“Where’s the Beef?”: Statistical Demand Estimation Using Supermarket Scanner Data

Fred H. Hays
University of Missouri—Kansas City

Stephen A. DeLurgio
University of Missouri—Kansas City

Abstract

This paper is a case study designed for students and instructors in managerial economics and intermediate price theory courses. It utilizes a publicly available database of monthly supermarket scanner data for various cuts of beef. Linear multiple regression models are used to estimate price, cross, and income elasticities of demand. A log-linear model is also used to provide direct elasticity estimates.

Keywords: (demand estimation, multiple regression analysis, scanner data, price elasticity, cross elasticity, income elasticity)

Background

Virtually all microeconomic principles textbooks discuss the concept of elasticity of demand, the responsiveness of quantity demanded to a change in some other variable such as the "own" price of a good (price elasticity), disposable income (income elasticity) or the price of a related good (cross elasticity). Generally the ensuing discussion includes calculation of point price elasticity with a few limited examples. In some texts there also may be examples of the ranges of price elasticity for various consumer items. In the basic course it is unusual to address the question of how elasticity is calculated from a statistical approach.

Managerial economics texts as well as some applied intermediate microeconomics texts take the discussion a step further by incorporating a summary of statistical applications of ordinary least squares regression to empirically estimate elasticity. A few limited data sets may be included either as examples or problems in the appendices or an accompanying course website. At times, these illustrations are contrived, leaving students, especially those in MBA or EMBA programs, to ask "how is this relevant in actual real world settings"? or "how did they come up with those elasticity estimates"?

This paper uses "real world" supermarket scanner data from a publicly available government website to generate elasticity estimates for various cuts of beef. This case study can be easily adapted for classroom use. It illustrates the calculation of "own" price elasticity, cross elasticity and income elasticity using a traditional simple linear multiple regression model. The paper also examines a multiplicative form for the model and estimates elasticity coefficients directly using log transformed data. We also consider the overall goodness of fit as well as the explanatory significance of individual regression estimates and the interpretation of the regression estimates.

Literature Review: Standing on the Shoulders of Giants

The current body of knowledge of demand theory, elasticity and statistical estimation techniques has been developed during the last century with sustained contributions from some of our greatest economics scholars. Some of the early contributions represented applications of demand theory to agricultural commodities. Indeed, the application of statistical measurement techniques to analyzing the elasticity of demand for beef dates to over 80 years ago (Schultz, 1924). Schultz (1935) also estimates elasticity of demand for beef using data for per capita consumption, deflated retail price and income using annual data from 1922-33.

There are several literature reviews encompassing these early works including (H. Working, 1925), (Ferger, 1932), (Ynmenta, 1939), (Stigler, 1954) and (Christ, 1985). These trace the progression and development of statistical demand analysis from the collection of social and accounting data and development of index numbers to the application of the concepts of probability, correlation and regression in estimating economic relationships including the calculation of various measures of demand elasticity.

More recent refinements address the appropriate form of estimating equations (linear, log transformed or generalized) (Chang, 1977), the dynamic properties of demand equations (Eales & Unnevehr, 1988) and the application of scanner data to estimation of demand functions (Capps, 1989). It is from this rich theoretical and empirical base that we are able to offer students a glimpse of the development of modern demand theory and estimation.

Data Sources and Issues

Supermarket scanner data of prices and quantities for various types of beef and poultry are available in Excel at http://www.retail-lmic.info/CD/Downloads/Beef.xls (There are additional time series for many additional cuts of meat available beyond those used in this paper. Data is available for short ribs, roast, round steak, sirloin, stew meat, T-bone, top loin and ground beef, among other cuts). The monthly data from the Economic Research Center of the US Department of Agriculture in cooperation with the Livestock Marketing Information Center (LMIC) covers January 2001 to December 2007.
(http://www.lmic.info/meatscanner/meatscanner.shtml
Scanners were introduced in supermarkets in the mid-70's, although the use of consistent and reliable scanner data dates to the late 1970's in statistical studies. Capps (1989) estimates that scanner data are available for 35,000 to 40,000 items in retail food stores.

Although many different income time series are available, we use per capita disposable personal income data that are available through subscription to Economagic. (http://www.economagic.com) Appendix 1 contains a spreadsheet with quantity and price data for three cuts of beef (Chuck, Porter House and Ribeye) plus data for chicken prices and disposable income, all on a monthly basis. This data was imported from Excel using a data query procedure into SPSS where it was analyzed using a multiple regression procedure.

A Conventional Linear Demand Model

We initially utilize a standard linear multiple regression model of the form:

$$Q_x = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + e_i \quad [1]$$

Qx= an index of beef quantities (base year =2001);
α= constant (equals quantity of X when all other variables =0)
$X_1 = Px_1$, the "own" price of a given type of beef
$[B_1 = \Delta Qx/\Delta Px_1]$
$X_2 = Px_2$, the price of a related good, chicken
$[B_2 = \Delta Qx/\Delta Px_2]$
X_3= measure of disposable (after-tax) income (Inc)
$B_3 = \Delta Qx/\Delta Inc$
X_4= trend variable (1,2,3………n)
e_i =error term

The quantity variable Q_x is an index of quantities for different cuts of beef using a base year of 2001=100. The index is based on supermarket scanner data for quantities purchased in pounds for each cut of beef.

Our initial analysis uses quantities and prices per pound for chuck roast, a relatively inexpensive cut of beef. Price elasticity is the percentage change in quantity demanded given a percentage change in the "own" price of the good. B_1 is the slope indicating how much quantity changes with a unit change in price of the good itself. This is not price elasticity. B_1 must be multiplied by the price/quantity value at a specific point on a demand curve to arrive at price elasticity. As the ratio of P/Q changes along the demand curve, so does the elasticity. "Own" price elasticity has a negative sign since there is a downward sloping demand curve and therefore an inverse relationship between P and Q.

The cross-elasticity of demand measures the responsiveness of the percentage change in quantity of one good to the percentage change in the price of a related good. The empirically determined sign of the cross-elasticity measure is important since positive signs indicate substitute goods while negative signs denote complementary goods. To measure cross-elasticity of demand we initially use the price of chicken per pound. Because of the abundance of available data on different cuts of beef, it is also possible to measure the cross elasticity between cuts (for example, between chuck roast and perhaps rib eye or Porterhouse steak).

Income elasticity measures the responsiveness of a percentage change in the quantity consumed of a good to a percentage change in the real disposable per capita income. Income elasticity values less than zero are inferior goods whereby consumers choose to reduce purchases with an increase in income. Normal goods have positive income elasticities.

Following the rationale of Schultz (1935) we include a simple trend variable with a value of 1, 2, 3…..n over the time series.

Table 1 contains the estimated regression coefficients and associated test statistics for initial model of chuck roast demand.

Table 1
Estimated Regression Coefficients
The Demand for Chuck Roast

Variable	Coefficient	t statistic	Significance
Chuck Price	-50.427	-3.051	.004
Chicken Price	-51.783	-1.111	.272
Disposable Income per capita	-.054	-2.367	.022
Trend	1.667	2.450	.018
Adj. R2 Durbin- Watson	.275 2.421	F=6.122	.000
Mean Q chuck roast =107.55	Mean P chuck roast = $2.47	$E = B_1 * P_x / Q_x$ $= -50.427*2.47/107.55$	“Own” price elasticity= -1.167

$$E\ (Chuck) = B_1*P_x/Q_x = \Delta Q_x/\Delta Px_1*P_x/Q_x$$
$$= -\ 50.427*2.47/107.55 = -1.167 \qquad [2]$$

As shown in the last row of Table 1 and equation [1], the estimated price elasticity of demand for chuck roast equals the estimated coefficient for B_1 multiplied by the mean value for the price of chuck ($2.47) divided by the mean index value for quantity of chuck (107.55). The resulting elasticity -1.167 has the expected negative sign and is statistically significant at the .004 level. It indicates that a 10% change in price will be associated with an inverse change in quantity demanded of 11.67%, an elastic response.

Cross elasticity for chuck roast measures the percentage change in the quantity of chuck roast associated with a percentage change in the price of a related good. In this case our initial related good is the price of chicken per pound, which is postulated to be a substitute good. The low value for the t-statistic and the associated low confidence level create significant doubt about the relationship.

Using the mean quantity of chuck roast of 107.55 and a mean price of chicken of $1.695 and the estimated coefficient of -51.783 we arrive at a cross elasticity of demand between chuck roast and chicken of -5.754 using the same procedure of equation [2]. The negative sign is consistent with a complementary good. This is contrary to intuitive expectations that chuck and chicken are substitute goods. The t-statistic of -1.111 is not statistically significant casting doubt on the result.

Mean Q chuck =107.55	Mean P chicken = $1.695	E= B_2*P/Q = -51.783*1.695/107.55	Cross elasticity= -5.754 (Complement)

Income elasticity is the relationship between the percentage change in quantity associated with a percentage change in income. The real per capita disposable (after tax) income is used as an income measure. The mean value for Q is again 107 .55 while the mean value for per capita disposable income is $9,252.85. Similar to equation [2], the estimated regression coefficient of -.054 is multiplied by the ratio of income to Q to arrive at the income elasticity estimate of -4.646. Values for income elasticity of less than zero are considered to be inferior goods. The t-statistic is -2.367 and a statistically significant confidence level of .022.

Mean Q chuck =107.55	Mean Disposable Per Capita Income =$9,252.85	E= B_3*Inc/Q = -.054*9252.85/107.55	Income elasticity= -4.646 (e< 0= inferior good)

Porter House and Rib Eye Steaks

We next examine two other cuts of beef that are generally considered to be higher quality than chuck roast---Porter House and Rib Eye Steaks. The same estimation procedure is repeated with the results summarized in Table 2.

Table 2
Estimated Elasticities
Porter House and Rib Eye Steaks

	Price Elasticity	Cross Elasticity with Chicken	Income Elasticity
Porter House Steak R^2= .64 (F=32.32; p=.000)	-2.568 (.000)	12.677 (.199) Substitute	.406 (.738) Normal good (necessity)
Rib Eye Steak R^2=.35 (F=9.304; p=.000)	-2.140 (.000)	18.274 (.005) Substitute	.266 (.009) Normal good (necessity)

(Significance levels of original regression coefficients are shown in parentheses)

Porter House Steaks. The results for Porter House steaks are mixed. The "own" price elasticity estimate of -2.568 indicates that a 10% increase in the price per pound for Porter House steak results in a 26% reduction in quantity demanded. The sign, as expected, is negative, and the estimated coefficient is significant at a 99.9% confidence level. The higher elasticity for Porter House Steaks seems consistent with theories related to the purchase of higher priced goods.

The sign of the cross elasticity measure is positive, indicating that chicken and Porter House steak are substitute goods. The income elasticity is a positive .406, indicating a normal good. However, the estimated coefficient is not statistically significant. The model explains 64% of the variation in Q and has a highly significant F value. This model was fitted without a trend variable.

While not developed here, the Porter House relationship has some statistically significant serial correlation. While serial correlations were corrected through an exact maximum-likelihood procedure, the resulting coefficients were not statistically significantly different than those reported in Table 2. We see that this data lends itself to a realistic discussion of the underlying assumptions of ordinary least squares and the impact of violations to those assumptions. These discussions are part of classroom use of these estimations but are not included here because of space limitations.

Rib Eye Steaks. The interpretation of Rib Eye elasticities are identical to those of Porter House steaks. In contrast to Porter House steaks, there were not statistically significant serial correlations in the residuals of its ordinary least squares estimations.

A Log-linear Model

Many textbooks in managerial economics or applied microeconomics will consider non-linear demand models of the form given below:

$$Q_x = \alpha P_x^{B1} P_o^{B2} Y^{B3}$$

Where: P_x = "own price" of good
P_o = price of a related good
Y = disposable income

The original equation above may be logarithmically transformed to:

$$\log Q_x = \log \alpha + B_1 \log P_x + B_2 \log P_o + B_3 \log Y$$

One of the properties of this log transformed function is that the estimated coefficients are direct estimates of elasticity for their corresponding variables (for example, B_1 is the "own" price elasticity of demand). The coefficient requires no further manipulation to represent elasticity. This approach was recognized by Moore (1924).

Table 3 presents the results for the demand for chuck roast. The results are consistent with the elasticity estimates from the linear model. The price elasticity is -1.353 with the expected sign and a statistically significant relationship. Cross elasticity indicates substitute goods, although the relationship is not statistically significant. The income elasticity also is not significant and the relationship changes from a weak inferior good to a weak normal good. While the adjusted R^2 is low at .236, the F statistic is significant at p= .001.

Table 3
Direct Estimates of Elasticity of Demand for Chuck Roast
From a Multiplicative Demand Function

Price elasticity	Cross elasticity	Income elasticity
-1.353	.038 Substitute	.072 Necessity
Significance (.001) Adj.R^2= .236 F= 6.564 p=.001	(.952)	(.901)

Summary and Conclusions

Elasticity of demand is an important concept for business managers and policy-makers to understand. It is frequently dismissed as "too theoretical" and "lacking 'real world' relevance" because students are not provided with "live" datasets with which to experiment in applying elasticity concepts. The advent of publicly available monthly supermarket scanner data allows students to see that concepts can be easily applied to solve "real world" problems. This case study is designed to permit students to see not only results but also the challenges associated with these applications. These types of analyses expose students to important theoretical and methodological problems related to effective demand estimation.

References

Anderson, Eugene W. & Shugan, Steven M. (1991) Repositioning for Changing Preferences: The Case Study of Beef versus Poultry, *Journal of Consumer Research*, 18 (2), 219-232.

Brester, Gary W. & Wohlgenant, Michael K. (1993) Correcting for Measurement Error in Food Demand Estimation, *Review of Economics and Statistics*, 75 (2), 352-356.

Capps, Oral, Jr. (1989) Utilizing Scanner Data to Estimate Retail Demand Functions for Meat Products, *American Journal of Agricultural Economics,* 70 (3), 750-759.

Chang, Hui-shyong (1977) Functional Forms and the Demand for Meat in the United States, *Review of Economics and Statistics*, 59 (3), 355-359.

Christ, Carl F. (1985) Early Progress in Estimating Quantitative Economic Relationships in America, *American Economic Review*, 75 (6), 39-52.

Eales, James S. & Unnevehr, Laurian J. (1988) Demand for Beef and Chicken Products: Separability and Structural Change*, American Journal of Agricultural Economics,71 (3)* , 521-532.

Farnham, Paul G. (2005) *Economics for Managers*, Upper Saddle River NJ: Pearson Education, Inc.

Ferger, Wirth F. (1932) The Static and Dynamic in Statistical Demand Curves, *Quarterly Journal of Economics*, 47 (1), 36-62.

Hirschey, Mark. (2006) *Managerial Economics*, Stamford CT: Thomson Higher Ed.

McGuigan, James R., Moyer, R. Charles & Harris, Frederick H.deB. (2005) *Managerial Economics: Applications, Strategies, and Tactics with Economic Applications*. Stamford CT: Thomson Higher Ed.

Moore, Henry Ludwell (1926) Partial Elasticity of Demand, *Quarterly Journal of Economics*, 40 (3), 393-401.

Schultz, Henry (1924) Statistical Measurement of the Elasticity of Demand for Beef, *Journal of Farm Economics*, 6 (3), 254-278.

Schultz, Henry (1935), Interrelations of Demand, Price and Income, *Journal of Political Economy*, 43 (4), 433-481.

Stigler, George, (1939) The Limitations of Statistical Demand Curves, *Journal of the American Statistical Association*, 34 (207), 469-481.

Stigler, George, (1954) The Early History of Empirical Studies of Consumer Behavior, *Journal of Political Economy*, 62 (2), 95-113.

Working, E.J., (1927) What Do Statistical "Demand Curves" Show?, *Quarterly Journal of Economics*, 41 (2), 212-235.

Working, Holbrook, (1925) The Statistical Determination of Demand Curves, *Quarterly Journal of Economics*, 39 (4), 503-543.

Yntema, Theodore O. (1939) Henry Schultz: His Contributions to Economics and Statistics, *Journal of Political Economy*, 47 (2), 153-162.

Appendix 1
Dataset for Estimating the Demand Elasticity for Beef

Year	Month	ChuckQ	CkuckP	PHQ	PHP	RibEQ	RibEP	ChickP	DiscInc
2001	1	120	$2.28	53	$6.04	74	$7.02	$1.61	8664.6
2001	2	76	$2.61	81	$5.37	79	$7.16	$1.68	8689.8
2001	3	102	$2.12	60	$5.74	71	$7.33	$1.57	8711.7
2001	4	106	$2.41	65	$6.93	112	$7.38	$1.73	8710.7
2001	5	87	$2.39	92	$5.95	113	$6.47	$1.74	8716.5
2001	6	94	$2.11	157	$5.24	89	$7.14	$1.63	8732.4
2001	7	97	$2.66	149	$5.39	146	$7.02	$1.69	8733.2
2001	8	79	$2.50	133	$5.54	120	$6.28	$1.71	8727.7
2001	9	138	$2.39	97	$6.28	120	$7.57	$1.79	8738.5
2001	10	129	$2.30	113	$5.43	106	$6.72	$1.65	8743.1
2001	11	77	$2.69	78	$5.46	76	$7.73	$1.75	8751.7
2001	12	94	$2.44	123	$4.81	93	$7.71	$1.79	8769.6
2002	1	135	$2.23	87	$5.49	83	$6.74	$1.65	8789.9
2002	2	120	$2.45	73	$5.54	136	$6.87	$1.72	8814.2
2002	3	131	$2.24	88	$5.76	119	$7.11	$1.68	8840.2
2002	4	104	$2.41	108	$5.28	114	$7.00	$1.67	8879
2002	5	144	$2.24	82	$6.05	151	$6.70	$1.63	8889.3
2002	6	139	$2.08	148	$5.34	248	$7.08	$1.66	8907.6
2002	7	98	$2.42	155	$5.10	201	$6.54	$1.73	8894.8
2002	8	112	$2.47	45	$6.88	168	$6.46	$1.73	8892.7
2002	9	184	$2.34	97	$5.67	165	$6.72	$1.76	8898.8
2002	10	98	$2.33	82	$5.65	105	$6.89	$1.68	8910.3
2002	11	109	$2.35	65	$6.38	88	$7.69	$1.62	8920.8
2002	12	135	$2.35	53	$6.22	108	$7.44	$1.83	8945.3
2003	1	132	$2.22	63	$5.86	182	$6.65	$1.70	8976.3
2003	2	142	$2.38	59	$6.72	150	$6.72	$1.67	9011.5
2003	3	128	$2.45	55	$7.12	144	$7.51	$1.68	9053.2
2003	4	94	$2.53	49	$6.54	121	$7.80	$1.84	9067.9
2003	5	93	$2.34	92	$6.70	161	$7.40	$1.81	9124.5
2003	6	126	$2.48	88	$6.91	280	$6.64	$1.74	9163.4
2003	7	83	$2.40	65	$7.33	170	$6.68	$1.71	9188.2

2003	8	109	$2.54	92	$6.70	165	$7.73	$1.82	9214.2
2003	9	112	$2.46	41	$7.75	106	$8.20	$1.80	9243.7
2003	10	83	$2.86	42	$7.22	104	$8.32	$1.80	9271.7
2003	11	90	$2.90	56	$7.57	91	$9.74	$1.79	9339.4
2003	12	77	$3.03	30	$8.40	63	$9.87	$1.85	9374.8
2004	1	86	$2.83	41	$6.65	67	$9.92	$1.61	9435.9
2004	2	165	$2.55	51	$6.94	111	$8.48	$1.61	9487.9
2004	3	109	$2.25	25	$7.40	118	$7.51	$1.64	9530.7
2004	4	89	$2.37	34	$7.45	107	$8.27	$1.66	9571.8
2004	5	125	$2.27	54	$7.51	125	$8.27	$1.66	9628.3
2004	6	108	$2.28	33	$8.60	117	$8.86	$1.66	9642.8
2004	7	88	$2.52	34	$8.54	119	$8.55	$1.70	9689
2004	8	101	$2.64	41	$8.36	143	$8.87	$1.85	9739.8
2004	9	85	$2.64	51	$6.98	100	$7.48	$1.72	9759
2004	10	113	$2.61	45	$6.88	82	$8.20	$1.63	9858.4
2004	11	85	$2.69	27	$7.19	73	$8.59	$1.68	9924.9
2004	12	90	$2.67	31	$7.08	70	$8.10	$1.64	10291
2005	1	113	$2.62	38	$6.89	89	$8.34	$1.59	10025.2
2005	2	98	$2.70	39	$7.05	84	$7.90	$1.55	10072.9
2005	3	86	$2.82	30	$7.80	105	$7.88	$1.61	10122
2005	4	93	$2.66	38	$7.88	111	$8.92	$1.64	10145.1
2005	5	127	$2.51	55	$7.68	159	$7.60	$1.60	10180.6
2005	6	83	$2.47	40	$8.17	141	$7.99	$1.65	10231.5
2005	7	94	$2.44	39	$8.06	156	$7.18	$1.62	10268.8