AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

DEMAND ESTIMATION IN THE PRESENCE OF AN UNOBSERVABLE PRODUCT ATTRIBUTE: THE CASE OF MINT QUALITY

Daniel Toro-Gonzalez

Graduate Research Assistant, School of Economic Sciences,

Washington State University

Jia Yan

Assistant Professor, School of Economic Sciences,

Washington State University

Karina Gallardo

Assistant Professor, School of Economic Sciences,

Washington State University

Jill J. McCluskey

Professor, School of Economic Sciences,

Washington State University

Contact author: Daniel Toro-Gonzalez, School of Economic Sciences, PO Box 646210, Hulbert Hall 323e, Washington State University, Pullman WA 99164-6210, Tel: 509-338-5372.

Abstract

This article estimates the demand for mint-flavored gum products using grocery store sales data and accounting for consumers' valuation of quality. Unobserved product attributes, such as flavor quality, are important elements to consider when estimating the demand for gum. The estimation results suggest that gum is an inelastic product. A positive relationship between willingness to pay and unobserved quality was identified, implying that gum industry should be able to command a premium for higher quality mint flavored products.

Keywords: Quality Differentiation, Unobserved Product Attributes, Demand Estimation, Gum

1. Introduction

Chewing gum is one of the best performing segments within the confectionery market, and the global market for gum is forecast to reach US$20.7 billion by the year 2015 (GIA, 2011). The industry is characterized by its product innovation, focused on novel and unique flavors, new ingredients, different product shapes, varied colors, and distinctive packaging techniques (GIA, 2011). The production of gum requires ingredients such as a gum base, sweeteners and a variety of flavors. The gum base is usually a standard mix of synthetic latex and natural rubber extracted from the Sapodilla trees. This ingredient is not a source of product differentiation since only two different gum textures are commercialized, chewing and bubble gum, with 83% and 17% of the market share, respectively. The case of sweeteners is similar, just two categories can be identified, sugar and sugar-free gum, with a share of 42% and 58% respectively. Flavor is the major source of product differentiation. This is reflected in the high number of products available in the market with differentiated flavors.

Mint is the flavor that displays the largest market share, accounting for approximately 50%, followed by fruit flavored gum with 19% (Nielsen, 2005), of total market share. Mint oil is an important product to the Pacific Northwest (Washington, Oregon, and Idaho), which is responsible for 83% of the United States spearmint production and 50% worldwide.[1] However, the share of mint oil supplied by U.S. producers has been falling in recent years. Mint oil imports now account for approximately 25% of this market. Dealers buy cheaper and lower quality oil from China and India and then blend these oils with the more expensive high quality

[1] http://www.farwestspearmint.org/history.htm

U.S. oil to accommodate each gum manufacturer standard.[2] From 1993 to 2010, the number of acres harvested dropped 37% and the price per pound dropped 23%.[3]

The supply chain of mint oil involves three parties: (1) the mint oil producers who sell mint oil to the dealers, (2) the dealers who mix different mint oils to generate the final flavoring oil mixtures, and (3) gum manufacturers who buy the oil mixtures from dealers to produce gum. For each flavor of mint gum (e.g. "doublemint," mint splash, or cool mint), the gum manufacturer demands a specific mixture of mint oils from the dealer. This specific mint oil mixture is the result of blending mint oils from different qualities, which are measured in terms of the presence of oil components such as limonene, menthone, purene, esthers, among others. The less mixture of oils from different growing regions used in the final product leads to the higher flavor quality which is measured in terms of strength and duration.[4]

An investigation of the elasticity of substitution between high quality domestically produced oil and low quality imported oil is important to the U.S. mint oil industry, as their market share has been decreasing with the increase of imports and they have been losing negotiating power with gum manufacturers. However, it is not possible for the researcher to observe or measure differences in the quality of the mint oil mixed to produce each gum flavor, simply because each gum recipe is private information for the manufacturing firms. One alternative is to analyze the consumers' price elasticity for mint gum accounting for product heterogeneity and unobservable (to the econometrician) product characteristics such as flavor quality.

[2] Based on personal communications with Rod Christensen, President of Far West Spearmint Oil Administrative Committee, August 8, 2010.
[3] Change from 1993 to 2010. United States Department of Agriculture, National Agricultural Statistics Service.
[4] Typically, mint oil contains menthofuran. This substance reduces palatability. Oils that contain high levels of menthofuran are considered lower quality, whereas oils with lower level of this substance are considered high quality. Personal communication with mint dealer company representative, August 8 2010.

We assume that consumers have a predetermined ranking of mint flavors based on their previous consumption and their own preferences. Based on the differences in flavor profiles across products, consumers know which flavor delivers the highest utility in consumption. It is necessary to highlight that although the differentiation is purely subjective, the differences in flavor profiles for each gum product are real rather than merely perceived by the consumers. For a discussion of price competition in a setting when product differentiation is purely subjective, see Tremblay and Polasky (2002). From the literature on product differentiated oligopolistic models, we know that unobservable product attributes generate endogeneity of prices, leading to biased estimators. The present study considers the unobserved attributes surrounding flavor quality and compares the results using different estimation strategies.

Only a few studies have been conducted on brand choice and products using mint oil as flavoring ingredient, such as in gum and other mint flavored products[5]. For toothpaste, previous studies include Kaya et al. (2010), Gutierrez (2005), Yang et al. (2005) and Shin (2008), and for chewing gum Chung and Szymanski (1997). All of these studies examine factors that affect brand selection. However, no study has analyzed mint flavored product choice in the presence of unobservable product quality.

The aim of this study is to estimate the demand for mint-flavored gum accounting for the existence of unobservable flavor quality attributes. We postulate that demand for gum can be depicted by a discrete-choice model in an oligopoly context in which prices are endogenously determined by price-setting firms. We also consider the existence of product characteristics unobservable (to the econometrician), but fully considered by gum manufacturers when setting prices and by consumers when purchasing the products. We use 2005 quarterly aggregated gum

[5] An estimate of 45% of the total U.S. mint oil production is devoted to the production of gum and the same percentage to the production of toothpaste. The remaining 10% in used in other confectioneries, pharmaceutical applications, flavor for liqueurs, and aroma therapy (Mint, 2000).

retail sales data, arranged in a three-dimensional panel of quantities and prices for 49 contiguous states (Nielsen, 2005).

The article is organized the following way. In the next section, the data is described. Following that, we present two different modeling strategies: the multinomial logit (MNL) and the nested logit (NL) model. In the subsequent section, we introduce the estimation results and comparisons across the different modeling strategies. Next, counterfactual scenarios are simulated to estimate the changes in market shares derived from changes in the consumer's valuations of the unobservable attributes of the products. Finally, implications are discussed and conclusions are drawn.

2. Data

The database consists on bubble and chewing gum household purchases obtained from the AC Nielsen Homescan survey aggregated by brand for each state. Summary statistics for our data set are presented in Table 1. We define each state as an independent market with information for 2005 four quarters (each quarter is composed of 3 months)[6]. The definition of the level of aggregation that constitutes a market was made based on the information availability for each brand within the geographical aggregation. Although information for the state and county levels are both available, in most of the cases, there are not enough individuals to statistically represent each county. On the other hand, for each state, there is sufficient individual-level information.

The chewing gum industry is highly concentrated with three firms (Wrigley, Cadbury[7], and Amurol) dominating the market (see Table 2). The three-firm concentration ratio (CR3) is

[6] We choose the year 2005 because it contains more information about product attributes and consumers characteristics than any other year available in the survey.
[7] Cadbury is a subsidiary of Hershey.

almost 90%. The Herfindahl-Hirschman Index (HHI), calculated with considering "other producers" as not having a significant share, is 4364. According to the Federal Trade Commission (FTC), an HHI above 2500 indicates high concentration. We do not have information about most of the small producers, but we know that all of the "other firms" have less than one percent of the market and do not to affect the HHI substantially.[8]

In this case, in spite of the existence of many gum varieties and brands (67 in the dataset), most of them are produced by a few number of firms. Ninety seven percent of gum sales were grouped in 66 different brands, the remaining 3% of the sales observations were gum products sold by brands with a frequency of less than 100 observations and they were aggregated into one label as "other brands." Thus, the total of number of brands in the sample is 67.

Table 3 presents a description of the gum sales data by flavor. The average price (in dollars per unit) paid is 82 cents with a standard deviation of 34 cents. Mint flavored gum accounts for 51% of the total sales share. The highest average price is paid for sour flavored gums, $1.20, followed by variety[9] flavor. The prices for mint, fruit, and spice are substantially lower than variety, sour, and other flavors in Table 3[10]. In terms of the form or shape of the product, 94% of the gum is sold as pieces or sticks, the average price paid for the consumers is 87 cents for gum in pieces form and 75 cents for gum in stick form.

The mint flavor category (51% of the sample) is divided into three sub categories, peppermint (19.6%), spearmint (13.8%) and other, not recognizable or artificial types of mint (17.9%).[11] According to the U.S Department of Agriculture (Economic Research Service, 1995),

[8] In fact, the calculated HHI considering "other producers" as just one firm with 11% of the market share is 4481.
[9] Variety stands for packages of assorted flavors.
[10] The category "other" corresponds to those products with no available information about their flavor and for which was not possible to assign one by product name, such as: Gourmet, Holiday Stripe, Island Squeeze, Mystery Magic, and Radical Red.
[11] This category includes types of mint agglomerate products flavors, such as "Arctic Chill" and "Cool Frost" in which it was not possible to determine clearly which type of mint was used.

"In the United States, the taste of peppermint is preferred to spearmint. As a result, peppermint has more end uses than does spearmint. Peppermint is the number-one mint used in chewing gum, which is the most important use of mint," (page 11). Peppermint gum has a higher average price per unit than spearmint, 84 cents versus 78 cents (see Table 4). Classifying by texture, the data consists of 83% chewing gum and 17% bubble gum. With respect to the sugar content, almost 60% of the purchases registered in the sample are sugar free.

Based on the product diversity in our data, consumers' preferences are varied with respect to observable characteristics, such as flavor and form. Unfortunately, we as econometricians cannot observe all flavor quality attributes especially those in terms of strength and duration. Consequently, in order to model the consumers' behavior, we rely on the subset of observable attributes in addition to our knowledge of the existence of other unobserved attributes. The AC Nielsen survey data allow us to observe socioeconomic characteristics, including household size, age, income level, presence of children in the household, and marital status. These socioeconomic characteristics are used as covariates in the estimations.

3. Empirical Framework

For purposes of estimation, we define the combination of brand and flavor as a product. This definition allows us to divide the set of differentiated products into subsets of homogeneous products. Examples of brands with their flavors are: peppermint, used in brands such as Wrigley's Doublemint, Wrigley's Extra, and Wrigley's Freedent. Spearmint is used in brands such as Wrigley's spearmint or Care-Free Koolerz. Fruit flavors are used in brands such as Wrigley's Juicy Fruit, Adams Bubblicious, or Adams Dentyne Tango. Spicy flavors are used in

brands as Wrigley's Big Red or Adams Dentyne Fire. Also, we control for other elements such as the form of the gum and size of the package.

The use of a discrete choice model for estimation is advantageous for several reasons. First, it allows us to utilize aggregate-level information. Specifically, in the context of the demand estimation in presence of unobserved product attributes, considering that unobserved product attributes are correlated with prices, endogeneity of prices can be addressed in an aggregate discrete choice model. However, our main purpose is to examine whether and how the observed price variation can be attributed to unobserved product attributes rather than to understand consumers' preference heterogeneity. In this sense, the present study does not account for any individual consumer preferences. Second, using a discrete choice model allows us to solve the dimensionality problem by projecting the products onto characteristics space. Third, it provides a tractable link between consumer theory and econometrics, allowing one to study markets with differentiated products in a structural model framework. Finally, since the estimation must account for non observable product-specific characteristics or demand factors, the model allows for the possibility of prices being correlated with unobserved demand factors.

We start by assuming the indirect utility of consumer i, for product j depends on the characteristics of the product and the consumer, $U(x_j, \xi_j, p_j, v_i; \theta_d)$, where x_j and ξ_j are the observed and unobserved product characteristics, respectively. The price of each product is represented by p_j. Consumer-specific terms affecting utility are v_i and θ_d. The vector x_j represents the observable characteristics including flavor, if the product is sugar free, the form of the product, texture, and size. Vector ξ_j represents the unobserved product characteristics or the product attributes that the econometrician cannot measure or observe, but producers consider

when setting their prices, and the consumers take into the account to make their choices, e.g. the average quality of each product derived from the individual consumer's valuation.

Consider a specification for the log indirect utility function where the unobserved consumer specific taste parameters are captured by the error terms:

$$u_{ij} = x_j\beta - \alpha p_j + \xi_j + \epsilon_{ij} \tag{1}$$

The random term ξ_j can be interpreted as the mean of consumers' valuations of all the unobserved product characteristics, and the error term (ϵ) represents the distribution of consumer preferences around ξ_j. However, we must consider that gum manufacturers know their own product characteristics, including those that are unobserved to the econometrician. Factors, such as higher quality ingredients, are considered to estimate their production costs, and they likely use this information to set the prices of their products. In this way, product prices (p_j) are likely to be correlated with those unobservable product attributes (ξ_j).

Another source of market price variation is the competition among products. Greater variety of products in a specific market implies intense price competition, affecting the prices of the products negatively. Recent literature has incorporated the product assortment as an endogenous variable to the model (see Draganska, et al. 2008, Draganska, et al., 2009, Mazzeo 2002, Seim 2006, and Allender, et al. 2010). Taking advantage of the modeling strategy developed by this literature, we assume product assortment is exogenous and it affects the price.

Other strategies that affect product pricing may include, for example, the introduction of "private labels" which are brands developed by individual retailers as a strategic tool to compete with national brands. This strategy seems to increase retailers' profits and national brand prices

(Bontemps, et al., 2008, and Cotterill, et al., 2000). However, there is no record of private labels in the case of gum products.

3a. Multinomial and Nested Logit Models Setup

Assuming that the error term is independently and identically distributed (*i.i.d.*) across products and consumers as an "extreme value" distribution, we can represent the traditional market shares multinomial logit (MNL) model in the usual way:

$$s_j = \frac{e^{\delta_j}}{\sum_{k=0}^{J} e^{\delta_k}} \tag{2}$$

The functional form characterized in Equation 2 is the closed-form of the MNL model, representing the probability of choosing good j among all other goods, including an outside good. Where $\delta_j = x_j\beta - \alpha p_j + \xi_j$ stands for the mean utility for product j, and $k = 0$ is the outside good that represents the consumer's expenditure in any other goods but gum. In this case, we use the state population as reference for potential market for the product to define the outside good.

Following Berry (1994), demand can be estimated by "inverting" the market-share equation to find the implied mean levels of utility for each good. A feature of the method is that allows for estimation by traditional instrumental variables (IV) techniques. By normalizing the mean utility of the outside good to zero and assuming the relation between observed and predicted market shares is invertible, we can represent this relation in a linear form as:

$$ln(s_j) - ln(s_0) = \delta_j = x_j\beta - \alpha p_j + \xi_j \quad (3)$$

On the left hand side of the equation is the observed market share of each product *j* relative to the outside good, and on the right hand side is the mean utility for product *j*.

As we pointed out previously, gum manufacturers know their own product characteristics, including those unobserved for the econometrician, and they use this information to set the prices of their products. In this way, product prices (p_j) are likely to be correlated with those unobservable product characteristics (ξ_j), so the explanatory "observable" variables are not completely exogenous to the model, specifically the price, generating an identification problem due to endogeneity.

An additional potential problem with the MNL approach is that assumes that the probability of each alternative is related equally with the probability of other alternatives, this is the independence of irrelevant alternatives (IIA) assumption does not hold. An approach to this problem is allowing consumer tastes to be correlated across products j in a restricted fashion by using the nested logit (NL) modeling approach.

Preserving the assumption that tastes are distributed via extreme value, but allowing consumer tastes to be correlated across products j in a restricted way, we also setup a nested logit (NL) model. We can then group gum products into exhaustive and mutually exclusive sets according to their flavor $g = 0, 1, \ldots, 6$, where the outside good $g = 0$, is assumed to be the only member of group 0[12]. If we denote the set of products in group g as $\mathcal{g}_g$, for product $j \in \mathcal{g}_g$, the indirect utility of consumer i can be represented by:

[12] The rest of the categories were assigned as: fruit (g=1), mint (g=2), other (g=3), sour (g=4), spice (g=5) and variety (g=6).

$$u_{ij} = \delta_j + \zeta_{ig} + (1 - \sigma_g)\epsilon_{ij}, \tag{4a}$$

Where $\delta_j = x_j\beta - \alpha p_j + \xi_j$ and ϵ_{ij} is i.i.d. extreme value. For consumer i, the variable ζ is common to all products in group g and has a distribution function that depends on σ_g, with $0 \leq \sigma_g < 1$. Parameter σ_g measures similarity of products within each group. As the parameter σ_g approaches one, the within-group correlation of utility levels goes to one, and as σ_g Approaches to zero, the within group correlation goes to zero.

We can interpret Equation 4a as a model involving random coefficients ζ_{ig} only on group-specific dummy variables. That is, if d_{ig} is a dummy variable equal to one if $j \in g_g$ and equal to zero otherwise, we can rewrite Equation 4a as:

$$u_{ij} = \delta_j + \sum_g d_{ig}\zeta_{ig} + (1 - \sigma_g)\epsilon_{ij} \tag{4b}$$

Thus, we can derive an analytic expression for mean utility levels similar to the MNL model represented by Equation 3 with just one additional term for each group as[13]:

$$\ln(s_j) - \ln(s_0) = \delta_j \equiv x_j\beta - \alpha p_j + \sum_{g=0}^{6} \sigma_g \ln(\bar{s}_{j/g}) + \xi_j \tag{5}$$

The new element compared with Equation 3 is the natural log of the within-group market share ($\bar{s}_{j/g}$). Using the NL model represented by Equation 5, the estimates of β, α, and σ_g can be obtained from a linear instrumental variables (IV) regression of differences in log market shares on product characteristics, prices, and the log of the within-group share.

[13] For details refer to Berry (1994).

However, as in the case of prices in MNL model specification, since the within-group share is also related with the unobserved characteristics via consumer preferences, $\bar{s}_{j/g}$ is endogenous, suggesting the need for additional exogenous variables that are correlated with the within group share but not with the unobserved product valuation. In both specifications, the MNL given by Equation 3 and the NL given by Equation 5, the error term (ξ_j) is a structural component of the model and represents the average consumer valuation of the unobserved product attributes such as quality.

Summarizing, the differences between models is that the NL model relaxes the assumptions of the MNL model. The basic idea of the NL model is to extend the MNL model in order to allow groups of alternatives to be similar to each other in an unobserved way; that is, to have correlated error terms (Heiss, 2002). In addition, as highlighted by Berry (1994), the NL model may be preferred when the researcher wants to model substitution effects depending only on predetermined classes of products, as it is the case in this project.

Basically, both the two models describe how market shares are generated from modelers' mind. Evaluated at the "true" value of parameters, the difference between the predicted (from the models) and observed market share depends only on the unobserved product attributes. These models can then be identified if we have a set of instruments, conditional on which the mean of the unobserved product attributes is zero. The relevance of the two models is that in both cases MNL and NL, Equations 3 and 5, the specification is linear in parameters, allowing the use of traditional instrumental variables (IV) to eliminate endogeneity.

3b. Instruments

Recall that the endogeneity originates in the relation between prices and the product specific characteristics. We use a set of instrumental variables (IV) that are related with the prices but

not with the unobserved characteristics captured by the error term (ξ_j). As instruments, we use two groups of variables: the prices of the same products in other markets (Hausman, 1996) and the distance from the production plants to account for the geographical location of each market. In this sense, prices of brand j in different markets will be correlated due to the common marginal costs, but they will be uncorrelated with the market-specific valuations of the product[14]. The distances from the production plants as IV for prices are proxy variables for the transportation costs, which are determinant of the supply function, so it is related with the prices but not with the unobservable market-specific valuation of the product[15]. Assuming the use of the correct instrumental variables the conditional mean of the error term equals zero given the repressors, so we can interpret the coefficients as the structural parameters of the model, see Cameron, et al. (2005). In the case of the NL model, where the within-group share is also endogenous, we use the within number of products as a variable of the market structure or the degree of competition. The market structure is correlated within each group, but not with the unobservable attributes[16].

By construction, the number of products is inversely correlated with the within market shares, but uncorrelated with the market-specific valuation of the product. From economic theory, we can support the use of number of products within the market as instruments by framing the situation as a sequential-decision game. The decision made by each firm on the number of products is made before the realization of consumer preferences. At this stage, firms do not know consumers' preferences. Hence, the number of products is not related with the

[14] We try different sets of instruments like: market structure (number of firms), rival products characteristics and wages.
[15] We estimate a first stage regression using other prices and distances as determinants of prices for each product. In both cases the coefficients were statistically significant.
[16] Other instruments evaluated were rival product characteristics, some demographics, and advertisement expenditure. Even though the results were not very different from what we present, the tests show that these other instruments where weaker than those we use.

consumers' valuations of the unobservable product attributes. Even though this approach allows consumer tastes to be correlated across products, giving more reasonable substitution patterns compared to the MNL model, the grouping of products or the choice of sets is made a priori without any basis in theory or empirical support, but instead is justified by the market segments we are interested in analyze, according to flavor. Given the set of instruments in both models, we have more moment conditions than parameters to be estimated, so it is a case of over-identified parameters that we estimate using the two stages least squares (2SLS) estimator.

4. Estimation Results

In this section, we present the results of the estimations of the MNL and NL models. The results for the initial benchmark estimation are presented in column 1 in Table 5, these coefficients correspond to an estimation ignoring the non-orthogonal relation between covariates and error term, specifically the relation between prices and unobservable product attributes. The coefficient for price is not statistically different from zero, which suggests that the demand for gum is completely inelastic to prices, which is inconsistent with economic theory. This anomaly is likely caused by an endogeneity problem, and could be explained if unobserved product quality is considered. As in Trajtenberg (1989), prices appear to have a positive, or in this case nil, effect on consumers.

From the estimation results, it is also possible to see the positive marginal utilities generated by some of the product characteristics such as flavor, sugar and texture. In the case of the size, the marginal utility is negative, suggesting that consumers prefer to carry smaller packages of gum. The second column in Table 5 presents the coefficients for the MNL model represented by Equation 3, now controlling for the presence of endogenous variables by using

the prices of other products and the distances from the production plants as IVs, specifically to address the endogeneity of prices (column 2). In this estimation, the results are more compatible with economic theory. The coefficient for price is negative, as expected, and statistically significant.

Column 3 presents the estimation results for the NL model represented by Equation 5. As in the previous case, we use prices of other products and the distances from the production plants as IVs to address the endogeneity of prices, but we also use the within-market number of products as instrument for the within-market shares to allow for product heterogeneity. In this case, the coefficient on price increases in absolute value when the model incorporates the IVs. An interpretation of this finding is that products with higher unmeasured quality characteristics sell at higher prices. These results suggest that ignoring the correlation between price and the demand error can lead to findings of upward sloping demand curves and other anomalies. In this case, we notice that the coefficient for price in the first estimation is not statistically different from zero. In contrast, when endogeneity is accounted for with the use of IVs, the coefficients are negative and statistically significant, which is consistent with economic theory.

In terms of the coefficients of the observable product characteristics, when we account for endogeneity, the form of the product, the size, and the flavor are statistically significant in explaining the market shares. If the gum is in the form of a stick, the market share decreases with respect to the other forms of products. If the size of the product decreases, the market share of the product increases. If the product is mint flavored, the probability of a higher market share also increases. In terms of consumers' socioeconomic characteristics, the income level and the age both decrease the probability of purchasing the good (see Table 5). The decrease in gum consumption with the income level might be related to substitution effects with respect to other

type of non-gum breath freshener products, such as Althoids or Mentos, which are not included in the sample.

One can also understand the importance of unobservable characteristics by examining the fit of the logit demand estimation. The explanatory power of the model increases after product heterogeneity is allowed by the nesting strategy. In the case of the MNL-IV model, the R^2 statistic is 30%. For the NL-IV, the R^2 statistic increases to 52%, reducing significantly the percentage of the variance in mean utility levels associated with the unobserved product characteristics (see Table 5).

In Table 6, we present results indicated similarities of products within each group. Recall that the parameter $0 < \sigma_g < 1$ measures this similarity. When the parameter is closer to one, the products within the group are more similar. When it is closer to zero, the products within each group are more heterogeneous. The empirical results support the theory, and all the coefficients are within the unit interval. Additionally, products with fruit, mint, or even spice flavor, are closer to each other within their group than the products classified as sour, variety, and other. This group of categories is expected to be the most heterogeneous because in each of the products in this category different flavor profiles are present. For example, sour gums can have fruit, mint or even spice flavor. In the case of the variety, or assorted packages, the parameter value is half way between the two extremes. We calculate the elasticities for different flavor profiles. We find that mint-flavored gum is slightly more inelastic (-0.1003) than non mint (-0.1203), but the difference in the coefficients is not statistically significant.

With respect to income elasticities, the parameters are small for all flavor categories, but statistically different from zero, on the order of 0.00023 and 0.0004 for mint and non mint flavor categories, respectively. The small, almost nil income elasticity, supports the fact that in the case

of gum products, the budget constraint is not binding. We used household income for the estimation, which is high compared with the price of gum products.

After controlling for unobservable product characteristics, just in the case of mint flavored products, we observe some correlation between product prices and the product unobserved quality in the case of mint, 21%. This is the consumers' willingness to pay increase with the quality of mint flavor (see Figure 1). The study of gum demand using this approach, contrast with the larger computational burden of the Berry, Levinsohn, and Pakes (BLP) model (Berry, et al., 1995). The NL model may be preferred when the researcher wants to model substitution effects depending only on predetermined classes of products (Berry, 1994), such as flavors, in this case.

4a. Counterfactual Exercises

In this section, using the estimation results, we answer hypothetical questions about the market. Counterfactual scenarios can help us to understand the role of the unobservable product attributes such as quality on the market shares. We partition the sample according to the consumer's quality valuation in two sub groups: high-quality goods (HQ) and lower-quality goods (LQ), according to the distribution around the mean of the consumers' valuations of the unobserved product characteristics. We find that the actual (and predicted) market shares for the first group in average is 3.5%, meanwhile for the second group is 1.2%. In terms of the predicted market shares, considering changes in the factors that affect the consumer's quality valuation, e.g. The quality of the inputs used in the production process, we note the following: If the products are the same, the consumer's quality valuation is the average for all products, then the market shares of the industry are more homogeneous, around 1.8% for each product.

Consider a case not as extremely homogeneous as the previous one. If the perceived differentiation by the consumers for HQ products reduces by half, and at the same time the perceived differentiation by the consumers for LQ products doubles, the market shares of the industry are more homogeneous, 2.3% for HQ and 1.3 for LQ.

5. Conclusions

In this study, we estimate demand for mint-flavored gum products with retail sales data while accounting for consumers' valuation of unobserved quality and other product attributes. We find that the unobserved product attributes are important to consider when estimating the demand for heterogeneous products. With a nested logit model, almost 47% of the variance in mean utility levels is associated with the unobserved product characteristics such as flavor quality. In terms of the parameter estimates, we find that the presentation of the product in terms of its form is important to consumers. Smaller packaged products have an advantage in terms of their market shares, and socio-economic characteristics such as income level and age, both decrease the probability of purchasing the good.

Our estimation results suggest that mint-flavored gum is more inelastic to changes in price than other flavors, and there exists an important variability in the valuation for quality among gum products. Given that mint-flavored gum is more inelastic to changes in prices than other flavors and the positive relationship identified between willingness to pay and unobserved quality, mint gum industry should be able to command a premium for higher quality product. The finding that U.S. consumers are willing to pay a premium for higher quality products is useful information for the U.S. mint oil industry, as they compete with cheaper foreign imports and lose negotiating power with gum manufacturers. Even though the tradeoff between lower

costs and higher quality is not going to disappear, consumer preferences for high quality products seems to guarantee the existence of a significant market share for high quality mint oils in the U.S. gum market.

Bibliography

Allender, W. J., & Richards, T. J. (2010). Market Diversion and Market Power: California Eggs. *Review of Industrial Organization*, 36 (1), 37-58.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics,* 25 (2), 242-262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63 (4), 841-890.

Bontemps, C., Orozco, V., & Réquillart, V. (2008). Private Labels, National Brands and Food Prices. *Review of Industrial Organization*, 33 (1), 1-22.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics Methods and Applications*. (C. U. Press, Ed.) New York.

Chung, S.-W., & Szymanski, K. (1997). Effects of Brand Name Exposure on Brand Choices: An Implicit Memory Perspective. *Advances in Consumer Research*, 24, 288-294.

Cotterill, R. W., & Putsis, W. P. (2000). Market Share and Price Setting Behavior for Private Labels and National Brands. *Review of Industrial Organization*, 17 (1), 17-39.

Draganska, M., Mazzeo, M., & Seim, K. (2009). Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *Quantitative Marketing and Economics*, 7 (2), 105-146.

Draganska, M., Misra, S., Aguirregabiria, V., Bajari, P., Einav, L., Ellickson, P., et al. (2008). Discrete Choice Models of Firms' Strategic Decisions. *Marketing Letters*, 19 (3-4), 399-416.

Economic Research Service, U.S.D.A. (1995). *Peppermint and Spearmint: An Economic Assessment of the Feasibility of Providing Multiple-Peril Crop Insurance.* Available at http://www.rma.usda.gov/pilots/feasible/PDF/mint.pdf

Global Industry Analysts, Inc. (2011). *Gums: A Global Strategic Business Report.* Available at: http://www.prweb.com/pdfdownload/8067444.pdf , Retrieved November 30, 2011.

Gutierrez, B. P. (2005). Determinants of Toothpaste brand choice in urban Philippines. *Philippine Management Review*, 12, 45-71.

Heiss, F. (2002). Structural choice analysis with nested logit models. *Stata Journal*, 2 (3), 227–252.

Kaya, T., Aktaş, E., Topçu, İ., & Ülengin, B. (2010). Modeling Toothpaste Brand Choice: An Empirical Comparison of Artificial Neural Networks and Multinomial Probit Model. *International Journal of Computational Intelligence Systems*, 3 (5), 674-687.

Mazzeo, M. J. (2002). Competitive Outcomes in Product-Differentiated Oligopoly. *The Review of Economics and Statistics*, 84 (4), 716–728.

Mint, A. (2000). *American Mint.* Retrieved from American Mint: http://www.mintstills.com/peppermintspearmint.html

Nevo, A. (2000). A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand. *Journal of Economic and Management Strategy*, 9 (4): 513-48.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica ,* 69 (2), 307-42.

Nielsen, A. (2005). Homescan Data.

Rasmusen, E. (2007). The BLP Method of Demand Curve Estimation in Industrial Organization. Available at http://www.rasmusen.org/published/blp-rasmusen.pdf

Seim, K. (2006). An Empirical Model of Firm Entry with Endogenous Product-Type Choices. *The RAND Journal of Economics*, 37 (3), 619-640.

Shin, S. (2008). *Disentangling Preferences, Inertia and Learning in Brand Choice Models*. PhD dissertation . Rochester, New York: University of Rochester.

Trajtenberg, M. (1989). The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners. *Journal of Political Economy* , 97, 444-479.

Tremblay, V. J. & Polasky, S. (2002). Advertising with Subjective Horizontal and Vertical Product Differentiation. *Review of Industrial Organization* , 20 (3), 253-265.

Yang, Z., Zhou, N., & Chen, J. (2005). Brand Choice of Older Chinese Consumers. *Journal of International Consumer Marketing*, 17 (4).

Table 1 Definition of Variables and Summary Statistics

Variable Name	Type	Description	Mean	Std. Dev.	Min	Max
Price	Numerical	Per unit	0.824939	0.347378	0.005	2.89
Units	Numerical	Number of units of the item purchased.	1.527541	1.211985	1	29
Product Characteristics						
Form	Categorical	Pieces, sticks or other.	1.587074	0.603887	1	3
Flavor	Categorical	Mint, fruit, spice, variety, sour, other flavors.	2.230704	1.701977	1	6
Mint Flavor	Categorical	Peppermint, spearmint, and other	1.049917	1.173701	0	3
Texture	Categorical	1 if chewing gum and 0 if bubble gum	0.828782	0.376701	0	1
Brand	Categorical	Brand code	106535.4	65655.44	4817	196550
Producer	Categorical	Name of the producer				
Volume	Categorical	Size of the package in ounces.	17.15474	17.68429	1	380
Sugar Content	Dummy	1 if Sugar-Free	0.585412	0.492653	0	1
Coupon	Dummy	1 if Coupon	0.100455	0.300606	0	1
Household Characteristics						
Household Size	Numerical	Number of individuals in the household.	2.867244	1.447626	1	9
Income Level	Categorical	Income intervals begin with annual incomes under $5,000 and the highest interval is $100,000 and over.	19.41256	5.642835	3	27
Children	Dummy	1 If children under 18 in the household	0.400851	0.490073	0	1
Marital Status	Dummy	1 If married	0.678326	0.467121	0	1
Age	Categorical	Ages interval starting at 25, top is 65+	6.47998	1.835488	1	9
Other Variables						
Projection61k	Numerical	Expansion factor.	3184.664	3427.413	139	31230
Producers	Categorical	Code for each producer				
Region	Categorical	East, west, south, central.	2.74584	0.915828	1	4
State	Categorical	49 contiguous states.	27.8706	16.32448	1	56

Source: AC Nielsen Homescan survey

Table 2 Market Shares by Quarter (2005)

Producer	Quarter				Average
	1	2	3	4	
WRIGLEY'S	60.0	60.0	58.0	55.0	58.3
HERSHEY	25.0	28.0	29.0	29.0	27.8
AMUROL	1.8	1.5	1.8	2.5	1.9
OTHER PROD	14.0	11.0	12.0	14.0	12.8
C3	**86.8**	**89.5**	**88.8**	**86.5**	**87.9**

Source: AC Nielsen Homescan dataset expanded by projection61k.

Table 3 Price Distribution by Flavor

Flavor	Freq.	Price	S.D.
Mint	51%	0.79	0.28
Fruit	19%	0.75	0.30
Spice	12%	0.78	0.35
Variety	4%	1.1	0.57
Sour	2%	1.2	0.42
Other	13%	1.0	0.42
Total	100%	0.82	0.34

Source: AC Nielsen Data, calculations by the authors.

Table 4 Differences by Mint Flavor

Mint Types	Share	Avg. Unit. Price	S.D. Unit Price
Peppermint	19.54	0.84	0.31
Spearmint	13.85	0.78	0.27
Other Mint	17.92	0.74	0.26
No Mint	48.58	0.86	0.40

Source: AC Nielsen Data, calculations by the authors.

Table 5 Gum Demand Estimation Results

Covariates	(1) MNL		(2) MNL-IV		(3) NL-IV	
Price	-0.030		-0.266	***	-0.130	*
	(0.022)		(0.079)		(0.068)	
Form (Piece)	0.008		0.269	***	0.055	
	(0.024)		(0.067)		(0.064)	
Form (Stick)	0.011		-0.163	**	-0.112	*
	(0.038)		(0.074)		(0.065)	
Size	-0.029	***	-0.120	***	-0.088	***
	(0.010)		(0.026)		(0.023)	
Sugar-Free	0.161	***	0.001		-0.024	
	(0.041)		(0.050)		(0.044)	
Texture	0.184	***	0.444	***	0.006	
	(0.026)		(0.058)		(0.062)	
Mint	0.316	***	0.752	***	0.723	***
	(0.057)		(0.053)		(0.071)	
Incomes	-0.007		-0.093	***	-0.072	***
	(0.007)		(0.022)		(0.019)	
Household Size	0.016		0.022		0.022	
	(0.011)		(0.030)		(0.025)	
Children	-0.008		-0.032		-0.034	
	(0.011)		(0.032)		(0.027)	
Age	-0.036	***	-0.076	***	-0.053	***
	(0.009)		(0.024)		(0.020)	
Observations	4403		4403		4403	
R^2	0.118		0.305		0.529	

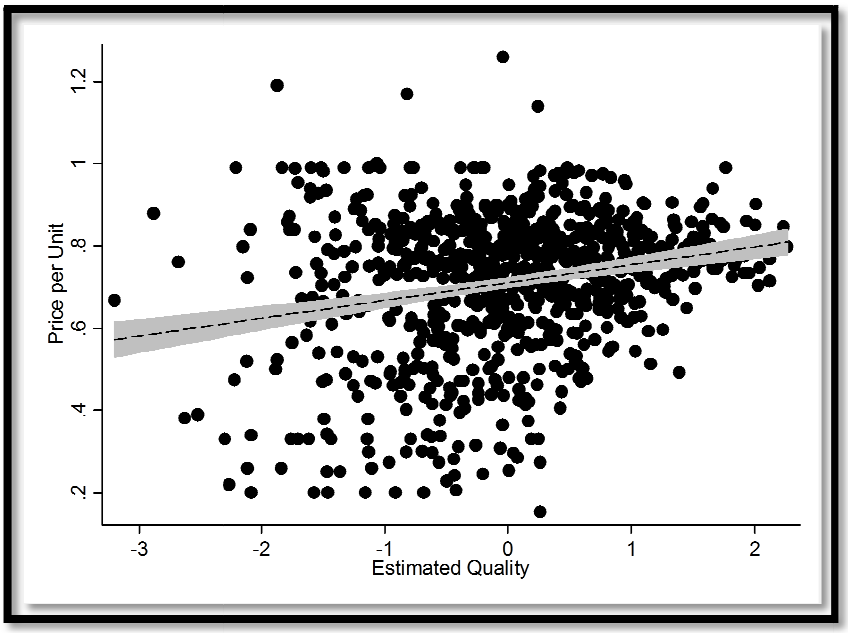
Legend: * p<.1; ** p<.05; *** p<.01, controlling for number of firms, producers, time period, and region. Standard Errors in parenthesis. Calculations by the authors.

Table 6 Products Similarity within Groups

Parameters	**NL-IV**	
σ(g=Fruit)	0.148	**
	(0.072)	
σ(g=Mint)	0.156	**
	(0.064)	
σ(g=Other)	0.290	***
	(0.076)	
σ(g=Sour)	0.303	***
	(0.101)	
σ(g=Spice)	0.152	*
	(0.083)	
σ(g=Variety)	0.268	***
	(0.078)	

Legend: * p<.1; ** p<.05; *** p<.01, controlling for producers, time period, and region. Standard Errors in parenthesis. Calculations by the authors.

Figure 1 Prices and Estimated Quality: Mint Flavor



Source: AC Nielsen Data, calculations by the authors.