1981 APPLEBAUM SCHOLARSHIP WINNER
MS. DONNA K. JOURDAN

ELASTICITIES ESTIMATED FROM GENERAL PRODUCE SCANNER DATA

by

Donna K. Jourdan
Research Assistant
Department of Agricultural Economics
Texas A&M University

Ms. Jourdan received her Bachelor of Science degree in Agricultural Economics from Purdue University and her M.S. Degree in Agricultural Economics from Texas A&M University, August 1981. Ms. Jourdan's course work included the areas of microeconomic theory, marketing, econometrics, statistics and operations research. She is now a Systems Engineering Representative with IBM, Houston, Texas.

Introduction

Traditional analysis of consumer demand has been dependent upon aggregate annual estimates of consumer purchases. These time series data are often historical relationships that do not always represent current market conditions and are usually too generalized for application to product-specific decision making problems (Purcell, Raunikar and Elrod, 1966). The desire for a more detailed source of information regarding the relationship between consumer prices and the quantities of products they purchase has caused the evolution of continuing consumer panels and consumer surveys. Both of these methods provide data that are reflective of current market conditions and detailed in regard to items purchased. While these methods allow for demand estimates to be caluclated for specific products, they are expensive methods of data collection and are not completely accurate (Kinnear and Taylor, 1979). There now is an alternative source of data for product-specific demand estimization, after discounting these traditional sources. The new alternative source, scanner data, records actual consumer purchases of individual items at specific prices by electronic scanning checkout systems in retail food stores.

The development of this data source is attributable to the implementation of the Universal Product Bar Codes (UPC) by the retail food industry. The latest statistics (1980) on optical scanning checkout systems indicates there are over 3,000 stores in the United States scanning most items (FMI, 1980). This represents a 66 percent increase in scanner installations over 1979. The Food Marketing Institute (FMI) estimates 1981 monthly installations of new scanners at 170 to 190 stores per month. This means there will be approximately 5,000 systems in place by the end of 1981, representing 15 percent of the industry's total supermarkets. The vastness of this data source is realized when one recognizes that there are over 18,000 items available in current retail supermarkets (Gowens, 1979).

The implication of the availability of daily sales volume and price information has been recognized by the marketing research industry. The following data suppliers--NABSCAN, NIELSON, SAMI and TRIM--now provide this information to manufacturers and advertisers. MRCA and Management Science Associates provide consumer panels where purchase information is collected in scanning stores (Management Science Associates, 1980).

Most managers using scanning checkout systems are utilizing their information to realize "hard" savings--improve albor scheduling, reduce "out of stock" situations and reduct shrink due to mispriced items. To fully utilize the information from these systems, pioneering efforts are now being made by economic researchers in the area of "soft" savings--price elasticity estimation, advertising and featuring effectiveness analyses. The potential for these savings are just beginning. Scanner data is an economical, flexible, sensitive source of data that should be utilized by those undertaking economic research and management decision making.

Statement of the Problem

The problem addressed in this research was to estimate short run own-price and cross-price elasticities of demand for groups of specific retail cuts of beef based on actual purchases. Own-price elasticity of demand is a numeric relationship of the quantity demanded to the price of that product. Past estimation procedures have relied upon historical aggregated time series or consumer panel information. Elasticity estimates based on scanner data will differ from previous studies because these studies utilized data based on aggregate product consumption and annual times series data. The nature of annual time series data tends to "smooth out" variability over the time period the data are collected (Purcell, Raunikar and Elrod). Evaluation of the results of this research endeavor were compared to those found by Green, Hassan and Johnson, Wilson, Brandow, and George and King. All of these studies, except Wilson's, used annual time series and in some cases also cross-sectional data (George and King) in their estimation procedures. Their elasticity estimates were also for commodity groups. The Wilson study, which used purchase intentions of specific products, generated elasticity estimates which were of larger magnitude than those based on annual data, but appeared reasonably consistent. Due to the difference in models, data bases, estimation procedures and aggregation of commodities, a cautious comparison of the estimates was made.

Objectives

The main objective of this study was to estimate an own-price and cross-price elasticity of demand for a group of related beef cuts from actual purchase data.

Other specific objectives of this study were:

1. Develop a demand estimation of model that would produce reliable short run elasticity estimates and could be easily modified for continued applications.

2. Compare the elasticity estimates from this study with those of previous analyses, recognizing the magnitude and direction of difference and judging its congruency with existing estimates.

3. Explore this source of data as reliable for application in future economic research endeavors.

The importance of this study is dependent upon statistically significant and economically reasonable price elasticity of demand estimates. The retail food industry and consumer demand are dynamic by nature. This dynamic characteristic will force the estimation model to only produce accurate short run estimations. The short run limitation for the application of the model should not discount its unique value to agricultural economics and retail food management.

Demand theory specifies that the quantity consumed of a particular commodity is a function of its price, the prices of other commodities and income. The model used in this analysis is a single equation multiple regression model. The single equation is formulated for a specific retail cut of beef or group of related cuts. The equation estimates the direct price and other relevant cross-price elasticities. The effect of all other variables is implied to be zero (George and King.). The model is of the following functional relationship:

$$(1) \quad Q_{Bi} = f(P_{Bi}, P_{Pk}, P_{OB}, Y_i, W, Z, F, A, I)$$

Q_{Bi} = Per customer consumption of the specific retail cut(s). (Pounds per 1,000 customers).

P_{Bi} = Price of the specified group of retail cuts ($/lb.)

P_{Pk} = Price of pork (weighted average price in $/lb. of all pork cuts).

P_{OB} = Price of other beef (weighted average price in $/lb. of all other beef cuts).

Y = Discrete variable for high or low (relative household) income level of the store's neighborhood (store effect).

W = Discrete variable for quarter in study (first quarter - weeks 1-15; second quarter - weeks 16-30).

Z = Discrete variable for whether item(s) were featured in weekly newspaper advertising.

A = Number of square inches of advertising in newspapers (combined weekly the advertising in urban and suburban newspapers. Four levels: (1) total store advertising (competitor); (2) total store advertising (chain); (3) total meat advertising (competitor); (4) index of chain's meat advertising compared to total chain's store advertising.

I = Interactions (seven) between: price and income; price and other beef price; price and featuring; and advertising (levels 1 and 2) and featuring.

Scanner data can be classified as primary data. It is information "collected specifically for the purpose of the investigation at hand" (Churchill). Scanner data also has properties similar to cross-sectional and time series data. This is because the observations are made over time while also being made over geographical areas. It is a flexible data base that can accomodate varied economic investigations.

The source of scanner data used in this study is from four stores in an independent chain located in Houston, Texas. With stores located in the affluent western and northern suburbs of the city, the 15-store chain attracts many middle and upper income shoppers (Bell). The chain places great emphasis on its meat and produce, through display counters designed to draw attention. The stores average 33,000-35,000 square feet in size and have netted a 50-60 percent increase in volume over the past year. That was the result of the purchase of the four largest Handy Andy stores. The chain maintains an eight percent share of the Houston market. This is a two percent increase over last year and makes it the leader among local chains. Korger, Safeway and Weingarten together control 80 percent of the market.

The chain was chosen for this study since their stores have equipment available to generate labels enabling fresh meat products to be electronically scanned. Other justifying factors in this decision were the size and reputation of their stores, the active market in which they are located and the cooperativeness of the management.

The unit of analysis is sales per customer per week. There are 30 weeks in the study (May 11, 1980 through November 29, 1980). In this study, the groups of beef cuts for which elasticities were estimated were selected from over 500 items recorded in the meat department by the scanning checkout system. To determine the specific list of beef cuts, three studies were consulted: Degner, Wilson and Progressive Grocer (February 1978). The decision of which retail cuts were used in the analysis was also determined from the amount and regularity of the sales recorded for that cut in the scanner data and the variability of their price during the time period of the analysis. The retail cuts of beef were grouped according to standard primal categories.

The model includes the variables of price of the beef cuts, price of pork, income and advertising. The price parameters are the average of the prices each week for all items in each of the respective categories. The prices are consisent across the four stores. Relative income per household was measured as a discrete variable. It is the average income per household of the particular trade area in which the store is located (high, medium, low). This number will be the same for all weeks in the study. Three of the stores in the study were located in upper-middle class areas and one store was in a blue collar area.

The advertising variable was monitored in the Houston Chronicle, Houston Post and suburban newspapers for the areas the five stores in the study are located. The decision to monitor only print advertising was based on information published in Supermarket News. It was reported that retail food advertising in Houston is strongly oriented towards newspapers. An average of 80 percent of store advertising budgets goes to the pring media (Bell). The Houston Chronicle circulation strength is 90 percent daily in the Houston metro area--where 83 percent of retail food sales are. For these reasons it was considered that newspaper advertising had the most significant effect on purchase behavior. The unit of analysis is the number of square inches because of varying sizes of columns of newspring between suburban newspapers and the major Houston newspapers. The food ads appearing in the urban and suburban newspapers were identical. Because of this, the advertising measure reflects only the single appearance of the ad in the urban and suburban publication. The advertising information was collected on the basis of total store advertising and total meat department advertising appearing in urban and suburban papers for the chain and all competitors combined on a per week basis.

Results

As mentioned previously in this paper, scanner data has innumerable applications in the area of "soft" savings. This study is an example of this type of scanner data application. The results from estimating elasticities for one group of retail cuts (roasts) is discussed here. The "roast" category includes all retail roast cuts from the loin, chuck, round and brisket.[1] Ordinary least squares was used to estimate the parameters of equation (1). The relationship among the variables are expected to be multiplicative rather than additive. Therefore, a linear in logarithms equation was estimated.[2] The model generated the following results (Table 1).

Since this model is linear in logarithms, the parameter estimates are the elasticity values. All of the variables in the model were significant, but the only elasticities that are interpretable are the ones associated with other beef price and total meat advertising (competitor). Both of these values indicate that the variables have a positive effect on the sales of roasts. The interpretation of the other beef price elasticity is: for a 1 percent change in the price of alternative beef cuts to roasts there will

Table 1. Estimates from model.

Parameter	Parameter Estimate	t Test	Significance
Intercept	-1.36		
Other Beef Price	5.36	2.68	***
Pork Price	-3.82	-3.02	**
Total Meat Advertising Competitor (TMAD)	0.79	2.98	**
Week	0.30	2.17	**
Interaction (Roast Price and TMAD)	-0.45	-2.82	***
Interaction (Roast Price and the Chain's Roast Features)	0.22	2.65	**
Durbin-Watson (a) = 1.64		R^2 = .32***	

***Significant at 99%
**Significant at 95%
*Significant at 90%

be a 5.36 percent change in the same direction in the sales of roasts. The interpretation of the advertising elasticity is similar in that a 1 percent change in advertising by competitors will result in a .79% change in the same direction in the sales of the chain's roasts. This value is based on the mean number of square inches of competitor advertising used in this analysis. There were 3,751 square inches of competitor advertising used in this analysis. This seems reasonable since many households rely on newspaper ads to compare product prices among their shopping alternatives. The elasticity associated with pork price has a negative value which is not reasonable. It implies that an increase in the price of pork will result in a decrease in the sales of roasts. Previous to this model, a stepwise regression was run with all the variables hypothesized to be significant in the relationships. This procedure indicated the following variables were insignificant: the price of roasts; total store competitor advertising; the chain's meat and total store advertising; weekly features of roasts; and the relative household income level.

Other statistical tests were performed on this model. The coefficient of determination (R^2) represents the proportion of the total variability of sales of roasts over time that is accounted for by the explanatory variables. The value shown in Table 1 is R^2 = .32. This means that the independent variables in the model explained 32 percent of the variation of sales.

Another statistical test, the Durbin-Watson test, was used to test for the absence of autoregression. Autoregressive disturbances occur when the disturbance occurring at one observation point is correlated with a distrubance of another observation. A disturbance is defined as a summary of random and independent factors, that are not measurable, but enter into the relationship under study (Kmenta). The value calculated in the model was d = 1.64. It was concluded that the model was not autoregressive.

Implications

Interpretation of results from scanner data in economic research or in management decision making for "soft" savings must be done cautiously. One must make sure they know the limitations of the data (i.e., lost information when the item is not scanned) and account for them in their analyses. The results of this research indicated that more information combined with a longer time period of observations would have improved the elasticity estimates from this model.[2] It is believed that the R^2 value would be increased if variables accounting for the amount of shelf space allocated to roasts and a variable indicating the store image relative to its competitors were used in the model. A strong indication for addition of variables to the model is the appearance of a negative elasticity value on pork and relatively low R^2. It must also be cautioned that this model can only be used for short term decision making, since the retail food industry is a dynamic atmosphere.

The acceptance and application of sophisticated, technical equipment by the retail food industry must be met with proven decision making models which utilize the data potential of these systems. Historically, retail food store management have been reluctant and skeptical to adopt formal decision making models. They have chosen to rely on established and relatively unsophisticated management practices and knowledge from experienced personnel. Not discounting this experience as inaccurate, but to combine this knowledge with the wealth of information available from scanner data and incorporate both into a decision making model can only serve to improve the efficiency of the retail food industry. Scanner data is an important data source that should be better utilized by management.

FOOTNOTES

[1]When an item was not scanned, the sale was not recorded by specific item. It was decided to delete those observations where less than 40 pounds per 1,000 customers were recorded, before applying the model.

[2]Other mathematical forms of the equation were also applied (linear and semi-log linear), but did not yield as satisfactory results.

REFERENCES

"A First-Time Look at Meat Movement Through the Scanner's Eye: How 340 Meat Items Sell." Progressive Gròcer. February 1979.

Bell, Sally. "Areas Supermarkets Vie on Pricing." Supermarket News. Sec. 2, August 18, 1980.

________. "Local Supermarket Business Being Shaken Up Ever Since Grand Union's Entry Into The Region." Supermarket News. Sec. 2, Aug. 18, 1980.

Brandow, G. E. Interrelations Among Demands for Farm Products and Implications for Control of Market Supply. Penn. Agr. Exp. Sta. Bul. 680, 1961.

Churchill, Gilbert A. Marketing Research Methodological Foundations. Hinsdale: Dryden Press, 1976.

Degner, Robert. "Impact of Price Specials on Estimates of Retail Meat Prices." M.S. Thesis, Texas A&M University, 1970.

"4th Quarter 1980 Scanning Fact Sheet." Washington, DC: Food Marketing Institute.

George, P. S. and G. A. King. "Consumer Demand for Food Commodities in the United States with Projections for 1980." Giannini Foundation Monograph No. 26. Univ. of Calif. Div. of Agri. Sci.

Gowens, Robert L. "Data Processing Alternatives for Independent Grocers." Food Dist. Rsch. Society: Oct., 1979.

Green, Richard, Zuhair A. Hassan, and S. R. Johnson. "Alternative Estimates of Static and Dynamic Systems for

Canada." Amer. J. Agri. Econ. 60(1978).

Kmeta, Jan. Elements of Econometrics. New York: McGraw-Hill, 1971.

Kinnear, Thomas C., and James R. Taylor. Marketing Research, An Applied Approach. New York: McGraw-Hill, 1979.

Purcell, Joseph C., Robert Raunikar and J. C. Elrod. Consumption and Expenditure Analysis for Daily Products, Fats and Oils in Atlanta, Georgia. Ga. Agr. Exp. Sta. Bul. N.S. 51, 1966.

Wilson, Winston. "Consumer Survey Data as a Basis for the Derivation of Retail Price and Income Elasticities of Demand." M.S. Thesis, Texas A&M University, 1969.