All Tuna is not Created Equal: The Existence of Quantity Surcharges Due to Product Differentiation.

#117955

AAEA Annual Meetings 2004

Hayley Chouinard
Assistant Professor
School of Economic Sciences
Washington State University
PO Box 646210
Pullman, WA 99164-6210
Chouinard@wsu.edu

Jill McCluskey
Associate Professor
School of Economic Sciences
Washington State University

David Sprott
Associate Professor
Marketing
Washington State University

May 2004

Abstract: Quantity surcharges exist when unit prices are higher for larger packages. We hypothesize that various sizes of goods are differentiated products, and this explains some surcharges. Estimating a random-coefficients logit demand model, we examine own and cross elasticities to determine the level of differentiation between products with different size packages.

Key words: quantity surcharge, product differentiation, random-coefficients logit.

JEL codes: L110, D120

All Tuna is not Created Equal: The Existence of Quantity Surcharges due to Product Differentiation

Introduction

Consumers often have strong expectations about the relative prices of products found is different size packages. Many compare the per unit price of products in hopes of finding a quantity discount (Granger and Billson 1972, Manning, Sprott and Miyazaki 1998, Nason and Della Bitta 1983). Quantity discounts occur when the unit price of a brand's larger size package is less than the unit price of the same brand's smaller size package. For example, when the price per ounce of a brand's twelve-ounce can of tuna is 16.5 cents while the price per ounce of the six-ounce can is 20 cents, a quantity discount exists. Price conscious consumers approve of these quantity discounts and may react negatively to the opposite situation known as a quantity surcharge. A quantity surcharge exists if the unit price of a brand's larger size is greater than the smaller size.

Various studies suggest between 16 and 34 percent of products available in two or more package sizes found in retail grocery outlets exhibit a quantity surcharge (Sprott, Manning and Miyazaki, 2003). Prices of canned tuna often exhibit quantity surcharges. When consumers discover quantity surcharges, previous research finds many consumers respond negatively toward the brand or the retailer. Consumers may believe the retailer has engaged in deceptive pricing practices or has eliminated a preferred course of action (purchasing the larger package) for the consumer, and may decrease the likelihood of purchasing the surcharged brand or shopping in that retail outlet (Manning, Sprott, and Miyazaki, 1998).

Cost differentials and price setting practices have been presented as reasons for quantity surcharges. It may be more expensive to refrigerate larger packages of some goods, which results in cost based quantity surcharges. Some suggest (without empirical support) retailers may price discriminate against consumers who do not notice quantity surcharges (Agrawal, Grimm and Srinivasan 1993, Gupta and Rominger 1996). Alternatively, retailers may not intentionally set prices that result in quantity surcharges. These retailers may actively compete with other retailers on specific sizes of fast moving items such as six-ounce cans of tuna and drive the price of that package size down, which results in a quantity surcharge for the larger package size product (Sprott, Manning and Miyazaki 2003).

In this paper, we propose an additional theory for the existence of quantity surcharges. Goods sold in different package sizes may represent different products to consumers. We hypothesize that various package sizes of a good are actually differentiated products, and consumers should not expect the price per unit of various sizes to be the same or smaller for larger packages. Tuna packaged in a twelve-ounce can may not be the same product as tuna found in a six-ounce can. In other words, two, six-ounce cans of tuna, may not be equivalent to one, twelve-ounce can. Consumers may view these as imperfect substitutes. For example, the various sized packages may require different storage options, both before and after the can is opened or the use of the product may differ for a given quantity of the good.

We address the question of whether the goods where quantity surcharges exist are actually differentiated products. We focus our empirical investigation on canned tuna. Examining the estimated own and cross elasticities allows us to determine the level of differentiation between canned tuna with different size packages. If the cross-price elasticities of a good packaged in two sizes are close to zero, we conclude each package size represents a differentiated product.

Understanding the causes of surcharges provides useful information about retail pricing and consumer behavior. All the previous work with quantity surcharges has suggested that surcharges occur based strictly on retailer behavior. We suggest an explanation that is consumer driven. This may eliminate the negative effects of incorrect presumptions of the causes of quantity surcharges. Retailers may adjust their pricing behavior and consumers may react differently with respect to quantity surcharges. The remainder of this paper is organized as follows. Related work is discussed in the next section. A model of the system of demand for tuna follows. The next section includes a description of the data used to estimate the relevant elasticities. The expected empirical results and plans for future work are presented in the final section.

Related Work

The study of product differentiation has advanced over the last several decades in order to better define markets and provide insight into pricing behavior. Many innovations in the measurement of differentiated products have occurred. However, the degree of product differentiation continues to rely on the examination of the cross-price elasticities of demand between products. Often a discrete choice model is used to provide estimates of these elasticities. Conditional logit models, based on McFadden (1973), have been applied to several of these problems. Many logit models include a restrictive assumption where the substitution between products depends exclusively on the market share and not the similarity of the products. This occurs because all the regressors, including price, are erroneously assumed to be exogenous. This endogeneity results in substantial biases in elasticity estimates (Villas-Boas and Winer 1999).

Recent work in mixed logit models has focused on the problems and methods to account for this endogeneity while investigating market power, innovation and product differentiation. These models begin with random utility models where utility is composed of a mean level of utility from consuming a product, and a deviation from the mean. The deviation from the mean depends on the interaction between consumer preferences and product characteristics. Some of the product characteristics are unknown to the researcher. Thus, from the researcher's point of view, prices are endogenous. Berry (1994) examines such a model of discrete choice of product differentiation. He uses instrumental variables to account for the endogeneity of prices. This technique is applied to the automobile industry in the work of Berry, Levinsohn and Pakes (1995). They generate own and cross-price elasticities for several models of automobiles. They find that substitution is more likely for vehicles with similar characteristics.

Bresnahan, Stern and Trajtenberg (1997) apply similar techniques to determine the level of market segmentation and innovation in the personal computer market. They define principles of differentiation, which characterize a notion of product similarity, and then estimate the degree of market segmentation attributable to these principles. They determine that the substitutability between frontier and non-frontier products, and brand-name reputation allow segmentation. Thus, high rates of entry affect only those products in similar segments.

Nevo (2001) uses a random-coefficients logit model to estimate the price-cost margins for ready-to eat cereal. He estimates a brand-level demand system to obtain demand elasticites then used to identify market power from product differentiation, multi-product pricing, and price collusion. Nevo (2000) describes in detail the estimation of demand systems based on a random utility model and verifies the accuracy of the cross-price elasticity estimates.

Hindman Persson (2002) applies the random coefficients logit model to the demand for health inputs. She examines the choice of sanitation inputs in households and the welfare implications from a change in the price of different sanitation inputs. This application is an appropriate application of the model as accurate estimation of the cross-price elasticities drives the results.

Model

To estimate the demand for canned tuna and the associated cross-price elasticities and determine if various package sizes are differentiated products, we use a random-coefficient logit model (Nevo 2000, 2001). This approach improves the accuracy of estimation of cross-price elasticities, since it explicitly models heterogeneity in the population. Estimating the parameters governing the heterogeneity allows us to calculate unrestricted cross-price elasticities, which characterize product differentiation.

The characterization of the demand system and the choice of estimation techniques are especially important in this context, as more restrictive logit models impose structure on the cross-price elasticities. The restrictive models include the assumption that substitution between brands occurs in proportion to market shares, regardless of brand characteristics. For example, if six ounce cans of Chicken of the Sea light tuna packed in water and six ounce cans of StarKist tuna packed in oil have similar market shares, then substitution from six ounce cans of Bumble Bee light tuna packed in water will be the same for the Chicken of the Sea and StarKist tuna.

The random-coefficient logit model does not force substitution patterns to be functions of market share by allowing prices to be correlated with the econometric term. In this model, products are defined by a set of characteristics that influence demand. Producers and consumers observe all the product characteristics. However, the econometrician only observes some of the characteristics. From the econometrician's point of view, the error term captures the unobserved characteristics. The unobserved characteristics influence the price of the product, and prices are endogenous. This endogeneity requires the use of instrumental variables to obtain consistent estimates of the demand system parameters.

Therefore, it is desirable to model a system where choices are correlated. Ideally, this correlation should be a function of product and consumer characteristics. Substitution patterns between products will then be similar for similar products, and consumers with similar demographics will have exhibit similar choice behavior. Such a system more accurately describes selection behavior and generates better estimates of cross-price elasticities.

In order to develop a demand system with the desired characteristics, we begin with a conditional indirect utility as a function of observed and unobserved characteristics. The utility of consumer, i, from consuming product j in time t is written as

(1) $$U(x_{jt}, \xi_{jt}, p_{jt}, \tau_i; \theta).$$

Here, x_{jt} and ξ_{jt} are observed and unobserved product characteristics respectively, p_{jt} is the price of the products, τ_i is the individual characteristics, and θ represents the unknown parameters. The observed characteristics for canned tuna include: packed as solid, packed as chunk, packed in water, packed in oil, and light. The unobserved characteristics include all types of unknown systematic shocks to demand.

Consumer taste parameters are modeled using the individual characteristics, τ_i. These characteristics consist of demographics, D_i, and unobserved additional characteristics, v_i. Neither type of characteristics is observed, but some information about the distribution of demographics is known. This information can be used to provide information about the distribution of the random coefficients. The demographic variables include race, household size, education level, employment and presence of children in households.

The indirect utility can be written in a quasi-linear form as a function of all the variables and parameters:

(2) $$u_{ijt} = \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, v_i, D_i; \theta_2) + \varepsilon_{ijt}$$

where, $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$, and $\mu_{ijt} = [-p_{jt}, x_{jt}]'*(\prod D_i + \sum v_i)$.

Here, β and α are linear parameters, and $\prod$ and $\sum$ are nonlinear parameters. This formulation of the utility separates the mean utility, δ_{jt}, from the mean-zero heteroskedastic deviation from the mean, $\mu_{ijt} + \varepsilon_{ijt}$. The deviation from the mean captures the effects of the random coefficients.

When we impose the independently and identically distributed extreme-value distribution assumption on ε_{ijt}, correlation between choices will occur between products with similar characteristics, and consumers with similar characteristics will have similar rankings.

The own and cross price elasticities of the market shares are

(3) $$\eta_{jkt} = \frac{\partial s_{jt} p_{kt}}{\partial p_{kt} s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) d\hat{P}_D^*(D) dP_v^*(v) & \text{if j = k,} \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} d\hat{P}_D^*(D) dP_v^*(v) & \text{otherwise,} \end{cases}$$

where,

$$s_{ijt} = \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{K} \exp(\delta_{kt} + \mu_{jkt})}$$

is the probability of individual *i* purchasing product *j*. These patterns of substitution depend on price sensitivity, not functional form, and substitution between brands will depend on product characteristics not market shares. The flexibility of this model provides accurate measures of the cross-elasticity between products. However, this model does not have an analytic closed form solution. In the full random-coefficients model, the demand system is solved numerically.

Data

We use data from the Dominick's Finer Food grocery store chain located in the Chicago area. The data was collected in cooperation between Dominick's and the James M. Kilts Center for Marketing in the Graduate School of Business at the University of Chicago. For this analysis, a market is defined as activity in a specific store in a specific week. We examine ten stores over four weeks, thus we consider forty markets. The ten stores are located in various neighborhoods across Chicago. The four weeks run from May 31, 1990 through June 27, 1990.

The products examined included canned tuna from three brands: Chicken of the Sea, StarKist, and Bumble Bee. Each brand offers canned tuna in several sizes approximately 3.25, 6, 9.25 and 12 ounces. They offer canned tuna in various sizes and various types of packing characteristics. We use 25 products in this analysis.

The dependent variable in the estimation is the market share of the product. To determine the market share we consider that in 1990 the US per capita consumption of canned tuna reached 3.7 pounds. This equates to 1.14 ounces of canned tuna consumption per person per week. The total available market for canned tuna in each store in each week is the number of customers in the store each week multiplied by 1.14. The market share for each product equals the total ounces of the product sold divided by the total available market.

The price of the products is recorded for each store and week. Although each market consists of the same grocery chain, there is considerable variation of prices by market across stores and across weeks. The price is measured per ounce.

The observed product data include product characteristics found by examining the labels. These characteristics include: packed as a solid, packed as chunks, packed in water, packed in oil, and light. Product specific dummy variables are also included to capture any factors that influence utility that are not found in the observed characteristics, and to capture the characteristics that do not vary by market.

The demographics are based on the store specific information. The data comes from the US Government 1990 census for the Chicago metropolitan area. Market Metrics processed the data to generate demographic profiles for each store. The demographic variables include: the percent of blacks and Hispanics, the average household size, the percent of college graduates, the percent of working women with full time employment, the percent of working women with children under the age of 5, and the percent of working women with children between the ages of 6 and 17.

To obtain consistent estimates, instrumental variables must be used to account of the endogeneity of prices. The instruments must be correlated with the prices, but uncorrelated with the error term. We use prices of the products at other Dominick stores in the Chicago area during different weeks. These prices are correlated with the original prices, but will not include the unobserved characteristics that lead to the endogeneity.

Expected Empirical Results

To estimate the demand for canned tuna and the associated cross-price elasticities and determine if various package sizes are differentiated products, we use a random-coefficient logit model. The estimation will generate elasticites for each product that we can compare to examine the degree of product differentiation. Based on the law of demand, we expect the own price elasticities to be negative for all the products. These elasticities can help verify that we obtain reasonable results. We expect the cross-price elasticities to be positive as we expect the goods to be substitutes. We also expect the cross-price elasticities to be substantially less than one. As the cross-price elasticity approached zero, the products become less good substitutes. The smaller the cross-price elasticities, the more support we gain for our theory that the various sizes of canned tuna are different products.

References

Agrawal, Jagdish, Pamela Grimm and Narasimhan Srinivasan (1993), "Quantity Surcharges on Groceries," *Journal of Consumer Affairs*, 27, 335-356.

Berry, Steven T. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25(2), 242-261.

Berry, Steven, James Levinshohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841-890.

Bresnahan, Timothy F., Scott Stern and Manuel Trajtenberg (1997), "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the late 1980s," *RAND Journal of Economics*, 28, s17-s44.

Granger, C.W.J and A. Billson (1972), "Consumers' Attitudes Toward Package Size and Price," *Journal of Marketing Research*, 9, 239-248.

Gupta, Omprakash and Anna Rominger (1996), "Blind Man's Bluff: The Ethics of Quantity Surcharges," Journal of Business Ethics, 12, 305-315.

Hindman Persson, Therese (2002), "Welfare Calculations in Models of the Demand for Sanitation," *Applied Economics*, 34, 1509-1518.

Manning, Kenneth C., David E. Sprott and Anthony D. Miyazaki (2003), "Consumer Response to Quantity Surcharges: Implications for Retail Price Setters," *Journal of Retailing*, 74(3), 373-399.

McFadden, Daniel (1973), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. P. Zarembka. New York: Academic Press.

Nason, Robert and Albert Della Bitta (1983), "The Incidence and Consumer Perceptions of Quantity Surcharges," *Journal of Retailing*, 59, 40-54.

Nevo, Aviv (2001), "Measuring Market Power in the Ready-To Eat Cereal Industry," *Econometrica,* 69(2), 307-342.

Nevo, Aviv (2000), "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9(4), 513-548.

Sprott, David E., Kenneth C. Manning and Anthony D. Miyazaki (1998), "Grocery Price Setting and Quantity Surcharges," *Journal of Marketing*, 67, 34-46.

Villas-Boas, Miguel and Russell Winer (1999), "Endogeneity in Brand Choice Models," *Management Science*, 45(10), 1324-1338.