shock is:

$$\phi_{ijt} = \frac{exp(\gamma_{jt} + \tau_{ijt})}{1 + exp(\gamma_{jt} + \tau_{ijt})}, \tag{5}$$

where γ_{jt} captures mean awareness utility, and τ_{ijt} captures the consumer specific information heterogeneity. In the full awareness demand model setting advertising may be specified as a complementary consumption attribute by adding $\gamma_{jt} + \tau_{ijt}$ to u_{ijt} in equation (1). The specific functional form for the utility from advertising component, γ_{jt}, is:

$$\gamma_{jt} = \Gamma(g^a_{jt}), \tag{6}$$

where,

$$\Gamma(g^a_{jt}) = \begin{cases} \alpha^a log(1+g^a), & \text{if } g^a \geq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

This particular functional form is required due to the S-shape of the logit sales response function, whereby the share function exhibits increasing returns to indirect utility for $\delta_{ij} + \mu_{ijt} < 0$. Under the specification defined in (7), we test whether α^a is strictly concave in goodwill stock g^a.[5] The functional specification for the augmented goodwill stock g^a is discussed in detail below. The information utility component τ_{ijt} is defined as:

$$\tau_{ijt} = \frac{\Gamma(g^a_{jt})}{\alpha^a}(\rho D_{it} + \upsilon_i \sigma_a), \ \upsilon_i \sim N(0,1), \tag{8}$$

where υ_i is unobserved consumer heterogeneity; D_{it} is again observed characteristics translated into preferences by ρ; and σ_a captures the scale of the distribution characterizing information heterogeneity. Combining each component described above and integrating over the market of consumers, market demand share is:

$$s_{jt} = \int_{B_{jt}} s_{ijt} dF_D(D) dF_\upsilon(\upsilon), \tag{9}$$

where $F(\cdot)$ denotes the respective distribution function and B_{jt} is the consumer specific set parameters and variables that induce the purchase of product j in market t.

3.2 Advertising

Advertising stock captures the dynamic carry-over effects of advertising's impact on demand. In other words, today's advertising has a lasting effect that carries over to the next period and beyond. This carryover effect is modeled as a distributed lag of advertising,

$$g_{jt} = \sum_{k=1}^{\infty} \lambda^k \Psi(A_{j,t-k}), \tag{10}$$

[5] We also test the fit of this specification versus a linear specification and a quadratic specification at the estimation phase.

where, $\Psi(\cdot)$ is a nonlinear goodwill production function. We assume that $\Psi(0) = 0$ and is a non decreasing function of advertising proliferation, A_{jt}. Firms produce goodwill by adding to the existing stock to generate an augmented goodwill stock,

$$g^a_{jt} = g_{jt} + \Psi(A_{jt}). \tag{11}$$

Augmented goodwill enters the consumer utility function directly as defined above in equation (7). Augmented goodwill stochastically depreciates overtime according to the following law of motion:

$$g_{j,t+1} = \lambda g^a_{jt} + \nu_{j,t+1} = \lambda(g_{jt} + \Psi(A_{jt})) + \nu_{j,t+1}. \tag{12}$$

$\lambda \in (0, 1)$ is a geometric decay factor and ν_{jt} is a mean zero shock included to capture the idiosyncratic aspects of advertising that are not captured by the data and not observed by firms. For example, ν may capture the impact of aspects that are particular to a television ad campaign's effectiveness or temporal variation in television audience composition, whereas A_{jt} only measures the reach and frequency of an ad for a particular product in a market as measured by GRPs. An expansion of equation (12) yields:

$$g_{jt} = \sum_{k=1}^{\infty} \lambda^k \Psi(A_{j,t-k}) + \omega_{jt}, \tag{13}$$

where $\omega_{jt} = \sum_{k=0}^{\infty} \lambda^k \nu_{j,t-k}$. The marginal return to advertising differs across time due to the carry-over effect. This fact implies that pulsing advertising policies may be adopted by firms, a fact that is consistent with many industries and is observed in our data for many brands in the RTE cereal industry. We apply the goodwill production function, $\Psi(A) = log(1 + A)$ if $A > 0$ and 0 otherwise, which is suggested by Dubé et al. (2005).

3.3 Profit

The demand for product j is a function of market size M and expressed $Q_{jt} = Ms_{jt}$, moreover it is determined by product characteristics, prices, advertising, and goodwill levels, as well as the vector of demand shocks:

$$Q_{jt} = Q_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t, \xi_t), \tag{14}$$

where, $\mathbf{g}_t = (g_{1t}, \ldots, g_{Jt})$ is a state vector that contains existing advertising stock levels and observed product attributes. The remaining components are advertising, $\mathbf{A}_t$, prices, $\mathbf{P}_t$, and demand shocks, ξ_t. The vector of demand shocks includes the shock to goodwill, ν_{jt}, which captures consumer response to adverting copy unobserved by firms before airing the advertisement. Each market's profit for firm j is:

$$\pi_{jt} = (p_{jt} - c_{jt})Q_{jt}(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t, \xi_t) - E_{jt}. \tag{15}$$

c_j is a constant marginal cost of production, E_{jt} is Firm F's advertising expenditure. Because prices and advertising are set before the goodwill shock is realized, firms maximize expected per period profits,

$$\pi_{jt} = \pi_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t) = \int (p_{jt} - c_{jt})Q_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t, \xi_t) f(\xi_t) d\xi_t - E_{jt}, \tag{16}$$

recognizing that ξ is *i.i.d.*. The multi-product firm maximizes profits jointly over each product in its portfolio.

$$\Pi(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t) = \sum_{j \in G_F} \pi_j(\mathbf{g}_t, \mathbf{A}_t, \mathbf{P}_t), \tag{17}$$

where G_F is the set of products in firms F's portfolio.

3.4 Firms

Firms play the following advertising game. At the start of a period the state of the market, $\mathbf{g}_t$, is observed by all firms. Upon observing the state vector firms make pricing and advertising decisions, $\sigma_j(\mathbf{g}_t) = (P_{jt}, A_{jt})$, for each product in their portfolio. In this game, firms' decisions rely only on payoff relevant state variables. Given product demand and current goodwill levels, the firm has all the necessary information to determine current and future sales. This implies that $\mathbf{g}_t$ contains all the necessary payoff relevant information. Once firms observe the state vector and firms choose prices and advertising levels the demand shocks, ξ_t, are realized and profits are determined.

The strategy profile vector $\sigma = (\sigma_1, \ldots, \sigma_N)$ contains the price and advertising decisions of all N firms for each of their products. The expected discounted profits for firm F in state $\mathbf{g}_t$ under strategy profile σ are:

$$V_F(\mathbf{g}_t|\sigma) = \mathbb{E}\left[\sum_{s=t}^{\infty} \beta^{s-t}\Pi_F(\mathbf{g}_s, \sigma_F(\mathbf{g}_s))|\mathbf{g}_t\right]. \tag{18}$$

The firm level objective is to maximize this stream of expected profits by choosing a strategy profile σ_F. To maximize the stream of profits in (18) the firm needs to know the state variables $\mathbf{g}_t$ and consequently the strategy profiles $\sigma(\mathbf{g}_t)$. Equation (12) states that the goodwill stock g_{jt} evolves according to a first order Markov process whose transition density is $p(g_{jt+1} \cdot |g_{jt}, A_{jt})$. Recalling that the goodwill shocks are *i.i.d.*, the Markov transition density of the state vector is:

$$f(\mathbf{g}_{t+1}|\mathbf{g}_t, \mathbf{A}_t) = \prod_{j=1}^{J} f(g_{j,t+1}|g_{jt}, A_{jt}). \tag{19}$$

Because firms make pricing and advertising decisions based solely on the current state vector, time dependent strategies are not considered because they would complicate an already difficult equilibrium computation. Firm F's Markov strategy is $\sigma : \mathbf{g} \rightarrow \sigma_F(\mathbf{g}) = (P_F(\mathbf{g}), A_F(\mathbf{g}))$. Each firm sets its policy, σ_F, conditional on the strategy profile of competing firms $\sigma_{-j} = (\sigma_1, \ldots, \sigma_N)$. The Markov strategy along with equation (19) fully define a firms value function as stated in equation (18).

3.5 Equilibrium

Equilibrium is described by the value function in a dynamic programming problem with strategic interactions. Each firm has a value function that satisfies the following Bellman equation:

$$V_F(\mathbf{g}_t|\sigma) = sup_{\tau\in\mathbb{R}^2_+}\left\{\Pi_F(\mathbf{g}, \tau, \sigma_{-j}(\mathbf{g})) + \beta\int V_F(\mathbf{g}'|\sigma)f(\mathbf{g}'|\mathbf{g}, \tau, \sigma_{-j}(\mathbf{g}))d\mathbf{g}'\right\} \tag{20}$$

Firms choose advertising levels and prices for their products, so the supremum is taken with respect to $\tau = (P_j, A_j)$, in $\mathbb{R}^2_+$. This Bellman equation includes the competing firms strategy profile σ_{-j} on the right hand side, which is firm F's best guess at the competing strategy profile and defines the firms best response to σ_{-j}. A Markov perfect equilibrium (MPE) of the dynamic game is a list of strategies, $\sigma^* = (\sigma^*_1, \ldots, \sigma^*_N)$ such that no firm can deviate from their action σ^*_n in any subgame that starts at state $\mathbf{g}$.[6] We restrict our attention to pure strategies due to the computational difficulty of determining equilibria in a more general

[6]For a more complete and concise treatment of the MPE concept see Maskin and Tirole (2001).

model that might include mixed strategies.

Define: A Markov perfect equilibrium is a Markov strategy profile σ^* such that:

$$V_j(\mathbf{g}|\sigma^*) \geq \pi_j(\mathbf{g}, \tau, \sigma^*_{-j}(\mathbf{g})) + \beta \int V_j(\mathbf{g}'|\sigma) f(\mathbf{g}'|\mathbf{g}, \tau, \sigma^*_{-j}(\mathbf{g})) d\mathbf{g}' \tag{21}$$

for all unilateral deviations $\tau = (A_j, P_j)$, states $\mathbf{g}$, and firms j.

Ericson and Pakes (1995) and Doraszelski and Satterthwaite (2003) explore general conditions for the existence and uniqueness of an equilibrium solution. Our empirical work relies on the existence of an equilibrium solution for a particular set of demand parameter estimates. If we are able to numerically compute an equilibrium its existence is established; Dubé et al. (2005, p. 115) argue along similar lines. Officially, we cannot determine, with absolute certainty, whether the equilibrium is a unique ideal point on the Pareto front. Validity of the solution is supported in two ways: one, we start the numerical algorithm for locating equilibria at different points to see if it converges to the same equilibrium; two, subsequent to finding a potential equilibrium we use the global optimization technique called simulated annealing to determine whether the current solution can be improved upon.

4 Identification

One of the reasons we apply the Berry et al. (1995) demand modeling framework is because it allows us to employ linear instrumental variable estimation techniques to control for the potential endogeneity of product attributes. The other reason we apply their framework is because we do not observe consumer level ad exposure and frequency. To identify the consumer distribution of the marginal utility of advertising stock and other product attributes, we combine aggregate market level data with micro level data recording consumer purchases.

Much like BLP, our identification strategy relies on the fact that we can separate mean market utility from consumer specific deviations and use a linear instrumental variable estimator to recover mean preference parameters. We further exploit the separability and specify two different sets of moment conditions to identify mean preferences and consumer specific deviations separately.

To the extent that price and advertising levels are set by firms upon observing demand factors unobserved by the econometrician, time varying product characteristics may be correlated with demand shocks. For example, firms may set price and advertising policies conditional on merchandizing activities and exogenous demand factors. This correlation generates endogeneity bias in demand parameter estimates if not properly controlled for.

4.1 Price Endogeneity

The principle identifying assumption is that the demand unobservables are mean independent of an exogenous set of instruments, z:

$$E[\xi_{jt}(\Theta_0)|z] = 0, \tag{22}$$

for a demand shock, ξ_j, evaluated at the true parameter values Θ_0. We do not observed ξ_{jt} but firms and consumers do, which leads to the endogeneity bias described above. To control for persistent geographic differences in demand we include a full set of metropolitan market fixed effects. To control for price endogeneity caused by time varying demand unobservables, Berry et al. (1995) show that variables that shift profit margins are valid instruments. In particular equation (15) indicates that profit is a function of demand and equation (4) specifies demand as a function of profit attributes, consequently BLP's reasoning applies here.

The full set of instruments, z, include geographic market fixed effects, product characteristics (also appearing in x), production input cost variables (not appearing in x), variables that measure advertising costs, and any linearly independent function of z. We test the validity of the moment restriction in (22) as well as the validity of z for explaining variation in price. These tests are conducted at the stage of estimation and they are found in the section that discusses demand results.

4.2 Advertising Endogeneity

According to the model we introduce in the previous section, Advertising endogeneity presents itself in two forms. First, advertising levels are potentially correlated with demand shocks, specifically as described in our pricing and advertising model. Advertisers may increase or decrease advertising with positive demand shocks. Second, our model suggests negative goodwill shocks would be correlated with high levels of advertising. We provide an institutional reason why exogenous variation in the observed advertising levels exist and at the stage of estimation we formally test whether it presents itself in bias parameter estimates when not controlled for; our test yields results consistent with the institutional facts.

One institutional feature of advertising contracts are “make good” clauses. Advertising contracts specify the amount of GRPs to be delivered in a week, delivered GRPs depend on the actual reach and frequency of the aired ads. “Often the contracted amount of GRPs are not delivered entirely and, consequently, the television stations ‘make good’ in subsequent periods by delivering the residual promised GRP levels”(Dubé et al., 2005, p.122). Make-goods appear in our data in the form of low levels of advertising at the end of a particular ad campaign. Figure 1 is a Histogram of strictly positive GRPs across our data. The large number of low advertising levels indicates that make-goods are observed frequently throughout out data. This source of exogenous variation in advertising levels enables us to identify advertising demand response independent of the strategically chosen advertising levels determined in our model.

4.3 Heterogeneity Parameters

To identify the parameters that capture the heterogeneity of consumer preferences in the market we sample households from our panel using survey weights provided in the home-scan data. The home-scan panel tracks household purchases as well as well as household demographic information. Using this information, we can “trace out” the distribution of consumer preferences (Berry & Haile, 2010; Fox, Kim, Ryan, & Bajari, 2011). To identify the parameters that characterize this distribution, $\Theta_h = \Omega, \Sigma, \rho, \upsilon$, we begin with the density of consumer choices,

$$f(y_{ijt}|\delta, \Theta_h, x, D) = \int \prod_t \prod_i \prod_j s_{ijt}(\delta, \Theta_h, x_{jt}, D_i)^{y_{ijt}} dF(\upsilon_i). \tag{23}$$

Here s_{ijt} is defined as in equation (4) and $y_{ijt} = 1$ if consumer i chooses product j in period t and $y_{ijt} = 0$ otherwise. This density is a likelihood and we specify micromoments that result from the first order conditions, they are defined as:

$$E[\partial log(s_{ijt})/\partial\Theta_h] = 0. \tag{24}$$

$\partial log(s_{ijt})/\partial\Theta_h$ is precisely the score function associated with the likelihood of purchases conditional on mean preferences. These micromoments differ from those specified by Petrin (2002), and others who apply his method, such as Goeree (2008), who assert that the deviation in predicted probability of choice from observed choice is mean independent of consumer and product attribute covariates:

$$E[(y_{ijt} - s_{ijt})Dx] = 0. \tag{25}$$

In contrast, the moment condition expressed in (24) exploits the full structure of the probability model that our demand system applies and gains the efficiency and well behaved estimation objective of a logit maximum likelihood estimation approach without restricting the functional relationship between demand shocks and instruments, defined in equation (22) - the way a full maximum likelihood estimation approach would.

5 Estimation Approach

To estimate our demand model we apply the Mathematical Programming Equilibrium Condition (MPEC) framework of Dubé et al. (2009). The complete set of parameters estimated are $\theta = \{\beta, \Sigma, \kappa, \rho, \sigma_a, \lambda, \sigma_\nu\}$. Our estimation objective function is a generalized method of moments (GMM) objective composed of two sets of moments. The first set of moments are the BLP macromoments which apply the mean independence of demand shocks and exogenous product attributes and marginal cost shifters, with the added condition that predicted market shares must equal observed market shares. The second set of moments specify the score functions, with respect to the heterogeneity parameters, implied by the consumer level share (choice) model, (4). The specified score is a function of the observed choices and product attributes,conditional on the market mean utilities, δ and γ, as determined by the first set of moments. Next we explain each in turn, and then wrap-up the section with a short discussion of the computational estimation approach.

5.1 BLP Macromoments

Under weak regularity conditions on the density of consumer unobservables, the existence of a unique mean utility that satisfies the observed market shares has been established by Berry (1994), which is extended by Goeree (2008) for the limited information model we employ. The unique mapping of mean utilities into shares allows us to use the condition that observed market shares must equal predicted market shares when consumer utility parameters are estimated. We solve for the δ implied by,

$$S_t^{obs} - s_t(\delta, \Theta) = 0, \tag{26}$$

where S_t^{obs} is the vector of observed market shares and s_t is the vector of predicted market share, based on our model. Conditional on the constraint implied by (26), the aggregate demand shock is:

$$\xi_{jt} = \delta_{jt}(S, \Theta) - x_j'\beta. \tag{27}$$

These unobserved demand shocks are used to build the macromoment conditions defined in equation (22).

5.2 Consumer Choice Micromoments

The micromoments are derived from the consumer choice model implied by the demand specification. After integrating out unobserved heterogeneity, the density of a household's choices over time is given by,

$$L_i(y_i|X, p, g^a, \Theta) = \int \prod_{t=1}^{T} \prod_{j=1}^{J} s_{ijt}(X, D, p, \delta, \gamma, \Theta)^{y_{ijt}} dF(\nu_i) \tag{28}$$

$y_i = (y_{1i}, \ldots, y_{ti})$ and $y_{ijt} = 1$ if household i chooses brand j in market t and 0 otherwise. The moments implied by this likelihood model are given by the score for the heterogeneity parameters, first the derivative of the likelihood with respect to the consumption utility heterogeneity parameters, $\theta^m_{cons} = [\Omega, \Sigma]$,

$$\mathbb{S}(\theta^m_{cons}) = \sum_{\mathcal{S} \in C_j} K_s p^{\mathcal{S}}_{ijt} \left(\frac{\partial \mu_{ijt}}{\partial \theta^m_{cons}} - \sum_{k \in \mathcal{S}} p^{\mathcal{S}}_{ikt} \frac{\partial \mu_{ijt}}{\partial \theta^m_{cons}} \right), \tag{29}$$

where $K_s = \prod_{l \in \mathcal{S}} \phi_{ilt} \prod_{r \notin \mathcal{S}} (1 - \phi_{irt})$ is the probability of product j salience and $p^{\mathcal{S}}_{ijt} = \frac{exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k \in \mathcal{S}} exp(\delta_{kt} + \mu_{ikt})}$ is the probability of choosing j conditional on being aware. Evoking the identity $K_s \frac{\partial log(K_s)}{\partial \tau_{ijt}} \equiv \frac{\partial K_s}{\partial \tau_{ijt}}$, the derivative of the likelihood with respect to the information utility heterogeneity parameter, $\theta_{info} = [\rho, \sigma_a]$, is:

$$\mathbb{S}(\theta_{info}) = \sum_{\mathcal{S} \in C_j} p^{\mathcal{S}}_{ijt} K_s \left[\sum_{l \in \mathcal{S}} (1 - \phi_{ijt}) \frac{\partial \tau_{ijt}}{\partial \theta_{info}} - \sum_{r \notin \mathcal{S}} \phi_{ijt} \frac{\partial \tau_{ijt}}{\partial \theta_{info}} \right]. \tag{30}$$

The first order conditions in equations (29) and (30) for Maximum Likelihood Estimation (MLE) make it a GMM estimator (Newey & McFadden, 1994, p.2118). We define these as consumer choice micromoments.

$$E[\mathbb{S}(\theta_{cons})|\delta,\gamma] = 0 \tag{31}$$

$$E[\mathbb{S}(\theta_{info})|\delta,\gamma] = 0 \tag{32}$$

One major benefit of linking the micromoments to the macromoments by conditioning on δ and γ is that both the heterogeneity parameters, $[\theta_{cons}, \theta_{info}]$, and the mean preference parameters β and κ will not suffer from endogeneity bias because we condition on ξ, which is contained in δ.

5.3 GMM Estimation Method

Many previous studies using BLP demand models rely upon the nested fixed point approach to estimate the random coefficients logit model (Berry et al., 1995; Nevo, 2001). The nested fixed point estimation approach is made up of two distinct parts. First practitioners use contraction mapping to find the mean utility that makes observed share equal to predicted shares. Then they estimate the density of preference parameters in a subsequent step using a generalized method of moments (GMM) estimation approach. We choose an alternative MPEC approach.

The MPEC approach recast estimation of the random coefficients logit as a mathematical programming problem with equilibrium constraints. The idea is to specify the GMM objective function and minimize it with the constraint that observed market shares equal predicted market shares. This presents a challenging nonlinear constrained optimization problem which requires the practitioner to use state-of-the-art constrained optimization tools, we use $KNITRO^{®}$ through the $TOMLAB^{®}$ optimization environment. Dubé et al. (2009) document several numerical concerns for the nested fixed point approach typically applied in the literature, and demonstrate that the constrained optimization approach is uniformly preferred for computational efficiency and accuracy. Conlon (2010) extends the MPEC approach to empirical likelihood estimation of a dynamic demand problem. We extend the MPEC approach to a limited awareness demand setting and include micromoments that incorporate consumer purchase data.

Our estimation algorithm solves the following mathematical program:

$$\min_{(x)} \quad g'Wg \tag{33}$$

s.t.

$$\begin{aligned}
s_{ijt} &= \sum_{\mathcal{S} \in C_j} \prod_{l \in \mathcal{S}} \phi_{ilt} \prod_{r \notin \mathcal{S}} (1 - \phi_{irt}) \frac{exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k \in \mathcal{S}} exp(\delta_{kt} + \mu_{ikt})} \\
S_{jt}^{obs} &= \frac{1}{n} \sum_{i} s_{ijt}(\delta, \gamma | \mu, \tau) \\
g &= \sum_{\forall j,t} (\delta_{jt} - x'_{jt}\beta) z_{jt} \\
&\sum_{\forall i,j,t} \mathbb{S}(\mu, \tau | \delta, \gamma) = 0.
\end{aligned}$$

Where W is the a two-step optimal GMM weighting matrix, and $\mathbb{S}(\mu, \tau|\delta, \gamma)$ is defined by (29) and (30).

6 Demand Results

This section discusses results from the demand model estimates. We present parameter estimates from several basic logit model specifications, including OLS and instrumental variable (IV) approaches, and a micro-BLP model to provide a baseline for comparison. Next we analyze parameter estimates for full and limited awareness demand models that include an advertising stock effect, each specified with and without parameter heterogeneity. All models use 3,776 market observations, based on 59 biweekly periods, 8 products, and 8 markets.[7] Models with micro moments sample 400 consumers from each DMA, based on Nielsen Homescan panel weights. Purchase decisions for these households are observed across the same 59 biweekly periods, 8 products, and 8 markets for a total of 1,510,400 household product choice observations.

6.1 Estimation Results for Logit OLS, IV, and micro-BLP Demand Models

Table 3 shows results for various specifications of the simple logit OLS and IV demand models. In all models *Price* is the price coefficient and *Advertising* is the marginal utility of goodwill. *Calories*, *Fiber*, *Sugar*, and *Protein* are the marginal utilities of product specific fixed characteristics. All specifications

[7] Our data has a total of 76 biweekly periods but we lose 5 periods due to lags in advertising expenditure and 12 additional periods due to the advertising specification in the model, thus resulting in 59 biweekly periods of data for demand estimation.

include DMA fixed effects. A quick look at this table reveals that key parameter estimates (price and advertising) are highly significant in every model as standard error estimates testify.

Columns (*i*), (*ii*), and (*iii*) exhibit estimates of the simple logit without instrumentation for prices. From these three columns one will observe that inclusion of product characteristic variables: calories, fiber, sugar, and protein, removes upward bias on price sensitivity. Alternatively the inclusion of an advertising variable removes a downward bias on price. Columns (*iv*), (*v*), (*vi*), and (*vii*) exhibit estimates from various logit specifications estimated with different sets of instrumental variables. These estimates demonstrate that instrumenting for price reduces the downward endogeniety bias. Model (*iv*) includes only last period's price as an instrument, while column (*v*) also includes last period's advertising expenditure as an instrument. The next two logit specifications, columns (*vi*) and (*vii*) include an expanded instrument set that consists of lag price and six periods of advertising expenditure.

Hansen's J statistic testifies to the degree of instrument exogeniety in columns (*v*), (*vi*), and (*vii*). An upper tail test fails to reject the null hypothesis in columns (*vi*) and (*vii*), that the instruments in these specifications are uncorrelated with the model error. First stage regression statistics testify to instrument relevance. First stage F-stats and R^2 indicate that the instruments are relatively strong and valid and explain a reasonable percentage of variation in the endogenous pricing variable.

To test whether 'make-goods' provide the exogenous variation required to treat advertising as exogenous, column (*viii*) shows results for the same specification as column (*vii*) with the exception that advertising is now treated as an endogenous variable. Parameter estimates between columns (*vii*) and (*viii*) are very similar, in fact a Hausman test for endogeneity of the advertising variable yields a chi^2 value of 0.59 failing to reject the null hypotheses that there is no systematic differences between the coefficient estimates for the two specifications. This supports our argument that advertising does not require treatment as an endogenous variable in this econometric model, and the observed exogenous variation in Figure 1 is sufficient to identify the advertising utility parameters.

The final four columns of Table 3 present estimates for the standard random coefficients logit specification commonly referred to as BLP, specified with the additional micromoments in the same fashion as the limited information demand model we estimate. We include these specifications to benchmark our merger simulation results and our demand estimation results, to past literature that investigates the RTE cereal market. One specification includes GRPs and the other does not. The specification that includes

GRP benchmarks our merger simulation; this would be the specification that a practitioner using extant methods would use given the available data. We include the specification without advertising to benchmark our results against past demand studies of the RTE cereal market. Although we do not present elasticity tables for these models, these specifications admit own price elasticities between -2 and -4.5, which is similar to those found in Nevo (2000) and Nevo (2001).

6.2 Estimation Results for Limited Information Demand Models and Full Information Demand Model with Advertising Stock

Table 4 shows the results for the limited information specification demand models and the full information random coefficients demand model with advertising stock. The key parameter estimates on price and advertising are denoted as $Price$ and α^a. These are highly significant in each model. Again, the coefficient estimates for calories, fiber, sugar, and protein are the marginal utilities of product specific fixed characteristics, where protein and calories are generally not statistically different from zero. All models in the limited information framework also include DMA fixed effects.

Columns (i) and (ii) are the estimates for the simple limited information demand model, without consumer heterogeneity. The key difference is that column (ii) includes the advertising stock effect, and thus estimation of λ and σ_a. Including the advertising stock reduces the price sensitivity while increasing the impacts of advertising. Similar to the simple logit models, Hansen's J provides evidence that the instrument set is exogenous in the limited information demand models.

The full information heterogeneous consumer random coefficients specification includes advertising stock and is provided as a benchmark to compare to the similar limited information model. The price coefficient is centered at -4.671 with a standard deviation of 1.164. This implies that 95% of the consumers have marginal price sensitivities between -6.999 and -2.343. The advertising coefficient is centered at 0.544 with a standard deviation of 0.084, resulting in a range of advertising sensitivities between 0.376 and 0.712. In computing these parameters we assumed the advertising stock effect parameters λ and σ_a are equal to the limited information demand model with advertising stock and without consumer heterogeneity.

The parameter estimates for the heterogeneous consumer random coefficients specification of the limited information model appears in the final columns of Table 4. The distribution of the price sensitivity coefficient is centered at -4.168 and has a standard deviation of 0.354, which implies that 95% of the consumers have marginal price sensitivities between -4.876 and -3.460, consistent with the law of demand. The advertising coefficient is centered at 2.102 with a standard deviation of 0.022, yielding a range of 2.058 to 2.146 within 2 standard deviations of the mean. With the advertising decay rate, λ=0.495, a single unit of advertising today will have less than 1% effect on goodwill after two months. Since the same instrument sets and same number of parameters in the specification of the full information and limited information models, an informal test of fit compares the J statistics. Comparing J statistics for the two models the limited information model appears to fit the data better suggesting that it may be more appropriate for this empirical setting. In general the limited information specification offers greater flexibility because of its superior ability to shift the demand curve in and out.

Table 5 displays own price elasticities along with mean cross price elasticities for each advertising inclusive heterogeneous consumer demand models we estimate. The results here indicate the direction of the estimation bias due to the advertising specification. When advertising stock is included in the micro-BLP model price sensitivity is reduced. The price sensitivity is further reduced under the limited information specification. The specification itself does not guarantee this result, it is an empirical matter. Draganska and Klapper (2012) provide evidence of bias in the other direction and argue that the direction of this bias should be expected, however this would only be true in a simple model with two choices, an inside good with limited awareness and an outside good. Goeree (2008) on the other hand finds bias in the same direction as our results.

Figure 2 and 3 plot expected demand response to advertising for the final two models in columns three and four of Table 4. An examination of these figures offers a direct test of the sales response concavity to advertising. As we expect, the sales response to advertising for every product is globally concave. Sales response concavity to advertising is particularly important because it is consistent with the pulsing of advertising observed in the data, and makes pulsing a potential equilibrium strategy. One additional interesting result of these plots is that Post Honey Bunches of Oats advertising appears to generate the largest response. This may suggest that they run particulary strong ad copy or that consumers of Honey Bunches of Oats are particulary fond of their advertising campaigns.

7 Market Simulations

In this section we simulate a merger in the taste enhanced wholesome segment of the RTE cereal market. The section first explains how we compute equilibrium. We conduct five simulations altogether. The first three benchmark our demand models with a price equilibrium only simulation; treating advertising exogenously. In the last two simulations advertising is fully endogenized in the joint pricing/advertising decision. We compare the simulations to identify the impact of incorporating advertising, thus illustrating the importance of including advertising in the analysis. We also formally investigate the fit of the simulated data for each model to the observed data.

7.1 Markov Perfect Equilibrium Computational Approach

To compute MPE in price and advertising, we introduce an approach that formulates the dynamic programming problem as a mathematical programming problem. The intuition for the problem is no different than nested fixed-point approaches applied in the past. The benefit of formulating the dynamic problem as a mathematical program with optimality conditions is that we can mix smooth optimization approaches with non smooth optimization approaches to find equilibria. The program is formulated as:

$$
\begin{aligned}
\max_{(p_t, A_t)} \quad & \textstyle\sum_f V^F \\
\text{s.t.} \quad & \\
& \text{c1}: \quad g_{t+1} = \lambda[g_t + \Psi(A)] + \nu_i \\
& \text{c2}: \quad p - c = (O^F \times_{elt} \Omega(p, A)) s(p, A) \\
& \text{c3}: \quad \Pi_A(p_{t+1}, g_t) = E[\beta \Pi_A(p_{t+1}, g_{t+1})]
\end{aligned}
\tag{34}
$$

This problem specifies maximization of each firms objective stated in equation (20), subject to the equation of motion and the competitive price and advertising reaction conditions under MPE behavior. The first constraint, c1, is the equation of motion. The second constraint, c2, is the static optimal markup condition given the firms' portfolio ownership matrix O^F. Simulating a merger implies a change to the market portfolio ownership structure defined by matrix O^F. $\Omega(p, A)$ is the demand response to price matrix with common element $\frac{\partial s_i}{\partial p_j}$.[8] We assume that the marginal cost of production remains constant with and without the merger, while the unit cost of advertising is allowed to change, which allows us to recover advertising cost reducing efficiencies due to a merger. Constraint three, c3, is the Euler equation defining the intertemporal optimal advertising condition. The integral in c3 is computed by Monte Carlo simulation. We directly solve for the equilibrium implied by the set of MPE constraints: c1, c2, and c3.

We use the first 12 periods of data to initiate advertising goodwill levels. The goodwill levels at the end of these 12 periods establish the initial state vector for simulating the MPE from period 13 on. Results are simulated under the assumption that the discount factor β is equal to .9916 which corresponds to an annual interest rate of 11%.

7.2 Merger Simulation

The merger simulation that we present in this section is motivated by a current business relationship between Kellogg's and Kashi. In 2000 Kellogg's purchased Kashi and states that they operate the brands independently following their original business philosophy.[9] It thus becomes an interesting empirical question to analyze two things. One, which equilibrium pricing and advertising conduct seems more plausible given the data we observe? And two, given that the brands are independently managed, if Kellogg's were to fully internalize the Kashi brand into it's portfolio, how would pricing and advertising conduct change? This exercise also provides some general insight into the ramifications of incorporating advertising into a competitive equilibrium analysis as well as the implication of demand model specification on equilibrium conduct. We focus our analysis on the Chicago market, due to the extensive computational exercise we conduct.

The first 12 biweekly periods are used to establish baseline advertising stock levels. Then we analyze a hypothetical merger occurring at the end of the twelfth period between Kashi and Kellogg's by computing the optimal advertising and pricing with and without the merger in the next 47 periods. The merger is anticompetitive if the post merger prices are significantly higher than prices without the merger. Note that with these brands in the merged entity it is possible for some prices to go up and one or more to go down creating a welfare tradeoff if consumers choose different brands.

[8] Element $O^F_{k,j} = 1$ if the firms have both k and j in their portfolio, 0 otherwise.

[9] Kashi Website: `www.kashi.com/meet_us/history`.

7.3 Merger Simulation of Dynamic Price and Advertising Impacts

In order to assess the overall effect of the merger, Table 6 presents five panels that summarize the equilibria implied for each merger simulation. The table displays average 16 ounce prices for each brand for the entire simulation period. The first panel of the table displays results for the benchmark analysis that considers static Nash-Bertrand pricing conduct, treats advertising as exogenous to the pricing game, and applies the full information demand specification from the first micro-BLP model in Table 3.[10] As one might expect the proposed merger results in a price increase for non-merged products, a modest increase for products in Kellogg's portfolio, and a five to six cent per standard box price increase for the Kashi product, when managed under Kellogg's banner. The second panel in Table 6 records the same simulation given the full information micro-BLP model inclusive of advertising stock (goodwill). Including advertising stock reduces price sensitivity which results in a larger price impact under the hypothetical merger, particularly for products involved in the merger. The third panel in Table 6 displays the merger price impact when the limited information model is applied to the price only merger simulation. This specification results in further reduced price sensitivity. Not surprisingly the price impact of the hypothetical merger is even larger, significantly more so for products under management of the merged firm. Results from these three simulations document the impact of demand specification on a price only analysis.

The first panel of Table 7 is the first of two simulations that internalize dynamic advertising effects as well as pricing. This simulation applied the full information micro-BLP model with advertising stock. Recall that this model specifies advertising as a complementary consumption attribute. Here we note that the proposed merger continues to result in upward price pressure, in fact more so than any of the supply models that treat advertising exogenously. This result follows from the conventional wisdom that the merger reduces competition, increases margins, and in turn increases the marginal benefit of advertising. As a result we see higher advertising levels in Table 7 and higher price levels than any of the price only simulations. A finding consistent with the theory of Dorfman and Steiner (1954) and Friedman (1983).

The second panel of Table 7 presents the price and advertising simulation result under the limited information demand model. The results indicate that upward pricing pressure under this model is less than under the full information model from the previous simulation. In addition Table 8 documents a decrease in advertising. This equilibrium suggests the contrasting view of advertising competition wherein a pris-

[10]This model and analysis is equivalent to extant unilateral price only analysis such as (Nevo, 2000, 2001).

oners dilemma outcome emerges. Effectively, advertising diminishes due to the competition for consumer awareness, so when the opportunity for collusion presents itself, firms back off on product advertising and pass those cost saving onto the consumers in the form of less elevated prices under the merger. This result is consistent with the Grossman and Shapiro (1984) findings that there is too much advertising because products are too similar so the beneficial effect of improved matching is outweighed by the wasteful effect of competing for consumer attention. If there was a sufficient degree of product differentiation we would not see the substantial decrease in advertising, which implies a more efficient allocation of advertising under a more collusive market structure. The contrasting results highlight an important implication of model specification in this empirical setting.

Table 7 reports observed average prices and Mean Integrated Square Error (MISE) computed using kernel density estimates of the price densities. The MISE statistic measures the affinity between the observed pricing distribution and simulated distributions, the closer the statistic is to zero the greater the degree of affinity. We have also provided the kernel density estimates for the pricing distributions and conducted Kolmogorov-Shmirnov tests to determine whether either of the simulated densities are significantly different from the observed distribution. Figure 4 illustrates that the limited information model generates a price distribution that matches the observed data closer than the full information model. The Kolmogorov-Shmirnov tests reject the simulated pricing distribution based on the full awareness persuasive model at the 10 percent level under either market structure. In contrast, we fail to reject the limited awareness informative model under either market structure. At the product level we find that prices simulated from the limited awareness model fits the empirical distribution better. We also find that the post acquisition model fits the empirical distribution slightly better although we cannot reject one in favor of the other from a statistical standpoint. As a practical matter, we conjecture that the limited information model provides an extra degree of flexibility that allows for more demand curve rotation than the full information model. As an empirical matter, the evidence supports the limited information framework in the taste enhance whole segment of the RTE cereal market, implying that the insights of Grossman and Shapiro (1984) are plausible here. It should also be noted that these results square with the findings of Tenn et al. (2010), who investigate promotions in the premium ice cream market.

Comparing the limited information model with price and advertising competition to the Nevo (2000) type merger simulation we see that the price impacts are larger when advertising is incorporated. This result is owed to a dominant specification bias effect. As support, when we compare price impacts for the merged products under the limited information demand model, we see that the price increase for the price only model is larger than the price and advertising model. This result indicates that incorporating advertising in the merger analysis mitigates the price impact due to the proposed merger. The smaller impact can be attributed to efficiencies in advertising costs and reduced demand due to lower advertising levels. Turning ones attention to Table 8, advertising cost efficiency is verified upon inspection, wherein we notice that the merger results in less advertising and consequently less advertising cost, which is passed through to prices. This result stems from the fact that cross advertising effects are negative and adequately large on average, which squares with the Grossman and Shapiro (1984) insights described above.

Our results demonstrate that the inclusion of advertising into this merger analysis mitigates the price impact for the products involved in the proposed merger as well as the overall price level in the TEW segment of the RTE cereal market. The larger implication of this finding, that extends beyond merger analysis, is that advertising should not be ignored in the evaluation of equilibrium behavior for two reasons. One implication is specification bias, the direction of which is an empirical matter. The second reason is the incorporation of advertising in the equilibrium analysis, which is also an empirical matter.

8 Conclusion

This paper introduces an approach for incorporating simultaneous and dynamic pricing and advertising strategies into market equilibrium analysis. We began by specifying demand models that include the dynamic process of advertising stock. Then we outlined the dynamic firm behavior and equilibrium concept. Methodologically, the paper advances a new econometric approach that specifies moments based on the consumer level share model and uses consumer data within an MPEC estimation problem to identify heterogeneity distribution parameters and strongly tie the consumer choice model to the aggregate demand system. We subsequently present a direct computation method to solve Markov perfect equilibria for firms marketing multiple products in the same market. These advances allow us to measure price impacts of mergers in a more realistic equilibrium framework that includes pricing and advertising decisions.

Using demand estimates, we specify and compute equilibria for pricing and advertising games with and without a hypothetical merger. To benchmark our results, we simulated a hypothetical merger with a series of demand models based on alternative theories of advertsing's effect. When we conducted a price only merger simulation, results indicated that incorporating the dynamic role of advertising in the RTE cereal industry generate larger price impacts due to estimation bias. The simulations that incorporate advertising dynamics offers the insight that price increases due to a merger are larger, however endogenizing advertising increases the upward pricing pressure of a merger in a full information model however it reduces the price impact of the proposed merger under the limited information model. Model fit diagnostics indicate that the limited information framework is more plausible for the segment of the RTE cereal market that we study. Given the empirical tractability of the limited information framework the key result documents and measures the degree of advertising efficiencies attributable to the proposed merger.

This paper demonstrates the implications of advertising specification in a demand model, as well as the importance of including it in an equilibrium analysis conducted for product markets where advertising is a key component of marketing strategy. Future work might investigate vertical interactions between manufacturers who develop advertising strategies and retailers who sell the products or might might look closer at the content of advertising copy to determine the sales response of different communication messages.

References

Becker, G. S., & Murphy, K. M. (1993). A Simple Thoery of Advertising as a Good or Bad. *Quarterly Journal of Economics*, *108*, 941–64.

Benkard, C. (2004). A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft. *Review of Economic Studies*, *71*(3), 581–611.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, *63*, 841–889.

Berry, S., & Haile, P. (2010). Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers. *Cowles Foundation Discussion Paper No. 1718.*

Berry, S. T. (1994). Estimating Descrete Choice Models of Product Differentiation. *RAND Journal of Economics*, *25*(2), 242–262.

Chintagunta, P., & Dube, J. (2005). Estimating and Stock Keeping Unit Level Brand Choice Model That Combines Household Panel Data and Store Data. *Journal of Marketing Research*, *42*(3), 368–379.

Conlon, C. T. (2010). A Dynamic Model of Costs and Margins in the LCD TV Industry. *Working Paper*.

Deneckere, R., & Davidson, C. (1985). Incentives to form coalitions with Bertrand competition. *Rand Journal of Economics*, *16*, 473–486.

Doraszelski, U., & Markovich, S. (2007). Advertising Dynamics and Competitive Advantage. *RAND Journal of Economics*, *38*(3), 557.

Doraszelski, U., & Satterthwaite, M. (2003). Foundations of Markov-perfect industry dynamics: Existence, purification, and multiplicity. Tech. rep..

Dorfman, R., & Steiner, P. (1954). Optimal Advertising and Optimal Quality. *The American Economic Review*, *44*(5), 826–836.

Draganska, M., & Klapper, D. (2012). Choice Set Heterogeneity and the Role of Advertising. *Journal of Marketing Research*.

Dubé, J.-P., Hitsch, G. J., & Manchanda, P. (2005). An Empirical Model of Advertising Dynamics. *Quantitative Marketing and Economics*, *3*, 107–144.

Dubé, J., Fox, J., & Su, C. (2009). Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation. *NBER Working Paper*.

Einav, L., Leibtag, E., & Nevo, A. (2010). Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?. *Quantitative Marketing and Economics*, *8*(2), 207–239.

Ericson, R., & Pakes, A. (1995). Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies*, *62*(1), 53–82.

Fox, J., Kim, K., Ryan, S. P., & Bajari, P. (2011). The Random Coeffcients Logit Model is Identified. *Journal of Econometrics*.

Friedman, J. (1983). Advertising and oligopolistic equilibrium. *The Bell Journal of Economics*, *14*(2), 464–473.

Gasmi, F., Laffont, J., & Vuong, Q. (1992). Econometric Analysis of Collusive Behavior in a Soft-Drink Market. *Journal of Economics & Management Strategy*, *1*(2), 277–311.

Gerstner, E., & Hess, J. (1991). A Theory of Channel Price Promotions. *The American Economic Review*, *81*(4), 872–886.

Goeree, M. (2008). Limited Information and Advertising in the US Personal Computer Industry. *Econometrica*, *76*(5), 1017–1074.

Grossman, G., & Shapiro, C. (1984). Informative Advertising with Differentiated Products. *The Review of Economic Studies*, *51*(1), 63.

Hausman, J., Leonard, G., & Zona, J. (1994). Competitive Analysis with Differenciated Products. *Annales d'Economie et de Statistique*, 159–180.

Martin, S. (2002). *Advanced Industrial Economics* (Second edition). Blackwell Publishers.

Maskin, E., & Tirole, J. (2001). Markov Perfect Equilibrium I. Observable Actions. *Journal of Economic Theory*, *100*(2), 191–219.

Naples, M. (1979). *Effective frequency: The relationship between frequency and advertising effectiveness.* Association of National Advertisers, New York.

Nerlove, M., & Arrow, K. J. (1962). Optimal Advertising Policy Under Dynamic Conditions. *Economica*, *29*, 129–42.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, *69*(2), 307–342.

Nevo, A. (2000). Mergers with Differentiated Products: The Case of the Ready-To-Eat Cereal Industry. *RAND Journal of Economics*, *31*(3), 395–421.

Newey, W. K., & McFadden, D. (1994). *Large Sample Estimation and Hyptohesis Testing*, pp. 2113–2245. Amsterdam: North-Holland. Newey, Whitney K. and McFadden, Daniel (eds).

Pakes, A., & McGuire, P. (1994). Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model. *The RAND Journal of Economics*, 555–589.

Petrin, A. (2002). Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy*, *110*(4), 705–29.

Shum, M. (2004). Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market. *Journal of Economics & Management Strategy*, *13*(2), 241–272.

Slade, M. (1995). Product Rivalry with Multiple Strategic Weapons: An Analysis of Price and Advertising Competition. *Journal of Economics & Management Strategy*, *4*(3), 445.

Tenn, S., Froeb, L., & Tschantz, S. (2010). Mergers when Firms Compete by Choosing Both Price and Promotion. *International Journal of Industrial Organization*, *28*(6).

Table 1: Descriptive Statistics of Share, Price Per Pound, Advertising Penetration (Gross Rating Points) and Advertising Expenditure Per GRP

	Pct Share		Price/lb		14-day GRPs		ADV Expend/GRP	
Brand	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Kashi GoLean Crunch!	1.21%	0.82%	2.78	0.58	91.00	165.27	93.71	70.46
Kellogg's Raisin Bran	2.53%	1.18%	1.90	0.44	0.01	0.10	259.38	145.76
Kellogg's Raisin Bran Crunch	1.06%	0.76%	2.28	0.64	235.33	201.14	187.63	396.29
Kellogg's Smart Start	0.94%	0.70%	2.71	0.72	0.15	0.80	121.62	76.53
Kellogg's Special K Red Berries	1.42%	0.96%	3.17	0.64	68.94	137.36	154.93	96.03
Private Label Raisin Bran	1.24%	0.83%	1.76	0.49	0.00	0.00	-	-
Post Raisin Bran	1.44%	1.10%	1.71	0.57	0.00	0.00	-	-
Post Honey Bunches of Oats	4.38%	1.95%	2.24	0.37	354.54	246.13	99.77	82.40
Total Pct. Share of 8 Brands	14.21%							

Note: GRP is the Gross Rating Points from television advertising, defined as the sum of the number of telecasts (minutes, messages) among households (persons) in a frequency group times the percent of households (persons) accounted for in the respective frequency group.
- indicates no advertising took place for the respective brand
The mean value is the average across the entire panel which has 76 biweekly periods for 8 DMAs.
Source: Author's Calculations

Table 2: Percent Advertising Frequencies

Brand	Mean[1]	Std. Dev.	Min	Max
Kashi GoLean Crunch!	38.2%	0.0%	38.2%	38.2%
Kellogg's Raisin Bran	1.3%	0.0%	1.3%	1.3%
Kellogg's Raisin Bran Crunch	80.4%	0.5%	80.3%	81.6%
Kellogg's Smart Start	10.5%	0.0%	10.5%	10.5%
Kellogg's Special K Red Berries	26.6%	0.6%	26.3%	27.6%
Private Label Raisin Bran	0.0%	0.0%	0.0%	0.0%
Post Raisin Bran	0.0%	0.0%	0.0%	0.0%
Post Honey Bunches of Oats	93.3%	1.1%	92.1%	94.7%

Note: This table displays percent advertising frequencies across all 11 markets.
Percent advertising frequency is the percent of biweek periods that at least one advertisement appeared.
The mean value is the average across the entire panel which has 76 biweekly periods for 8 DMAs.
Source: Author's Calculations

Table 3: Logit OLS and IV Demand Model Estimation Results

Parameter	(i)	Logit OLS (ii)	(iii)	(iv)	(v)	Logit IV (vi)	(vii)	(viii)	Random Coefficient mean	Random Coefficient s.d.	Random Coefficient mean	Random Coefficient s.d.
Price	-1.744	-2.238	-1.946	-3.908	-6.330	-7.595	-6.806	-6.666	-6.536	1.164	-6.927	0.372
	(0.146)	(0.185)	(0.183)	(1.519)	(1.009)	(0.729)	(0.849)	(0.862)	(0.882)	(0.018)	(0.739)	(0.019)
Calories		0.310	-1.332	0.457	2.693	4.153	3.121	3.077	2.300	0.481	3.580	0.023
		(0.350)	(0.398)	(1.411)	(0.953)	(0.602)	(0.819)	(0.815)	(0.845)	(0.008)	(0.605)	(0.007)
Fiber		-0.837	-0.928	-1.005	-1.089	-1.114	-1.112	-1.098	-1.074	0.048	-1.091	0.103
		(0.083)	(0.084)	(0.105)	(0.104)	(0.107)	(0.102)	(0.103)	(0.098)	(0.011)	(0.103)	(0.011)
Sugar		-2.421	-2.631	-2.202	-1.648	-1.343	-1.555	-1.567	-1.787	0.253	-1.532	0.171
		(0.147)	(0.146)	(0.365)	(0.264)	(0.218)	(0.242)	(0.242)	(0.245)	(0.009)	(0.216)	(0.009)
Protein		-0.263	-0.081	0.171	0.481	0.598	0.540	0.509	0.410	0.114	0.462	0.169
		(0.092)	(0.095)	(0.224)	(0.173)	(0.163)	(0.161)	(0.165)	(0.160)	(0.012)	(0.161)	(0.012)
Advertising	1.607		1.099	0.793	0.424		0.333	0.304	0.293	0.159		
	(0.110)		(0.111)	(0.250)	(0.184)		(0.155)	(0.159)	(0.113)	(0.022)		
R^2	0.099	0.199	0.214									
Test of Overidentifying Restrictions												
Number of Restrictions					1	6	6	5	6		6	
Hansen's J					3.769	8.078	5.745	5.113	4.651		6.044	
p-value					0.0522	0.2325	0.4524	0.4022	0.5893		0.4183	
First-stage Regression Statistics												
Price R^2				0.377	0.382	0.376	0.389	0.376	0.389		0.376	
Price F-test				43.094	53.057	33.671	21.251	33.671	21.251		33.671	
Prob>F				0.000	0.000	0.000	0.000	0.000	0.000		0.00	
Advertising R^2								0.679				
Advertising F-test								128.841				
Prob>F								0.000				
Hausman Test for Endogeneity of Advertising (Ho: difference in coefficients not systematic)												
chi^2							0.590					
Prob>chi^2							1.000					
Instruments				Lag Price	Lag Price	Lag Price	Lag Price	Lag Price	Lag Price		Lag Price	
					Advertising Expenditure (t-1)	Advertising Expenditure (t+2:t-3)	Advertising Expenditure (t+2:t-3)	Advertising Expenditure (t+2:t-3)	Advertising Expenditure (t+2:t-3)		Advertising Expenditure (t+2:t-3)	

Source: Author's Calculations
Notes: Dependent variable is $ln(Sjt) - ln(Sot)$.
Demand estimation is based on 59 biweekly periods, 8 products, and 8 markets which results in 3,776 market observations.
Micro moments are then estimated with 400 consumer observations for the same 59 biweekly periods, 8 product, and 8 market which results in 1,510,400 consumer observations.
All regressions include DMA fixed effects.
Asymptotically robust $s.e.$ computed by the delta method are reported in parentheses.

Table 4: Demand Model Estimation Results

Parameter	Lim. Inf. (i)	Lim. Inf. /w Adv. Stock (ii)	Full Inf. Adv. Stock Random Coefficients mean	s.d.	Lim. Info. /w Adv. Stock Random Coefficients mean	s.d.
Price	-6.042	-4.140	-4.671	1.164	-4.168	0.354
	(0.922)	(1.723)	(1.364)	(0.018)	(1.455)	(0.050)
Calories	2.408	0.097	-0.033	0.482	0.010	0.021
	(0.890)	(1.991)	(1.465)	(0.008)	(1.629)	(0.019)
Fiber	-1.099	-1.062	-1.046	0.048	-1.071	0.058
	(0.099)	(0.096)	(0.095)	(0.011)	(0.094)	(0.030)
Sugar	-1.791	-2.271	-2.287	0.253	-2.305	0.163
	(0.252)	(0.458)	(0.356)	(0.009)	(0.385)	(0.024)
Protein	0.424	0.258	0.255	0.114	0.239	0.164
	(0.163)	(0.212)	(0.186)	(0.012)	(0.192)	(0.032)
α^a	1.111	2.136	0.544	0.084	2.102	0.022
	(0.552)	(1.073)	(0.173)	(0.016)	(0.982)	(0.057)
λ		0.497	0.497		0.495	
		(0.157)	(0.157)		(0.050)	
σ_a		0.300	0.300		0.087	
		(0.278)	(0.278)		(0.056)	
Test of Overidentifying Restrictions						
Number of Restrictions	6	5	5		5	
Hansen's J	4.550	2.804	3.799		2.980	
p-value	0.603	0.730	0.579		0.703	
Instruments	Lag Price	Lag Price	Lag Price		Lag Price	
	Advertising Expenditure (t+2:t-3)	Advertising Expenditure (t+2:t-3)	Advertising Expenditure (t+2:t-3)		Advertising Expenditure (t+2:t-3)	

Source: Author's Calculations
Notes: Dependent variable is $ln(Sjt) - ln(Sot)$.
Demand estimation is based on 59 biweekly periods, 8 products, and 8 markets which results in 3,776 market observations.
Micro moments are then estimated with 400 consumer observations for the same 59 biweekly periods, 8 product, and 8 market which results in 1,510,400 consumer observations.
All regressions include DMA fixed effects.
Asymptotically robust *s.e.* computed by the delta method are reported in parentheses.

Table 5: Demand Model Elasticity Estimates

	Full Information		Full Information with Ad Stock		Limited Information with Ad Stock	
Brand	Own Price	Mean Cross Price	Own Price	Mean Cross Price	Own Price	Mean Cross Price
Kashi GoLean Crunch!	-5.308	0.209	-3.494	0.127	-1.621	0.020
Kelloggs Raisin Bran	-5.402	0.175	-3.618	0.108	-1.090	0.028
Kelloggs Raisin Bran Crunch	-5.429	0.207	-3.596	0.127	-1.336	0.012
Kelloggs Smart Start	-5.404	0.211	-3.568	0.129	-1.594	0.013
Kelloggs Special K Red Berries	-5.320	0.205	-3.510	0.125	-1.827	0.025
Private Label Raisin Bran	-5.527	0.193	-3.697	0.120	-1.038	0.012
Post Raisin Bran	-5.533	0.194	-3.701	0.120	-1.007	0.011
Post Honey Bunches of Oats	-4.727	0.139	-3.124	0.085	-1.231	0.064

Source: Author's Calculations

Table 6: Merger Impact Analysis on Price in Chicago DMA

Simulation 1: Micro BLP with Exogenous Advertising

	Observed	Simulated Price			
	Price	Without Merger	With Merger	Increase Due to Merger	Pct Increase
Brands of the Merging Firms					
Kashi GoLean Crunch!	2.7191	2.6664	2.7101	0.0437	1.64%
Kelloggs Raisin Bran	1.8106	1.8833	1.8942	0.0109	0.58%
Kelloggs Raisin Bran Crunch	2.0854	2.2224	2.2332	0.0109	0.49%
Kelloggs Smart Start	2.4110	2.6472	2.6581	0.0109	0.41%
Kelloggs Special K Red Berries	2.7311	3.0595	3.0704	0.0109	0.36%
Brands in the Other Firms					
Private Label Raisin Bran	1.6682	1.7101	1.7101	0.0000	0.00%
Post Raisin Bran	1.5183	1.7113	1.7114	0.0001	0.00%
Post Honey Bunches of Oats	2.0031	2.2037	2.2038	0.0001	0.00%

Simulation 2: Micro BLP with Exogenous Advertising Stock

	Observed	Simulated Price			
	Price	Without Merger	With Merger	Increase Due to Merger	Pct Increase
Brands of the Merging Firms					
Kashi GoLean Crunch!	2.7191	2.6168	2.6772	0.0604	2.31%
Kelloggs Raisin Bran	1.8106	1.8604	1.8749	0.0145	0.78%
Kelloggs Raisin Bran Crunch	2.0854	2.1994	2.2139	0.0145	0.66%
Kelloggs Smart Start	2.4110	2.6243	2.6388	0.0145	0.55%
Kelloggs Special K Red Berries	2.7311	3.0366	3.0511	0.0145	0.48%
Brands in the Other Firms					
Private Label Raisin Bran	1.6682	1.6614	1.6614	0.0000	0.00%
Post Raisin Bran	1.5183	1.6819	1.6819	0.0001	0.00%
Post Honey Bunches of Oats	2.0031	2.1743	2.1743	0.0001	0.00%

Simulation 3: Limited Information with Exogenous Advertising Stock

	Observed	Simulated Price			
	Price	Without Merger	With Merger	Increase Due to Merger	Pct Increase
Brands of the Merging Firms					
Kashi GoLean Crunch!	2.7191	2.6410	2.7069	0.0659	2.49%
Kelloggs Raisin Bran	1.8106	1.7604	1.7750	0.0146	0.83%
Kelloggs Raisin Bran Crunch	2.0854	2.1273	2.1420	0.0146	0.69%
Kelloggs Smart Start	2.4110	2.5537	2.5683	0.0146	0.57%
Kelloggs Special K Red Berries	2.7311	2.9556	2.9702	0.0146	0.49%
Brands in the Other Firms					
Private Label Raisin Bran	1.6682	1.6886	1.6887	0.0000	0.00%
Post Raisin Bran	1.5183	1.6199	1.6200	0.0001	0.00%
Post Honey Bunches of Oats	2.0031	2.1078	2.1079	0.0001	0.00%

Source: Author's Calculations

Prices are for a 16oz package of cereal.

Table 7: Merger Impact Analysis on Price in Chicago DMA

Simulation 1: Full Information Micro BLP with Endogenous Advertising

		Simulated Price					
	Observed Price	Without Merger	MISE	With Merger	MISE	Increase Due to Merger	Pct Increase
Brands of the Merging Firms							
Kashi GoLean Crunch!	2.7191	2.6716	0.0052	2.8470	2.0276	0.1755	6.57%
Kelloggs Raisin Bran	1.8106	2.0129	0.6832	2.0447	0.0571	0.8302	1.58%
Kelloggs Raisin Bran Crunch	2.0854	2.3519	3.6594	2.3837	4.9573	0.0318	1.35%
Kelloggs Smart Start	2.4110	2.7768	0.7275	2.8068	0.8449	0.0381	1.15%
Kelloggs Special K Red Berries	2.7311	3.1891	0.0297	3.2209	0.0489	0.0318	1.00%
Brands in the Other Firms							
Private Label Raisin Bran	1.6682	1.7013	.0009	1.7013	0.0007	0.0000	0.00%
Post Raisin Bran	1.5183	1.7259	1.1395	1.7285	1.2766	0.0026	0.15%
Post Honey Bunches of Oats	2.0031	2.2183	0.7890	2.2209	0.8760	0.0026	0.12%
Segment	2.1527	2.2794	0.1024	2.3137	0.1200		
Kolmogorov-Shmirnov p-value		0.0703		0.0959			

Simulation 2: Limited Information with Endogenous Advertising

	Observed	Simulated Price					
	Price	Without Merger	MISE	With Merger	MISE	Increase Due to Merger	Pct Increase
Brands of the Merging Firms							
Kashi GoLean Crunch!	2.7191	2.6009	0.0968	2.6634	0.0426	0.0625	2.40%
Kelloggs Raisin Bran	1.8106	1.7184	0.2021	1.7310	0.1745	0.0126	0.74%
Kelloggs Raisin Bran Crunch	2.0854	2.0843	0.5256	2.0973	0.3945	0.0130	0.62%
Kelloggs Smart Start	2.4110	2.5114	0.2188	2.5241	0.1774	0.0127	0.51%
Kelloggs Special K Red Berries	2.7311	2.9133	0.0066	2.9257	0.0056	0.0132	0.42%
Brands in the Other Firms							
Private Label Raisin Bran	1.6682	1.6490	0.0192	1.6491	0.0191	0.0001	0.00%
Post Raisin Bran	1.5183	1.5726	0.5893	1.5724	0.6010	-0.0002	-0.01%
Post Honey Bunches of Oats	2.0031	2.0668	0.6338	2.0066	0.6425	-0.0002	-0.01%
Segment	2.1527	2.0825	0.0228	2.0969	0.0195		
Kolmogorov-Shmirnov p-value		0.7479		0.8452			

Source: Author's Calculations

Prices are for a 16oz package of cereal.

Table 8: Merger Impact Analysis on Advertising in Chicago DMA

Micro BLP with Endogenous Advertising Stock

	Percentage Change in Average Advertising GRP	Percentage Change in Strictly Positive Advertising	Advertising Cost per Unit
Brands of the Merging Firms			
Kashi GoLean Crunch!	100.60%	0.00%	0.0276
Kelloggs Raisin Bran	2.88%	0.00%	0.3361
Kelloggs Raisin Bran Crunch	-0.81%	0.00%	0.3361
Kelloggs Smart Start	1.99%	0.00%	0.3361
Kelloggs Special K Red Berries	-0.33%	0.00%	0.3361
Brands in the Other Firms			
Private Label Raisin Bran	-	-	-
Post Raisin Bran	20.66%	0.00%	0.0155
Post Honey Bunches of Oats	45.54%	0.00%	0.0155

Limited Information with Endogenous Advertising Stock

	Percentage Change in Average Advertising GRP	Percentage Change in Strictly Positive Advertising	Advertising Cost per Unit
Brands of the Merging Firms			
Kashi GoLean Crunch!	-92.92%	3.23%	0.0276
Kelloggs Raisin Bran	-60.32%	-58.33%	0.3361
Kelloggs Raisin Bran Crunch	-64.19%	-33.33%	0.3361
Kelloggs Smart Start	-79.64%	-11.11%	0.3361
Kelloggs Special K Red Berries	-100.00%	-100.00%	0.3361
Brands in the Other Firms			
Private Label Raisin Bran	-	-	-
Post Raisin Bran	-32.43%	0.00%	0.0155
Post Honey Bunches of Oats	-10.28%	-2.78%	0.0155

Source: Author's Calculations

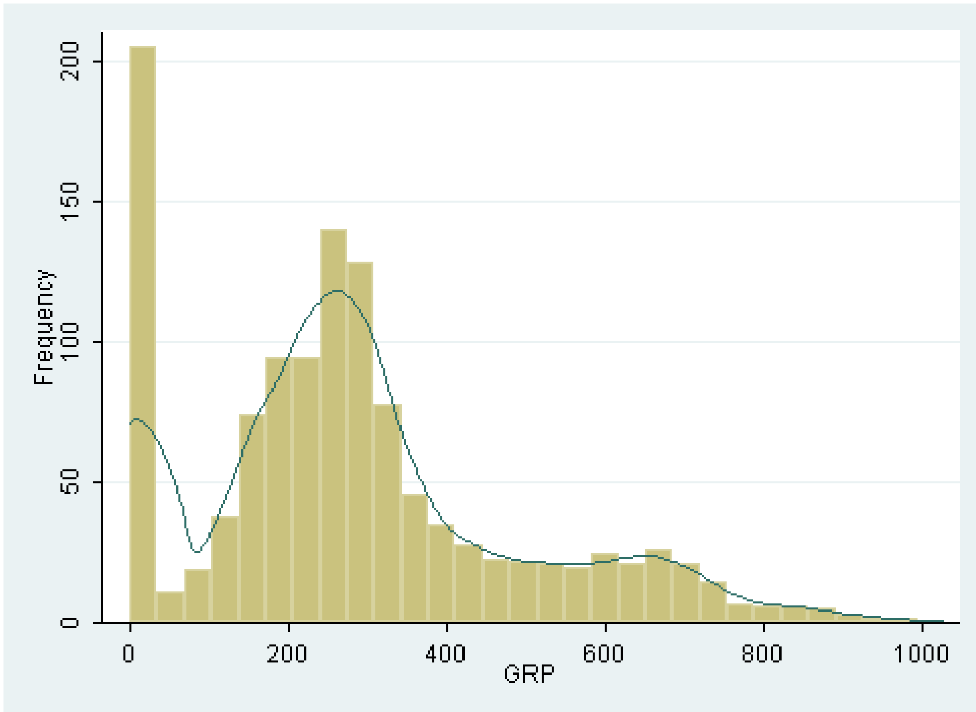

Figure 1: Frequency Distribution of Strictly Positive GRP levels

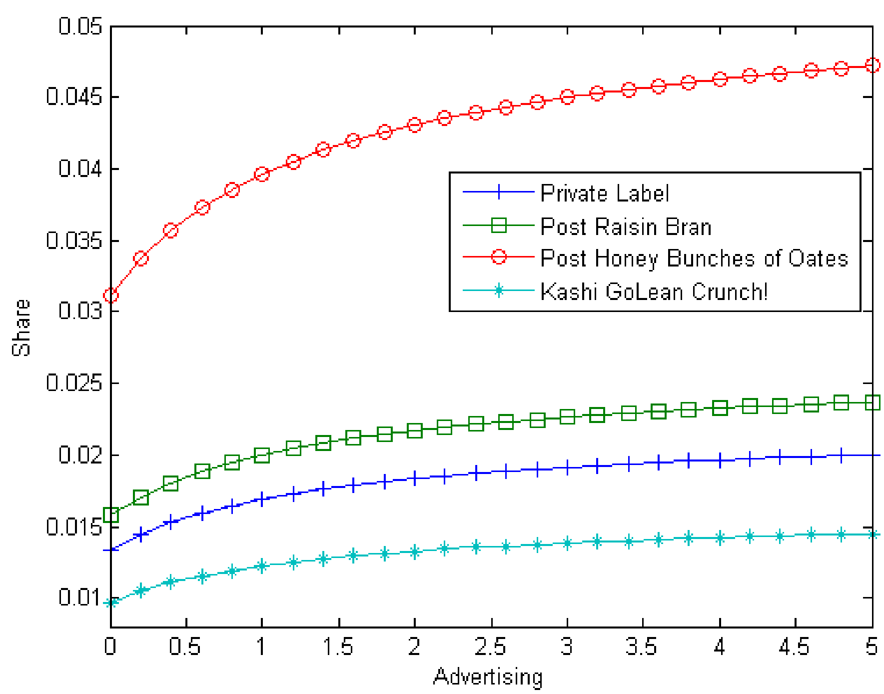

Figure 2: The Relationship between Share and Advertising: Private Label, Post, and Kashi

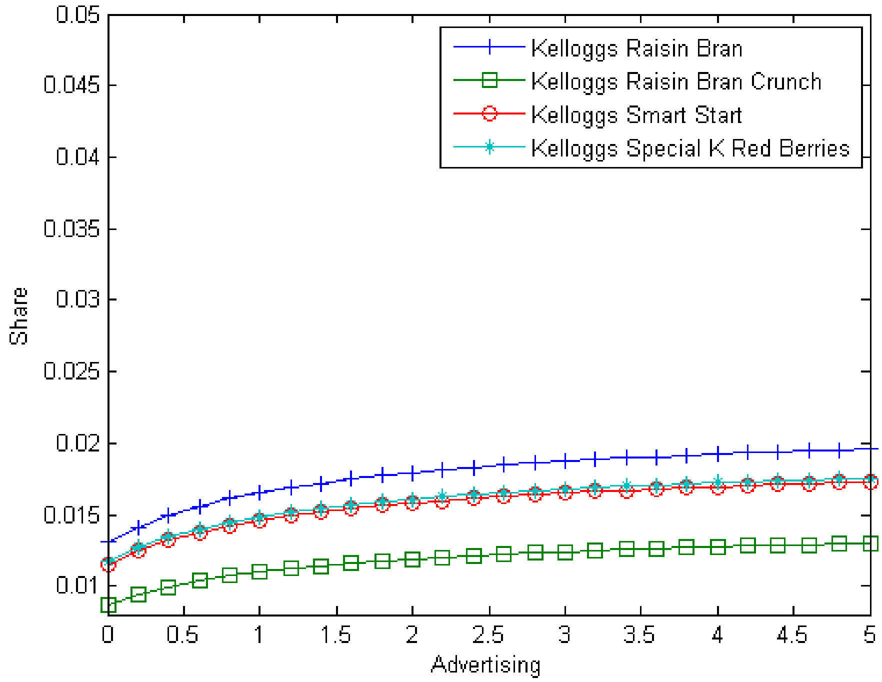

Figure 3: The Relationship between Share and Advertising: Kellogg's

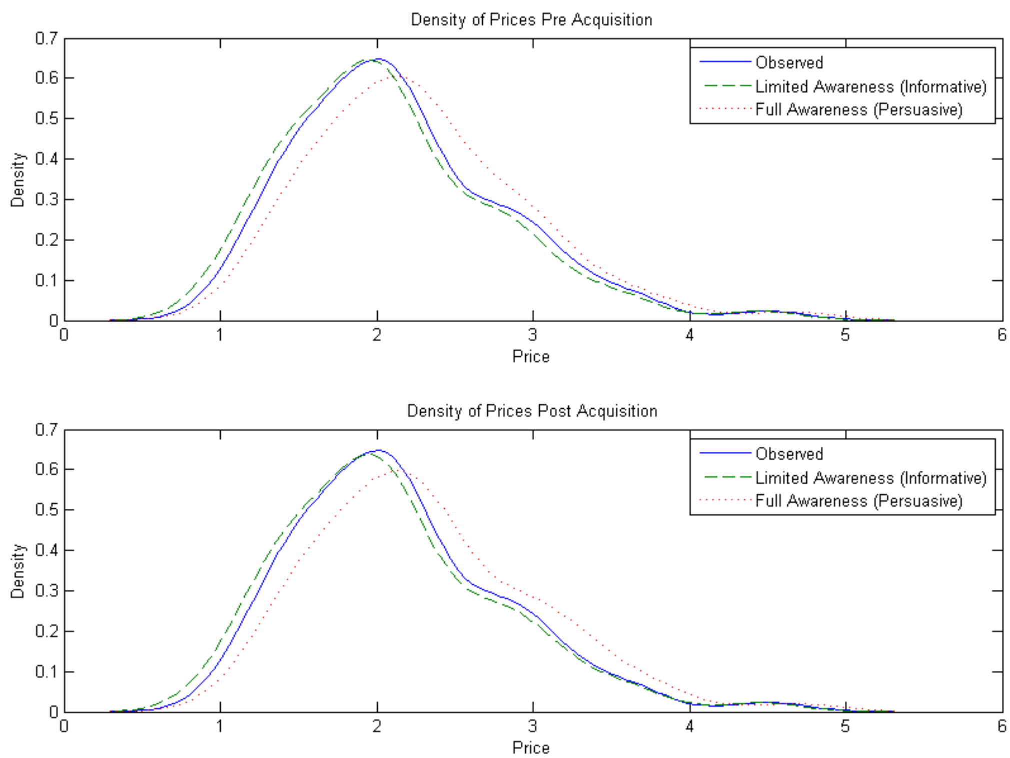

Figure 4: Density of Observed and Simulated Prices