Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market

MATTHEW SHUM
Johns Hopkins University
Baltimore, MD 21218
mshum@jhu.edu

In differentiated product markets where consumer preferences are characterized by brand loyalty, an important role for advertising may be to overcome brand loyalty by encouraging consumers to switch to less familiar brands. Using a scanner panel dataset of breakfast-cereal purchases, I find evidence consistent with the hypothesis that advertising counteracts the tendencies of brand loyalty toward repeat purchasing. Equivalently, advertising reduces switching costs in this market. Furthermore, counterfactual experiments demonstrate that in markets with brand loyalty, advertising is an attractive and effective option—relative to alternative promotional activities, such as price discounts—of stimulating demand for a brand.

1. Introduction

Brand loyalty in consumer preferences can be a significant source of incumbent advantage in many differentiated product markets because it builds up switching costs, which makes consumers reluctant to try new brands. In these markets, a potentially important role for advertising may be to counteract the tendencies of brand loyalty by encouraging consumers to "switch" to newer, less familiar brands.

This consideration goes against the grain of an argument dating back to Bain (1956) and Comanor and Wilson (1974) (hereafter B/CW) that advertising fosters perceived product differentiation among otherwise very similar brands. For example, Bain (1956) wrote that "product

I am grateful to James Lattin (Stanford GSB) and Ronald Cotterill (University of Connecticut Food Marketing Policy Center) for providing me with the data used in this paper. I thank the editor, Dan Spulber, as well as a coeditor and two referees, for their careful reading and suggestions on previous drafts. I also thank Anand Bodapati, Tim Bresnahan, Andrea Coscelli, Greg Crawford, Nancy Gallini, Amil Petrin, Rob Porter, Peter Reiss, Frank Wolak, and participants in the NBER Summer Institute 1999 for their comments and advice.

Journal of Economics & Management Strategy, Volume 13, Number 2, Summer 2004, 241–272

differentiation is propagated by [. . .] advertising and sales-promotional efforts designed to win the allegiance and custom of the potential buyer" (p. 114).

But Shapiro (1982) countered that

> [CW] correctly identify brand loyalty [. . .] as the critical issue [. . .] in markets of the type they are looking at, i.e., where advertising is important. The problem arises when they try to attribute brand loyalty to advertising expenditures alone. A very different conclusion emerges if brand loyalty is attributed instead to . . . consumer experience, with advertising being a method of *overcoming* brand loyalty (p. 6).

In this paper I evaluate the validity of Shapiro's (1982) argument for the breakfast-cereals industry by investigating whether advertising reduces brand loyalty in this market. This market appears, at first glance, to be a textbook case of the B/CW argument. The advertising intensity of the breakfast-cereals market is extraordinarily high: The advertising-to-sales ratio for the Grain Mills Products industry [Standard industrial classification (SIC) 2040, the bulk of which is cereals] is about 1.2 times the average value for the food sector and is about 3.5 times higher than the average value for all industrial sectors.[1] At the same time, there are a very large number of brands of cereals available at any one time (218 distinct brands appear in my dataset). These two characteristics of the industry—high advertising intensity, and substantial product differentiation—by themselves tend to justify the traditional B/CW arguments that advertising sustains perceived product differentiation among the competing brands.

One potential explanation for these trends is provided in the literature on advertising's role in informing consumers, either directly (cf. Butters, 1977; Grossman and Shapiro, 1984; Stigler, 1964) or indirectly (via "signaling"; cf. Kihlstrom and Riordan, 1984; Milgrom and Roberts, 1986; Nelson, 1974) about brand attributes and/or prices. These models, however, may have a difficult time explaining cereal advertising, since it is the well-established brands such as General Mills' Cheerios or Kellogg's Frosted Flakes that are the most advertised. Motivated by Shapiro's remark, I consider an alternative interpretation of these advertising trends: If consumers' preferences are characterized by *brand loyalty*, advertising—even for as well-known a brand as Cheerios—must

1. These figures are derived from the *Advertising Ratios and Budgets* publication of Schonfeld and Associates.

be maintained in order to convince consumers loyal to rival brands to switch.[2]

The concept of brand loyalty is connected closely to that of switching costs so that the question of whether advertising overcomes brand loyalty is analogous to one regarding whether advertising reduces switching costs. Therefore, this paper complements a large theoretical literature on the competitive effects of switching costs in oligopolistic industries (see, for example, Klemperer 1984, 1985), as well as a more recent empirical literature measuring the extent of switching costs in specific markets [see Chen and Hitt (2001) for a study of online brokerages, and Stango (2002) for an examination of credit card markets].

In this paper, I employ a scanner panel dataset to estimate household-level cereal brand-choice models in which advertising's effects depend on a household's brand loyalty, as measured by its recent purchases of particular brands. The results enable me to quantify how much advertising for a given brand reduces the (implicit) switching costs that households incur in trying a brand they have not purchased recently.

In the next section I describe some of the existing literature on this subject and introduce my dataset. In Section 3, I present the cereal brand-choice model used in this analysis and discuss important specification and identification issues. Section 4 contains the estimation results and discussion thereof. Section 5 reports results from counterfactual experiments that examine the market-level implications of my findings. Section 6 concludes.

2. Background: Existing Literature and Data Description

While brand loyalty is a standard component of many brand-choice models [especially in the empirical marketing literature, cf. the seminal paper by Guadagni and Little (1983)], I believe that this paper is the first to examine its market-level consequences, as well as the implications of advertising that potentially reduce brand loyalty. Several recent papers explore the degree of market power and product differentiation in the breakfast-cereals market (cf. Cotterill and Haller, 1997; Hausman,

2. Indeed, the *New York Times* has reported a remark by a marketing director for Philip Morris' Polish operations that "I can do much more about switching of brand (*sic*) [than whether a person smokes or not]. . ." (p. A1). While the physically addictive nature of cigarettes potentially intensifies brand loyalty in that market in a manner different from the cereals market, this statement illustrates the importance that marketers attach to advertising's role in promoting switching behavior.

1996; Kiser, 1998; Nevo, 2001). These papers do not focus on the role of advertising in the breakfast cereals market, or on its effects on the dynamics of purchase decisions.

The closest antecedent to my work is Ackerberg (2001), who reports strong evidence of advertising's disproportionately larger effects on households who previously have not purchased a brand. He uses a scanner dataset for the yogurt market, which includes detailed advertising exposure data gleaned from television viewing logs. Furthermore, by focusing on the case of the entry of a new brand of yogurt, Ackerberg can attribute the differential effects of advertising between households who have and have not purchased recently the brand to informational effects. Since such an interpretation would not be as appropriate for the cereals market, where the established brands maintain high advertising levels, I avoid all informational interpretation of the findings in this paper.[3] However, the essential empirical design in this paper is the same as Ackerberg's paper: Households are separated into two groups, depending on whether or not they have purchased recently a given brand. The effect of advertising on a household's purchases of this brand are assumed to differ across these two groups, where the hypothesis of interest—that advertising overcomes brand loyalty—implies that advertising should have larger effects on the purchases of the households who have not purchased recently the brand.

2.1 Data Description

I employ a detailed household-level scanner dataset [collected by Information Resources, Inc. (IRI)], which tracks the cereal purchases of 1,010 households in six supermarkets in the Chicago metropolitan area on a weekly basis from June 1991 to December 1992.[4]

As any even casual observer of the cereal markets is aware, there are a large number of breakfast-cereal brands available to the consumer: In my dataset, around 110 brands of cereals were found in the supermarket cereal aisle during the average week and, for each of these brands, at least two box sizes typically were available. Due to the large number of brands, I aggregate up to the top 50 brands in my dataset

3. See also Deighton, Henderson, and Neslin (1994). The question of *how* advertising encourages switching is beyond the scope of this paper. Psychological hypotheses that posit that advertising "cues" or reminds consumers of past experiences may be plausible in this market, especially given the large number of competing cereal brands. Alba, Hutchinson, and Lynch (1991) contains a general description of this literature.

4. This data were recorded by scanners at the supermarket checkout counters and constitutes a (small) part of a unique multicategory market basket database in the Stanford Business School. See Bell and Lattin (1998) for additional details.

and combine all the other brands into a composite 51st brand.[5] Since my focus is on brand choice, I aggregate across all different box sizes in defining each brand so that I do not distinguish between, for example, 12-ounce. and 20-ounce. boxes of Shredded Wheat. On the other hand, "umbrella extensions" of a brand name are classified as distinct brands so that, for example, Cheerios and Honey-Nut Cheerios are classified as two distinct brands. This aggregation procedure resembles that used in Hausman (1996) and Nevo (2001).

Table I presents summary characteristics for the 51 brands of cereals used in my analysis. The top-50 brands accounted for just over 75% of all purchases in my data. Comparing columns 5 and 6 of Table I shows that in-sample market shares (calculated from purchases in my dataset) are very close to national market shares. I classify the top-50 brands into family (13 brands), adult (25 brands), and kids' (12 brands) segments, taking as a guide the classification scheme in Hausman (1996).

Column 6 shows that more than half of the top 50 brands existed prior to 1983. Five entered between 1983 and 1988, and 13 after 1988. Furthermore, IRI's 1995 *Marketing Fact Book* indicates that 49 of the top-50 brands (the sole exception being Quaker's Popeye brand) still existed in 1995. Therefore, while there is substantial product entry into and exit out of the cereals market,[6] the set of top-selling cereals has remained quite stable over long periods of time. The third column in Table I summarizes the average transaction prices for each brand.[7] A given brand's price varies across both stores and weeks.

The main advertising data employed in this analysis are quarterly aggregate (i.e., national) brand-level advertising expenditures data from leading national advertisers (LNA). Column 5 of Table I presents across-time averages of quarterly advertising expenditures for each brand. The most highly advertised brands are well-established brands, such as Cheerios and Frosted Flakes. Family cereals are advertised the most but are cheaper than both adult and kids' cereals. While the advertising numbers include expenditures on 10 media, almost all of the advertising dollars were spent on broadcast television advertising, including both national and local television.

5. Since stores vary in the nontop-50 brands they carry, the composition of this 51st brand therefore varies across households as well as over time, depending on in which stores the households choose to shop.

6. Hausman (1996) notes that from 1980 to 1992, approximately 190 new brands were introduced on a basis of about 160 existing brands.

7. On average, the transactions prices (which are computed net of consumers' coupon savings) are markedly lower than the shelf prices, on an order exceeding 10%. While most of the results I present here were obtained using transactions prices, I also have estimated the model using shelf prices to gauge the robustness of the results.

TABLE I.
BRAND CHARACTERISTICS

	Name	Average Transaction Price ($/lb)	Average Ad Expense[c]	National Market Share[d]	Sample Market Share[e]	Exist in 1983? In 1988?[f]	Feature[g]	Display[h]
1	KG[a] Corn Flakes (F[b])	1.81	7.109	5.1	5.67	xx	0.040	0.058
2	GM[a] Cheerios (F)	3.16	7.287	4.8	4.38	xx	0.038	0.043
3	KG Rice Krispies (F)	2.96	6.034	3.8	4.04	xx	0.039	0.057
4	KG Frosted Flakes (F)	2.52	7.867	4.5	3.82	xx	0.019	0.023
5	KG Raisin Bran (F)	2.34	5.591	3.2	2.73	xx	0.034	0.045
6	GM Total (A[b])	3.61	3.926	1.8	2.36	xx	0.017	0.015
7	GM HoneyNut Cheerios (F)	3.14	4.030	2.7	2.26	xx	0.016	0.017
8	KG Special K (A)	3.48	3.531	1.3	2.16	xx	0.000	0.000
9	PT[a] Grape Nuts (A)	2.14	6.740	2.9	2.12	xx	0.001	0.001
10	NB SpoonSize ShdWt (A)	2.81	0.025	1.2	2.08	xx	0.035	0.002
11	QK[a] 100% Natural (A)	2.24	1.612	1.0	1.96	ox	0.011	0.009
12	KG Frosted Mini Wheats (A)	2.62	6.106	2.8	1.84	xx	0.015	0.009
13	KG NutriGrain (A)	2.87	2.508	0.8	1.55	xx	0.012	0.011
14	KG Mueslix (A)	3.31	1.975	0.8	1.53	oo	0.017	0.016
15	GM Wheaties (F)	2.55	2.257	1.4	1.52	xx	0.056	0.157
16	PT Raisin Bran (F)	2.23	4.361	1.9	1.46	xx	0.041	0.062
17	RL[a] Muesli (A)	3.34	0.215	0.4	1.26	oo	0.019	0.016
18	KG Corn Pops (F)	3.51	3.198	1.0	1.46	xx	0.023	0.019
19	GM Raisin Nut Bran (A)	2.98	1.659	1.1	1.35	ox	0.011	0.007
20	GM Basic 4 (A)	3.27	2.510	0.8	1.31	oo	0.022	0.019

21	GM Cocoa Puffs (K[b])	3.46	2.097	0.6	1.28	xx	0.012	0.006
22	GM Golden Grahams (K)	3.24	2.953	1.0	1.24	xx	0.029	0.032
23	GM CinnToast Crunch (K)	3.36	2.963	1.2	1.23	ox	0.024	0.026
24	KG Froot Loops (K)	3.53	3.110	1.9	1.20	xx	0.024	0.022
25	KG Low Fat Granola (A)	2.68	2.327	0.9	1.17	oo	0.024	0.053
26	GM Trix (K)	3.96	3.236	1.2	1.13	xx	0.023	0.027
27	GM Triples (A)	2.33	3.036	0.9	1.12	oo	0.032	0.045
28	KG Crispix (A)	3.28	3.225	1.2	1.12	xx	0.019	0.033
29	GM Kix (K)	3.67	3.801	1.2	1.08	xx	0.034	0.009
30	GM Lucky Charms (K)	3.45	3.079	1.4	1.08	xx	0.021	0.013
31	GM AppleCinn Cheerios (F)	3.02	3.120	NL	1.06	oo	0.014	0.000
32	KG Cracklin Oat Bran (A)	3.19	2.279	0.9	1.06	xx	0.008	0.007
33	NB Big Biscuit ShdWt (A)	2.79	0.000	0.8	0.99	oo	0.025	0.013
34	PT Honey Bunches of Oats (A)	2.85	3.749	1.1	0.95	oo	0.040	0.034
35	PT Great Graines (A)	2.90	2.648	0.5	0.89	oo	0.030	0.026
36	GM Otml Raisin Crisp (A)	2.71	1.641	0.5	0.97	ox	0.012	0.004
37	QK Oat Squares (A)	2.43	1.472	0.8	0.94	ox	0.012	0.015
38	RL Rice Chex (A)	3.40	0.875	0.8	0.89	xx	0.047	0.080
39	GM Total Raisin Bran (A)	3.00	1.874	0.4	0.89	ox	0.026	0.010
40	KG Product 19 (A)	3.38	1.408	0.4	0.89	xx	0.016	0.010

(Continued)

Table I.
Continued

	Name	Average Transaction Price ($/lb)	Average Ad Expense[c]	National Market Share[d]	Sample Market Share[e]	Exist in 1983? In 1988?[f]	Feature[g]	Display[h]
41	KG Apple Jacks (K)	3.64	1.465	0.7	0.84	xx	0.030	0.023
42	QK Capt Crunch (K)	2.55	1.714	1.8	0.83	xx	0.052	0.066
43	NB Shredded Wheat (A)	2.82	2.925	0.5	0.80	xx	0.021	0.000
44	PT Fruity Pebbles (K)	3.21	1.710	0.8	0.83	xx	0.040	0.040
45	GM Clusters (F)	3.14	1.425	0.9	0.78	ox	0.019	0.002
46	KG Cinnamon MiniBuns (F)	2.75	0.002	0.8	0.76	oo	0.029	0.043
47	KG Double Dip Crunch (A)	3.01	1.454	0.6	0.73	oo	0.022	0.034
48	GM MultiGrain Cheerios (F)	3.34	2.520	NL	0.75	oo	0.039	0.018
49	PT Honeycomb (K)	3.40	2.567	0.7	0.74	xx	0.022	0.023
50	QK Popeye (K)	1.77	0.000	NL	0.67	oo	0.001	0.004
51	Basket of All Other Brands	2.68	0.645	34.3	24.29		0.441	0.627

[a] KG: Kelloggs; GM: General Mills; PT: Post (Phillip Morris); RL: Ralston; QK: Quaker Oats.
[b] F: family segment; A: adult segment; K: kids' segment.
[c] Quarterly expns, $mill. **Source:** Leading National Advertisers (1990–1993). Avg. 1991:ii—1993:ii.
[d] Shares of national cereal volume, 1992. **Source:** IRI. NL: Brand not listed in source.
[e] Share of total in-sample purchases. Source: Author's calculations.
[f] xx: exist in both years; ox: exist in 1988, not in 1983; oo: exist in neither year. Source: IRI.
[g] =1 if brand was featured in newspaper ad during a given (store-week); avg. over 823 store weeks.
[h] =1 if brand had feature display in store during a given (store-week); avg. over 823 store weeks.
[i] Sum of average quarterly advertising expenditure for all the nontop-50 brands.

TABLE II.
HOUSEHOLD DATA: SUMMARY STATISTICS

Variable	Definition	N	Mean	StdDev	Min	Max
YGCHIL	=1 if kids under age 18 present[a]	1,010	0.29	0.45	0	1
FAMSIZE	Number of persons in household[b]	1,010	2.61	1.37	1	6
LOGINC	Log(family income/10,000)	1,010	0.97	0.81	−0.69	2.08

[a]The data do not allow for a distinction between households with only very young children and those with only teenagers.
[b] Top-coded at six.

While the brand-level quarterly aggregate LNA data is the best advertising data that generally is available for the breakfast-cereals market,[8] its main drawback is the lack of household-level variation in advertising exposure. Since these data constraints naturally raise questions regarding identification of model parameters and robustness of the estimates, I address both issues in greater detail below.[9]

Columns 8 and 9 of Table I contain sample averages of store-promotional variables related to each brand. These are dummy variables (1) DISPLAY, which is equal to one if a given brand was promoted via a display [10]; and (2) FEATURE, which is equal to one if a brand was featured in newspaper circulars for the store. Both of these covariates vary across stores, brands, and weeks; furthermore, since different households shop at different stores, they also vary across households.

Table II contains definitions and summary statistics of the household-level covariates employed in this analysis. Since the focus of this paper is on brand choice, I abstract away from frequency of shopping issues by aggregating all shopping activity up to a weekly basis and model households' weekly brand choices conditional on their making a shopping trip.[11] Using this aggregation scheme, an average

8. To my knowledge, however, there is no scanner dataset that includes such individual-level advertising data [similar to the data employed in Ackerberg (2002)] for the cereals market. In a recent study on the cereals market, Cotterill and Haller (1997) used *weekly* aggregate advertising data, but this data are not available generally to researchers.

9. Furthermore, in using advertising dollars as a measure of advertising exposure, I am assuming that the "price" of a message is the same over all brands and over all time. Aggregation to a national level, as well as the fact that most cereals primarily employ broadcast advertising, justifies this assumption. Most variation in broadcast advertising costs is over the different times of the day (the "dayparts").

10. IRI distinguishes among several types of displays: lobby, aisle (front, end, back), and specialty/shipper, but I do not distinguish among different types of displays in constructing the DISPLAY variable.

11. The sample selection involved in ignoring-weeks where households do no shopping introduces bias into the parameter estimates of my brand-choice model only if, *conditional on the included covariates*, there is systematic correlation between the frequency of purchase and the brands bought—if frequent shoppers always tend to buy certain brands.

household made a shopping trip during 39.14 out of the 52 weeks in the sample and purchased cereal during 10.36 weeks. There are weeks in the sample in which a household makes multiple purchases, i.e., purchases of more than a single brand. In the empirical specifications, I model each of these within-week purchases as independent events.[12]

Comparison of the demographics and cereal-purchase frequencies from the scanner data used in this study with the *Consumer Expenditure Survey* (a representative consumption survey sample of US households collected by the Bureau of Labor Statistics) shows that the households in my dataset tend to be older than the average US household.[13] Furthermore, a comparison of cereal purchases suggests that there might be a substantial incidence of missing purchases in the scanner dataset (arising from households purchasing cereal at nonscanner equipped stores). While these problems do not affect the interpretation of the individual-level results, they are controlled for in the aggregate-level counterfactual experiments by introducing appropriate household weights, which calibrate each sample household according to its representativeness in the US population.

I measure brand loyalty by households' past purchases. Specifically, I construct an indicator variable PASTUSE_{iht}, which is equal to one if household h bought brand i in the w weeks preceding purchase occasion t.[14] The empirical results reported in this paper were obtained

---

Since the covariates include family size, I believe I have controlled for the most potentially important source of this type of correlation: households with kids who purchase cereal most often and whose purchases tend toward kids' cereals.

12. Households purchased more than one brand in roughly one-quarter of all (household-weeks) in which purchase occurred. By ignoring the multiple purchase dimension, I may be abstracting away from important across-brand synergies that may characterize preferences in this market. However, some auxiliary calculations failed to find signs of across-brand synergies in demand patterns that would require modeling the multipurchase decision. In particular, I looked for signs of a likely type of demand synergy: that families with children may tend to buy "bundles" of adult and kids' cereals versus one single brand. I found no evidence of this. See Hendel (1999) and Dube (1999) for an extension and applications of the discrete-choice demand framework to handle multiple choices.

13. Apparently IRI included more older households in their sample because they sought households they deemed to have a low probability of moving out of the sample area. I thank Anand Bodapati for this insight.

14. An alternative approach would be to allow loyalty to be a "stock" that fades over time [as in many marketing brand-choice models, starting with Guadagni and Little (1983)]. However, initial conditions pose nontrivial problems in these models, as does determining the appropriate "decay rate" for the experience stock. Moreover, the possibility of unobserved purchases (as discussed previously) leads to measurement error in the PASTUSE variable. In earlier versions of this paper, I included results from specifications that controlled for this measurement error by explicitly parameterizing the probability of unobserved purchases as a function of household and brand characteristics. Since the model became quite cumbersome computationally without yielding noticeably different results, I have not included them here.

TABLE III.
SWITCH PROPENSITIES IN THE RAW DATASET[a]

w	%SWITCH	NumObs
2	82.7	22363
4	70.8	21946
8	57.6	21136
12	50.5	20200
24	38.5	17560

[a]Computed using all purchases of a top-50 brand in my sample.

from models that assume $w = 12$.[15] In Table III, I compute how often households in my dataset "switched" to brands they had not purchased recently. The "%SWITCH" column gives the percentage of observed purchases that represents purchases of brands a household did not purchase in the most recent w weeks, for different values of w. For $w = 12$, households switched 50.5% of the time.

The dataset employed in my analysis includes the observations from January–December 1992, totaling 44.224 observations. The observations from June–December 1991 were used to initialize the loyalty variable PASTUSE.

As a first glimpse of the main empirical exercise undertaken in this paper—quantifying the differences in advertising's effects across households that are and are not loyal to a given brand—I aggregate the purchases of each brand across all households up to the quarterly level (the same frequency as the advertising data) and examine the across-(brand/quarter) correlations between advertising levels and the proportion of purchases of each brand undertaken by households who are not "loyal" to it (where loyalty is measured with the PASTUSE variable, for both $w = 12$ and $w = 4$). As Table IV shows, results from linear regressions (both with and without brand fixed effects to control for brand-specific unobservables) indicate positive correlation for both $w = 12$ and $w = 4$. However, not all of these effects are statistically significant (the 12-week effect without brand-fixed effects is significant at the 13% level, and the four-week effect with brand effects is only significant at a 20% level). Moreover, while this positive correlation is consistent with the hypothesis that advertising encourages switching behavior, one cannot adopt this causal interpretation based on these correlations alone. For these reasons, I next examine results from brand choice models.

15. I also estimated a specification assuming $w = 4$ as a robustness check. Since the results did not change much relative to the reported results, I do not consider them here.

Table IV.
Advertising and Switching: Correlations from Raw Dataset[a]

Loyalty Measure	12 Weeks	4 Weeks
Without Brand-Fixed Effects	0.0068	0.0154
	(0.0045)	(0.0040)
With Brand-Fixed Effects[b]	0.01524	0.0072
	(0.0072)	(0.0057)

[a]Slope coefficients from a linear regression of $\frac{NOTLOY_{it}}{TOTPURC_{it}}$ on $NADV_{it}$. $NOTLOY_{it}$: quarter t purchases of brand i by non-loyal households; $TOTPURC_{it}$: total quarter t purchases of brand i; and $NADV_{it}$: quarter t national advertising for brand i. Loyalty is measured over both a 12- and a 4-week period. Standard errors in parentheses. Computed using all purchases of a top-50 brand in my sample.
[b]Excluded brand is brand 50.

3. Empirical Model

Following much of the existing empirical literature on brand choice, I derive the expressions for the purchase probabilities from a discrete-choice model of household-level brand choice. On each purchase occasion, household h chooses among $I = 50$ cereal brands. In addition, it can choose to purchase no brand of cereal at all: This is denoted option "0". Household h chooses the alternative $i (i = 0, \ldots, 50)$, which provides the highest indirect utility:

$$\max_{i \in [0,\ldots,50]} U_{iht}. \tag{1}$$

Taking a random-utility approach, I assume that the utility $U_{iht} = V_{iht} + \eta_{iht}$, where V_{iht} is a deterministic component and η_{iht} a random component of utility. The latter is observed by households when they make their choices but is unobserved by the econometrician.[16]

In order to accommodate brand loyalty in household preferences, I allow both V_{iht} and η_{iht} to depend on whether household h has purchased

16. The choice model specified here is myopic, because I do not consider the possibility that households are cognizant that by choosing a brand today, they become loyal to it in the future. Such a model would resemble the rational addiction model of Becker and Murphy (1988). Ackerberg (2002) and Erdem and Keane (1996) have estimated dynamic models of a household's brand-choice decision allowing for households to learn gradually (in Bayesian fashion) about the quality of different brands. This dynamic programming approach would be infeasible computationally given the large number of brands in the cereals market (Both of the aforementioned studies modeled only consumers' choices between a small number of brands).

brand i prior to period t. I employ the following specification for V_{iht}:

$$\begin{aligned}
V_{iht} &= X_i'\beta_0 + \delta_{\mathbf{1}} * \text{PASTUSE}_{iht} + (\alpha_1 + \delta_{\mathbf{2}} * \text{PASTUSE}_{iht} + \beta_3 * X_h) * p_{it} \\
&\quad + (\alpha_2 + \delta_{\mathbf{3}} * \text{PASTUSE}_{iht} + \beta_4 * X_h) * \text{ADV}_{it} \\
&\quad + (\alpha_3 + \delta_{\mathbf{4}} * \text{PASTUSE}_{iht}) * \text{ADV}_{it} * p_{it} \\
&\quad + X_{ih}'\beta_1 + X_{2ht}'\beta_2, \qquad\qquad i = 1, \ldots, 50, \ \forall t \\
V_{51ht} &= \alpha_{51} + (\alpha_1 + \lambda_1) * p_{51ht} + (\alpha_2 + \lambda_2) * \text{ADV}_{51ht} \\
&\quad + (\alpha_3 + \lambda_3) * p_{51ht} * \text{ADV}_{51ht} + \lambda_4 * \log N_{51ht}, \qquad \forall t \\
V_{0ht} &= 0, \qquad \forall t.
\end{aligned} \tag{2}$$

For brands $i = 1, \ldots, 50$, I allow brand loyalty (via PASTUSE) to affect utility in several ways as measured by the δ parameters. δ_1 and δ_2 capture the difference in the intercept and slope (with respect to price) of a brand's utility, while δ_3 and δ_4 capture the differential effects of advertising on the intercept and slope of the utilities for households loyal to a particular brand.[17] The four δ parameters, as well as the α parameters associated with price and advertising, are the focus of the empirical estimation.

The variable p_{it} denotes the price of brand i during purchase occasion t. In order to maintain parsimony in the specification, I assume that the two promotional variables affect brand choice behavior in a proportionate degree to national advertising (as measured by NADV_{it}, which denotes the national brand-level advertising expenditure variable summarized in column 4 of Table I). More precisely, I assumed that household h's exposure to brand i's advertising during week t, denoted ADV_{iht} in (3), is a linear combination of NADV_{it} and the two promotional variables:[18]

$$\text{ADV}_{iht} = \text{NADV}_{it} + \zeta_1 * \text{FEATURE}_{iht} + \zeta_2 * \text{DISPLAY}_{iht}. \tag{3}$$

X_i consists of brand dummies for each brand. X_{ih} consists of interactions of household-specific demographics FAMSIZE_h and LOGINC_h

17. This specification assumes that advertising only has a contemporaneous effect. However, there is a large amount of empirical evidence that the effects of advertising persist for a period after a consumer is exposed to the ad. I have run a version of Model A with advertising stock levels defined as in Stern (1996) (i.e., The advertising stock in period t equals the advertising flow in period t plus 0.8 times the advertising flow in the previous period), with very little changes in the results. This may not be surprising since the advertising data already is aggregated to a quarterly level and, indeed, since the evidence cited in Little (1979) appears to indicate a very small "impulse response" to an ad after three months (p. 642).

18. I also estimated specifications where the promotional variables entered independently and interacted with PASTUSE. Since the results qualitatively were very similar to the reported results, I do not discuss them further.

with brand-specific characteristics FAMSEG_i, ADULTSEG_i, KIDSSEG_i (These are the segment-specific dummy variables), SUG_i, and FAT_i. Furthermore, β_3 and β_4 are coefficients which capture the differential effects of price and advertising, respectively, on households with different demographic variables X_h. In the baseline specification, X_h consists of FAMSIZE_h and LOGINC_h.

X_{2ht} contains regressors specific to purchase occasion t. Since cereals are a "stocked" item in most households, it seems natural to assume that, at a weekly level, the utility a household derives from a cereal purchase depends (probably inversely) on the stock available. These purchase frequency dynamics are captured by the indicator variable PREVPURC_t, which is equal to 1 if a purchase of cereal occurred in week $t - 1$.[19]

In the expression for V_{51ht} in (2), p_{51ht} and ADV_{51ht} are, respectively, the store sales-weighted price and advertising averages for the brands included in the composite, and N_{51ht} is the number of brands included in the composite.[20] The parameters λ_1 to λ_4 are specific to the utility from the 51st brand. The first three of these parameters allow the coefficients α_1, α_2, and α_3, respectively, to differ for this composite brand, since the covariates included here are averages and therefore are different from the covariates used for the other brands. The λ_4 parameter captures the inherent differentiation between the brands in this composite.[21]

For the noncereal purchase utility, I set $V_{0ht} = 0$, across all households h and weeks t. This requires the assumption that the price of the outside good—which in principle includes all other commodities purchased during a shopping trip—stays constant over the sample period, so that $p_{0t} = 0, \forall t$. This assumption is justified by noting that there is very little monthly variance in the CPI during the relatively short time-span of my data: From June 1991 to June 1993, the Chicago-area nondurables CPI rose only about 7%.

19. An indicator of purchase in the previous week captures the dynamics of purchase frequency only rudimentarily; an attractive alternative would have been total *amount* of cereal purchased in the previous week. However, employing a binary variable such as PREVPURC facilitates simulating households' purchase histories, which is done in order to perform the counterfactual experiments detailed following. With a continuous variable, simulating purchase histories also would involve specifying a process for the *amount* purchased, which is not a component of the brand-choice model specified here.

20. This latter covariate is included in the specification to capture the attractiveness of this composite brand due simply to the number of brands included [this covariate's function is discussed further in McFadden (1978) and Ben-Akiva and Lerman (1985:254–260) .

21. I do not define loyalty for the composite 51st brand because it is not clear what an appropriate definition of loyalty with a basket of brands is.

3.1 Nested Logit Model

I consider nested logit discrete-choice models where brands a given household has and has not purchased recently are placed in separate nests. The likelihood function corresponding to any observed purchase is derived in the Appendix.[22] An important insight in Cardell (1997) is that the nested-logit model can be interpreted as a random-effects model where there are unobserved utility components common to the brands grouped in the same nest. Precisely, the nesting structure implies that the unobservable term η_{iht} has an error components structure consisting of an alternative-specific component that is independent over all alternatives in the choice set, as well as a nest-specific component that affects only the alternatives within a common nest.

For the case where brands are classified into nests depending on a household's past purchases of them, these common components can be attributed to unobserved heterogeneity in a household's preferences for brands not accommodated explicitly via the included covariates. Furthermore, these common components induce correlation among the random ηs for the brands a household has purchased recently, leading to closer substitutability between two brands a household recently has purchased.

3.2 Identification and Endogeneity Issues

The important parameters of the empirical exercise—the δ's and α's—are identified from variation in brand choices across time and households. Specifically, the price parameters (α_1 and α_3) in a panel discrete-choice setting are identified from the extent that households purchase brands during weeks when they are "expensive" relative to their competitors. Analogously, brand loyalty is associated with lower (higher) price sensitivity if households are more (less) likely to purchase brands they have purchased recently during weeks when they are relatively expensive. Similarly, advertising's interaction with brand loyalty is identified by variation in households' willingness to purchase highly advertised brands depending on having purchased those brands recently.

To this point, I have maintained an assumption that the unobservables η_{iht} are independent across h and t and are uncorrelated with the included covariates. This assumption may not hold due to

22. Several alternative specifications were tried but ultimately were rejected. This included models in which brands were nested on the basis of brand segment (i.e., FAMILY, ADULT, and KIDS nests), as well as nesting the nonpurchase option apart from all other alternatives.

the potential endogeneity of price and advertising, which has been a concern in existing applications of discrete-choice demand models (cf. Berry, 1994; Nevo, 2001; Stern, 1996). By including a complete set of 51 brand dummies, I control for unobserved brand-specific factors that may affect firms' pricing or advertising choices. However, given the panel nature of the data, it is possible that there are additional unobservables (varying over time) that simultaneously affect a brand's demand as well as its price or advertising. Since my measure of national advertising varies only across quarters and brands, it is difficult to introduce additional unobservables without adversely affecting the identification of the effects of advertising, which is the focus of this paper.

Therefore, I attempt to assess the extent of potential endogeneity problems by less formal means. An analysis of variance on the advertising data, controlling for both brand and quarter effects, showed that brand effects accounted for roughly 81% of the variation, while the quarter effects accounted for a negligible 1.5%. Therefore, most of the roughly 20% of the variation not accounted for by the brand effects—the variation that, given the inclusion of brand fixed effects, identifies the advertising coefficients in the models I estimate—seems due to variation in advertising expenditures across brands and over time. These conclusions are confirmed graphically in Figure 1, where I plot the four quarterly advertising expenditure observations for each of the brands.[23] The graphs illustrate clearly that there are no seasonal trends in advertising that appear systematic across brands so that unobserved seasonal trends are not a likely source of endogeneity. Nevertheless, the fact remains that, with only aggregate advertising data, one cannot disentangle the effects of advertising from unobserved brand-quarter effects that may affect firms' advertising decision.

3.3 Unobserved Heterogeneity

Allowing for unobserved heterogeneity is an important aspect of my analysis for two reasons. First, in panel discrete-choice models including lagged dependent variables (such as the PASTUSE variable employed here), a potentially serious inferential problem arises because the effects of brand loyalty (or state dependence) and unobserved individual-specific effects observationally are equivalent, in the sense that time-persistent effects observed by households but unobserved by the econometrician may lead to persistence in an agent's choices over time

23. I thank a referee for this suggestion.

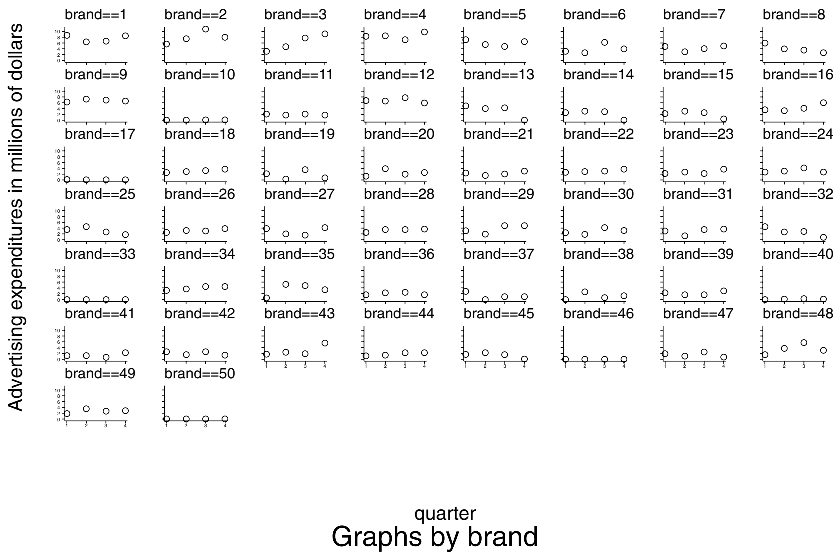

FIGURE 1. QUARTERLY ADVERTISING EXPENDITURES BY BRAND
Notes: *Y*-axis: advertising expenditures, in millions of dollars (each tick represents \$2 mills.); *x-axis*: quarter of 1991, with 1 = January–March, 2 = April–June, 3 = July–September, 4 = October–December.

that the econometrician may attribute spuriously to state dependence (see Heckman, 1981, 1991 for a discussion). Second, Moulton (1986) recognized that estimated standard errors may be biased when random-group effects are present in empirical models that pool aggregate and micro-level data. These insights are relevant to this analysis because I combine national advertising data with the household-level purchase dataset.

To address these two sets of issues, I accommodate unobserved heterogeneity via random effects in all of my empirical models. Both specifications include unobservables at the household, brand, and quarter levels that address both the spurious state-dependence issues raised by Heckman as well as the additional problems with pooling micro and macro data raised by Moulton. By assumption, the random effects are distributed independently of the covariates, including advertising. Thus, inclusion of these random effects does not control for the potential endogeneity problems discussed in the previous section. The specific details of the unobserved heterogeneity specifications, including a precise statement of the required stochastic assumptions, is given in Section A.2 of the Appendix.

4. Estimation Results

Results from two specifications of the brand choice model are given in Table V. In the following discussion, standard errors are enclosed in square brackets [· · ·].

Model A is a nested-logit specification with random effects at the household, brand, and quarter levels to control for unobserved heterogeneity. (Details of the random-effects specification are given in Section A.2 of the Appendix.) The coefficient on PASTUSE, δ_1, is large and is significant (2.08 [0.132]), suggesting that brand loyalty or, equivalently, switching costs are high in this market (a precise quantification of this will be given following). The advertising coefficient, α_2, is modest in magnitude but is marginally significant (0.048 [0.026]).

The estimate of δ_2 (0.470 [0.045]) indicates that households that recently have purchased a brand are less sensitive, *ceteris paribus*, to price. The coefficients on the interactions terms of PASTUSE with advertising are more problematic to interpret. The negative estimate of δ_4 (−0.06 [0.009]) implies that advertising, in conjunction with past use, increases households' price sensitivities. However, this negative effect is so large in magnitude that the net marginal effect of advertising on the purchase probabilities of loyal households (as given by $\alpha_2 + \delta_3 + (\alpha_3 + \delta_4) * \text{price}$) is negative. While this finding is unanticipated, one should keep in mind that the parameters measure the *marginal effects* of advertising at the household level: Indeed, I shall show that once we aggregate over time and across households, advertising has a positive effect on aggregate demand, despite the negative effects on loyal households implied by these results.

The coefficient on PREVPURC is negative (–0.30 [0.23]), which is consistent with a stockpiling story. The point estimate of the nesting parameter σ (for households without children) is 0.311 [0.214]. Finally, the standard deviation on the brand-quarter random effect, σ_c, is estimated imprecisely.

Table VI contains calculations that illustrate what these estimates imply about the extent of brand loyalty in this market.[24] This table demonstrates that households are much more likely to repurchase brands they have purchased recently: For example, recent purchase of a brand raises the median household's purchase probability from 0.26% to 5.66%, which is about a 20-fold increase. More telling are calculations

24. Since the quantities calculated in Tables VI and VII always do not have analytic solutions, a simulation approach was used to derive the standard errors. Specifically, the quantities were recalculated for a number of draws from the asymptotic distribution of the estimated parameters, and the reported standard errors are the standard deviations of these quantities over the simulated draws.

TABLE V.
MODEL ESTIMATES[a]

Variable	A: Estimate	A: StdErr	B: Estimate	B: StdErr
α_1: Price	−0.704	0.044	−0.712	0.044
α_2: ADV[b]	0.048	0.026	0.042	0.026
α_3: Price*ADV	−0.0009	0.0078	−0.0009	0.0077
δ_1: PASTUSE*constant	2.083	0.132	2.242	0.141
δ_2: PASTUSE*price	0.470	0.045	0.211	0.048
δ_3: PASTUSE*ADV	0.002	0.025	−0.151	0.027
δ_4: PASTUSE*price*ADV	−0.064	0.009	−0.011	0.010
ζ_1: [c] FEAT	−0.008	0.066	−0.085	0.071
ζ_2: DISP	0.191	0.087	0.272	0.107
PREVPURC	−0.029	0.023	−0.029	0.023
σ (Households without Young Children)	0.311	0.214	0.305	0.212
σ (Households with Young Children)	0.387	0.009	0.387	0.009
Unobs'd Heterogeneity Params:[d]				
First Component:				
π	0.302	0.017	0.303	0.037
θ^H	1.711	0.046	1.711	0.046
Second Component:				
μ_b (coef. on PASTUSE(t = 0))[e]			1.474	0.021
σ_{10} (Family Segment, No Young Children)	0.391	0.023	0.444	0.024
σ_{11} (Family Segment, Young Children)	0.269	0.035	0.435	0.031
σ_{20} (Adult Segment, No Young Children)	0.558	0.019	0.605	0.021
σ_{21} (Adult Segment, Young Children)	0.428	0.033	0.367	0.035
σ_{30} (Child Segment, No Young Children)	0.278	0.033	0.281	0.037
σ_{31} (Child Segment, Young Children)	0.066	0.047	0.077	0.042
Third Component:				
σ_d	0.089	0.557	0.012	0.008
Brand-Fixed Effects	yes		yes	
Household-brand random effects	yes		yes	
Brand-Quarter Random Effects	yes		yes	
Interactions with/LOGINC[f]	yes		yes	
Interactions with/FAMSIZE[g]	yes		yes	
Estimate $\lambda_1 - \lambda_4$[h]	yes		yes	
LogL:	−65948.24		−64783.93	
M (Number of simulated draws)	10		10	

[a]Unobserved heterogeneity via household-brand and household-quarter random effects (see Section A.2 for details).
[b]ADV is an index of national advertising and store-level promotional variables, cf. (3).
[c]See (3) for the definition of this parameter.
[d]For definition of parameters, see Section A.2.3 of the Appendix.
[e]For definition of this parameter, see (8).
[f]LOGINC interacted with price, advertising, segment dummies, and nutritional characteristics.
[g]FAMSIZE interacted with price, advertising, segment dummies, and nutritional characteristics.
[h]These are parameters specific to the utility from the 51st composite brand. See (2).

TABLE VI.
ADVERTISING AND BRAND LOYALTY: ADVERTISING REDUCES SWITCHING COSTS[a]

Using Model-A Estimates						
	At Observed Ad Levels		At 75% of Observed Ad Levels		⇒ Difference	
	Pt. est.	Std Error[b]	Pt. Est.	Std Error	Pt. Est.	Std Error
Summarized across All Brands						
$d_i \mid \text{EXP} = 1$	0.0566	0.0118				
$d_i \mid \text{EXP} = 0$	0.0026	0.0015				
Switching cost s_i	4.33	0.720	5.01	0.481	−0.68	0.466
Implied Switching Costs for 10 Selected Brands						
KG Corn Flakes	3.64	0.836	3.96	0.762	−0.31	0.229
Cheerios	3.66	1.041	4.12	0.839	−0.46	0.289
Rice Krispies	3.93	0.804	4.28	0.715	−0.35	0.189
Frosted Flakes	3.64	1.020	4.00	0.858	−0.36	0.254
KG Raisin Bran	3.87	0.620	4.35	0.678	−0.48	0.286
Grape Nuts	3.73	0.698	4.21	0.765	−0.48	0.289
Nutrigrain	4.14	0.786	4.97	0.482	−0.83	0.477
Froot Loops	4.54	0.874	5.32	0.568	−0.78	0.502
Lucky Charms	4.62	0.913	5.31	0.586	−0.69	0.539
Product 19	5.38	0.660	5.61	0.447	−0.23	0.395

[a]Using estimates from models A and C. d_i: purchase probability Switching cost: defined as monetary amount $\$s_i$ such that $d_i(p_i, \text{EXP} = 1) = d_i(p_i - s_i, \text{EXP} = 0)$. Purchase probabilities and implied switching costs evaluated at median values of the household characteristics and average prices and advertising for each brand.
[b]Computed using simulation (see footnote 31).

of the implicit switching costs that a household must be paid in order to purchase a given brand to which it is not loyal with the same probability as a household that is loyal to this brand. These costs (labeled s_i) also are reported in Table VI. The average switching cost across all brands is \$4.33: This implies a very strong effect of brand loyalty, since it is larger than the price of any brand.[25] Table VI also contains similar switching cost calculations for 10 major brands in my analysis.

In the right-most two columns of Table VI, I reduced the advertising levels for each brand by 25% and recomputed the switching costs the median household implicitly would incur by switching from a brand to which it is loyal to a brand to which it is not loyal. Comparing the figures in the right-most two columns to those in the first two columns, we see that a decrease in advertising levels raises switching costs. For example,

25. On the other hand, this is not inconsistent with the large number of coupons for "free samples" dispensed in this industry.

the average switching cost rises by \$0.68 [0.47] across all brands; this decrease is \$0.46 [0.29] for Cheerios. These results are consistent with Shapiro's idea that advertising reduces brand loyalty in this market.

One important assumption of the previous model is that the the random effects are distributed independently of all the included covariates. This is disturbing because one of the covariates is PASTUSE, a lagged purchase indicator, and it is plausible that the (household-brand) random effects may be correlated with the value of PASTUSE at the beginning of the estimation sample.[26] In *Model B*, I amend Model A by allowing the mean of the (household-brand) random effect to depend on PASTUSE_{ih0}. Section A.2.1 of the Appendix, contains details of this specification.

The estimated parameter μ_b, the coefficient on PASTUSE_{ih0} in the mean of the household-brand random effect, is large in magnitude (1.474 [0.021]), indicating a strong dependence (as one would expect) between a household's initial choices of brands and those for which it has a strong unobserved preference. Furthermore, the maximized log-likelihood function for this specification is markedly higher than that for Model A (–64,783.93 versus –65,948.24), indicating much better fit.

Furthermore, we also see that some of the estimates of the key δ parameters also change in this specification: δ_2, which measures the decrease in price sensitivity for loyal households, becomes smaller (0.211 [0.141]); δ_4, which captures the interaction effect of advertising with price and past use, also is smaller in magnitude and is statistically indistinguishable from zero. However, δ_3 has become large in magnitude and negative (–0.151 [0.027]) so that advertising's marginal effect on loyal households remains negative, as in the previous specifications. Given the similarity of the results across Models A and B, I focus on the Model-A results in performing the sets of counterfactual experiments that conclude this paper.[27]

5. Market-Level Implications: Advertising versus Price Discounts?

Up to this point, this paper has focused on measuring the differential effects of advertising on loyal and nonloyal households and the various specifications—from the raw data correlations in Table IV to the brand-choice models—and all indicate that advertising plays a role

26. In what follows, I use the shorthand notation PASTUSE_{ih0} to denote the value of the past purchase indicator for household h and brand i at the beginning of the sample period.

27. For convenience, I do not include the counterfactual results for Model B in the paper, because they are qualitatively similar to the Model-A results.

in reducing switching costs to brands a household has not purchased recently. Next, I consider counterfactual experiments to explore whether the large observed advertising expenditures are justified relative to alternative means—such as price discounts—for stimulating demand among consumers loyal to rival brands.

In particular, I consider an alternative promotional strategy whereby the advertising expenditure for each brand i is reduced unilaterally by 25% from the observed levels but whereby producer i offers price discounts so that brand i's market share remains unchanged under this alternative strategy. More formally, for each brand i, I simulate its aggregate demand [28] and solve for the discounted price P_i^* such that

$$\begin{aligned} &\mathit{MktShare}\big((1-\Delta) * \mathrm{ADV}_i,\ \mathrm{EXP}_{i0}=0,\ P_i^*\big)\\ &\quad = \mathit{MktShare}(\mathrm{ADV}_i,\ \mathrm{EXP}_{i0}=0,\ P_i), \end{aligned}$$

where $\Delta = 0.25$ denotes the percentage reduction in the advertising levels.

If firms' advertising expenditures are justified from a profit-maximizing point of view, this alternative promotional strategy strictly should be inferior to the observed strategy so that the savings from the alternative strategy (the reduction in advertising expenditures) should not exceed the costs, which are the foregone profits from offering price discounts:

$$\begin{aligned} &\underbrace{\Delta * \mathrm{ADV}_i^*}_{\text{Savings from ad reduction}}\\ &\quad < \underbrace{\big(P_i - P_i^*\big) * \mathit{MktShare}\big((1-\Delta) * \mathrm{ADV}_i,\ \mathrm{EXP}_{i0}=0,\ P_i^*\big)}_{\text{Foregone profits from price discount}}, \end{aligned}$$

where P_i denotes the average observed price of brand i (from the data).

Table VII shows the average estimated price discount across all brands, as well as individually for the same 10 brands as in the earlier tables. On average, across the 48 brands for which positive ad levels were observed, the price discount would "cost" about \$3.77 [4.08] million (= 4.58–0.81) more than the reduction in advertising expenditures. A similar finding holds across all brands, and the table reports the figures for 10 of the brands. Thereby, these simulations indicate that firms indeed are rational in the sense that even at the observed levels of

28. In order to calculate aggregate demand for each brand, I simulate 52-week purchase histories for each of the 1,010 households in my sample, assuming that they make one shopping trip per week. I aggregate these histories over households in order to estimate long-run (in-sample) aggregate purchases for each brand, weighing each household by a sample weight estimated using the Consumer Expenditure Survey, which captures its representativeness in the US population. See the discussion in Section 2.1.

TABLE VII.
ADVERTISING VERSUS PRICE DISCOUNTS: SIMULATED COUNTERFACTUALS USING MODEL-A RESULTS

	25% Ad Reduction ($mill)	Foregone Profits ($mill)
Summarized across All Brands	0.81	3.66 (3.61)
	For 10 Selected Brands	
KG Corn Flakes	1.78	5.38 (4.58)[a]
Cheerios	1.82	3.96 (2.27)
Rice Krispies	1.51	6.51 (6.98)
Frosted Flakes	1.97	6.37 (3.66)
KG Raisin Bran	1.40	8.39 (5.07)
Grape Nuts	1.69	2.24 (2.07)
Nutrigrain	0.63	3.36 (2.00)
Froot Loops	0.78	1.68 (1.02)
Lucky Charms	0.77	2.47 (1.81)
Product 19	0.35	4.94 (3.94)

[a]Standard errors calculated via simulation (see footnote 31).

advertising, it is more efficient to stimulate demand by advertising rather than through price discounts.

In addition, the finding that the forgone profits from a decrease in advertising are positive across all brands suggests that advertising's net effect on *aggregate* demand is positive: As we discussed previously, the results that advertising's marginal effect on loyal households is negative is problematic, but the results here confirm that advertising's net aggregate effect is positive, which one should expect.

A caveat of the above results is that, when considering brand i, I assume that the price and advertising of all rival brands $j \neq i$ are fixed across the two scenarios. By not allowing the competitive response of brand i's competitors to brand i's entry tactics, the above results might overestimate both the demand-enhancing effects of brand i introductory advertising as well as its price discounts, resulting in an ambiguous effect on my estimates of ADV_i^* and of the foregone profits.[29,30]

29. I have resimulated these counterfactuals assuming, in turn, a 10% cut in prices and a 10% increase in advertising by the competing brands, as a crude approximation to modeling these brands' competitive response to a new brand's entry. The magnitudes of the simulated quantities do not differ in any noteworthy way. In addition, I also ran versions of the experiments that allowed prices to vary (with a standard deviation of 0.1*price) across households, brands, and time. The results did not change much from the reported quantities.

30. Similar results obtain from counterfactuals where I increase each brand's advertising, and compare the whether these extra costs are balanced out by the increase in revenues that the firm could get by raising prices. For example, for Corn Flakes, a 50% increase in

6. Conclusions

In this paper, I confirm that an important effect of advertising in the breakfast-cereals market is to encourage "switching" behavior at the household level, which overcomes brand loyalty by persuading households to try brands they have not purchased recently. While the analysis in this paper has been confined to the breakfast-cereals market, these results can shed light more generally on a potentially important effect of advertising in differentiated product markets in which households' preferences are characterized by brand loyalty. Numerous marketing studies suggest that brand loyalty is a robust feature of preferences across a variety of frequently purchased consumer product markets.

The conclusiveness of my results is limited by the availability of only national brand-level advertising data. It will be of interest to see whether my findings—especially the problematic results that advertising's marginal effects on the purchase probabilities of loyal households is negative—continue to hold with more disaggregate advertising data.

These demand-level findings raise some interesting implications for advertising's effects on market structure in this industry. Since, in a dynamic market setting, brand loyalty can lead to substantial incumbent advantage, one implication of my findings is that advertising may be an attractive and effective option for entrant firms in their bid to overcome incumbent advantage due to brand loyalty. Furthermore, the finding that advertising reduces switching costs could imply that advertising facilitates entry of a new brand into a market populated by consumers loyal to the incumbent brand so that fewer brands may exist in the cereals market in the absence of advertising.[31] Fleshing out these issues would require much more detailed modeling of the supply side than has been attempted in this paper and is a goal of future research.

Appendix

A.1 Nested Logit Model Likelihood Function

Let $\mathcal{E}_{ht}$ denote the set of brands $i \in \{1, \ldots, 50\}$ that household h is loyal to on purchase occasion t, and let $\mathcal{N}_{ht}$ denote the other brands. Then the

---

advertising would cost \$3.55 million, while the rise in revenues from the accompanying price increases would generate only \$1.99 million.

31. For empirical work on these issues, see Scott-Morton (2000) for a study of how the advertising of branded products affects the entry of generic firms for the pharmaceutical industry.

likelihood of a purchase of a brand $i \in \mathcal{E}_{ht}$ is

$$l_{iht} = \frac{\exp\Big(V_{iht} - \sigma_h \log\Big[\sum_{j\in\mathcal{E}_{ht}} \exp(V_{jht})\Big]\Big)}{\exp(V_{0nt}) + \exp(V_{51ht}) + \Big[\sum_{j\in\mathcal{E}_{ht}} \exp(V_{jht})\Big]^{1-\sigma_h} + \Big[\sum_{j\in\mathcal{N}_{ht}} \exp(V_{jht})\Big]^{1-\sigma_h}}. \tag{4}$$

Similarly, the likelihood of a purchase of brand $i \in \mathcal{N}_{ht}$ is

$$l_{iht} = \frac{\exp\Big(V_{iht} - \sigma_h \log\Big[\sum_{j\in\mathcal{N}_{ht}} \exp(V_{jht})\Big]\Big)}{\exp(V_{0nt}) + \exp(V_{51ht}) + \Big[\sum_{j\in\mathcal{E}_{ht}} \exp(V_{jht})\Big]^{1-\sigma_h} + \Big[\sum_{j\in\mathcal{N}_{ht}} \exp(V_{jht})\Big]^{1-\sigma_h}}. \tag{5}$$

The likelihood of a purchase either of the outside good ($i = 0$) or of the composite 51st brand ($i = 51$) is

$$l_{iht} = \frac{\exp(V_{iht})}{\exp(V_{0ht}) + \exp(V_{51ht}) + \Big[\sum_{j\in\mathcal{E}_{ht}} \exp(V_{jht})\Big]^{1-\sigma_h} + \Big[\sum_{j\in\mathcal{N}_{ht}} \exp(V_{jht})\Big]^{1-\sigma_h}}. \tag{6}$$

I allow σ_h, which parameterizes the larger substitutability within than across nests, to vary for households that contain children (i.e., YGCHIL = 1). Note that as $\sigma_h \to 0$, the model becomes a nonnested multinomial logit model.

A.2 Controlling for Unobserved Heterogeneity

The usual stochastic assumptions that ensure consistency of the random-effects approach must be strengthened for the specifications estimated in this paper, since they include lagged dependent variables (PREVPURC and PASTUSE) among the covariates.

Let $\Omega_h \equiv \{\omega_{ihq}, i = 1, \ldots, 50, q = 1, \ldots, 8\}$ denote the sequences of unobservables associated with household h's sequence of purchases. The usual random-effects approach is based on the assumption that ω_{ihq} is *i.i.d.* over i, h, and q and also is distributed independently of the included covariates. In the presence of the lagged dependent variables PREVPURC and PASTUSE, however, the likelihood function for a household's observed purchases also depends on the initial conditions $(y_{-1}, \ldots, y_{-12})$, which are the purchases of a given brand in the 12 weeks prior to the beginning of the sample.

In order to implement the random-effects approach, I make the assumptions that (i) the initial values $(y_{-1}, \ldots, y_{-12})$ are exogenous scalars; and that (ii) the joint distribution of the unobserved effects $F_h(\Omega_h \mid y_{-1}, \ldots, y_{-12})$ does not depend on the values of $(y_{-1}, \ldots, y_{-12})$. Essentially, these assumptions ensure that the distribution of the unobserved heterogeneity parameters Ω is invariant to the initial values of (PASTUSE_{iht}, $i = 1, \ldots, 50$) and PREVPURC_t.[32] Given these assumptions, the resulting random-effects logit log likelihood function involves an integral over the joint distribution of Ω_h for all the observations pertaining to household h:

$$L = \sum_h \log \left\{ \int \left[\prod_t^{T_h} \prod_{i=0}^{51} l_{iht} d_{iht} \mid \Omega_h \right] dF(\Omega_h) \right\}, \tag{7}$$

where d_{iht} is an indicator for whether household h bought brand i at time t.

A.2.1 Details of Random Effects Specification. Here, I describe the assumptions made on the distribution of the household-, brand-, and quarter-specific unobservables ω_{ihq} in Model A. I consider a household-brand-quarter correlated random effect ω_{ihq} that has three components:

$$\omega_{ihq} = \omega_h^a + \omega_{ih}^b + \omega_{iq}^b.$$

The first component, ω_h^a, is a household-specific effect that captures differences in unobserved tastes for cereal across households (analogous to the ω_h defined in the previous section). The second component ω_{ih}^b captures heterogeneity in unobserved household-brand effects, which can induce spurious state dependence, as discussed by Heckman. The final component ω_{iq}^c controls for brand/quarter-specific unobservables that are the random-group effects, which, as Moulton (1986) pointed out, can lead to inferential difficulties in empirical models where macro- and micro-data are pooled.

I assume that the random variables ω_h^a, ω_i^b, and ω_{iq}^c are mutually independent across all triples (h, i, q). Across households, ω_h^a is i.i.d. with a discrete distribution with two points of support:

$$\omega_h^a = \begin{cases} \theta^H & \text{with prob. } \pi \\ \theta^L & \text{with prob. } 1 - \pi. \end{cases}$$

32. However, as pointed out below, Model B avoids this assumption by allowing the distribution of ω_{ih} to depend on PASTUSE_{ih0}.

A zero-mean restriction requires that $\theta^L = \frac{-\theta^H \pi}{1-\pi}$ so that only π and θ^H are estimated.

The second component, ω_{ih}^b, is drawn independently (but not identically) across brands and households from a normal distribution:

$$\omega_{ih}^b \sim N\left(0, \sigma_{ih}^2\right), \quad \forall i, h.$$

I allow σ_{ih}^2, the variance of this random-effect distribution, to be different for households with and without young children and also to be different across brand segments.

The third component is assumed to be drawn *i.i.d.* across brands i and quarters q:

$$\omega_{iq}^c \sim N\left(0, \sigma_c^2\right).$$

Model B differs only from Model A in that I relax the orthogonality restrictions of the random effects with the PASTUSE variable by parameterizing the mean of ω_{ih}^b, the (household-brand) second component, to be a function of PASTUSE_{ih0}:

$$\omega_{ih}^b \sim N\left(\mu_b * \text{PASTUSE}_{ih0}, \sigma_{ih}^2\right), \quad \forall i, h. \tag{8}$$

A.2.2 Estimation Details. Both Models A and B were estimated via simulated maximum likelihood (SML), using 10 simulation draws. The reported standard errors for these specifications were obtained by inverting the outer product matrix of the numeric gradients of the (simulated) log-likelihood function and are valid asymptotically (cf. Gourieroux and Monfort, 1996 : ch. 3) assuming that the number of simulation draws $s \to \infty$ at a rate faster than the number of observations N.

One may worry about the validity of statistical inferences based on the reported standard errors given that I use only 10 simulation draws. To assess these issues, I reestimated (at much computational expense) Model A using $s = 200$. The converged log-likelihood function value increased by only 0.95%, while the average change in the parameter estimates and the standard errors relative to the estimates obtained using only 10 simulation draws were, respectively, only 2.9% and –2.6%. These small changes suggest not only that variation in the parameter estimates due to simulation may be rather small but also that the parameter estimates and standard errors are quite stable for an increased s, so that worries about the validity of statistical inferences based on the reported standard errors may be minimized.

TABLE VIII.
LONG-RUN PRICE ELASTICITIES: ESTIMATED FROM MODEL-A RESULTS[a]

Average long-run aggregate own-price elasticities

	Avg. ϵ_{ii}	StDev[b]		Avg. ϵ_{ii}	StDev		Avg. ϵ_{ii}	StDev
All Brands	−2.050							
Family Brands	−2.014		Adult brands	−1.992		Kids brands	−2.210	
Top-Five Family Brands			Top-Five Adult Brands			Top-Five Kids Brands		
Corn Flakes	−1.519	0.141	Total	−2.450	1.146	Cocoa Puffs	−2.273	1.005
Cheerios	−2.362	0.309	Special K	−2.373	0.963	GldnGrhms	−2.387	1.146
Rice Krispies	−2.330	0.427	Grape Nuts	−1.502	0.768	CinnTstCrnch	−2.381	1.086
Frstd Flakes	−2.060	0.370	SpnSzShdWt	−1.745	1.470	Froot Loops	−2.326	1.071
KG RaisinBn	−1.773	0.603	100% Natural	−1.737	0.950	Trix	−2.665	1.262

Cross-Price Elasticity Matrix for FAMILY Cereals[c]

Brand #	1	2	3	4	5	7	15	16	18	31	45	46	48
1: KG	−1.52	0.08	0.09	0.06	0.03	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.04
CornFl	0.14	0.05	0.05	0.04	0.03	0.03	0.02	0.01	0.01	0.01	0.01	0.00	0.01
2: GM	0.23	−2.36	0.03	0.07	0.04	0.03	0.00	0.01	0.04	0.02	0.00	0.01	0.05
Cheerios	0.75	0.31	0.17	0.15	0.09	0.09	0.07	0.05	0.07	0.06	0.02	0.02	0.06
3: KG	0.06	0.03	−2.33	0.07	0.07	0.06	0.04	0.04	0.04	0.03	0.00	−0.01	0.01
RiceKr	0.75	0.87	0.43	0.22	0.15	0.14	0.08	0.05	0.09	0.06	0.03	0.04	0.08

4: KG	0.13	0.07	0.23	−2.06	0.04	0.04	0.04	0.05	0.05	0.02	0.05	0.03	0.07
FrstdFl	0.75	0.88	0.89	0.37	0.25	0.23	0.15	0.11	0.15	0.11	0.08	0.06	0.15
5: KG	0.03	0.13	0.12	0.05	−1.77	0.10	0.02	−0.01	0.00	0.03	0.00	0.04	0.05
RaisinBn	1.35	1.65	1.46	1.32	0.60	0.51	0.41	0.42	0.41	0.29	0.21	0.19	0.41
7: GM	0.31	0.37	0.32	0.20	0.10	−2.15	0.08	0.03	0.06	0.05	−0.01	−0.03	0.12
HN Ch'os	2.59	2.83	2.63	2.17	1.40	0.67	0.54	0.38	0.64	0.36	0.17	0.12	0.59
15: GM	0.47	0.41	0.47	0.54	0.00	0.15	−1.72	−0.03	0.06	0.02	0.02	−0.03	0.02
Wheaties	4.03	4.48	4.47	3.55	2.42	2.37	1.12	0.82	1.13	0.63	0.50	0.46	1.05
16: PT	0.02	0.14	0.15	0.06	0.17	0.20	0.12	−1.50	0.17	0.13	0.01	0.06	0.16
RaisinBn	3.63	4.29	3.54	3.02	2.12	2.37	1.47	0.85	1.27	0.98	0.73	0.73	1.91
18: KG	0.55	0.38	0.61	0.41	0.09	0.11	0.15	0.03	−2.41	0.08	−0.08	0.01	0.07
CornPops	5.50	4.94	5.39	5.02	3.22	3.12	1.71	1.50	1.16	0.95	0.59	0.55	1.36
31: GM	0.40	0.37	0.34	0.60	0.33	0.11	0.05	0.13	0.04	−2.04	0.02	0.00	0.26
AC Ch'os	5.20	5.77	5.14	5.29	3.66	3.00	1.57	1.60	1.92	1.33	0.79	1.05	1.94
45: GM	0.36	0.33	0.25	0.27	0.31	0.12	0.23	0.20	0.11	0.06	−1.97	0.13	0.46
Clusters	10.34	11.87	10.58	10.25	8.26	6.79	4.39	4.44	4.00	3.45	1.41	2.89	7.36
46: KG	0.58	0.14	−0.09	0.10	0.24	0.37	−0.12	−0.14	−0.06	−0.18	−0.20	−1.98	0.17
MiniBuns	12.91	11.80	11.15	9.17	8.27	8.88	4.58	3.96	5.91	4.01	3.31	2.03	5.61
48: GM	0.19	0.04	0.18	0.15	−0.01	0.02	−0.04	0.15	0.11	0.17	0.11	0.12	−2.38
MG Ch'os	3.55	3.30	3.45	3.21	1.81	1.93	1.23	1.41	1.12	1.18	0.80	0.69	0.98

[a] See for details on computation of elasticities.
[b] Average and StDev calculated over 200 simulated household populations.
[c] First row is average value; second row is standard deviation.

A.3 Long-Run Price and Advertising Elasticities

In this section I present estimates of long-run aggregate brand-purchase elasticities calculated using the Model-A estimates. In calculating these elasticities, I start by simulating one-year (52-week) purchase histories for each of the 1,010 households in my sample. I aggregate these histories over households in order to estimate long-run (in-sample) aggregate purchases for each brand, weighing each household by a sample weight estimated using the Consumer Expenditure Survey, which captures its representativeness in the US population. Then I derive the long-run arc elasticities by calculating how these simulated aggregate purchases would respond to a specified percent increase in price:

$$\epsilon_{ij} \equiv \frac{\%\Delta D_i}{\%\Delta p_j} \approx \frac{(D_i(p_j * (1+\Delta_p); p_{k,k\neq j}) - D_i(p_j; p_{k,k\neq j}))/D_i(p_j; p_{k,k\neq j})}{\Delta_p},$$

where D_i and p_j are the long-run aggregate purchases for brand i and the price of brand j, respectively, and where Δ_p is a given percentage change in p_j. In the calculations reported here I set Δ_p and Δ_{adv} = to 0.1 (i.e., a 10% price increase).

Table VIII contains the calculated price elasticities using the Model-A results. Two hundred sets of elasticities were simulated, and the reported numbers are the average (and standard deviations) across these 200 estimates.

The top part of the table shows own-price elasticities, both for individual brands as well as averaged over all brands or brand segments. These elasticities are estimated quite precisely. Across all 50 brands, the average own-price elasticity is –2.05 and varies slightly across segments, with kids' cereals having slightly more elastic demand (with an average own-price elasticity of –2.21) and with adult cereals having the least elastic demand (with an average of –1.99). Among the family cereals, Cheerios and Rice Krispies are estimated to have the most elastic demand, while Total and Trix are, respectively, the adult and kids' cereals with the most elastic demand.

The bottom part of the table shows the cross-elasticity matrix for the 13 brands in the family segment. In contrast to the own-price elasticities, the cross-price elasticities are not estimated precisely at all, even after using 200 simulated populations to construct the elasticities.

References

Ackerberg, D., 2001, "Empirically Distinguishing Informative and Prestige Effects of Advertising," *RAND Journal of Economics*, 32, 316–333.

——, 2003, "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Examination," *International Economic Review*, 44, 1007–1040.

Alba, J., J. Hutchinson, and J. Lynch, 1991, "Memory and Decision Making," in T. Robertson and H. Kassarjian, *Handbook of Consumer Behavior*, eds., Prentice Hall, 1–49.

Bain, J., 1956, *Barriers to New Competition*, Cambridge, MA: Harvard University Press.

Becker, G. and K. Murphy, 1988, "A Theory of Rational Addiction," *Journal of Political Economy*, 96, 675–700.

Bell, D. and J. Lattin, 1998, "Shopping Behavior and Consumer Preference for Store Price Format: Why 'Large Basket' Shoppers Prefer EDLP," *Marketing Science*, 17, 68–88.

Ben-Akiva, M. and S. Lerman, 1985, *Discrete Choice Analysis*. Cambridge, MA: MIT Press.

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

Butters, G., 1977, "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies*, 44, 465–491.

Cardell, N.S., 1997, "Variance Components Structures for the Extreme Value and Logistic Distributions with Applications to Models of Heterogeneity," *Econometric Theory*, 13, 185–213.

Chen, P. and L. Hitt, 2001, "Switching Cost and Brand Loyalty in Electronic Markets: Evidence from On-line Retail Brokers," mimeo., Wharton School.

Comanor, W. and T. Wilson, 1974, *Advertising and Market Power*. Cambridge, MA: Harvard University Press.

Cotterill, R. and L. Haller, 1997, "An Econometric Analysis of the Demand for RTE Cereal: Product Market Definition and Unilateral Market Power Effects," University of Connecticut Food Marketing Policy Center, Research Report No. 35.

Deighton, J., C. Henderson, and S. Neslin, 1994, "The Effects of Advertising on Brand Switching and Repeat Purchasing," *Journal of Marketing Research*, 31, 28–43.

Dube, J., 2004, "Multiple Discreteness and Product Differentiation: Strategy and Demand for Carbonated Soft Drinks," *Marketing Science*, 23.

Erdem, T. and M. Keane, 1996, "Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15, 1–20.

Gourieroux, C. and A. Monfort, 1996, *Simulation-Based Econometric Methods*, Oxford, UK: Oxford University Press.

Grossman, G. and C. Shapiro, 1984, "Informative Advertising with Differentiated Products," *Review of Economic Studies*, 51, 63–81.

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

Hausman, J., 1996, "Valuation of New Goods under Perfect and Imperfect Competition," in T. Bresnahan and R. Gordon, eds., *The Economics of New Goods*. Chicago, IL: University of Chicago Press, 209–237.

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

Kihlstrom, R. and M. Riordan, 1984, "Advertising as a Signal," *Journal of Political Economy*, 92, 427–450.

Kiser, E., 1998, "Demand and Pricing in the Breakfast Cereals Industry," Ph.D. Dissertation, University of Wisconsin at Madison.

Leading National Advertisers (1990–1993), *Ad Summary*.

Little, J., 1979, "Aggregate Advertising Models: The State of the Art," *Operations Research*, 27, 629–667.

McFadden, D., 1978, "Modeling the Choice of Residential Location," in A. K. et al. eds., *Spatial Interaction Theory and Residential Location*, North Holland Pub. Co.

Milgrom, P. and J. Roberts, 1986, "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 94, 796–821.

Moulton, B., 1986, "Random-Group Effects and the Precision of Regression Estimates," *Journal of Econometrics*, 32, 385–397.
Nelson, P., 1974, "Advertising as Information," *Journal of Political Economy*, 82, 729–755.
Nevo, A., 2001, "Measuring Market Power in the Ready-to-Eat Cereals Industry," *Econometrica*, 69, 307–342.
Scott-Morton, F., 2000, "Barriers to Entry, Brand Advertising, and Generic Entry in the US Pharmaceutical Industry," *International Journal of Industrial Organization*, 18, 1085–1104.
Shapiro, C., 1982, "Advertising as a Barrier to Entry?," Federal Trade Commission Bureau of Economics Working Paper #74.
Stango, V., 2002, "Pricing with Consumer Switching Costs: Evidence from the Credit Card Market," *Journal of Industrial Economics*, 50, 475–492.
Stern, S., 1996, "The Demand for Pharmaceuticals," mimeo, Kellogg School.
Stigler, G., 1964, "The Economics of Information," *Journal of Political Economy*, 61, 213–225.