An Economic Analysis of Consumers' Purchasing Behavior for Breakfast Cereals

Yuanjing Jin, Eugene Jones Joyce Chen, and Abdoul Sam

Consumer price sensitivity is an issue of concern for most food marketers. This study uses 2006–2007 supermarket scanner data to estimate the price sensitivity of four groups of shoppers: higher-income, moderate-income, lower-income, and rural dwellers. Economic theory is used to hypothesize price sensitivity for three of these groups of shoppers. Higher-income shoppers are expected to show the highest level of price sensitivity for national brands and lower-income shoppers are expected to show the lowest level of price sensitivity for these brands. Price sensitivity for rural dwellers is not hypothesized, but the level of income for these dwellers is between that of moderate- and lower-income shoppers. The results for national brands of cereals show own-price elasticities for moderate-income, lower-income, and rural shoppers to be larger (absolute value) than those for higher-income shoppers; these relationships do not hold for private label brands. Furthermore, private label brands are shown to be strong substitutes for national brands, but national brands are weak substitutes for private label brands.

Kellogg's introduced the first breakfast cereal, Corn Flakes, in 1906, and cereals continue to be a breakfast food staple for many Americans. Brands and product varieties have proliferated over the years and cereal manufacturers have used heavy advertising and promotion to encourage and persuade consumers to try these products (Rudd Center for Policy and Obesity 2009). As overweight and obesity became serious health issues for our nation, cereal manufacturers faced mounting criticism for producing products that are high in fat, salt, and sugar, and then targeting these products to children (Brownell and Horgen 2004). In response to this criticism, these firms have introduced products with lower amounts of fat, salt, and sugar (Hellmich 2009). This study examines the purchasing behavior of consumers for all brands and varieties of cereals.

Supermarket scanner data are used in two time-series cross-sectional models to estimate own-price, cross-price and expenditure elasticities for several brands of breakfast cereals within seven grocery stores. One of these models segments breakfast cereals by manufacturer; the other, by product type. Cereal manufactures and corresponding brands for Model 1 are General Mills, Kellogg's, Post, Quaker Oats, and private label. Product types for Model 2 are sweet, mainstream, and healthy; each type is offered as national brands and private labels. The data show that national brands of cereals are higher priced than private label brands. Indeed cereal prices follow the order as listed for Model 1, declining in magnitude from General Mills to private label. For cereal types, as used in Model 2, prices are highest for healthy cereals and lowest for sweet cereals. This pattern holds for both national brands and private labels.

Six of the seven stores used in these analyses were selected by income levels within defined geographic areas of Columbus, Ohio. A seventh store was selected from a rural area, based mainly on sparseness of population density (Table 1). Consumer purchasing behavior with respect to income is the primary motivation for this study. To this end, the stores are grouped into four classes: higher-income (two stores); moderate-income (two stores); lower-income (two stores); and rural (one store). With respect to income, the rural store falls somewhere between the income levels of the moderate- and lower-income stores, but no purchasing behavior is hypothesized for these residents because their travel distance for shopping differs sharply from that of other residents. The rest of this paper is outlined as follows: Section 1 provides some general characteristics of the residents surrounding each store; Section 2 gives a non-technical description of the two models and estimation procedures; Section 3 provides a discussion of the findings; and Section 4 ends the paper with conclusions.

Jin is a former Graduate Student, Jones is Associate Professor, and Chen and Sam are Assistant Professors, Department of Agricultural, Environmental and Development Economics, The Ohio State University Columbus.

Table 1. Socioeconomic Characteristics of Ohio, Franklin County, and Seven Supermarkets (%).

	State of Ohio	Franklin County	Higher-income stores		Moderate-income stores		Lower-income stores		Rural store
			Store 1	Store 2	Store 3	Store 4	Store 5	Store 6	Store 7
Income									
Less than $20,000	22.0	20.2	8.9	9.1	14.5	15.5	27.3	33.6	18.5
$20,000 to $49,999	37.8	36.9	23.1	26.9	42.4	35.9	36.4	43.2	35.5
$50,000 to $99,999	30.4	31.4	35.1	37.7	36.2	36.5	26.2	19.4	35.6
$100,000 & more	9.8	11.5	20.9	17.3	2.9	5.9	5.1	2.0	4.0
Race									
White	84.9	75.5	90.4	95.4	71.1	91.2	81.0	43.9	87.1
Black	11.4	17.6	1.6	1.5	21.5	3.3	5.0	49.5	10.4
Asian	1.2	3.1	6.2	2.1	2.7	2.9	10.0	1.1	0.7
Other	2.6	3.8	1.8	1.0	4.7	2.7	4.1	5.6	1.8
Education									
High School	53.1	41.4	27.9	38.7	40.3	21.2	17.7	58.3	59.2
College	39.5	48.0	56.1	49.5	53.0	55.8	56.5	36.5	36.9
Graduate	7.4	10.7	16.1	11.8	6.7	23.0	25.8	5.2	3.9
Household Size									
One	27.3	40.0	20.2	17.2	31.4	35.5	38.3	35.1	22.5
Two	33.3	32.6	31.2	37.2	33.9	37.8	36.4	30.6	35.9
Three or more	39.2	27.5	48.6	45.6	34.8	26.7	25.4	34.3	41.7
Children under 18	25.4	25.1	30.5	25.3	21.2	16.1	10.9	23.1	19.2

General Characteristics of Store Shoppers

It is well known that inward and outward migration occurs within most communities. Yet given an aging population and a declining economy, net migration within the past decade has slowed for most communities. This suggests that socioeconomic characteristics captured in the 2000 Census are likely to be fairly representative of current conditions. To this end, the 2000 Census data in Table 1 are used as proxy for shoppers who patronize the stores in this study. More specifically, 2006–2007 scanner data are used for this study and these data are hypothesized to represent shoppers with the approximate income levels described in Table 1.

With the exception of Store 7, data shown in Table 1 represent residents who live within a three-mile radius of each store. For Store 7, a wider radius of six miles is used. The income classifications of higher, moderate, and lower are derived from the distribution of households within the listed income classes. The two income classes offering the most visible evidence for the described segmentation are "less than $20,000" and "$100,000 and above." For the higher-income stores, an average of nine percent of households has incomes below $20,000; 19.1 percent have incomes of $100,000 or more. As we move from higher- to moderate- and to lower-income classes, the percentage of household with incomes below $20,000 increases significantly, while the percentage of households with incomes above $100,000 decreases drastically.

With respect to diversity, there is very little racial diversity among residents of the higher-income stores; considerably more among residents of the moderate-income stores; and an even higher level among residents of the lower-income stores. Education levels among the three income groups are somewhat at odds with the generally expected positive relationship between education and income. This is explained by an anomaly in the selection of stores for this study. One of the moderate-income stores and one of the lower-income stores are located in census tracts that are within close proximity of The Ohio State University. More directly, these two stores are within census tracts with large numbers of graduate and professional students who hold advanced degrees, but have limited incomes because they are pursuing even higher degrees. Household size characteristics suggest that residents who pa-

tronize higher-income stores are more likely to be from traditional family structures (mother, father, and at least one child) than are residents who patronize the other three classes of stores. A similar conclusion about family structure can be drawn from the number of children under 18 (Jin 2009).

Model Development

Many coefficients are derived from the specified models for this research, but the coefficients of particular interest for this paper are own-price elasticities for higher-, moderate-, lower-income and rural shoppers. These elasticities are measured with respect to space and time, and therefore a time-series cross-sectional model is viewed as most appropriate (Pindyck and Rubinfeld 1998). Various model specifications are possible, but the error components model has been shown to be the most robust (Fuller and Battese 1974). The general form of this model is:

$$(1) \quad Y_{qr} = \sum_{s=1}^{V} X_{qrs} B_s + u_{qr}, \quad q = 1, 2, ..., N;$$
$$r = 1, 2, ..., T,$$

where N is the number of cross sections and T is the length of the time series for each cross section.

For estimation purposes, the above time-series cross-sectional regression model is expanded as

$$(2) \quad \log(Quantity)_{mit} = \sum_{k=1}^{6} \log(Price)_{itk}\beta_{mk} +$$
$$\sum_{j=1}^{3} AreaDummy_{jt}\log(Price)_{itm}\alpha_{mj} + \log(Sales_{it})\delta_m$$
$$\sum_{j=1}^{3} AreaDummy_{jt}\log(Sales_{it})\varphi_{mj} +$$
$$\sum_{l}^{6} StoreDummy_{lt}\varphi_{ml} + u_{mit}.$$

Observations used in the model cover 104 weeks; therefore the subscript t, week, ranges from 1 to 104. Seven stores are part of these analyses and therefore the subscript i, store, ranges from 1 to 7. Two private label and four national brands are included in Model 1; three private label and three national brands are included in Model 2. Hence the subscript m ranges from 1 to 6 in both models. Variables in the models are specified as follows: $\log(Quantity)_{mit}$ is the total quantity sold (ounces) for week t, store i, and

category m; logarithms are taken for each variable so $\log(Price)_{itk}$ is the log weighted average price of week t, store i, and category k. For Model 1 the subscripts identified as m and k range from 1 to 6; the six dependent variables represented by m are private label regular, private label value, Kellogg's, General Mills, Post, and Quaker Oats. For Model 2 the subscript m includes dependent variables national brand sweet, national brand mainstream, national brand healthy, private label sweet, private label mainstream, and private label healthy. If $m = k$, β_{mk} is an own-price elasticity and if $m \neq k$, β_k is a cross-price elasticity.

For own-price elasticities, higher-income stores are used as the base and other own-price elasticities are calculated relative to this base. Specifically, own-price elasticites for moderate-income, lower-income, and rural shoppers are determined by adding the estimated differences to the base value. These differences are shown in equation 2 as $AreaDummy_{jt}$. The subscript j ranges from 1 to 3, capturing elasticity differences for lower-income, moderate-income, and rural shoppers. This $AreaDummy_{jt}$ is 1 if store i falls into area j and 0 otherwise. The coefficient α_{mj} is the coefficient estimate for own-price elasticity difference.

Total expenditure is a proxy variable included in the models to capture the effects of income. This variable is specified as $\log(Sales_{it})$, where i represents store i and t represents week; δ_m is a coefficient that provides an estimate of expenditure elasticity. $AreaDummy_{jt}$ is also used to estimate the differences in expenditure elasticity between higher-income areas and others.

To estimate differences in quantities sold by store, a dummy variable $StoreDummy_{lt}$ is included to allow for changes in the intercept. Store 1, a higher-income store, is used as the base store, and the subscript l ranges from 1 to 5. The coefficient φ_{ml} Is included to estimate store differences between store l and store i for category m. The error term is specified as u_{mit}, where the subscripts are for store i in week t and category m.

Prices are determined by expressing each product sale as a ratio of all product sales within a given product group. Specifically, weighted price for product group i in each period is

(3) $P_i = \Sigma_j W_{ij} P_{ij}$,

where $W_{ij} = (P_{ij} Q_{ij})/(\Sigma_j P_{ij} Q_{ij})$ and j denotes the products in the same group. Because each product group is a potential substitute for or complement with other product groups, all product groups are included in each equation.

As illustrated in Equation 2, both models used in this research are expressed in double-log form. Econometricians have shown that this form is most relevant for time-series data that are likely to have the statistical problems of autocorrelation or multi-collinearity. When normal econometric procedures are used to estimate equations without accounting for autocorrelation and multicollinearity, the estimated coefficients are generally unstable, non-significant, and often have algebraic signs that differ from theoretical expectations (Chatterjee and Price 1977). Thus the double-log form of this study is intended to minimize these econometric problems.

Discussion of Findings

Model 1

This discussion is somewhat general, as a page limitation will not allow the inclusion of tables with estimated results. Own-price elasticities for the six brands of cereals in higher-income areas have negative signs as hypothesized, and are statistically significant at the 0.01 level. On average, own-price elasticities for private label value, Kellogg's, and General Mills are greater than −2 (absolute value), while own-price elasticities for private label regular, Post, and Quaker are less than −2 (absolute value). These elasticities show that higher-income shoppers have the highest price sensitivity for the highest-priced national brands (General Mills and Kellogg's) and the lowest-priced private label brands (private-label value). Lower price sensitivity for Post, Quaker Oats, and private label regular cereals by higher-income shoppers could reflect price/quality tradeoffs in purchase decisions. That is, higher-income consumers may perceive a less than perfect relationship between price and quality.

Numerical differences in own-price elasticities for non-higher-income shoppers are captured by price dummy variables included in the model, and these coefficients are shown to be statistically significant at the 0.05 level; private label regular is statistically significant at the 0.10 level. Contrary

to expectations, own-price elasticity differences show that lower-income consumers are not the most price-sensitive for all cereal products. In fact, lower-income shoppers do not show the highest price-sensitivity for any of the national brands of cereals. Relative to moderate-income and rural shoppers, lower-income shoppers are shown to have the smallest own-price elasticities (absolute value) for General Mills and Post cereals. Statistically there is no difference between the price sensitivity of lower- and moderate-income shoppers for Kellogg's brands of cereals. Moreover, moderate-income shoppers have higher price sensitivity than do lower-income shoppers for Quaker Oats brands of cereals. These unexpected results could be due to several factors. Time constraints of shoppers, store locations, and changes in consumers' demographics since the 2000 census could be relevant factors.

Own-price elasticities for lower-income consumers, as hypothesized, are smaller in absolute value for private-label cereals than for national brands. Own-price elasticities for private-label regular and private-label value cereals are −1.37 and −1.19, respectively. By comparison, these elasticities average −2.47 for Kellogg's, General Mills and Post cereals and −2.68 for Quaker Oats cereals. These values suggest a 2.47 percent change in quantity purchases for each one percent change in price for Kellogg's, General Mills, and Post cereals. Purchases of Quaker Oats cereals would change by 2.68 percent for each one percent change in price. Less-elastic responses are shown for private label cereals, with private label regular cereals changing by 1.37 percent for each one percent change in price, and private label value cereals changing by 1.19 percent for each one percent change in price. As expected, these differences for private label cereals show that consumers are more sensitive to higher prices. Furthermore, to maximize the purchasing power of each dollar, these elasticities suggest that lower-income consumers are likely to purchase private- label value brands over private-label regular cereals.

Although moderate-income shoppers were hypothesized to be less price-sensitive than lower-income shoppers, results for Post and Quaker Oats cereals show that they are more price-sensitive. Adding the elasticity differences to the base for higher-income shoppers, elasticities of −2.67 and −2.84 are shown for Post and Quaker Oats cereals,

respectively. One factor that may explain these elasticities is a highly educated population with low incomes. That is, these are the characteristics of many shoppers surrounding one of the two moderate-income stores. Price sensitivity for rural shoppers was difficult to hypothesize, but the results show patterns that are characteristic of lower-income shoppers. For Kellogg's, General Mills, and private label regular cereals, rural residents have the highest level of price sensitivity. For Post and Quaker Oats cereals, moderate-income shoppers are shown to have the highest price sensitivity. This means that lower-income shoppers have the highest price sensitivity only for private label value cereals.

Model 2

As discussed for Model 1, these results are derived from a time-series cross-sectional model. Own-price elasticities for three national brands of cereals in the higher-income area have negative signs, as hypothesized, and they are all statistically significant at the 0.01 level. Relative to magnitude they are all greater than one in absolute value, suggesting elastic demands and strong price sensitivity. The highest level of price sensitivity is shown for national brand sweet and the lowest level of price sensitivity is shown for national brand healthy. This lower level of price-sensitivity for healthy cereals supports the argument that higher-income shoppers are willing to pay higher prices for healthy products. This elasticity for healthy cereals is not inelastic and it suggests that factors other than price are critical determinants of demand. With private label products having lower prices, the derived own-price elasticities for private label mainstream and healthy cereals show that higher-income shoppers are indeed price insensitive for these two product classes. Statistically insignificant own-price elasticity is derived for private label sweet cereals, perhaps reflecting the power of advertising and promotion. These own-price elasticities for national and private label cereals suggest that higher-income shoppers evaluate purchases for national and private label brands differently, even when products are part of the same functional class.

Differences in own-price elasticities for higher-income and other income groups are statistically more significant for sweet cereals than for any other

brands. For national brand sweet, all income groups are shown to have own-price elasticities that differ from those of higher-income shoppers. The largest difference is observed for moderate-income shoppers, with an own-price elasticity of −2.21 versus −1.75 for higher-income shoppers. By comparison, rural shoppers are shown to have the highest price sensitivity for private label sweet cereals, while the difference for moderate-income shoppers is not statistically significant. Also, it should be noted that higher-income shoppers are shown to have statistically insignificant own-price elasticity for private label sweet cereals. As a result, this makes interpretation of deviations from the own-price elasticity for higher-income shoppers problematic.

In addition to statistically significant own-price elasticity differences for both national and private label sweet cereals, three other own-price elasticities are shown to be statistically significant. Moderate-income shoppers are shown to have an own-price elasticity of −1.8 for national brand mainstream cereals, versus a value of −1.5 for higher-income shoppers. Similarly, moderate-income shoppers are shown to have an own-price elasticity of −1.9 for national brand healthy cereals, versus a value of −1.5 for higher-income shoppers. Lastly, moderate-income shoppers are shown to be price insensitive toward the purchase of private label healthy cereals, having an own-price elasticity of −0.3.

Summarizing the estimated own-price elasticity differences, it is clear that differences are most pronounced for both national brand and private label sweet cereals. Five of the six own-price elasticities measuring differences from the base group of higher-income shoppers are statistically significant. For both national brand and private label healthy cereals, two of six own-price coefficients are statistically significant; for mainstream cereals, just one coefficient is statistically significant. Relative to price sensitivity, the highest level is shown for national brand sweet and the lowest level for private label sweet. Focusing on healthy cereals, the results show that national brand healthy has the lowest level of elasticity while private label healthy has the highest level of elasticity. These estimates suggest that consumers have different perceptions of cereal brands. Perhaps higher-income consumers have high reservation prices for the healthy characteristics of breakfast cereal, while lower-income consumers have much lower reservation prices.

Conclusion

This study examines the purchasing behavior of four different groups of consumers for national brands and private label brands of breakfast cereals. Income levels were used to identify three of these groups, while sparseness of population was used to identify the fourth group. Key findings are own-price elasticities for moderate-income, lower-income, and rural shoppers are larger (absolute value) than those for higher-income shoppers for national brand, but not necessarily for private label brands; price sensitivity differences are observed among higher-income consumers and all other consumers, and market share purchases of national and private label brands are also significantly different for these groups; and private label brands of cereals are shown to be strong substitutes for national brands, but national brands are much weaker substitutes for private label brands.

References

Brownell, K. and K. B. Horgen. 2004. *Food Fight: The Inside Story of the Food Industry, America's Obesity Crisis, and What We Can Do About It*, Chicago: Contemporary Books.

Chatterjee, S. and B Price. 1977. *Regression Analysis by Example*. New York: John Wiley and Sons.

Fuller, W. A. and G. E. Battese. 1974. "Estimation of Linear Model with Crossed-Errors Structure." *Journal of Econometrics* 2:67–78.

Hellmich, N. 2009. "Kids Cereals: 85% More Sugar, 65% Less Fiber than Adult's Cereals." *USA Today* October 27.

Jin, Y. 2009. "An Econometric Model of Breakfast Cereals: Measuring Price Sensitivity by Brands and Consumer Characteristics." M.S. Thesis. The Ohio State University.

Pindyck, R. S. and D. L. Rubinfeld. 1998. *Econometric Model and Economic Forecasts, 3rd Edition*. New York: McGraw-Hill.

Rudd Center for Food Policy and Obesity. 2009. "Cereal Facts: Evaluating the Nutrition Quality and Marketing of Children's Cereals." Yale University.