ABSTRACT

GULSEVEN, OSMAN. A Hedonic, Hedonic Metric and Logistic Approach to Estimating Demand for Fluid Milk Products Using Micro Level Data. (Under the direction of Michael Wohlgenant.)

This study analyzes the market for functionally enhanced milk products. The models in each paper are tested using the AC Nielsen retail milk market demand data. The continuous product attributes that characterize each milk type are derived from USDA Nutritional Nutrient Database.

The first paper investigates the factors affecting the demand for soymilk products in the dairy market based on a two-stage hedonic model. In the first stage, the relationship between the prices of dairy products and the attributes of these products are exploited to derive the marginal implicit attribute prices. In the second stage, these marginal prices are used along with the information on households' demographic background to explain the demand for product attributes. Our results indicate that although the soymilk attribute has a negative value itself, since soymilk is lactose/cholesterol free (LFCF) and mostly organic, it has a higher price premium than other milk types. Moreover, as the negative value of soymilk attribute diminishes, consumers purchase more soymilk. Among the demographic variables, presence of children, education level and age of the household have significantly positive impact on soymilk demand.

In the second paper, we introduced the concept of Hedonic Metric (HM) approach as an approximation method to estimate the price elasticities in classical traditional models. In these models, the number of estimated parameters increases exponentially with the number of variables included in the model. One method suggested by Pinske, Slade, and Brett (2002) suggests using spatial distances as a neoclassical approach to estimate the elasticities. Rojas and Peterson (2008) applied the Distance Metric (DM) method to the beer market using different combinations of distances. Although their method is insightful, the choice of distances is ambiguous and depends on prior judgments about data and trial/error

methodology. Hedonic Metric (HM) method applied in this paper is practical and significantly reduces the number of parameters. The Hedonic Metric approach is compared with the Distance Metric approach to see which method gives better approximations to original LA/AIDS and RM models.

In the last paper, we applied a two-stage logistic model to analyze consumer attitudes towards specialty milk types. Although the shares of specialty milk products are low, demand for these products has been increasing rapidly over time. Logistic models that make use of hedonic prices enable us to understand how hedonic attribute prices and household characteristics affect demand for soymilk, cholesterol free/lactose free (CFLF) milk and organic milk. In the first stage of our estimation we exploit the relationships between milk prices and milk attributes to derive marginal implicit attribute prices. In the second stage, we use these implicit prices along with the demographic information on households to estimate the effects of these variables on households' decisions to purchase specialty milk types. Our results indicate that while CFLF attribute and soy attribute are complementary with each other, the organic attribute is a substitute between CFLF and soy. Moreover household demographics significantly affect purchase decisions. Minority households have a much higher probability of purchasing specialty milk types than white households.

A Hedonic, Hedonic Metric and Discrete Approach to Estimating Demand for Fluid Milk Products Using Micro Level Data

by
Osman Gulseven

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Economics

Raleigh, North Carolina

2008

APPROVED BY:

Dr. Barry Goodwin

Dr. Nick Piggott

Dr. Raymond Palmquist

Dr. John Monahan

Dr. Michael Wohlgenant
Chair of Advisory Committee

DEDICATION

To

my dear parents, Ilyas and Muazzez Gulseven

and

my lovely sister, Necibe Nur Tarkak

BIOGRAPHY

Osman Gulseven was born on August 4, 1981 as the son of Muazzez and Ilyas Gulseven. He lived his childhood in Ankara, capital of Turkey, with his younger sister Necibe Nur. Following his graduation from Ankara Ataturk Anatolian High School in 1998 he enrolled at Bilkent University with a full scholarship. He received his Bachelor of Science degree in Industrial Engineering in 2003. Upon earning his bachelor degree he was appointed as a teaching assistant in economics department at Middle East Technical University. In 2004 he was awarded a full scholarship to jump start his graduate school career at North Carolina State University Economics Department. While working on his PhD he earned his Master degree in Economics in 2006. During his PhD studies he was also an instructor of economics and he received the Certificate of Accomplishment in Teaching in 2008. His research interests concentrate on applied microeconomics, agricultural economics and industrial organization.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my dissertation chair Michael Wohlgenant for his valuable guidance throughout my research. His continuous support and creative suggestions made it much easier to accomplish this dissertation. He has always been a very pleasant person to work with. I would also like to thank my advisors: to Nick Piggott for showing me how to get the perfect data I need for my research, to Raymond Palmquist for his insightful comments about hedonic models, to Barry Goodwin for his expertise on the topic, and to John Monahan for being a great instructor who helped me to become an expert in statistical analysis. I would also like to thank Matthew Holt for continuously reminding us 'PhD is about not giving up'. It was a great pleasure to take Microeconomics courses with him. Surely these courses raised my curiosity in this area in which I ended up doing my dissertation. Xiaoyong Zheng's deep knowledge of industrial organization also deserves a great acknowledgement. I would also like to thank David Flath for his valuable career advice as the graduate program director.

Special thanks go to my old friends who have always been extremely supportive. Ozgun Ekici has not only been a great friend but also provided me excellent career guidance since high school years. I feel very lucky to share the same dorm room with him during the last year of my undergraduate education at Bilkent University. His determination to start graduate school career also proved to be my motivation. Our limitless discussions over phone provided me a much better perspective of economics. If I did not meet him, my life could have shaped in a totally different direction. I also wish to express my appreciation to my childhood friend Mihrac Akbas who showed me that education is just a step towards success. His financial success gave me a great insight on how real economy works; we all need risk-taking entrepreneurs who will take the first step and the rest comes itself. Among my best friends, Sercan Teleke has shown me a perfect example of hard-working graduate student. His ambition for published research deserves great respect. His true friendship during my research period has given me the fresh air I needed from my research.

Table of Contents

List of Tables

List of Figures

Chapter 1

1. A Hedonic Approach to the Demand for Soymilk Using Scanner Data

Introduction

The retail dairy market is characterized by many consumers purchasing varieties of dairy products, frequently purchasing multiple quantities. Because of this type of consumer behavior, we modified the classical hedonic models such that now we are able to describe the multiple-type, multiple-unit purchase behavior. The model is extended so that, each week the household chooses a variety of dairy products in order to maximize utility derived from the consumption of product attributes.

The soymilk and organic milk types have lower market shares than other milk types. However, as the prices of traditional dairy products have increased over time, the shares of the soymilk and organic milk have also increased. This trend is manifest in the AC Nielsen retail milk market demand data, which give purchase information for a panel of over thirty thousand participants between years 1998 and 2005. From this panel, we selected a core sample of about three thousand households who participated regularly in the panel during 2002-2005 periods. These data is merged with the product contents derived from USDA Nutrient Database, to get continuous product attributes including protein, fat, and carbohydrate content.

Unlike previous applications of hedonic models which use only cross sectional data, our access to AC Nielsen Homescan Panel enables us to aggregate the consumption for each week and for each location. By doing so we obtain significant variation in data which allows us to see the effects of household demographics on the amount of soymilk they purchase. We divided our data into 8 regional groups based on their region and whether the household is located in a large metropolitan area. Our final demographic panel includes significant amount of information about households' background for each region/week combination. Using this panel along with the hedonic prices derived from first-stage regression enables us to estimate demand for product attributes.

Ladd and Suvannunt (1976) show that the price of each good can be characterized by a set of nutritional attributes. If we define this attribute set as $x = [x_1, \dots, x_k]$ then the price of a good can be explained by the relative unit prices and quantities of characteristics. In this case, a functional relationship exists between the price of a good and its characteristic vector x such that $p = f(x) + \mu$ where μ is the error vector.

The marginal implicit prices of characteristics can be calculated as the partial derivatives of the hedonic functions.

$$\frac{\partial p}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} \; i = (1 \dots k) \tag{1.1}$$

If we define the product prices in log terms, an analogous functional relationship can be expressed as $\log p = f(x) + \mu$. In this semi-log form the partial derivative of the hedonic function is multiplied by the price of the product to get the hedonic price of each characteristic

$$\frac{\partial \log p}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} \; i = (1 \dots k) \; or \; \frac{\partial p}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} p \tag{1.2}$$

We used the semi-log form since it is reasonable to assume that the value of an attribute depends on the product it is located. Since we observe different product prices in each week, semi-log form also enables us to enrich our data by creating additional variation in attribute prices.

The first-stage hedonic regression results indicate that marketing variables based on product labels explain as much variation in the prices as the nutritional contents of the products. The soy attribute itself has a negative price suggesting that consumers are willing to avoid it. However, the price of organic label and LFCF attribute significantly dominate the negative effect of soy taste thereby explaining the price premium on soymilk. The amounts of cholesterol and sodium have negative prices which indicate that they are undesirable attributes.

Once we derive the attribute prices we can estimate the demand for these attributes using attribute prices along with household demographics. The second-stage demand equation for each attribute can be specified as

$$q_i = g(attribute\ prices, demographics) + error \qquad (1.3)$$

The results from the second-stage hedonic regression reveal that, as consumers get more concerned regarding the health effects of lactose/cholesterol and as they get more used to the soy taste, they buy more soymilk. Also, regional and demographic differences affect households' demand for soymilk products.

The next section in this paper includes an extensive literature review regarding the demand for soymilk, organic milk and also the effects of marketing on the demand for dairy products along with applications of hedonic models. In the model section, we derive the hedonic demand model for soymilk based on hedonic price theory. The detailed derivation of the data, sampling, aggregation issues, product attributes and household characteristics used in the estimation are explained in data section. The first-stage estimation results for hedonic prices and the second stage estimation results for attribute demands are explained in the results section. Finally, the paper concludes with a summary of findings and directions for future research.

Literature Review

As indicated by Palmquist (2003), hedonic models are widely used in property value models. In the first stage, prices of properties are usually explained in terms of property attributes. In the second stage, household characteristics are introduced along with the first-stage hedonic prices to estimate demand for attributes. This method is especially helpful in estimating the welfare effects of changes in attributes where explicit prices for attributes do not exist. The hedonic model can be used to estimate the value of clean air (Palmquist and Israngkura 1999), to identify land as a differentiated factor of production (R. Palmquist

1989), or to estimate the demand for characteristics of housing (R. Palmquist 1984). In retail markets, hedonic models can also be used to estimate unobserved attribute prices for differentiated products.

Ladd and Suvannunt (1976) applied the hedonic price analysis to consumer level prices for food items. Their model suggests that the retail price of a product is a weighted linear combination of the product attributes. Moreover, the utility of the consumer depends on the product attributes. Therefore, consumers maximize utility by selecting products that maximize the sum of utilities derived from each attribute. Using the national average retail price of 31 food items they estimate their model. The attributes of beef, veal, chicken, and dairy products are measured in terms of water, energy, protein, fat, carbohydrate, mineral and vitamin content. The results indicate that protein and potassium are highly appreciated by consumers whereas phosphorus and ascorbic acid are undesirable characteristics with negative shadow prices. Another paper by Ladd and Martin (1976) analyze a neoclassical cost minimizing firm model to test for the hypothesis that input prices are determined by the total value of input's characteristic attributes. Using this model to evaluate the current corn grading scheme, the authors find that there are some properties which don't affect the mixing scheme. They suggest that the current system needs to be updated in order to achieve sign optimality. Hager (1985) applies a similar hedonic model to the demand for meat products.

Nimon and Beghin (1999) apply a hedonic analysis to analyze the factors that affect retail prices in the organic cotton market. They estimate the hedonic prices for different characteristics using retail order apparel catalog data. Retail catalogs were used to obtain information about categorical variables such as item types, catalog types, gender of apparel, age range, type of cotton used and type of dye processed during the production along with the suggested retail prices. The model is specified as a semi-log hedonic model where prices are normalized by the amount of cotton included in the item:

$$\begin{aligned}\ln\left(\frac{price}{fiber\ content}\right) = \beta_0 &+ \beta_1 items + \beta_2 catalog + \beta_3 gender \\ &+ \beta_4(age\ categories) + \beta_5(dye\ types) + \beta_6 organic\end{aligned} \tag{1.4}$$

They find a significant willingness to pay for organic labeled apparels and significant discount for “no-dye” label. They also test for consumers’ attitudes towards health related concerns but find no evidence of price premium for health related claims in the apparel market.

An article by Nti and Larweh (2003) explored the sensory characteristics of flavored soymilk samples in Ghana. The authors applied a survey of 30 individuals to test for the acceptability of different soymilk flavors. By adding different flavors in different ratios, they evaluate the consumers’ response regarding sensory characteristics such as color, taste, mouth feel, and aroma. The addition of flavors at different concentration levels gives an estimate of the optimal level of flavor in the soymilk.

Glaser and Thompson (1998) utilize national level supermarket scanner data to analyze the demand for the organic and conventional frozen vegetables. Using an AIDS model their findings show a very small but increasing share of organic products in the frozen food vegetable market. The authors claim that as the market shares of organic frozen vegetables increase along with their availability in mainstream supermarkets, the price premiums decline.

Lohr (2001) suggests changes in income, lifestyle decisions and increasing food safety concerns are major factors affecting the demand for organic foods. The distinction between organic and non-organic food products is certification and labeling. Moreover, the price quality trade-off is based on the perception of term “organic” for consumers. Consumers need and want, to know what they are paying for. Strict certification and labeling requirements clarify this issue. Therefore informative labels are among the essential factors in the decision making process for organic consumers.

Factors affecting the acceptance of a new product depend on a variety of variables. The product attributes, prices and consumer specific characteristics as well as the interactions between those factors are essential in product acceptance. As Garretson and Burton (2000) suggests, product labeling can be a valuable source of information. It will not only explain many search characteristics such as nutritional values but also other characteristics such as health and environment related effects. People who have a strong perception of health and diet connection are more likely to try more innovative and nutrition enhanced food products.

Functional foods have been a popular area of interest as they are becoming increasingly popular among consumers. The perception of healthy diet induces consumers to become more concerned about their food choices. In their paper, Peng et al. (2006) experiment and analyze consumer perceptions of Conjugated Linoeic Acid (CLA) enhanced milk products. The authors apply a survey to test the hypothesis that consumers' acceptability of a CLA enriched dairy product is related to consumer characteristics such as diet-health perception, previous purchase decisions, and perception of CLA products (as well as the usual socio-demographic indicators). They find out that consumers who perceive a significant connection between health and diet are friendlier to CLA enrichment. Other variables such as beliefs in functional attributes, health claims, as well as previous purchase experience of functional foods increase the acceptability of CLA enriched products. Those who believe that dairy products have essential nutritional values have positive attitudes towards CLA enrichment. However, consumers who have already formed habits of buying a special type of milk are more likely to buy CLA-enriched milk products of that type. Since the perceptions of product characteristics are essential in purchase decisions, informing consumers about the health benefits of functional enhancements will greatly improve their attitude towards CLA-enriched dairy products. Once consumers experience functional enhancements, they will be more open minded and positive about new technologies. Moreover, the health conscious consumers who believe in the health benefits of dairy products are willing to pay more for functional enhancements.

Recent research by Chema et al. (2006) focuses on the functional benefits and acceptance of soy based dairy products. The authors study consumer views about the biotech foods with functional attributes. They interviewed 60 random households about their attitude towards soymilk, soy yogurt, dairy milk, and dairy yogurt. They find that consumers are willing to pay more for higher protein, increased calcium, and lower cholesterol. Their analysis indicates that consumers who are familiar with soybean products are willing to pay more for functional attributes, whereas those who are unfamiliar with soybean products consider soy as inferior on the basis of taste. As the authors point out, overcoming the perception of poor taste is essential to increase market share of soymilk and its derivatives. The majority of consumers are unfamiliar with the taste of soymilk and that might be the reason why they view it as inferior to dairy products.

Espinosa and Goodwin (1991) apply hedonic model to a cross-sectional time series data set of Kansas wheat characteristics. The analysis of pooled panel data indicates that wheat prices are not only determined by nutritional attributes such as protein content of wheat but also milling and dough characteristics. Moreover, location of wheat production also affects price. The authors suggest that regional differences reflect the significant variety in handling and processing facilities and differences in distances from major markets.

The choice of the functional form for the hedonic regressions depends on the specific application and the data source. Brachinger (2002) suggests using the double log form since it is widely used in the case of computer systems. However Nimon and Beghin use the semi-log form since it is widely accepted in hedonic literature. On the other hand, Ladd and Suvannunt use linear approach in their analysis. It is also possible to transform the data using Box-Cox transformations. However as noted by Cropper et al. (1988) the choice of Box-Cox transformed models can be highly misleading if there are missing attributes. Cropper, et al. suggests linear and semi-log models as the best options. In this paper, we use a semi-log model since that enables us to create extra variation among the attribute prices while making it easy to derive implicit prices.

Model

In classical hedonic demand theory, consumers are trying to maximize utility under the budget constraint. They choose a variety of products in order to achieve maximum utility. Utility derived from these products depend on the characteristics they provide. In our model, utility is also derived from product attributes such that consumers choose the bundle of goods that give them the best combination of utility maximizing attributes. Therefore the demand for products is a derived demand from product characteristics. Following the work of Ladd and Suvannunt, the economic model can be derived as follows:

Let's say there are k distinct characteristics and let x_j denote the characteristic $j, (j = 1 \dots k)$. We can think of x_j as the total amount of protein consumed or percentage of daily vitamin intake from a particular milk type. Now, let x_{ji} denote the j^{th} characteristic provided by one unit of product i. For example, if i is a specific type of milk, x_{ji} is the amount of protein in one unit of that milk type, or it can be the presence of an organic label. The total amount of attribute consumed will be a function of the products that contains this attribute and the amount of attribute per unit of these products such that $x_j = x_{j1}q_1 + x_{j2}q_2 + \cdots + x_{jm}q_m$ where q_i is the quantity of product i (i=1… m and j=1…k) The utility function is defined as $U = U(q_1, \dots, q_m)$ which can be rewritten in terms of product attributes such that $U = U(x_1, \dots, x_k)$. We can write the utility function as

$$U = U\left(\sum_{i=1}^{m} x_{1i}q_i, \sum_{i=1}^{m} x_{2i}q_i, \dots, \sum_{i=1}^{m} x_{ki}q_i\right) \tag{1.5}$$

Therefore the utility function can be denoted in terms of the relevant variables and parameters such that

$$U = U(q_1, q_2, \dots, q_m, x_{11} \dots x_{1m}, x_{21} \dots x_{2m}, x_{k1} \dots x_{km}) \tag{1.6}$$

Assuming that producers decide on the product ingredients beforehand, the amount of attributes in a product can be considered as exogenous to the consumers. After defining the utility function, the Lagrangian can be written as

$$L = U(x_1, \dots, x_k) - \lambda\left[I - \sum_{i=1}^{m} p_i q_i\right] \tag{1.7}$$

where I is the given income, and p_i is the price of the i^{th} product.

The Lagrangian above indicates that the consumer is deciding on the set of products that will provide the best combination of attributes to maximize utility. Taking derivatives of the Lagrangian expression with respect to the quantities, we get:

$$\frac{\partial L}{\partial q_i} = \frac{\partial U}{\partial x_1}\frac{\partial x_1}{\partial q_i} + \frac{\partial U}{\partial x_2}\frac{\partial x_2}{\partial q_i} + \cdots + \frac{\partial U}{\partial x_k}\frac{\partial x_k}{\partial q_i} - \lambda p_i = 0 \quad (1.8)$$

Solving for p_i gives

$$p_i = \left(\sum_{j=1}^{k} \frac{\partial U}{\partial x_j}\frac{\partial x_j}{\partial q_i}\right) \Big/ \lambda \quad (1.9)$$

where λ is the marginal utility of income such that $\lambda = {}^{\partial U}/_{\partial I}$. We can rewrite the prices in terms of marginal utility of income such that

$$p_i = \sum_{j=1}^{k} \left(\frac{\frac{\partial U}{\partial x_j}}{\frac{\partial U}{\partial I}}\right) \frac{\partial x_j}{\partial q_i} \quad (1.10)$$

The ratio of the marginal utility of the j^{th} characteristic to the marginal utility of the income gives us the marginal rate of substitution (MRS) between income and the j^{th} characteristic. Since all income is spend so that $income = expenditure$ ($I = E$), the MRS between I and x_j can be defined as the implicit or hedonic price paid for the j^{th} characteristic.

The partial derivative of attribute j with respect to product i, $\partial x_j/\partial q_i$ gives us the marginal change in the quantity of attribute j when the consumer utilizes one more unit of product i. We assume that the attribute per unit product is pre-determined by producers, so can state this relation as

$$\partial x_i/\partial q_j = x_{ij} \quad (1.11)$$

The MRS between income and attribute j can be calculated from the regression of product prices on the attributes per unit product such that

$$p_i = \sum_{j=1}^{k} x_{ij}\beta_j + \varepsilon = f_i(x) + \varepsilon \qquad \left(\frac{\partial f_i(x)}{\partial x_{ij}} = \beta_j\right) \quad (1.12)$$

where β gives the MRS.

If we specify the regression as above, the partial derivative of the hedonic regression, β_j, indicates how much the price of product i changes with respect to an additional unit change in characteristic x_{ij}. However, since our data are based on weekly time series, it is reasonable to assume the price of an attribute might differ between regions and can change over time. Therefore we specified our first-stage hedonic model as a semi-log model where prices of products are measured in log form:

$$p_i = \beta_a \prod_{j=1}^{k} \exp\,(\beta_j x_{ij}) \;\; or \ln p_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} \; where\; \beta_0 = \ln \beta_a \qquad (1.13)$$

In this form β_j is equal to the derivative of $\log p_i$ with respect to x_{ij} such that

$$\beta_j = \frac{\partial lnp_i}{\partial x_{ij}} = \frac{\partial p_i / p_i}{\partial x_{ij}} \; and \; \frac{\partial p_i}{\partial x_{ij}} = \beta_j p_i \qquad (1.14)$$

The coefficient β_j shows the marginal percentage change in price due to a unit change in an attribute. The semi-log form is convenient for the purpose of calculating the price premium in percentage terms. It also allows us to differentiate the hedonic prices for each product such that the price of the product also affects the prices of characteristics within the product. Once we derive the hedonic prices of each attribute over time such that $p_{x_{ij,t}} = \beta_j p_{i_t}$ we can use these prices along with the background information on consumers to estimate their demand for attributes in the second stage.

As Rosen (1974) indicates, demand for the product attributes is related to not only attribute prices but also socioeconomic characteristics of the households. Some attributes might give different utility to people with different backgrounds. Some might prefer one attribute to be more desirable than other. The consumer surveys indicate that people who prefer to consume cow milk believe that soymilk is inferior to cow milk. The amount of unhealthy nutrients such as cholesterol and sodium might be undesirable attributes for a person primarily concerned about health whereas a person primarily concerned about taste will get a higher level of utility from full fat milk. Income spent on dairy products might also be related to consumer characteristics. Therefore they should also be included in the second-stage regressions. Household size is among the important factors determining the amount of income spent on dairy products.

A household's utility function is viewed as representative of individual utilities within the household. We observe a variety of purchases on each trip to the store. Some in the family might favor soymilk while others might prefer traditional milk types. Different products can be consumed for different occasions. It might be the case that one type of milk is preferred for cooking whereas another type is consumed for breakfast. That might also be another reason why we observe multiple type purchases.

We estimated the second-stage regressions in the following form:

$$X_j = x_j(p_x, y, a) + u_j \tag{1.15}$$

where X_j is the amount of attribute consumed, p_x is the price set that includes the price of the attribute of concern and related hedonic prices, y is the household income and a is a vector of household demographics and u_j is the error term. In the property value literature, the second-stage equation is usually estimated at the household level. However, in our model, the prices are observed only when purchases are made. Whenever the consumer does not purchase a product we do not observe price. Because soymilk products are purchased only 3% of the time, we decided to follow an alternative approach. In the second stage, we derived the average weekly prices for each region. The demographic information is also used to form average prices for each region and time. Although our data include households who participate in the panel regularly, they purchase different amounts at different time periods. The panel nature of our data enables us to create variation among the independent variables.

Data

As indicated by the form of the hedonic model, in the first stage we regress price of the product on the attributes associated with that product. We desired to have an exhaustive set of attributes that include not only marketing variables but also the nutritional content of each product. In order to accomplish this objective, we combined AC Nielsen Homescan Panel with USDA – ERS Nutrient Database.

AC Nielsen Homescan Data is a panel data set in which each household records the purchases of products at the time of sales. Soon after making purchases the panel participants record their purchases by using a scanner and subsequently the data are uploaded to a database. The portion of data we obtained is related to dairy purchases made by 50223 households across the United States between 1998 and 2005. Because not all households participate regularly in the panel, we selected a sample of core households based on regular participation (at least 12 times a year) in the panel between the years 2002 and 2005. Our core data include approximately 3000 participants and are good representative of the total panel. These households reported a total of 525323 purchase occasions during 4 year period. A detailed comparison of core household data with the original AC Nielsen data regarding household demographics is shown in **Table 1.1.**

Scanner data possess information on prices, discounts, volumes, expenditures and purchase dates of all dairy products purchased by consumers combined with specific demographic information. Although detailed information on consumers is available from the Homescan panel date, the product attributes are limited to discrete marketing variables based on package labeling. Most discrete variables are binary such as organic label, soymilk, LFCF, and vitamin enrichment. There are also some non-binary yet discrete attributes such as container type, product type, product group, size, etc.

AC Nielsen panel data are limited to discrete product attributes, so information about attributes is enriched using the USDA-ERS Nutrient Database to obtain detailed information on product contents. Continuous nutritional contents such as protein, carbohydrate, fat contents, along with sodium and cholesterol contents are obtained from the nutrient database.

The USDA National Nutrient Database is based on an extensive set of research that is compiled from many published and unpublished resources and provides a reference to most food composition databases. It is possible to find very detailed scientific analyses of components for each product in the data. Even in abbreviated form, the components are exhaustive and not only include information about major food components such as protein, carbohydrate, fat, but also a detailed analyses of these components. Moreover the list of vitamins and minerals included in the database ranges from A to Zinc. However, although most of the vitamins included in the database are given in terms of milligrams per 100 gram (mg/100g), some vitamins such as vitamin B_{12} are measured in terms of micrograms per 100 gram (μg/100g) and some are measured in international units per 100 gram (IU/100g) such as vitamin A. To overcome differences in measurement units, all values for vitamins and minerals are transformed into percentages provided in a serving size (1cup=240ml) as indicated by Daily Recommended Intake[1] (DRI) values. The DRI values are based on Federal Drug Administration's reference values for nutrition labeling. **Table 1.2** lists the food labeling guide as suggested by FDA – CFSAN.

In order to overcome the correlation issue between vitamins and minerals, we initially combined them into two indexes based on DRI values. First we aggregated the vitamins into a single vitamin index. Next a similar approach is applied to the minerals to come up with a mineral index. Both of these indexes are based on recommended percentages of vitamin/mineral needs based on DRI. However, the high level of correlation between the vitamin and mineral indices is a problem. Therefore we decided to combine all minerals (other than sodium) and vitamins into a single vitamin/mineral index. The minerals included in this final index are calcium, iron, magnesium, potassium, and zinc; the vitamins are listed as vitamin A, vitamin B_2 (riboflavin), vitamin B_{12}, vitamin C and vitamin B_9 (folic acid).

The reason we excluded sodium from this index is because sodium is an important distinguishing component for consumers. Sodium and cholesterol cause health concerns among some consumers. In fact increased amounts have negative effect on the value of the product. Therefore, they are included as separate components for nutritional analysis.

The nutritional food attributes derived from nutrient database and the marketing attributes derived from the Homescan panel data characterize each milk product in hedonic space. The final set of attributes included in the first-stage analysis includes marketing

[1] *Center for Food Safety and Applied Nutrition (CFSAN) issues A Food Labeling Guide that is based on a diet of 2000 calorie intake. This guide suggests the recommended daily intake values of nutritional component. The list can be found at* http://www.foodsafety.gov/~dms/flg-7a.html

variables (i.e., soy dummy, lactose/cholesterol free (LFCF), organic claim, vitamin-mineral enhancement, promotion dummy, and container type) and nutritional variables (i.e., protein (g/serving), carbohydrate (g/serving), fat content (g/serving), percentage DRI index of cholesterol, sodium, and vitamin/minerals along with the servings per package). Packaging is an important attribute because it affects the total amount of attributes consumed per purchase and it also reflects consumer preference for size. We also included quarterly/weekly variables to compensate for inflation over time. However, as can be seen in the results section, a superior approach is to deflate each price by a Divisia price index of individual milk and soy products. **Table 1.3** gives a description of the product attributes.

In the second stage model, besides household income we also used information on household size, presence of children under 18, race, location, education level, and age/sex/employment/marital status of household leader. In the literature, there is no significant evidence to indicate the effects of household demographics have a significant effect on purchase decision. However, Thompson (1998) suggests that higher income households can afford more expensive products, and education might affect consumers' attitude towards healthy products. In our data, the frequency analysis of purchase type based on income suggests that, on average, households with higher income prefer skim milk whereas those with lower income prefer full fat milk. The average income of soymilk consumers is slightly lower than the sample average. Also, the number of children in the family seems to affect the household's purchase decision. Those who are single, living alone or with no kids under 18 seem be more likely to consume soymilk. Higher education is also associated with non-traditional types of milk purchases. There is no apparent relationship between the effect of household gender on milk purchase type.

The race variables indicate whether the household is White, Black, Asian, Hispanic or other. Lactose intolerance can be genetically transferred between generations. Moreover, in other parts of the world, soy is considered a significant source of protein. Therefore the coefficient on race variable measures whether different races purchase more or less soymilk than others. **Table 1.4** indicates that race of household significantly affects the choice of milk type.

Location of the household can also affect their purchase decision. The AC Nielsen data includes four US census regional variables for households. We enriched the variation in the data by creating another dummy variable based on whether the household lives in a populated metropolitan area. Combining the metropolitan variable with the region allows us to differentiate consumers based on eight different locations. (East/Central/South/West and also Metropolitan)

Results

For the first stage we estimated linear hedonic regressions for each quarter from the first quarter of 2002 until the last quarter of 2005 in order to see the trends in the hedonic prices. This linear model estimated as:

$$P_{it} = \sum_{j=1}^{n} \beta_{jt} x_{jit} \tag{1.16}$$

where p_{it} is the price per serving size of the i^{th} item at period *t*, β_{jt} is the implicit price of the attribute at period *t* and x_{jit} is the amount of the j^{th} attribute in product *i* at period *t*. Since our aim is to fully understand the effects of product attributes on price, we did not include constant terms in these preliminary regressions. Instead we used weighted least squares estimation where each product type is weighted by its market share.

Table 1.5 gives the results of these preliminary regressions for each quarter. For all quarters, the soy dummy has a negative effect on price, which indicates that, on average, the consumer is willing to avoid the soy taste. Being organic or LFCF adds significant value to the product. The hedonic prices of protein, carbohydrate and fat content are all positive. Among these major nutrients, the price of protein is significantly higher than others for twelve of the sixteen quarters. Vitamin-mineral enhancement index based on nutritional contents is a significantly desirable attribute whereas cholesterol content and sodium content reduce the value of the products as expected. Serving size has a negative effect on the price per serving. Convenience stores and neighborhood groceries have the highest prices whereas membership club stores have the lowest prices. There is a premium associated with convenience to shop for milk from a local store whereas there is a discount for prepayment for club membership. Moreover, milk in glass or carton containers costs significantly more than plastic containers. The cheapest dairy products are sold in the central region while households in the southern region pay the highest price. Interestingly the customers in the western region paid higher prices early in the sample but since the middle of 2003 they have paid lower prices.

In order to test the hedonic model we randomly selected a couple of milk products from local stores to see whether the assumptions of the model are consistent with our conjectures based on product attributes. **Appendix 1.1** gives a detailed hedonic price analysis for non-organic non-soy based LFCF milk. Although this model is a linear model without a constant, it gives reasonable results even for the extreme observations shown in **Appendix 1.1.**

On a regional basis we obtained similar results for implicit attribute prices. **Table 1.6** gives the results of the first-stage regressions by each region. As the number of servings increase, the product price per serving declines due to loss of convenience in storage and consumption. Soy label and promotion dummy reduces the final price paid, whereas LFCF, organic, and vitamin enhancement labels add positive values. Only in the nonmetropolitan west regions is the vitamin enhancement parameter is negative; this result might be spurious and due to correlation among related variables. Protein, carbohydrate, fat, and vitamin-mineral content have generally positive effects on price. However, for some regions these beneficial nutrients have negative values indicating possible correlation problems. We observe the same correlation problem for unhealthy attributes such as cholesterol content and sodium amount. Although these undesirable attributes have mostly negative values, for some regions they are estimated as positive. We also encounter this problem for container type and marketing channel. When we estimate the regressions by each location, some attributes have positive values at some locations whereas the same attributes can have negative values for other locations. The regressions by two regional variables (Metropolitan / Nonmetropolitan) solve this sign change issue to some extent, but the sign differences still exist for cholesterol content, container type, and distribution channel.

Because milk is a portable commodity, all regions were combined into one single market. The single market hedonic regression results are shown in **Table 1.7**. In these regressions, prices are measured in cents per serving. The signs for all attributes are consistent across regressions. We first estimate linear and semi-log regressions without directly adjusting prices for inflation. Instead, we initially implemented a trend variable for week to control for the effect of time. In the linear form, the parameter estimates give directly marginal implicit prices for each attribute. In the semi-log form the parameter estimates are multiplied by the final price of the product to obtain the implicit prices. The semi-log form is not only convenient but it also adds variation in the attribute prices. It gives us a unique attribute price for each unique product price. We also estimated using the first-stage regression by adding interaction terms between products attributes, however these terms were insignificant, and are not reported here.

Our results indicate that there is not much difference between using deflated prices based on the Divisia Index and using the week trend variable in the semi-log models. However, if we do not deflate the prices or if we remove the trend variable, not only the estimation power (R-square) value reduces significantly but also the coefficients are overestimated. Therefore in the first stage, we used a semi-log hedonic regression where prices are deflated using a Divisia price index

$$\log P_{it} = \sum_{j=1}^{k} \beta_{jt} x_{jit} \tag{1.17}$$

The dependent variable is the logarithm of product price at time *t* whereas the regressors are the amounts of attributes associated with the product.

Our results indicate that although soy has an undesirable taste with negative self attribute, its' price is significantly higher than other milk types because it is LFCF and mostly organic. Vitamin enhancement labeling and nutritional value are desirable attributes whereas consumers want to avoid cholesterol and sodium content. Also, among all the major nutrients, protein content is valued significantly more than fat or carbohydrate content. Products that are promoted are on average 18% cheaper than non-discounted products. Each additional serving in a package reduces the price by 3.5%. Compared to plastic packaging, glass, carton, box packages are more expensive; purchasing in pouch (bundled) reduces the price per serving. Also, compared to the grocery stores, convenience stores are slightly more expensive, while drug stores, mass merchandisers, supercenters, and club stores are relatively cheaper. In fact, club stores are 19% cheaper per serving than grocery stores. Since grocery stores and convenience stores offer convenience of location and faster customer service we observe higher prices in these stores.

We used the hedonic prices derived in the first stage estimation along with the demographic information on households to estimate their demand for soymilk and other specialty milk types. The data we used in the second stage are a panel data set based on the hedonic prices and household demographics for each location-week combination. We have 8 different regions and 208 weeks, a total of 1672 observations. In the second stage, the dependent variable is the consumption of a particular attribute per household which depends on hedonic prices and the demographic information of panel participants at this particular location during that time

$$\bar{X}_{lt} = f(\overline{Hedonic\ Prices_{lt}}, \overline{Participant\ Demographics_{lt}})$$

where $\bar{X}_{lt}$ is the average consumption per household at location l during week t. Demographics and hedonic prices are also measured in averages. Since we only observe prices of items that are purchased by a household for each purchase occasion, we do not know the rest of the prices that the consumer faces. Imputing values for these variables as averages of prices for each location during each week solves the issue of nonexistent prices. As can be seen in **Table 1.8** we initially started by excluding the demographic information.

Next, we added the demographic variables to the regressions. Including the demographic variables significantly increased the explanatory power of the second-stage regressions from an adjusted R-Square value of 0.283 to 0.457. However, many demographic variables are highly correlated with each other. The same correlation problem also exists in the first stage hedonic regressions. In order to mitigate the effects of multicollinearity, demographic variables included in the models were based on the results of stepwise regression analysis. The backward and forward variable selection methods gave a similar set of variables.

The regression results for soymilk purchase per household indicate that there is a positive relationship between quantity purchased and price of the soymilk attribute. Although that might appear incorrect, it simply means that the soymilk attribute has a negative value itself. Therefore an increase in the price of soymilk means a reduction in consumers' willingness to avoid soy taste. As consumers get used to the taste of soymilk, the marginal price of the soymilk becomes less negative. Thus, as a result of consumers' friendlier attitude towards soymilk, the purchased amount increases. The significant negative relationship between the LFCF attribute and purchased soymilk quantity is due to the fact that soymilk is naturally a LFCF product. An increase in the price of LFCF attribute increases the price of soymilk, which in turn reduces the average quantity purchased by households. Increase in average household income in the region has a negative effect on soymilk purchases. As can be seen in **Table 1.4**, average incomes of soymilk shoppers are slightly lower than average incomes of other consumers. Average household size, presence of kids, education level and age of household leaders all have positive effects on the amount of soymilk purchased. It seems parents do care about healthy alternatives for their children when it comes to dairy products. It is also interesting that as the percentage of Hispanic consumers in the population increases, demand for soymilk increases. As Scrimshaw and Murray (1988) point out, 51% of Hispanic population suffers from lactose maldigestion whereas this ratio drops to 21% among Caucasians.

Soymilk purchases constitute 57% of LFCF milk. The estimation results for the LFCF milk purchases show a similar pattern to the estimation results for soymilk, which is what we would expect. The demand for LFCF milk is positively related to the (negative) attribute price of soy and negatively related to the price of LFCF attribute. Also in regions where the ratio of more educated and older households are higher, the demand for LFCF milk per household increase. The effect of Hispanic household ratio is also positive. Since LFCF purchases constitute 75% of organic milk, the demand for organic milk is negatively related to the price of both organic attribute and also the LFCF attribute. The percentage of Hispanic population, the average household education, age, size all increase demand for organic milk.

Hedonic analyses were also conducted using fat content rather than quantities purchased. The demand for fat content is negatively associated with the price of fat, as expected. Our results also indicate that protein and carbohydrate content are substitutes for fat content. As the price of protein or carbohydrate increase demand for fat also increases. Moreover, household size has a positive effect on fat demand; however demand for fat is negatively correlated with the presence of children in the household.

Summary

In this paper we analyzed the factors affecting demand for alternative milk types including soymilk, organic milk and also functional enhancement labels such as being lactose/cholesterol free. Using AC Nielsen Homescan Panel along with the USDA Nutrient Database, we estimated a two-stage hedonic model. In the first stage we estimated hedonic prices for both marketing and nutritional attributes. The first-stage results indicate marketing variables based on product labels are as important as the nutritional contents of the products.

The soy attribute itself has a negative marginal implicit price, but since soymilk is naturally LFCF, and mostly organic, we observe a significant price premium on soymilk products. Also, vitamin/mineral enhancement labels affect the final prices in a positive way while promotional products cost 18% less. Protein, carbohydrate, fat contents are all positively valued. Among the major nutrients, protein has the highest value. Cholesterol and sodium content are undesirable attributes and significantly reduce the price of the milk. The industry might consider reducing these undesirable attributes since removing them could increase the value of their products. Also, when milk is purchased in pouch (bundled) containers, it costs less than the milk purchased in glass containers. The convenience of neighborhood grocery stores and convenience stores increases prices whereas club stores offer discounted prices even when compared to supercenters.

In the second stage models, we analyzed the relationship between demand for product characteristics and households' demographic information. Unlike previous approaches to hedonic analysis in food products, the rich nature of our data sources enabled us to exploit demographic variation in both cross sections and time series. We transformed our data into a weekly panel data based on eight regions. Using stepwise regression, we identified different estimators for each functional food type. The demand estimates for soymilk indicate that as consumers become more familiar with the soy attribute, they demand more soymilk. Households' education level increases demand for functionally enhanced food types. Also, older and larger families demand more of these functional enhancements. Compared to non-Hispanic households, the percentage of Hispanic households increases the demand for soymilk, LFCF milk and organic milk.

We estimated a continuous model to analyze the demand for product characteristics based on their implicit prices and household demographics. It is also possible to apply a probabilistic approach to analyze these relationships. A discrete approach where one estimates the probability of purchasing a special type of milk would be especially useful for marketing research. It is also possible to mix hedonic approach with the discrete choice method. Once we derive the marginal implicit prices in the first stage we can use these prices along with the consumer information to estimate their choice of milk type. Since milk types such as skim, 1%, 2% and full fat are differentiated based on fat content, the implicit price of fat can help us to identify consumer choices other than soymilk.

References

Bastian, C. T., D. M. McLeod, M. J. Germino, W. A. Reiners, and bB. J. Blasko. "Environmental Amenities and Agricultural Land Values: A Hedonic Model Using Geographic Information Systems Data." *Ecological Economics* 40, no. 3 (2002): 337-49.

Binkley, James, Abbott Sharon, Christine Wilson, and Kevin McNamara. "Determinants of Functional Food Consumption." *A workshop on the Use of Scanner Data in Policy Analysis.* Washington: USDA - ERS, 2000.

Brachinger, Hans Wolfgang. "Statistical Theory of Hedonic Price Indices." *Seminar of Statistics.* Fribourg: University of Fribour, 2002.

Brown, J.N., and Rosen H.S. "On the Estimation of Structural Hedonic Price Models." *Econometrica*, no. 50 (1982): 765-768.

Chema, S. Kambua, Leonie A. Marks, Josephy L. Parcell, and Maury Bredahl. "Marketing Biotech Soybeans with Functional Health Attributes." *Canadian Journal of Agricultural Economics* 54 (2006): 685-703.

Cropper, M, L Deck, and K McConnell. "On the choice of functional form for hedonic functions." *Review of Economics and Statistics*, no. 70 (1988): 668-675.

Espinosa, Juan A, and Barry K. Goodwin. "Hedonic Price Estimation for Kansas Wheat Characteristics." *Western Journal of Agricultural Economics* 16, no. 1 (1991): 72-85.

Garretson, J., and S. Burton. "Effects of nutrition facts panel values, nutrition claims, and health claims on consumer attitudes, perceptions of disease-related risks, and trust." *Journal of Public Policy and Marketing* 19, no. 2 (2000): 213-227.

Glaser, Lewrene K., and Gary D. Thompson. "Demand for Organic and Coventional Frozen Vegetables." *American Agricultural Economics Association Annual Meeting.* Nashville, Tennesse, 1998.

Hager, Christine. "PhD Dissertation." *Demand for Nutrient and Nonnutrient Components in Household Purchases of Red Meat, Poultry, and Fish Products Using a Hedonic Approach.* North Carolina State University, 1985.

Harris, Mike. "Properties of Scanner Data." *A Workshop on the Use of Scanner Data in Policy Analysis.* Washington: USDA - ERS, 2003.

Jensen, Helen H, and Huffman K. Sonya. "Demand for Enhanced Foods and the Value of Nutrition Enhancements of Food: The Case of Margarines." *A workshop on the Use of Scanner Data in Policy Analysis.* Washinton: USDA - ERS, 2003.

Ladd, George W., and Marvin B. Martin. "Prices and Demands for Input Characteristics." *American Journal of Agricultural Economics* (American Agricultural Economics) 58, no. 1 (Feb 1976): 21-30.

Ladd, George W., and Veraphol Suvannunt. "A model of Consumer Goods Characteristics." *American Journal of Agricultural Economics* (American Agricultural Economics Association) 58 (August 1976): 504-510.

Lohr, Luanne. "Factors Affecting International Demand and Trade in Organic Food Products." *University of Georgia Faculty Series*, February 2001.

Nelson, Phillip. "Information and Consumer Behaviour." *Journal of Political Economy* 78, no. 2 (1970): 311-329.

Nimon, Wesley, and John Beghin. "Are Eco-Label Valauble? Evidence from the Apparel Industry." *American Journal of Agricultural Economics* (American Agricultural Economics Association) 81, no. 4 (November 1999): 801-811.

Nti, Christina Antwiwaa, and Patience Mateko Larweh. "Production and Sensory Sharacteristics of Flavoured Soymilk Samples." *International Journal of Consumer Studies* (Blackwell Publishing Ltd) 27, no. 3 (June 2003): 181-184.

Palmquist, R.B. "Estimating the Demand for Characteristics of Housing." *Review of Economics and Statistics*, no. 66 (1984): 394-404.

Palmquist, R.B. "Land as a Differentiated Factor of Production: A Hedonic Model and its Implications for Welfare Measurement." *Land Economics*, no. 65 (1989): 23-28.

Palmquist, R.B., and Israngkura A. "Valuing Air Quality with Hedonic and Discrete Choice Models." *American Journal of Agricultural Economics*, no. 81 (1999): 1128-1133.

Palmquist, Raymond B. "Property Value Models." In *Handbook of Environmental Economics*, edited by Goren K Maler and Vincent Jeffrey. Elsevier, 2003.

Peng, Yanning, Gale E. West, and Cindy Wang. "Consumer Attitudes and Acceptance of CLA - Enriched Dairy Products." *Canadian Journal of Agricultural Economics* 54 (2006): 633-684.

Rosen, R. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, no. 82 (1974): 34-55.

Scrimshaw, N. S., and E. B. Murray. "The acceptability of milk and milk products in populations with a high prevalence of lactose intolerance." *American Journal of Clinical Nutrition* 48 (1988): 1079-1159.

Silver, Mick. "Scanner Data and Price Indices." *A workshop on the Use of Scanner Data in Policy Analysis.* Washington: USDA - ERS, 2003.

Thompson, Gary D. "Consumer Demand for Organic Foods: What We Know and What We Need to Know." *American Journal of Agricultural Economics* 80, no. 5 (1998): 1113-1118.

Table 1.1 Core Data Household Demographics vs. Entire Data Household Demographics

CORE DATA DEMOGRAPHICS

Household Income

HHIncome	Percent	Cumulative
3	0.51	0.51
4	0.45	0.96
6	0.45	1.4
8	1.4	2.81
10	2.26	5.07
11	4.56	9.63
13	7.43	17.06
15	6.38	23.44
16	7.46	30.9
17	7.46	38.36
18	7.24	45.6
19	7.14	52.74
21	11.42	64.16
23	9.76	73.92
26	15.91	89.83
27	10.17	100

Household Size

HHSize	Percent	Cumulative
1	20.98	20.98
2	42.6	63.58
3	15.78	79.37
4	12.79	92.16
5	5.23	97.39
6	1.63	99.01
7	0.61	99.62
8	0.19	99.81
9	0.19	100

Age and Presence of Children

Children	Percent	Cumulative
1	2.46	2.46
2	7.02	9.47
3	7.14	16.61
4	3.22	19.83
5	0.41	20.25
6	4.5	24.74
7	0.99	25.73
No Children	74.27	100

Household Race

HHRace	Percent	Cumulative
White	82.53	82.53
Black	8	90.53
Asian	2.2	92.73
Others	1.21	93.94
Hispanic	6.06	100

ENTIRE DATA DEMOGRAPHICS

Household Income

HHIncome	Percent	Cumulative
3	0.81	0.81
4	1.32	2.12
6	1.14	3.27
8	1.62	4.89
10	3.05	7.94
11	5.44	13.38
13	7.73	21.11
15	6.89	28
16	7.85	35.84
17	6.97	42.82
18	6.83	49.64
19	6.45	56.09
21	11.05	67.14
23	8.78	75.92
26	14.87	90.79
27	9.21	100

Household Size

HHSize	Percent	Cumulative
1	26.1	26.1
2	37.4	63.5
3	14.91	78.41
4	13.13	91.55
5	5.55	97.1
6	1.92	99.02
7	0.6	99.62
8	0.25	99.87
9	0.13	100

Age and Presence of Children

Children	Percent	Cumulative
1	3.71	3.71
2	6.44	10.15
3	7.95	18.11
4	3.33	21.44
5	0.57	22.01
6	4.3	26.31
7	0.87	27.17
No Children	72.83	100

Household Race

HHRace	Percent	Cumulative
White	78.83	78.83
Black	10.07	88.9
Asian	2.13	91.03
Others	1.82	92.85
Hispanic	7.15	100

Table 1.2 Food Labeling Guide

Food Labeling CFR References

Reference Values for Nutrition Labeling

(Based on a 2000 Calorie Intake; for Adults and Children 4 or More Years of Age)

NUTRIENT	UNIT OF MEASURE	DAILY VALUES
Total Fat	grams (g)	65
Saturated fatty acids	grams (g)	20
Cholesterol	milligrams (mg)	300
Sodium	milligrams (mg)	2400
Potassium	milligrams (mg)	3500
Total carbohydrate	grams (g)	300
Fiber	grams (g)	25
Protein	grams (g)	50
Vitamin A	International Unit (IU)	5000
Vitamin C	milligrams (mg)	60
Calcium	milligrams (mg)	1000
Iron	milligrams (mg)	18
Vitamin D	International Unit (IU)	400
Vitamin E	International Unit (IU)	30
Vitamin K	micrograms (µg)	80
Thiamin	milligrams (mg)	1.5
Riboflavin	milligrams (mg)	1.7
Niacin	milligrams (mg)	20
Vitamin B_6	milligrams (mg)	2
Folate	micrograms (µg)	400
Vitamin B_{12}	micrograms (µg)	6
Biotin	micrograms (µg)	300
Pantothenic acid	milligrams (mg)	10
Phosphorus	milligrams (mg)	1000
Iodine	micrograms (µg)	150
Magnesium	milligrams (mg)	400
Zinc	milligrams (mg)	15
Selenium	micrograms (µg)	70
Copper	milligrams (mg)	2
Manganese	milligrams (mg)	2
Chromium	micrograms (µg)	120
Molybdenum	micrograms (µg)	75
Chloride	milligrams (mg)	3400

Nutrients in this table are listed in the order in which they are required to appear on a label in accordance with 101.9© This list includes only those nutrients for which a Daily Reference Value (DRV) has been established in 101.9(c)(9) or a Reference Daily Intake (RDI) in in 101.9(c)(8)(iv).

Source: http://www.foodsafety.gov/~dms/flg-7a.html REV. Jan 30, 1998

Table 1.3 List of Product Attributes

Variable Name	Type	Explanation
Marketing Variables		Measurement Unit
MSoy	Marketing	Dummy (1 if Soymilk, otherwise 0)
MCFLF	Marketing	Dummy (1 if Cholesterol/Lactose Free, otherwise 0)
MOrgClaim	Marketing	Dummy (1 if Organic, otherwise 0)
MVitminIndex	Marketing	Dummy (1 if Vitamin Enhanced, otherwise 0)
MPromo	Marketing	Dummy (1 if the Product is on Promotion, otherwise 0)
MContainer	Marketing	Discrete (1 if Box, 2 if Carton, 3 if Glass, 4 if Plastic, 5 if Pouch)
Nutritional Variables		
NProtein	Nutritional	Grams per Serving
NCarboHydrt	Nutritional	Grams per Serving
NLipid_Tot	Nutritional	Grams per Serving
NPerCholestrl	Nutritional	Percentage of Maximum Allowed Intake Value per Serving
NPerSodium	Nutritional	Percentage of Maximum Allowed Intake Value per Serving
NPerVitMinIndex	Nutritional	Percentage of Vitamin/Mineral Value per Serving
Other Variables		
ServingsPackage	Continous/Discrete	Number of Servings Per Package
Week	Time - Week	Weekly Time Series
QTR	Time - Quarter	Quarterly Time Series
LChannel	Purchase Channel	Discrete (1 if Grocery, 2 if Drug Store, 3 if Mass Merchandiser, 4 if SuperCenter, 5 if Club, 6 if Convenience, 7 Others)

Table 1.4 Analysis of Demand for All Milk Types by Household Demographics

Analysis Variable	Total Sample	MOrgClaim		MCFLF		MSoy				MFat		
		0	1	0	1	0	1	0.00%	1.00%	2.00%	3.25%	SOY
Frequency	525292	512366	12926	498666	26626	510044	15248	147268	92825	172825	97126	15248
HHIncome	$55,369	$55,259	$59,727	$55,325	$56,198	$55,378	$55,093	$59,395	$57,591	$53,509	$50,494	$55,093
HHKids	25.58%	25.75%	18.93%	26.00%	17.68%	25.88%	15.48%	20.27%	26.06%	28.48%	29.59%	15.48%
HHSize	2.57	2.57	2.35	2.59	2.26	2.58	2.25	2.38	2.54	2.67	2.75	2.25
HHAge	7.17	7.17	7.26	7.15	7.56	7.16	7.58	7.38	7.16	7.10	6.93	7.58
HHCompAlone	17.20%	17.10%	20.93%	16.84%	23.80%	17.00%	23.71%	21.47%	15.93%	14.48%	15.73%	23.71%
HHCompMarried	71.65%	71.74%	68.17%	72.02%	64.76%	71.91%	63.05%	71.18%	72.67%	74.16%	68.28%	63.05%
HHCompNonRelated	2.85%	2.85%	2.72%	2.81%	3.57%	2.82%	3.89%	2.37%	2.49%	2.51%	4.35%	3.89%
HHCompRelated	8.31%	8.31%	8.18%	8.33%	7.86%	8.28%	9.35%	4.98%	8.91%	8.85%	11.64%	9.35%
HHEducation	4.16	4.15	4.26	4.15	4.28	4.15	4.25	4.36	4.24	4.05	3.93	4.25
HHEmpBoth	31.69%	31.67%	32.52%	32.06%	24.85%	31.88%	25.37%	30.01%	31.47%	34.66%	30.16%	25.37%
HHEmpFemale	16.52%	16.47%	18.37%	16.31%	20.36%	16.42%	19.86%	15.79%	16.03%	15.45%	19.46%	19.86%
HHEmpMale	22.85%	22.87%	22.41%	22.82%	23.48%	22.79%	25.07%	21.97%	22.16%	22.41%	25.31%	25.07%
HHEmpNone	28.94%	28.99%	26.70%	28.81%	31.31%	28.91%	29.69%	32.23%	30.34%	27.48%	25.08%	29.69%
HHLeaderBoth	74.58%	74.70%	70.02%	74.92%	68.19%	74.82%	66.57%	73.18%	74.81%	77.18%	73.11%	66.57%
HHLeaderFemale	8.13%	8.06%	10.90%	8.06%	9.45%	8.02%	11.82%	8.64%	8.25%	7.02%	8.64%	11.82%
HHLeaderMale	17.29%	17.24%	19.08%	17.02%	22.37%	17.16%	21.61%	18.18%	16.93%	15.80%	18.25%	21.61%
HHMaritalDivorced	9.40%	9.37%	10.74%	9.26%	12.18%	9.32%	12.15%	9.45%	8.46%	9.27%	10.04%	12.15%
HHMaritalMarried	71.81%	71.91%	68.17%	72.19%	64.76%	72.08%	63.05%	71.23%	73.07%	74.27%	68.51%	63.05%
HHMaritalSingle	10.62%	10.61%	11.25%	10.57%	11.59%	10.59%	11.67%	11.52%	10.32%	8.51%	13.14%	11.67%
HHMaritalWidowed	8.16%	8.12%	9.84%	7.98%	11.47%	8.01%	13.13%	7.80%	8.15%	7.95%	8.31%	13.13%
HHRaceAsian	2.28%	2.23%	4.50%	2.13%	5.16%	2.23%	4.07%	1.89%	2.82%	1.79%	2.96%	4.07%
HHRaceBlack	6.36%	6.16%	14.17%	5.89%	15.20%	6.06%	16.19%	3.08%	4.70%	6.07%	11.88%	16.19%
HHRaceHispanic	5.96%	5.98%	5.04%	5.95%	6.04%	6.02%	3.84%	4.65%	5.05%	6.92%	7.42%	3.84%
HHRaceOthers	1.30%	1.30%	1.27%	1.32%	0.99%	1.31%	1.01%	1.16%	1.34%	1.04%	1.99%	1.01%
HHRaceWhite	84.10%	84.33%	75.01%	84.72%	72.62%	84.38%	74.90%	89.21%	86.10%	84.18%	75.75%	74.90%
MSOY	2.90%	1.08%	75.27%	0.03%	56.74%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%
MOrgClaim	2.46%	0.00%	100.00%	0.64%	36.64%	0.63%	63.81%	0.58%	1.01%	0.48%	0.61%	63.81%
MCFLF	5.07%	3.29%	75.47%	0.00%	100.00%	2.26%	98.99%	3.87%	1.19%	2.21%	0.94%	98.99%

Table 1.5 Hedonic Analysis by Each Quarter

		QUARTER								
		1	2	3	4	5	6	7	8	9
Variable	Label	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
MUSDAOrg	Marketing USDA Organic Label	9.44623	8.34929	6.99552	6.52018	7.00173	8.18381	5.60971	5.73262	6.84121
MSoy	Marketing Soy Dummy	-4.14549	-5.35699	-6.12223	-4.48816	-5.88609	-7.86167	-8.66128	-8.45623	-7.7502
MVitminIndex	Marketing Vitamin Mineral Index	2.89932	3.79995	2.31896	3.58016	3.86372	3.70771	3.7402	4.34151	2.54779
MCFLF	Marketing Lactose Cholesterol Free	19.40998	19.22485	19.74704	19.31368	18.84445	19.45315	19.60871	18.2476	18.74273
NProtein	Nutrient Protein Content (g)	1.17078	0.59265	0.42669	0.89636	1.08347	0.59996	0.71684	0.80534	0.70285
NCarboHydrt	Nutrient Carb Content (g)	0.71473	0.76801	0.61865	0.64669	0.68346	0.61279	0.62103	0.70973	0.68751
NLipid_Tot	Nutrient Lipid (Fat) Content (g)	0.62345	0.61071	0.54655	0.53251	0.60086	0.64805	0.54412	0.4145	0.57429
NPerVitMinIndex	Nutrient Vitamin-Mineral Percentage Index	0.20309	0.58922	1.04902	0.64772	0.64557	0.82797	0.78057	0.78371	1.13278
NPerCholestrl	Nutrient Cholesterol DRI Percentage of Max	-0.33565	-0.3195	-0.25129	-0.22872	-0.2968	-0.34572	-0.22743	-0.13142	-0.27449
NPerSodium	Nutrient Sodium DRI Percentage of Max	-1.07697	-1.20653	-1.20723	-1.03796	-1.30212	-0.9834	-1.00836	-1.37404	-1.69784
SQuantity	Purchased Serving Size Quantity	-0.06687	-0.05419	-0.07928	-0.07787	-0.07347	-0.07692	-0.07619	-0.07158	-0.06532
LChannelID1	Location Channel Grocery	1.33691	1.1762	0.95754	1.23222	1.09884	1.56095	1.60641	1.69462	1.75023
LChannelID2	Location Channel Drug	0.51254*	0.51577*	0.35203*	0.68686	0.37027*	0.64691*	0.51782*	-0.27003'	0.51065*
LChannelID3	Location Channel SuperCenters	0.78636	0.76085	0.6422	0.73725	0.73705	1.78691	1.10831	0.84589	1.40576
LChannelID4	Location Channel Supermarket	0.55036	0.2759'	0.19483'	0.38123*	0.44077'	0.94577	0.60139	0.2154'	0.54087
LChannelID5	Location Channel Club	-0.31696**	-0.72796	-1.81263	-1.76774	-2.14844	-1.95095	-1.75894	-1.63816	-2.22328
LChannelID6	Location Channel Convenience	1.98189	1.53987	0.76258	0.96786	1.05685	1.21726	1.10053	1.26386	1.34133
MContainer1	Marketing Container Box	-7.69199'	-15.54731	0	0	0	0	0	0	14.27242
MContainer2	Marketing Container Carton	1.47364	0.81784'	0.92334**	-1.37513*	-2.32915	-1.5509*	-1.157**	-0.50128'	-0.39879'
MContainer3	Marketing Container Glass	15.55044	13.79684	14.08699	10.18473	10.37344	9.6379	9.14423	12.55439	11.39503
MContainer4	Marketing Container Pouch	-2.78943	-3.44179	-3.69058	-5.81495	-6.85565	-5.76727	-5.79059	-4.97668	-5.14163
LRegion2	Location Region Central	-1.70759	-1.92459	-1.85728	-1.93881	-1.81551	-1.81567	-1.90666	-2.37867	-2.62014
LRegion3	Location Region South	0.99917	0.65726	0.83296	0.8279	0.97461	0.94866	0.72168	0.49132	0.62946
LRegion4	Location Region West	0.96908	0.82597	0.56735	0.73652	0.54706	0.86097	1.29119	0.82881	0.67738

(*) Significant at 5% level. (**) Significant at 10% level. (') Insignificant.
All variables are significant at 1% level unless indicated otherwise.

Table 1.5 Hedonic Analysis by Each Quarter (Cont.)

		QUARTER						
		10	**11**	**12**	**13**	**14**	**15**	**16**
Variable	Label	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
MUSDAOrg	Marketing USDA Organic Label	4.78339	4.76323	6.39618	6.10341	6.60132	6.69051	7.20352
MSoy	Marketing Soy Dummy	-6.41584	-5.35309	-5.99603	-7.14368	-9.00444	-8.90704	-8.3017
MVitminIndex	Marketing Vitamin Mineral Index	1.55785	1.22577	1.19291	1.92633	2.16474	2.04438	1.75547
MCFLF	Marketing Lactose Cholesterol Free	16.65882	17.33883	18.64819	18.41594	17.83836	18.7661	18.42639
NProtein	Nutrient Protein Content (g)	0.66442	0.7846	1.06636	1.01956	0.66915	0.86763	0.97448
NCarboHydrt	Nutrient Carb Content (g)	0.6193	0.61835	0.66854	0.62619	0.6428	0.75723	0.72991
NLipid_Tot	Nutrient Lipid (Fat) Content (g)	0.6215	0.68689	0.59989	0.46169	0.72507	0.59555	0.66832
NPerVitMinIndex	Nutrient Vitamin-Mineral Percentage Index	1.05195	1.25998	1.05376	0.63933	0.79702	0.66643	0.76878
NPerCholestrl	Nutrient Cholesterol DRI Percentage of Max	-0.23043	-0.2669	-0.15691	-0.05148'	-0.36313	-0.19462	-0.25568
NPerSodium	Nutrient Sodium DRI Percentage of Max	-1.45839	-1.8445	-1.9188	-1.50999	-1.4195	-1.57247	-1.62057
SQuantity	Purchased Serving Size Quantity	-0.07791	-0.08143	-0.07849	-0.08104	-0.06963	-0.07747	-0.08137
LChannelID1	Location Channel Grocery	1.5437	1.1559	1.46513	1.83766	1.62234	1.71262	1.5633
LChannelID2	Location Channel Drug	-0.93534	1.21259	0.00237'	-0.21382'	-0.46968*	0.41746**	0.19981'
LChannelID3	Location Channel SuperCenters	0.7409	1.74819	1.58693	2.42943	1.63162	2.08317	1.8918
LChannelID4	Location Channel Supermarket	0.23212'	0.40479*	0.83791	0.8574	1.00001	1.09856	1.13141
LChannelID5	Location Channel Club	-2.05542	-2.13689	-1.77573	-1.0314	-1.42097	-1.41083	-1.60306
LChannelID6	Location Channel Convenience	0.88388	0.8261	1.40108	1.85002	1.56818	1.43738	0.85565
MContainer1	Marketing Container Box	9.77707	2.71198'	11.28142	12.15863	20.89229	15.06585	20.14433
MContainer2	Marketing Container Carton	4.15805	2.94408	1.62273*	4.78744	5.94648	4.84154	4.05741
MContainer3	Marketing Container Glass	15.76508	12.60429	12.25924	16.48781	17.69992	14.15308	14.45392
MContainer4	Marketing Container Pouch	-0.85326'	-2.01223	-3.45291	-0.39659'	0.28589'	-0.93064'	1.67055'
LRegion2	Location Region Central	-2.45626	-2.62382	-2.78525	-2.99844	-3.01161	-3.09566	-2.93072
LRegion3	Location Region South	0.45644	0.50916	0.19723	0.39602	0.51258	0.72096	1.01511
LRegion4	Location Region West	0.36179	0.72272	0.9081	1.07089	-1.03662	-0.87249	-0.85803

(*) Significant at 5% level. (**) Significant at 10% level. (') Insignificant.

All variables are significant at 1% level unless indicated otherwise.

Table 1.6 SemiLog Hedonic Regressions by Region

Dependent				**Regional**	**Metropolitan**									**LRegionMetropol**
Variable:	*Region1*	*Region2*	*Region3*	*Region4*	*MID=0*	*MID=1*			*NonMetropolitan*					*Metropolitan*
Log Price	*East*	*Central*	*South*	*West*			*East*	*Central*	*South*	*West*	*East*	*Central*	*South*	*West*
Label	Est.	Est	Est	Est	Est	Est	Est	Est	Est	Estimate	Est	Est	Est	Est
Intercept	1.865	1.865	1.963	3.158	2.219	2.243	2.472	1.720	2.514	3.039	1.823	2.448	1.676	3.308
Soy Dummy	-0.147	-0.147	-0.226	-0.499	-0.316	-0.274	-0.354	-0.096	-0.365	-0.577	-0.124	-0.054'	-0.170	-0.531
LFCF	0.712	0.712	0.541	0.578	0.699	0.632	0.664	0.826	0.550	0.639	0.715	0.585	0.546	0.580
Organic Claim	0.220	0.220	0.364	0.349	0.332	0.333	0.161	0.259	0.413	0.296	0.227	0.006	0.356	0.370
Vit-Min Index	0.162	0.162	0.221	0.0095'	0.115	0.169	0.0073'	0.141	0.121	-0.106	0.178	0.105	0.290	0.058
Promotion Dummy	-0.155	-0.155	-0.210	-0.102	-0.208	-0.169	-0.252	-0.186	-0.243	-0.123	-0.106	-0.174	-0.182	-0.066
Protein Content (g)	0.079**	0.098	0.072	0.0035'	0.059	0.067	0.0005'	0.101	0.046	-0.016**	0.087	0.105	0.086	0.014
Carb Content (g)	0.027	0.032	0.025	0.014	0.020	0.025	0.024	0.029	0.022	-0.0028*	0.026	0.043	0.026	0.026
Lipid Content (g)	0.026	0.00002'	0.024	0.014	0.029	0.013	0.031	-0.003*	0.015	0.055	0.021	0.0008'	0.028	-0.016
Cholesterol DRI %	-0.011	-0.011	-0.008	0.014	-0.008	0.005	-0.025	0.019	-0.005	-0.027	-0.005	0.0047**	-0.008	0.043
Sodium DRI %	-0.090	-0.090	-0.064	0.011	-0.037	-0.061	-0.065	-0.088	-0.067	0.042	-0.091	-0.153	-0.062	-0.005
Vit-Min DRI %	0.040	0.040	0.031	0.0026**	0.014	0.018	0.052	0.021	0.027	0.024	0.038	0.008	0.032	-0.020
# of servings	-0.023	-0.023	-0.029	-0.059	-0.033	-0.036	-0.024	-0.026	-0.036	-0.046	-0.023	-0.051	-0.026	-0.064
Quarter	0.016	0.016	0.016	0.011	0.017	0.015	0.018	0.016	0.018	0.013	0.016	0.018	0.016	0.009
Container Box	0.0187'	0.0188'	0.24'	0.116'	0.217'	0.053'	0.000	-0.022'	-0.124'	0.41*	-0.0073'	0.091'	0.296	-0.015'
Container Carton	-0.007	-0.007	-0.0019'	0.067	0.007	0.013	-0.012	0.033	-0.111	0.149	-0.004**	0.017**	0.040	0.010
Container Glass	-0.224	-0.224	0.379	0.780	-0.027*	0.473	-0.118	0.577	0.137	0.132'	0.530	0.534	0.427	0.712
Container Pouch	0.000	0.000	0.000	0.000	-0.448	-0.587	0.000	-0.390	0.000	0.000	0.000	-0.50*	0.000	0.000
Drug Store	-0.042	-0.042	-0.033	0.037	-0.060	-0.062	-0.100	-0.081	-0.022	0.0009'	0.016	-0.119	-0.046	0.053
MassMerchant	-0.056	-0.056	-0.017	0.025	-0.017	-0.028	-0.068	0.020	-0.0038'	-0.016'	-0.0465*	-0.027	-0.024	0.044
SuperCenter	-0.057	-0.057	-0.031	-0.084	-0.024	0.023	-0.067	-0.034	-0.070	-0.043	-0.033	-0.106	0.0028'	-0.092
Club Store	-0.160	-0.160	-0.197	-0.199	-0.158	-0.213	-0.143	-0.094	-0.176	-0.161	-0.163	-0.122	-0.202	-0.217
Convenience	-0.007	-0.007	-0.0034'	0.116	0.0024'	0.009	-0.074	0.049	0.0054'	-0.0161'	0.010	0.017	-0.0004'	0.208
Other Store	-0.071	-0.071	-0.156	-0.131	-0.120	-0.096	-0.138	-0.051	-0.189	-0.163	-0.060	-0.128	-0.142	-0.085
Adj R-Sq	0.516	0.373	0.494	0.594	0.408	0.513	0.502	0.349	0.499	0.478	0.530	0.451	0.504	0.693

(*) Significant at 5% level. (**) Significant at 10% level. (') Insignificant.

All variables are significant at 1% level unless indicated otherwise.

Table 1.7 Hedonic Regression Stage I Results

	Week Adjustment		No week parameter	
Dependent Variable	PPPCup	LogPPPCup	LogDefPPPCup	LogPPPCup
Time Adjustment	Week Trend	Week Trend	Deflated	None
Label	Estimate	Estimate	Estimate	Estimate
Intercept	-7.50685	2.2544	2.31653	2.32372
Marketing Soy Dummy	-7.76391	-0.29258	-0.29127	-0.28696
Marketing Lactose Cholesterol Free	19.35535	0.65619	0.66074	0.67561
Marketing organic Claim	9.10183	0.33252	0.33766	0.35081
Marketing Vitamin Mineral Index	4.79728	0.14614	0.14536	0.14458
Marketing Promotion Dummy	-3.23968	-0.19571	-0.18385	-0.12169
Nutrient Protein Content (g)	2.29756	0.06148	0.0642	0.07108
Nutrient Carb Content (g)	0.7441	0.02239	0.0228	0.02313
Nutrient Lipid (Fat) Content (g)	0.62913	0.02175	0.02142	0.02131
Nutrient Cholesterol DRI Percentage of Max	-0.17633	-0.00253	-0.0023	-0.00233
Nutrient Sodium DRI Percentage of Max	-1.80266	-0.05351	-0.05639	-0.06222
Nutrient Vitamin-Mineral Percentage Index	0.74765	0.01988	0.02013	0.02059
# of servings in a package	-0.79602	-0.03559	-0.03572	-0.03638
Week	0.02098	0.0012		
Marketing Container Box	4.91994	0.0921	0.08098*	0.09094
Marketing Container Carton	0.011'	0.01476	0.01179	0.003*
Marketing Container Glass	11.33017	0.35409	0.35696	0.36018
Marketing Container Pouch	-10.05194	-0.51911	-0.52486	-0.51375
Location Channel Drug	-1.22728	-0.06076	-0.05947	-0.0561
Location Channel MassMerchant	-0.51239	-0.0245	-0.02527	-0.02065
Location Channel SuperCenter	-0.71043	-0.02363	-0.02069	-0.00454
Location Channel Club	-3.09501	-0.19031	-0.18677	-0.16988
Location Channel Convenience	0.19588	0.0134	0.01573	0.0149
Location Channel Other	-1.89019	-0.10809	-0.10821	-0.10147
Adj R-Sq	0.5387	0.4664	0.4611	0.4246

(*) Significant at 5% level. (**) Significant at 10% level. (') Insignificant.
All variables are significant at 1% level unless indicated otherwise.

Table 1.8 Hedonic Regression Stage II Results

Parameter	Soymilk Equation 1 Estimate	Soymilk Equation 2 Estimate	Soymilk Equation 3 Estimate	CFLF Milk Equation 4 Estimate	Organic Milk Equation 5 Estimate	Fat Demand Equation 6 Estimate
Intercept	-0.45882	-12.5997	-1.044'	-11.07567	-4.68'	-654.29829
PMSoy	0.024*	0.03202	0.02973	0.0287*		
PMCFLF	-0.01527	-0.01191	-0.01341	-0.01**	-0.01521	-1.20314
PMOrgClaim	-0.01**	-0.0011'			-0.01336	
PMVitMinIndex	1.144*	1.38381	1.4272			
PMPromo	0.044	-0.01911'				
PNProtein	6.451*	3.37'				1378.75373
PNCarboHydrt	22.56969	22.94323	18.43069	9.3109	21.94995	1444.37083
PNLipid_Tot	8.88252	7.30713	6.90453		-3.29**	-3889.0675
PNPerCholestrl	116.857	116.10267	113.45158		-100.84277	-56234
PNPerSodium	9.62821	4.248'				
PNPerVitMinIndex	-19.27*	-25.76016	-21.87756		-28.60958	-7759.20131
HHIncome		-0.00001332	-0.00001157	-0.00001857	-0.00001295	-0.00336
HHSize		0.35013	0.34845	0.83066	0.3664	150.16486
HHKids		0.64**	0.64*			-147.58611
HHEducation		0.30914	0.32955	0.57187	0.2901	102.01314
HHAge		0.49552	0.53179	0.85715	0.32461	118.44099
HHCompMarried		11.81482	-4.08885	7.3983		
HHCompNonRelated		5.83668		9.9796	2.42734	
HHCompRelated		1.99361		3.17427	1.24279	-506.81093
HHCompAlone			-5.92			-881.57444
HHMaritalMarried		-5.687'				
HHMaritalWidowed		1.474*	1.57402			330.34832
HHMaritalDivorced		1.51428	1.48782			415.31583
HHMaritalSingle				-3.20606	-1.135*	
HHEmpBoth		1.083*				
HHEmpFemale		0.354**	-0.75658		-165.78*	
HHEmpMale		1.39775				
HHEmpNone			-1.23	-1.87997		-289.28541
HHLeaderFemale		4.41942				
HHLeaderMale		5.06044				
HHLeaderBoth			-4.73	-8.2535		-800.71963
HHRaceAsian		0.581**				
HHRaceBlack		-0.528**				-174.2506
HHRaceHispanic		1.83305	2.10006	4.17278	2.58898	322.68283
HHRaceOthers		-7.61886	-7.06802	-13.4791	-7.51228	-1531.30357
Adj R-Sq	0.2832	0.4572	0.4576	0.5092	0.4292	0.4996

Equation 1: Parameter Estimates (Dependent = SoyMilk Purchase per Household)
Equation 2: Parameter Estimates (Dependent = SoyMilk Purchase per Household)
Equation 3: Parameter Estimates (Dependent = SoyMilk Purchase per Household)
Equation 4: Parameter Estimates (Dependent = CFLF Milk Purchase per Household)
Equation 5: Parameter Estimates (Dependent = Organic Milk Purchase per Household)
Equation 6: Parameter Estimates (Dependent = Fat Purchase per Household)

(*) Significant at 5% level. (**) Significant at 10% level. (') Insignificant.

All variables are significant at 1% level unless indicated otherwise.

Figure 1.1 Dairy Market Overview

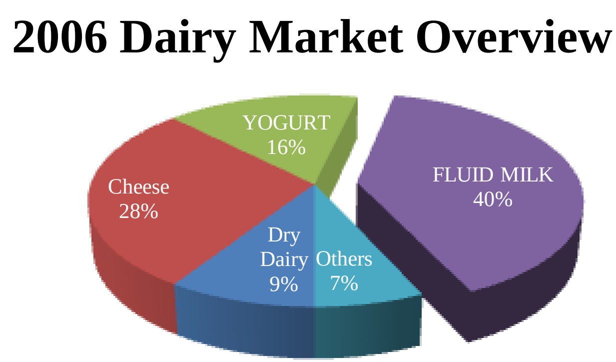


Figure 1.2 Average Shares in Milk Market

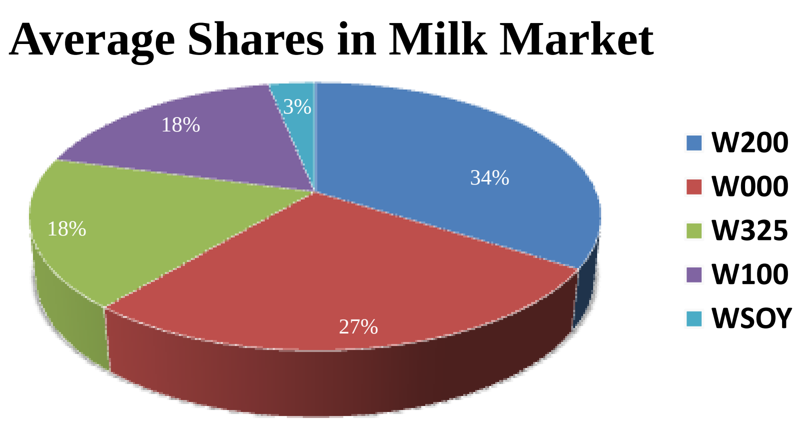
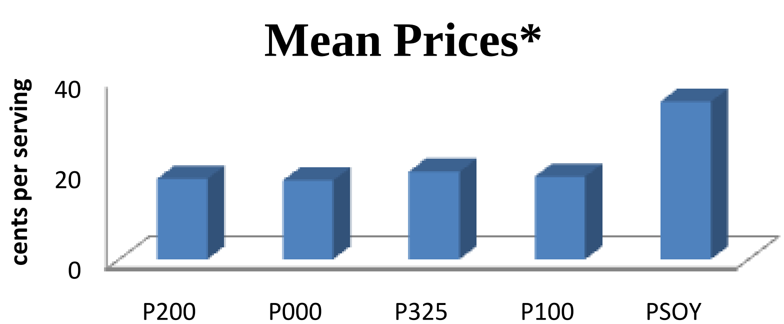


Figure 1.3 Average Prices in Milk Market



*Prices are measured in terms of cents per serving. A gallon of milk has 16 servings in a single package.

Appendix

Appendix 1 Applied Analysis

Applied Analysis

- ❖ **National Brand Whole Milk: 100% Lactose Free, Calcium Fortified. 64.00 fl oz**
- ❖ **Sold at a local grocery store in a Carton Container.**
- ❖ **Serving Size = 240 ml. Servings per Container = 8 Servings.**

Nutrition Facts

Serving Size 240 ml
Servings Per Container 8

Amount Per Serving	
Calories 150	Calories from Fat 70
	% Daily Value*
Total Fat 8g	**12%**
Saturated Fat 5g	**25%**
Polyunsaturated fat 0g	
Cholesterol 35mg	**11%**
Sodium 125mg	**5%**
Total Carbohydrate 12g	**4%**
Dietary Fiber 0g	**0%**
Sugars 12g	
Protein 8g	
Vitamin A	6%
Vitamin C	0%
Calcium	50%
Iron	0%
Vitamin D	25%

* Percent daily values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs:

Attribute	Content	Added Value	
		2002 - 1st Quarter	2005 - 4th Quarter
USDA Organic	X	0	0
Soy Based	X	0	0
Vit-Min Enhancement Label	✓	2.90	1.75
Lactose / Cholesterol Free	✓	19.41	18.43
Protein Content (g)	8	9.37	7.80
Carbohydrate Content (g)	12	8.58	8.76
Fat Content (g)	8	4.99	5.35
Vitamin Mineral DRI Index (%)	16	2.03	7.69
Cholesterol DRI Index (%)	5	-3.69	-2.81
Sodium DRI Index (%)	11	-5.38	-8.10
Purchased Serving Quantity	8	-0.53	-0.65
Purchase Location	Grocery	1.34	1.56
Container Type	Carton	1.47	4.06
Location of Household	South	1.00	1.02
Estimated Price Per Serving		41.49	44.86
Estimated Total Price		\$3.319	\$3.589
Actual Price in March 18, 2008		\$3.89	

Chapter 2

2. A Hedonic Metric Approach to Estimating Demand Using Micro Level Data

Introduction

A significant amount of empirical research concentrates on estimating price elasticities between products of similar types. In traditional models such as the Almost Ideal Demand Systems (AIDS) and the Rotterdam Model (RM), the time-series relationships between prices and market shares is exploited to estimate the own-price, cross-price and income elasticities. Although symmetry, adding-up, and homogeneity reduce the number of parameters to be estimated, as the number of goods gets large, the number of parameters to be estimated increases exponentially.

Hedonic models have been widely used to analyze the price differentials between goods with respect to differences in characteristics. Yet to date, little work has been done to link hedonic characteristics to price differentials in estimation of elasticities. In recent years random utility models have been a popular approach to estimating demand elasticities. In this approach, the utility maximizing consumer chooses the product that gives the maximum utility derived from the commodity attributes. However, this approach assumes that the consumer chooses only one of the alternatives among many choices. Random utility models that use the simulated maximum likelihood approach of Berry, Levinsohn, and Pakes (1995) to relax this assumption can explain multiple purchases yet they are very computationally complex to be estimated. A simpler distance metric approach by Pinske, Slade, and Brett (2002) uses spatial distances between locations to estimate price elasticities between locations. Rojas and Peterson (2008) apply this method to the retail beer market using alcohol content as the main distance measure along with different combinations. Although the authors claim that the results are not affected by their choice of distance, their choice of distances is ambiguous and depends on prior judgments about the data. The hedonic metric method applied in this paper alleviates this ambiguity while reducing the number of parameters to estimate. Most significantly, it is based on strong theoretical foundations.

The model's performance is evaluated through estimating elasticities in the retail milk market. The retail milk market is a dynamic market where there has been increased competition between traditional milk products and specialty products such as soymilk and organic milk. Although the share of soymilk and organic milk is very low compared to other milk types, they have a higher price premium and their share of the market has been increasing rapidly. This trend is evident in the data we use; A.C. Nielsen Homescan Data obtained from purchase information for a panel of over thirty thousand participants during the period 1998 to 2005. From this panel, a core sample of about three thousand households is selected since these households regularly participated in the panel from 2002 to 2005. These data are merged with the product contents that are derived from USDA Nutrient Database to get continuous product attributes, including protein, fat, and carbohydrate content. The differences in the attributes of milk varieties are exploited to create a single distance matrix based on hedonic distance. This distance is based on pair wise comparison of the amount of each characteristic in the product weighted by its hedonic price. The distances used in the distance metric approach are based on different combinations of organic percentages, average servings per package and fat content of each type.

The LA/AIDS and RM models are ideal for comparing the HM approach with the DM approach. The performance of distances is measured with respect to closeness to the original models. Both LA/AIDS and RM are well accepted and widely applied in the literature. Although both models are based on neoclassical theory that exploits the price - market share relationships between products, using different variables might give different results. Since the aim of this paper is to compare the HM and DM approaches, the performance of different distances in different models are of interest. Therefore both LA/AIDS and RM models are estimated. The results are surprisingly close to the original models for both approaches.

The next section in this paper includes an extensive literature review regarding different models and estimation techniques along with a discussion of the milk market. The details of data, the sampling, aggregation issues and product attributes used in creating distances are explained in the data section. The model section introduces the basics of LA/AIDS and RM Models. A detailed derivation of DM and HM models is also included in the model section. The estimated uncompensated price elasticities and income elasticities are compared in the results section. Finally, the paper concludes with a summary that also suggests directions for future research.

Literature Review

Hedonic theory suggests that the price of each good can be characterized by a set of its attributes. If we define this set as $x = [x_1, \dots, x_k]$ then the price of a good can be explained by the relative unit prices and quantities of its characteristics. In this case, a functional relationship exists between the price of a good and its' characteristics vector x such that $p = f(x) + \mu$ where μ is the error term. The implicit marginal prices of characteristics can be calculated as the partial derivatives of the hedonic functions:

$$\frac{\partial p}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} \; i = (1 \dots k) \tag{2.1}$$

While analyzing the implicit prices in food market, Ladd and Suvannunt (1976) define the characteristic vector as water, food energy, protein, fat, carbohydrate, ash, calcium, phosphorus, iron, sodium, riboflavin, potassium, vitamin A, thiamine, niacin, and ascorbic acid. Starting from this set and iteratively eliminating the insignificant ones, the characteristic vector is reduced to food energy, protein, carbohydrate, phosphorus, iron, potassium, riboflavin, and ascorbic acid. Interestingly their results show that phosphorus, iron, and ascorbic acid has negative prices which the authors explain as a result of unpleasant taste, texture, or odor.

Rosen (1974) developed a structural model to identify and estimate the unobservable attribute prices. His model is more general than that of Ladd and Suvannunt's because it also includes consumer characteristics in the characteristic vector. Over time, Rosen's hedonic model has become the most commonly used model in the property value literature. As Palmquist (2003) indicates, the model allows the econometrician to estimate a schedule of prices for differentiated goods. Hedonic regressions are also widely accepted in analyses of markets for consumer goods. Nimon and Beghin (1999) define the characteristics of apparels almost entirely in terms of dummies for item types, catalog types, gender, age, dye type and organic / nonorganic type.

In discrete choice modeling of good characteristics, the consumer is assumed to select the best product that gives him the highest utility among his choice set where the utility is defined as a function of the product attributes. (Cropper, et al. 1993). In this respect discrete choice models are similar to hedonic models. However, these models impose a good deal of structure on the consumer's preference function. In discrete choice models, the distribution of the error terms characterizes the estimation technique. McFadden's nested multinomial logit model (1978) is among the most popular discrete choice models. Davis and Wohlgenant (1993) applied this model to natural and artificial Christmas tree choices to estimate demand elasticities. They use a two-stage procedure where consumers first decide whether to purchase or not; given the decision to purchase next they decide on the type of the tree.

Anderson and Palma discussed the logit model with different applications. Their first paper discusses the multinomial logit as a model of product differentiation. They introduce a taste parameter that measures the response of consumer's utility to the product variety. This method provides a simple parameterization of diversity in consumer tastes (Anderson and De Palma 1992a). In another paper, the authors apply a nested logit model to multiproduct firms. According to their nested logit specification, products that are close substitutes are combined in a nest. They parameterize their model using two constants that captures the degree of substitution between the nests and within a single nest (Anderson and De Palma 1992b).

Cotterill and Dhar (2003) apply a similar analysis using the nested logit model on the Boston fluid milk market. They utilize monthly retail scanner data from Information Resources Incorporation (IRI) and other data from public sources. Their results indicate that competition within the store is higher than across stores. They assume oligopolistic competition structure which is maintained through brand labeling. The two-stage demand is modeled such that consumers first choose the retail store, and then they choose the brand of the milk. The characteristics of the milk are retail price, volume per unit, ratio of whole milk to skim/low fat milk, weighted average price reduction, and percentage volume merchandising. The demand analysis indicates that the ratio of skim/low fat to whole milk and weighted average price reduction are positively correlated with quantity demanded. The effect of price on quantity is negative as expected. After calculating the relevant elasticities, the authors conclude that consumers have a tendency to stick with their choice of store rather than the choice of brand. It is easier for a consumer to switch brands within the store rather than changing the store for the same brand. This indicates that stores have a local monopoly power in the neighborhood they are located.

Berry, Levinsohn, and Pakes (1995) (BLP) introduce the random parameters approach to account for consumer's preferences. The producer maximizes profits given the attributes and prices of its' products and prices of competitors' products. Estimation complexities lead them to apply simulated method of moments where an algorithm simulates the moment conditions. Although the model does not extend the single purchase discrete models to multiple purchases of different products, introduction of random coefficients produces more flexible substitution patterns.

Based on the BLP methodology, Igal Hendel (1999) relaxes the single choice among mutually exclusive alternatives assumption. Like the BLP methodology the consumer's perception for quality of products is measured by introducing random coefficients. The author presents a multiple-discrete choice model for the analysis of differentiated products. The model is applied to a firm's input demand for personal computers. It is assumed that each firm has a number of potential tasks that is related to firm's characteristics. Therefore it is possible to observe purchases of multiple types of computers since different computers are used for different tasks.

Another application of random coefficient discrete choice model can be found in Nevo (2000, 2001). In these papers, the author models the demand for ready-to-eat cereal products in terms of product attributes, heterogeneous consumer characteristics, and unknown (random) parameters. The data obtained from Information Resources, Inc. (IRI) is used to measure the market power of each cereal company in this market. The heterogeneity is modeled as a function of empirical nonparametric distribution of demographics.

Chan (2006) extends the BLP methodology of discrete choice analysis to the hedonic framework. The model utilizes different random parameters that create flexible substitution patterns which can characterize consumer behavior in the retail market. The author estimates the model using simulated method of moments that is applied to soft drinks markets. The model is flexible enough to characterize the multiple-type, multiple-unit purchasing behavior in the retail market. In this model, the consumer chooses a variety of products in order to maximize the utility derived from the consumption of product attributes subject to budget constraint. The estimation utilizes a two stage nested algorithm based on the simulated method of moments suggested by McFadden (1989).

An alternative to the BLP methodology is the DM modeling approach of Pinske, Slade, and Brett (PSB) (2002) based on the semi-parametric estimation that uses spatial distances between locations. This method is easier to estimate than the random coefficients discrete choice models and it allows for different types of competitive effects based on the products' location in characteristic space. PSB's model is simple enough to be estimated without simulations, yet flexible enough to characterize the substitution patterns between hundreds of locations. It accounts for both global competition where all companies compete with each other, and also local competition, in which products compete only with their closest substitutes. While implementing the product attributes, consumers are also allowed to have a taste for variety and thus they can purchase multiple products in multiple quantities. The model is estimated using a cross section of prices, locations, and other variables at terminals of the U.S. wholesale gasoline markets. The estimator utilizes measures of the distances between sellers, where the distance is measured in terms of sellers' location in geographic space. The authors assume that a strong relationship exists between these distance measures and the degree of competition between locations. To get the best distance measure the authors experiment with several distance measures. PSB first create continuous Euclidean distances between terminal locations. Next, they create binary discrete variables that are based on whether firms are nearest neighbors, whether they share a market boundary, and whether they share a common boundary with a third competitor. PSB found that the competition in the wholesale gasoline market is highly localized. Moreover if two terminals are nearest neighbors, the distance between them is not much important.

Another important result that PSB point out is that, although their estimates are based on a semi-parametric approach, similar outcomes can be obtained using a parametric model. Moreover, the authors suggest that pair wise distances based on product attributes can also be used to locate the products in the attribute space. Therefore it is possible to estimate the degree of competition based on these distances. Following PSB methodology, Rojas and Peterson (2008) analyze the demand for beer in the U.S. beer market using the DM approach within the Linear Approximated Almost Ideal Demand System (LA/AIDS) Model.

Rojas and Peterson specify the cross-price coefficients as functions of different distance measures. They also define the own-price and own-expenditure coefficients in terms of products' characteristics. The elasticities are recovered from the interaction of these distance coefficients and the prices. For the beer market a natural choice of primary distance is alcohol content (ALC) of each brand in the market. The authors also experiment with other distance combinations based on product coverage (COV) and the container size (SIZE). Using these distances authors come up with three single dimensional (ALC/COV/SIZE), three two-dimensional (ALC-COV/ALC-SIZE/COV-SIZE), and one three dimensional (ALC-COV-SIZE) distances. These continuous distance measures are scaled to be between 0 and 1 using the inverse of Euclidean distances. There are also many discrete distance measures based on whether the brands are in the same segment, whether they are produced by the same brewer, and whether they are national or local brands. Following PSB's suggestions they also include binary distances based on whether two brands are nearest neighbors in product space, or whether they share a common boundary. Considering tens of possible combinations of these distance measures, the authors first start with preliminary OLS regressions to determine the most significant continuous & discrete distance measures. An elimination method tests for different distance combinations to determine the best measure.

Data

In this paper we estimated the demand for fluid milk products using price and market share data for retail milk types derived from AC Nielsen Homescan Panel. Household level point of sale purchase data is aggregated to obtain weekly price, quantity and expenditure data for five different milk types. These milk types are classified according to their fat content such 2% fat, skim (0%), whole (3.25%), 1% fat, and also soymilk. The households that regularly participated in the panel for at least 12 times a year are selected to get the purchase behavior of core households over a four year period between 2002 and 2005. The demographics of the core household data resemble that of the entire sample contained in the AC Nielsen data. A detailed comparison of core household data with the original AC Nielsen data regarding household demographics can be found in **Table 2.1**. Although detailed information on consumers is available in the Homescan data, the product attributes are limited to discrete marketing variables based on package label. Most discrete variables are binary such as organic label, soymilk, lactose free, cholesterol free, vitamin enrichment, calcium enrichment. There are also some non-binary discrete attributes such as container type, product type, product group, size, etc. Because the AC Nielsen data is limited to discrete product attributes, information about attributes are enriched using USDA-ERS Nutrient Database to obtain detailed information on product contents. In particular continuous nutritional contents including protein, carbohydrate, fat amount, along with sodium and cholesterol contents are obtained from the nutrient database.

AC Nielsen Homescan Data

AC Nielsen Homescan Panel track households based on individual's retail purchases. We obtained a portion of this data related to dairy purchases made by 50223 households across the United States between 1998 and 2005. There is an issue of entrance/exit by many households because not all households participate regularly in the panel. Moreover, an important share of the households does not regularly report their purchases. To overcome these issues, a set of core households were selected based on regular participation (at least 12 times a year) in the panel between years 2002 and 2005. These households record their purchases by a scanner and subsequently the data is uploaded to AC Nielsen database. Scanner data possesses information on prices/discounts, volumes/expenditures and purchase dates of all dairy products purchased by consumers together with their demographic background.

The Homescan panel has been an important source of data for many researchers in industrial organization and empirical microeconomics area. It is also being used by the US Department of Labor, BLS, to calculate price indices based on hedonic regressions (Silver 2003). The rich nature of the data makes it possible to analyze brand-level demand and functional food consumption. Jensen and Huffman (2003) analyze the demand for enhanced foods and the value of nutritional enhancements in food products. In their paper, the authors model the consumers' choice set based on their economic, ethnic, and other socioeconomic characteristics with an application to the margarine market. Binkley et al. (2000) analyze the purchase patterns of consumers who buy certain categories of functionally enhanced foods. One of the major obstacles they encounter with Homescan data is the way that different terms might refer to the same type of functional enhancements. For example, RF (reduced fat) and LF (low fat) both refer to a reduction in fat content. Another issue is that, although the data is rich in content, one encounters many zero purchase observations. To overcome this issue we aggregated dairy purchases on a weekly pattern based on fat content.

USDA National Nutrient Database

The USDA National Nutrient Database is the major source of food composition data for food products consumed in United States. For almost any type of food product the Nutrient Database gives a detailed scientific analysis of components in the product. It also comes in an abbreviated form that is convenient for our analysis. Even in this abbreviated form, the components are almost exhaustive and not only include information about major food components such as protein, carbohydrate, fat, but also a detailed analysis of these components. For example fat content is disaggregated into saturated fatty acids (g/100g), monounsaturated fatty acids (g/100g), polyunsaturated fatty acids (g/100g), and also cholesterol (mg/100g) content. Moreover, the list of vitamins and minerals included in the database ranges from A to Zinc. Most of the vitamins included in the database are given in terms of milligrams per 100 gram (mg/100g) yet some vitamins such as Vitamin B12 are measured in terms of micrograms per 100 gram (μg/100g) and some are measured in international units per 100 gram (IU/100g) such as Vitamin A. To overcome the differences in measurement units, all values for vitamins and minerals are transformed into the percentages provided in a serving size (1cup=240ml) as suggested by Daily Recommended Intake[2] (DRI) Values based on FDA's reference values for nutrition labeling. **Table 2.2** lists the food labeling guide as suggested by FDA – CFSAN.

Since there is a high amount of correlation among vitamins, we combined them into a single vitamin index based on the average of individual vitamins provided in the products. This vitamin index is based on a weighted percentage of daily recommended dosage of vitamins as suggested by dietitians. A similar aggregation is performed on minerals to come up with a single mineral index. However, since nutritional enhancements increase both vitamin and mineral contents, there was an almost perfect correlation between vitamin and

---

[2] *Center for Food Safety and Applied Nutrition (CFSAN) issues A Food Labeling Guide that is based on a diet of 2000 calorie intake. This guide suggests the recommended daily intake values of nutritional component. The list can be found at http://www.foodsafety.gov/~dms/flg-7a.html*

mineral indexes. To overcome this problem, all vitamins and minerals other than sodium content were combined into a single vitamin-mineral index. Sodium and cholesterol are two important components listed in the product labels that might raise health concerns among consumers. In fact they might even have a negative effect on the value of the product. Therefore they are included as a separate component for nutritional analysis. Other nutritional components included in the hedonic analysis can be listed as protein, carbohydrate, and lipid (fat) content.

Product Attributes

Nutritional food attributes derived from the Nutrient Database and the marketing attributes derived from Homescan Data characterize milk products in hedonic space. The final attribute space includes marketing variables such as fat type, organic claim, soy dummy, promotion dummy, lactose/cholesterol free (LFCF), vitamin-mineral enhancement, and nutritional variables such as protein, carbohydrate, lipid (fat) content, percentage DRI index of cholesterol, sodium, and vitamin/minerals along with the servings per package. The mean values of these components by fat type can be seen in **Table 2.3.** As indicated by **Table 2.3**, soymilk has a much higher organic ratio compared to other types of milk. Soymilk is also promoted more than other milk types and it is lactose/cholesterol free. An interesting difference can be observed in average serving size per package. On average, soymilk comes in smaller packages than cow's milk.

Aggregated Data

In order to estimate the relationship between prices and quantities, the data have been aggregated into weekly purchases made by all core households. The data contain information on prices, quantities and expenditures on five different milk types between 2002 and 2005. The number of observations used in the estimation is equal to 52 (the number of weeks in a year) times 4 (years data available), equaling 208 observations. Average market share of 2% milk is highest with a 34% share, followed by skim milk with a 27% market share. Whole (3.25%) and 1% milk have 18% market shares on average. However, as shown in **Figure 2.1**, the market share of whole milk has been decreasing whereas the market share of 1% milk has been increasing. The price of soymilk per serving is almost double that of other milk types. Although this price premium has been decreasing over time, it is a distinguishing difference which might be due to differences in product attributes. Because soymilk is primarily organic, lactose free, cholesterol free, and consumed in smaller (more convenient) packages, it has more desirable attributes than other milk types. **Table 2.4** gives the summary statistics for each type of milk.

Figure 2.2 shows the price-quantity functions for each type of milk. As can be seen from **Figure 2.2**, there is strong negative relationship between the prices and quantity demanded for each type of milk as expected. One striking result that can be observed by comparing **Figure 2.1** with **Figure 2.2** is that as the price of soymilk decreases, the market share of soymilk rises. Price plots suggest that the prices of conventional milk types have trended upward over time. Moreover, these prices are highly correlated. The average price of whole milk is highest among these products and the price differential between whole milk and other types of conventional milks have been widening. That might be a factor causing reduction in market share of full-fat milk. Also the retail prices for non-soymilk products increased sharply following the USDA's decision to increase the base per-gallon price paid to dairy farmers by 50 cents (Nii 2004).

Distances

The data on prices and market shares are enough to estimate the price and income elasticities in the context of RM and LA/AIDS models. Although we use the same time series data for DM/HM methods, defining distances is the most crucial part of DM and HM approximations to the original models. In this paper, not only the cross-price elasticities are recovered from these distance coefficient estimates but also the own-price elasticities.

Since the most distinguishing characteristics of milk types are organic percentage, fat content, and container size, the distances based on these characteristics are used in DM estimation. The continuous distance measures can be single dimensional based on these attributes (Fat, Organic, Container) or two-dimensional based on pair-wise distances (Fat-Organic, Fat-Container, Organic-Container) or three dimensional (Fat-Organic-Container). The distance measures are calculated from the differences between milk types based on these measures. Following PSB's methodology, a discrete distance based on nearest neighbor (NN) concept is also introduced for two-dimensional and three-dimensional distances. Two products are nearest neighbors if they are next to each other in the attribute space. Note that the nearest neighbor concept is not symmetric, i.e., a milk type might be the nearest neighbor of another milk type, whereas the reverse might not be true. For example, in terms of fat content distance, the nearest neighbor of whole milk is 2% milk, whereas the nearest neighbor of 2% milk is soymilk. For the purpose of approximating own-price elasticities in DM, the contents based on market share (W), fat content (F), organic claim (O) and servings per package (S) are used to approximate these elasticities.

Similar to the DM approach, the hedonic metric approach proposed in this paper also uses distances between product attributes to approximate the cross-price and own-price elasticities. The distinction between these approaches is based on the distances used in the estimation. In the HM approach, the products are located in the hedonic space which characterizes them. To get the location of each product type, first implicit prices of the attributes that characterize the product are estimated using hedonic regression. Next, using these implicit prices, the value added for each attribute is calculated for each milk type. For a linear hedonic model, the estimated coefficients in attribute quantities give us hedonic prices. For a semi-log model, these coefficients are multiplied by the prices to obtain the hedonic prices.

The difference in the added values scaled by the prices of the attributes gives us the location of each product in hedonic space. Thus, the hedonic prices of the attributes scale the attributes according to their values. The nearest neighbor concept in HM approach is based on hedonic distances. For estimating the own-price elasticities, a different concept based on the sum of total pair wise distances between each product is proposed. Since the distance measure is in inverse form, distance acts as a total closeness index based on hedonic distance.

For example, because soy is the most unique product in hedonic space, it has the lowest closeness index among all milk types. This closeness index based on the inverse of hedonic distance is used along with the product's market share to estimate own-price elasticities for each milk type. We estimate four different hedonic regressions (Quantity Weighted Linear, Quantity Weighted Semi-Log, Non-weighted Linear, Non-weighted Semi-Log) so there are four different hedonic distance measures.

Model

The Rotterdam Model and Almost Ideal Demand System are among the most popular and widely accepted functional forms in empirical microeconomic analysis. The Rotterdam Model is derived by totally differentiating the Marshallian demand functions and substituting the Slutsky equation to derive the relationship between market shares and prices in a demand system such that for *n* goods:

$$\begin{gathered}
\bar{w}_i d \log q_i = b_i d \log \bar{X} + \sum_j c_{ij} d \log p_j \\
d \log \bar{X}_t = d \log X_t - \sum_j \bar{w}_t d \log p_{jt} \\
d \log p_{jt} = \log p_{jt} - \log p_{jt-1} \\
\bar{w}_t = 0.5(w_t + w_{t-1}) \\
d \log q_{it} = \log \mathrm{q_{it}} - \log \mathrm{q_{it-1}}
\end{gathered} \tag{2.2}$$

X_t refers to the total expenditure on products at time t. Theoretical restrictions can be imposed in the form of homogeneity, $\sum_j c_{ij} = 0\ \forall i$ symmetry, $c_{ij} = c_{ji}\ \forall i, j (i \neq j)$ and negativity $c_{ii} < 0\ \forall i$.

- Dividing the coefficient on differenced log price with the average own market share gives the compensated (Hicksian) price elasticity: $e_{ij} = c_{ij}/w_i\ \forall i, j$
- Dividing the coefficient on differenced log expenditure with the own market share gives the income elasticity: $e_i = b_i/w_i\ \forall i$

- Uncompensated (Marshallian) demand elasticities are recovered using the Slutsky equation in elasticity form: $e_{ij}^{m} = e_{ij} - e_i * w_j = (c_{ij} - w_j b_i)/w_i \; \forall i, j$

Using Shepherd's lemma and imposing the homogeneity and symmetry restrictions we derive the AIDS Model such that

$$\begin{gathered} w_i = \alpha_i + \sum_j c_{ij} \log p_j + \; b_i \; \log(x/P) \\ \text{where } \log P = \; \alpha_0 + \sum_j \alpha_j \log p_j + 0.5 \sum_i \sum_j c_{ij} \log p_i \log p_j \end{gathered} \tag{2.3}$$

Dividing the coefficient on total expenditure with the coefficient on own market share plus one gives the income elasticities: $e_i = b_i/w_i + 1 \; \forall i$

- Marshallian demand elasticities are derived as:

$$e_{ij}^{m} = \frac{c_{ij} - b_i\big(w_j - b_j \log(x/P)\big)}{w_i} - \delta_{ij} \; where \; \delta_{ij} = 1 \; if \; i = j \; and \; c_{ij} = 0 \; if \; i \neq j \; \forall i, j$$

- Using the Slutsky equation in elasticity form we can derive the Hicksian elasticities as:

$$e_{ij} = \; e_{ij}^{m} + e_i * w_j \; \forall i, j$$

The linearly approximated (LA/AIDS) version makes the estimation easier by converting the nonlinear form into a linear model such that

$$\begin{gathered} w_i = \alpha_i^* + \sum_j c_{ij} \log p_j + b_i \log(x/P^*) \\ \text{where } \log P^* = \sum_i w_i \log p_i \;\; and \; \alpha_i^* = \alpha_i - b_i \alpha_0 \end{gathered} \tag{2.4}$$

Moschini (1995) indicates that the Stone price index is not invariant to changes in units of measurement in prices. He suggests using different price indexes like Paasche price index, P^P such that $\log P^P = \sum_i w_i \log(p_i/p_0)$ where p_0 refers to the base prices. We normalized the prices by their means in order to correct our price index is for scaling invariance. Moreover, as Asche and Wessels (1997) show, by normalizing the prices to one at their mean values, we can use the same elasticity formulas for LA/AIDS model. In fact, the LA/AIDS estimation gives the same elasticities with the AIDS model at the point of normalization.

The elasticities are calculated in a similar fashion to the original version.

- Marshallian elasticities: $e_{ij}^{m} = \left(c_{ij} - b_i * w_j\right)/w_i - \delta_{ij}\ \forall i, j$
- Income elasticities: $e_i = 1 + b_i/w_i\ \forall i, j$
- Hicksian elasticities: $e_{ij} = \left(c_{ij} - b_i * w_j\right)/w_i - \delta_{ij} + (1 + b_i/w_i)\ w_j\ \ \forall i, j$

Since the elasticity results derived from both quadratic and linear AIDS model are very similar to each other, choosing one over the other does not affect the results. Therefore we decided to continue with the LA/AIDS model since the full AIDS model can become very complex especially when using distance/hedonic metric approximations.

Distance Metric Method

The distant metric method introduced by PSB is an approximation method to estimate the relationships between prices and market shares. Following their approach the cross-price coefficients $(c_{ij}\ \forall\ i, j)$ are specified as functions of distant measures between different products such that

$$c_{ij} = \sum_{l=1}^{L} \lambda_l d_{ij}^{l} \tag{2.5}$$

where L is the number of attribute spaces and d_{ij}^{l} is the distance between product i and j in space l. These distant measures are based on the product attributes and can be continuous (content), discrete (type) or both. The continuous distance measures are based on the inverse of Euclidean distance in attribute space between products. In this form, these measures range between 0 and 1. This closeness index refers to the closeness of these products such that a higher index (close to 1) implies closer products whereas a lower measure (close to 0) implies distant products. For an n-dimensional attribute space, the distance l between two products i, j can be defined as

$$d_{ij}^{l} = \frac{1}{1 + \sqrt{\left(\delta_{ij}^{1}\right)^2 + \left(\delta_{ij}^{2}\right)^2 + .. + \left(\delta_{ij}^{n}\right)^2}} \tag{2.6}$$

$$where\ \delta_{ij}^{k}\ is\ the\ distance\ measure\ in\ dimension\ k$$

For example, the closeness index between 2% milk and 1% milk based on a three-dimensional fat-organic-size (FOS) attribute space can be calculated as

$$d_{2\%,1\%}^{FOS} = \frac{1}{1 + \sqrt{\left(\delta_{2\%,1\%}^{FAT}\right)^2 + \left(\delta_{2\%,1\%}^{ORGANIC}\right)^2 + \left(\delta_{2\%,1\%}^{SIZE}\right)^2}}$$

Following Rojas and Peterson, the continuous distances between products are scaled by dividing the differences in contents with the maximum amount of content available in any product. For example, the fat distance between 2% Milk and 1% Milk is calculated as

$$\delta_{2\%,1\%}^{FAT} = \frac{Fat\ Content\ of\ 2\%\ Milk - Fat\ Content\ of\ 1\%\ Milk}{Fat\ Content\ of\ 3.25\%\ Milk}$$

The discrete distances can be based on the type of the product depending on whether they are in the same classification or not. In that form they are either equal to 1 (same classification) or 0 (different classification). They can also be based on whether the products are neighbors in any given product space or not. A product is the nearest neighbor of another product if it has the highest closeness index for a given attribute space. Note that while the continuous distances between products are symmetric ($\delta_{ij}^n = \delta_{ji}^n \rightarrow d_{ij}^l = d_{ji}^l$), the nearest neighbor concept is not symmetric. A particular milk type might be the nearest neighbor for another type whereas the reverse might not be true. For example, in the two-dimensional organic-fat space the closest neighbor (highest index) of 2% milk is 1% milk, whereas in the same attribute space the closest neighbor of 1% milk is skim milk. By construction, the discrete distance measures are also normalized to one such that for a given attribute space, the sum of distance measures for each product type equals one.

The own-price coefficients are also specified in terms of product attributes. However, this time instead of the distances, the actual amount of product attributes interact with the own-price coefficients.

$$c_{ii} = \beta_0 + \sum_{j=1}^{C} \beta_j \chi_{ji} \qquad (2.7)$$

β_0 is the constant coefficient on the own-prices and χ_{ji} is the content of characteristic j in product i that interacts with the price.

Incorporating these parameter approximations to the RM model gives us the following empirical distance metric approximated RM model

$$\begin{aligned}\bar{w}_i d\log q_i &= b_i d\log \bar{X} + \beta_0 \,\mathrm{dlog}\, p_i \\ &+ \sum_{j=1}^{C} \beta_j \chi_{ji} \,\mathrm{dlog}\, p_i + \sum_{i\neq j}\sum_{l=1}^{L} \lambda_l d_{ij}^{l}\, d\log p_j \;\; \forall i,j\end{aligned} \tag{2.8}$$

Since different approximation techniques are used for own-price and cross-price coefficients, the corresponding elasticities are calculated based on the approximation technique.

- The Hicksian own-price elasticity for RM Model, $e_{ii} = \left(\beta_0 + \sum_{j=1}^{C} \beta_j \chi_{ji}\right)/w_i \;\forall\, i$ whereas the Hicksian cross-price elasticities can be calculated as $e_{ij} = \left(\sum_{l=1}^{L} \lambda_l d_{ij}^{l}\right)/w_i$ $\forall\, i \neq j$.
- Income elasticities are calculated as the original RM Model as $e_i = b_i/w_i \;\forall i$
- The Marshallian elasticities are recovered using the Slutsky equation:

$e_{ii}^{m} = \left(\beta_0 + \sum_{j=1}^{C} \beta_j \chi_{ji}\right)/w_i - b_i \;\forall i$ and $e_{ij}^{m} = \left(\sum_{l=1}^{L} \lambda_l d_{ij}^{l} - b_i w_j\right)/w_i \;\forall\, i \neq j$

Applying the distance metric approximation to the LA/AIDS Model results in the following empirical form

$$w_i = \alpha_i^{*} + \beta_0 \log p_i + \sum_{j=1}^{C} \beta_j \chi_{ji} \log p_i + \sum_{i\neq j}\sum_{l=1}^{L} \lambda_l\, d_{ij}^{l} \log p_j + b_i \log(x/P^{*}) \tag{2.9}$$

where P^{*} and α_i^{*} are same as original LA/AIDS model.

The elasticities are calculated by substituting the original own-price and cross-price coefficients with the approximated forms.

- Marshallian Elasticities can be recalculated as follows:

  $e_{ii}^{m} = \left(\beta_0 + \sum_{j=1}^{C} \beta_j \chi_{ji}\right)/w_i - b_i - 1 \;\forall i$ and $e_{ij}^{m} = \left(\sum_{l=1}^{L} \lambda_l d_{ij}^{l} - b_i w_j\right)/w_i \;\forall\, i,j$
- Income elasticities: $e_i = 1 + b_i/w_i \;\forall\, i$
- Hicksian elasticities can be derived using the Slutsky equation in elasticity form.

Hedonic Metric Method

The hedonic metric (HM) method proposed in this paper is an alternative approximation to the distance metric (DM) method. Similar to the DM approach, the elasticities between differentiated products are approximated using the distances between products. However, in the DM approach these distances are based on specific attributes which can be multidimensional (Fat-Organic-Size) or single dimensional (Fat). The HM approach allocates each product in the multidimensional hedonic space. In order to create distances between products first a hedonic regression is estimated to get the hedonic prices of each attribute.

According to the hedonic theory, each consumer is trying to maximize own utility that depends on the product attributes. Therefore, the consumer selects products that will maximize the sum of utilities derived from each attribute. Based on the hedonic model the price of each good can be characterized by the set of its attributes. Defining this set as $x = [x_1, \ldots, x_k]$, the functional relationship between the price of a good and its' characteristics vector x can be stated as $p = f(x) + \mu$ where μ is the error vector.

If the relationship is assumed to be linear, then the price of each good can be derived as the sum of the attribute values. Thus the retail price of the product is equal to the sum of monetary values of product attributes. The total value of each attribute is equal to the quantity of the attribute multiplied by the implicit price of that attribute. This implies

$$P_i = \sum_{j \in J} x_{ji}\beta_j + E_i + \in_i \; J \; is \; the \; set \; of \; product \; attributes \qquad (2.10)$$

where x_{ji} is the amount of attribute j in product i and E_i is the unique characteristic of the product.

Regarding the hedonic regression above, the implicit prices of characteristics can be calculated as the partial derivatives of the hedonic functions.

$$\frac{\partial p}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} = \beta_j \;\; \forall j \qquad (2.11)$$

Thus, in the linear model, the coefficients on the attributes give us the hedonic prices for these attributes. The value added for each attribute is calculated by multiplying the implicit prices with the attribute quantities. For product i, the value added from attribute j can be calculated as $v_{ij} = x_{ji}\beta_j$. If the price attribute relationship is assumed to be in semi-log form, then instead of the price, the log-price of the product is defined in terms of attributes such that

$$\log P_i = \sum_{j \in J} x_{ji}\beta_j + E_i + \epsilon_i \tag{2.12}$$

In this form, the implicit price of the attribute is calculated by the multiplying the coefficients on attributes with the price of the products

$$\frac{\partial p}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} P_i = \beta_j P_i \ \ \forall\, j, i \tag{2.13}$$

The value added term also accounts for the price of the product $v_{ij} = x_{ji}\beta_j P_i$. The semi-log form implies that the same amount of attribute can have a higher value if it is located in a product with a higher retail price.

In this paper we have applied the hedonic metric method based on both linear and semi-log models. The multidimensional hedonic space is based on both discrete marketing attributes and continuous nutritional attributes along with the size variables. The marketing variables used in the hedonic regression are organic claim, soy dummy, promotion, cholesterol free – lactose free dummy, and vitamin enrichment index. The nutritional attributes are the amounts of protein (g), carbohydrate (g), fat (g), percentage of cholesterol (%DRI), sodium (%DRI), and vitamin-minerals (%DRI).

In both RM and LA/AIDS models the differences in the added values for each product return the hedonic distance in terms of a single attribute. Combining the sum of these price-weighted attribute distances and rescaling them to be between 0 and 1 gives us the continuous hedonic distance matrix. The nearest neighbor concept introduced as a discrete distance is based on this hedonic distance matrix. The nearest neighbor matrix along with the hedonic distance matrix approximate the cross-price coefficients such that

$$c_{ij} = \lambda_h d_{ij}^h + \lambda_{nn} d_{ij}^{nn} \quad (2.14)$$

where ${d_{ij}}^h$ refers to the hedonic distance and ${d_{ij}}^{nn}$ refers to the distance based on nearest neighbor concept.

Own-price coefficients are also approximated by interacting product attributes with own-prices. These coefficients are estimated based on each products average market share and inverse of the hedonic distance vector (i.e., closeness index)

$$c_{ii} = \beta_0 + \beta_1 \chi_i^s + \beta_2 \chi_i^c \quad (2.15)$$

where χ_i^s refers to the market share and χ_i^c is the closeness index of product i. The average market share is employed as a proxy to differentiate between market coverage and availability of milk types.

The HM approximation to the RM model can be specified as

$$\begin{aligned} \bar{w}_i d \log q_i = b_i d \log \bar{X} + \beta_0 d \log p_i + (\beta_1 \chi_i^s + \beta_2 \chi_i^c) d \log p_i \\ + \sum_{i \neq j} \left(\lambda_h d_{ij}^h + \lambda_{nn} d_{ij}^{nn}\right) d \ln p_j \;\; \forall i, j \end{aligned} \quad (2.16)$$

- In this form, the Hicksian own-price elasticity for RM Model is, $e_{ii} = \beta_0 + \beta_1 {\chi_i}^s + \beta_2 \chi_i^c / w_i \,\forall\, i$ whereas the Hicksian cross-price elasticities can be calculated as $e_{ij} = \left(\lambda_h d_{ij}^h + \lambda_{nn} d_{ij}^{nn}\right) / w_i \,\forall\, i \neq j$.
- Income elasticities are calculated as in the original RM Model as $e_i = b_i / w_i \,\forall i$
- The Marshallian elasticities are recovered using the Slutsky equation:

  $e_{ii}^m = (\beta_0 + \beta_1 {\chi_i}^s + \beta_2 \chi_i^c)/w_i - b_i \,\forall i$ and $e_{ij}^m = \left(\lambda_h d_{ij}^h + \lambda_{nn} d_{ij}^{nn}\right) - b_i w_j / w_i \,\forall i \neq j$

For the LA/AIDS model after applying the HM Approximation we get the following form

$$\begin{aligned} w_i = \alpha_i^* + \beta_0 \log p_i + (\beta_1 \chi_i^s + \beta_2 \chi_i^c) \log p_i \\ + \sum_{i \neq j} \left(\lambda_h d_{ij}^h + \lambda_{nn} d_{ij}^{nn}\right) \log p_j + b_i \log \,(x / P^*) \end{aligned} \quad (2.17)$$

where P^* and α_i^* are same as original LA/AIDS model. The elasticities are calculated by substituting the original own-price and cross-price coefficients with the approximated forms.

- Marshallian own-price elasticities: $e_{ii}^{m} = (\beta_0 + \beta_1\chi_i^s \ + \beta_2\chi_i^c)/w_i - b_i - \ 1 \ \forall i$

  Marshallian cross-price elasticities: $e_{ij}^{m} = \left(\lambda_h d_{ij}^h + \lambda_{nn} d_{ij}^{nn} - b_i w_j\right)/w_i \ \forall i, j$
- Income elasticities: $e_i = 1 + b_i/w_i \ \forall \ i$
- Hicksian elasticities can be derived using the Slutsky equation in elasticity form.

Results

Original Models

The original models are simultaneously estimated using a system of equations for 2%, skim, full, and 1% milk. The equation for soymilk is not included in estimation to prevent singularity. For RM model the cross-price, own-price and income coefficients for soymilk is derived from theoretical restrictions such that

$$\begin{gathered} c_{Soy,Milk_i} = c_{Milk_i,Soy} \ \forall i \\ c_{Soy,Soy} = -\left\{c_{Soy,2\%} + c_{Soy,Skim} + c_{Soy,Full} + c_{Soy,1\%}\right\} \\ b_{soy} = 1 - (b_{2\%} + b_{Skim} + b_{Full} + b_{1\%}) \end{gathered} \tag{2.18}$$

For the LA/AIDS model the coefficients for soymilk is obtained from similar restrictions:

$$\begin{gathered} c_{Soy,Milk_i} = c_{Milk_i,Soy} \ \forall i \\ c_{Soy,Soy} = -\left\{c_{Soy,2\%} + c_{Soy,Skim} + c_{Soy,Full} + c_{Soy,1\%}\right\} \\ b_{soy} = -(b_{2\%} + b_{Skim} + b_{Full} + b_{1\%}) \end{gathered} \tag{2.19}$$

Durbin-Watson statistics for each milk type indicates there is no autocorrelation in the residuals. There are 22 (10 own-price coefficients, 4 cross-price coefficients, 4 expenditure coefficients, 4 constants) independent parameters to be estimated in the models. The Log Likelihood value for RM model is 2738.944 whereas for LA/AIDS model it is slightly higher and equal to 2741.829

In both RM and LA/AIDS models, all own-price elasticities and expenditure elasticities are statistically significant and the elasticity estimates are very close to each other. The only exception is the expenditure elasticity for soymilk. Soymilk has the highest expenditure elasticity in the RM model yet it has the lowest expenditure elasticity in the

LA/AIDS Model. As can be seen from **Table 2.5**, soymilk has the highest own-price elasticity followed by skim milk, 2% milk, 1% milk and whole milk. It is interesting to observe that although whole milk has the highest average price among the dairy based milk types; it has the lowest own-price elasticity. One possible explanation might be the loyalty of the whole milk consumers to their choice of milk type. Moreover, although each milk type has different expenditure elasticities, none of them are statistically significant from 1.

Distance Metric Approximations

All continuous distances along with the nearest neighbor distances for Fat-Organic, Fat-Size, and Fat-Organic-Size are included in the full models. The nearest neighbor (NN) for Organic-Size dimension is not included since it allocates milk types in the same way they are allocated in the Fat-Size dimension. Therefore the full model estimates 23 parameters (8 continuous distance measures, 3 measures, 4 own-price interactions, 4 constants, and 4 expenditure coefficients).

The full models give the lowest AIC and BIC scores, however none of the estimated approximation parameters are significant. Moreover, the mean own-price elasticity for soymilk is significantly different than the LA/AIDS original model. For the full DM approximation to RM model we fail to reject the equivalence of own-price elasticities for all milk types at the 95% significance level.

In both RM and AIDS models, the cross-price elasticities that measure the effect of soymilk prices on the market shares of 1% milk, and full-fat milk ($e^{m}_{1\%,Soy}$ *and* $e^{m}_{Full\,,Soy}$) are also statistically different from the original elasticities between these products. Moreover, the high amount of correlation between distance measures suggests us to use only the most important characteristics as distances. To identify the most distinguishing distances we estimated the DM approximation to RM model using only single continuous distances. The results from these estimations in **Table 2.6** suggest that all distances have positive coefficients as expected.

Since dairy based milk types are classified according to their fat content, closeness in fat content space is included in the estimation. Organic percentage is also another distinguishing attribute along with the average serving size per package that identifies the milk type in the attribute space. A variety of different nearest neighbor distance combinations have been estimated to get the closest approximations to the original models. However, most of these estimations cause singularity among parameters as a result of the high correlation between continuous and discrete distance measures. Three different final distance measures that are used in estimation for both RM and LA/AIDS models can be listed as *Fat, Organic-Size, and NN for Fat-Organic-Size* space (version 1; *F-OS-NNFOS*), *Fat, Organic-Size, and NN for Fat-Organic* space (version 2; *F-OS-NNFO*) and *Fat, Organic, and NN for Fat-Organic* space (version 3; *F-O-NNFO*). In these forms the number of parameters to be estimated is reduced to 15 parameters (2 continuous distance measures, 1 nearest neighbor measure, 4 own-price interactions, 4 constants, and 4 expenditure coefficients).

The estimated Marshallian and expenditure elasticities along with a summary of estimation results for Distance Metric approximations to RM model can be found in **Table 2.7** whereas **Table 2.8** give the results for LA/AIDS Model. In all versions of DM models, the estimated expenditure elasticities match almost perfectly with the original models.

In all three versions of DM-RM models, all own-price and expenditure elasticities are significant and the estimated expenditure elasticities match almost perfectly with the original models. The own-price coefficient terms based on market share, fat content, and organic claim are not statistically significant whereas the constant price coefficients are significantly negative indicating the negative relationship between own-prices and market shares.

Similar to the original results, soymilk has the highest uncompensated own-price elasticity, yet the order of the rest of the own-price elasticities differ significantly. In the original model, own-price elasticity for soymilk is followed by that of skim milk, 2% milk, 1% milk, and full-fat milk whereas in the DM approximation version 1 and version 2, own-price elasticity for soymilk is followed by that of 1% milk, skim milk, full-fat milk, and 2% milk. The DM method underestimates the own-price elasticities for 2% milk and skim milk and it overestimates the cross-price elasticities for full milk and 1% milk. In version 1 and version 2, the coefficient on the inverse of the distance in OS space is positive and statistically significant indicating that the closer are the products in OS space, the higher will be the cross-price elasticities between them. In version 3, the order of own-price elasticities can be ranked as soymilk, followed by 1% milk, full fat milk, 2% milk and skim milk. Although DM method approximates the own-price elasticities for 1% milk and soymilk very closely, it underestimates these elasticities for 2% milk, and skim milk and overestimates that of full fat milk. In version 3, the two-dimensional OS space is replaced with single dimensional organic space, yet the coefficient on the organic distance is also significantly positive indicating the positive relationship between distances in organic percentage space and substitutability of the milk types. Between these three versions, based on the AIC and BIC the version 3; *F-O-NNFO* is the preferred choice since it has the lowest information criteria.

The LA/AIDS model is also approximated using the same three versions of distance measures based on *F-OS-NNFOS* space, *F-OS-NNFO* space and *F-O-NNFO* space. Similar to the DM – RM model, all own-price and expenditure elasticities are significant and almost all elasticities including cross-price terms are within the 95% confidence interval of original elasticities in DM-LA/AIDS model. Only the own-price elasticity for soymilk in the *F-OS-NNFOS* space is slightly higher than the 95% confidence interval for soymilk in the original LA/AIDS model.

One striking difference among the estimation results between DM-RM and DM-LA/AIDS is that the significance of the organic-size distance in the DM-RM model is replaced with that of the fat distance in all versions of DM-LA/AIDS model. For DM-LA/AIDS model the Euclidean distance based on fat content is the dominant distance and significantly positive implying that the closer are the products in fat content, the higher will be the cross-price elasticities between them.

The DM-LA/AIDS model makes a better approximation of the own-price elasticities in the sense that the own-price elasticity of soymilk is highest, followed by skim milk, 1%

milk, 2% milk, and full milk. In the LA/AIDS form the DM approximation almost gives perfectly matching results for the own-price elasticities in all three forms. Since the dominant continuous distance is the same in all three versions of DM-LA/AIDS, the elasticity estimates do not differ much between each model. Moreover, the calculated AIC and BIC values are very close to each other, yet in contrast to the DM-RM results, the model based on *F-OS-NNFO* space has the lowest AIC and BIC values among all versions of DM-LA/AIDS models.

Hedonic Metric Approximations

The monetary values of product attributes are initially estimated by both linear and semi-log hedonic regressions using quantity weighted regressions. For each purchase occasion, the error terms are weighted with the number of servings per package. However, it is also of interest to see how the results differ when we do not include the servings per package and estimate the attribute values per package. Considering weighted and non-weighted options we estimated four hedonic systems; weighted semi-log (version 1), weighted linear (version 2), non-weighted semi-log (version 3) and non-weighted linear (version 4).

Table 2.9 gives the parameter estimates from the hedonic regressions. Regardless of the choice of model the most distinguishing positive attribute is the LFCF label followed by organic claim. Soy attribute is highly influential in the weighted regressions. Although in the non-weighted regressions this effect is reduced, it is still significant and negative. A vitamin mineral enhancement label has a positive effect on the price. If a product is discounted we expect to see an average reduction of 1-2 cents in price per serving (or 8-16 cents per half gallon package). Among the nutritional attributes, protein has the highest value followed by carbohydrate and fat content. Carbohydrate has almost twice as high hedonic prices in non-weighted regressions than fat whereas the difference diminishes in size weighted regressions. Cholesterol and sodium content have significantly negative effect on the product values. In all models vitamin and mineral content is highly valued whereas an increasing size reduces the price per serving. This makes sense since smaller packages offer the convenience for storage and consumption whereas larger packages are usually preferred by larger households.

The estimated uncompensated price elasticities and expenditure elasticities along with summary of estimation results for hedonic metric approximation to RM Model can be found in **Table 2.10** whereas **Table 2.11** gives the results for the HM-LA/AIDS Model. Regardless of the hedonic regression version, in all HM models, the estimated expenditure elasticities match almost perfectly with the original models. Moreover, in HM approximated models the number of parameters to be estimated are reduced to only 13 parameters (1 hedonic continuous distance measure, 1 nearest neighbor measure based on hedonic distance, 3 own-price interactions, 4 constants, and 4 expenditure coefficients). The non-weighted hedonic regression models give the lowest AIC and BIC scores, however the parameter estimates differ significantly between HM-RM and HM-LA/AIDS Models.

In the HM-RM models, all own-price elasticities and expenditure elasticities are significant and their signs are as expected. Moreover, the coefficient on the inverse of the hedonic distance terms is negative indicating that products closer in hedonic space are close substitutes. The substitution effect declines as the distance between products in hedonic space decreases. The price interaction coefficient on the hedonic uniqueness term is also negative in all models other than non-weighted linear model. This implies more unique product types have higher own-price elasticities. This is true in the case of own-price elasticity for soymilk in the RM model since soymilk has the highest uncompensated own-price elasticity among all milk types in the RM model. In the linear hedonic based models, the own-price elasticity for soymilk is highest followed by that of 1% milk, full-fat milk, skim milk and 2% milk. In the semi-log hedonic based models, the own-price elasticity of full-fat milk is slightly higher than that of 1% milk yet the order of other own-price elasticities are not different than linear models. Similar to the DM-RM results, the HM-RM models have a tendency to underestimate the own-price elasticities for 2% milk, and skim milk and overestimate that of whole milk. However, in the size-weighted hedonic based HM-RM models, all of the estimated elasticities at the mean values are within the 95% confidence interval of original RM results. This indicates that HM based approximation to the RM model performs very well in estimating elasticities.

HM-LA/AIDS models also give significant results for all own-price elasticities and expenditure elasticities. The expenditure elasticities for all different versions match almost perfectly with the original model; only the own-price elasticity for soymilk is underestimated. However, calculated mean elasticities of all dairy based milk types fits into the 95% confidence interval of the original models. HM-LA/AIDS version 4 has the lowest AIC and BIC among all versions and it has the highest own-price elasticity of soymilk (-0.599), yet the own-price elasticity of soymilk does not fit in the 95% confidence interval. Although soymilk has the lowest own-price elasticity in all versions, the order of other own-price elasticities differs between different versions. In version 3 and version 4, skim milk has the highest own-price elasticity followed by that of whole milk, 2% milk, and 1% milk. In version 2, 2% milk has the highest own-price elasticity followed by that of skim milk, full fat milk and 1% milk. In version 1, full-fat milk has the highest own-price elasticity whereas 1% milk has the lowest.

Summary

In this paper we compared the DM method with the HM method regarding their performance in approximating the elasticities estimated by the RM and LA/AIDS models. In our estimation we used AC Nielsen Homescan Data, which record household level purchases, and USDA-ERS Nutrient Database, which provides detailed nutritional facts about individual products. Combining these sources and aggregating consumption to a weekly basis gives us time-series quantity-price and market share data that are used to estimate demand for different milk types. In both RM and LA/AIDS models, the estimated elasticities other than the own-price elasticity for soymilk almost match perfectly with each other. Soymilk has the highest own-price elasticity in the RM model yet it has the lowest price elasticity in the LA/AIDS model. The uncompensated cross-price elasticities indicate that the soymilk prices do not affect the market shares of dairy based milk types, yet the inverse might not be true.

The DM approximation based on all possible combinations of distance measures give non-conforming estimates for some cross-price elasticities which might be due to high correlation between some measures. Therefore the models are approximated using each candidate distance measure to determine the most distinguishing milk characteristics. After rigorously considering all possible distance combinations three different distance measure combinations are used: F/OS/NNFOS, F/OS/NNFO, and F/O/NNFO. The results based on these distances are very similar to each other. However, the order of estimated own-price elasticities based on DM approximation is different than the original models. The significant inverse distance coefficient is positive suggesting that closer product types are closer substitutes.

The HM approximations are based on hedonic distances calculated as the sum of the pair wise differences in the value added by each attribute for each product type. Therefore, it eliminates the need to search for significant characteristics and has a stronger foundation than the DM method. The calculated elasticities are very similar to the original ones. In fact all mean elasticities other than the own-price elasticity for soymilk in DM-LA/AIDS model fit into the 95% confidence interval of original estimation results. The coefficient on the hedonic uniqueness parameter is negative in the DM-RM models suggesting unique products have higher own-price elasticity. The same coefficient in DM-LA/AIDS model is positive suggesting uniqueness reduces the own-price elasticity. Although that might be seen as a contradiction, it simply reflects the superiority of HM approximation in capturing differences in calculated own-price elasticities in the different models.

References

Anderson, Simon P, and Andre de Palma. "Multiproduct Firms: A Nested Logit Approach." *The Journal of Industrial Economics* 40, no. 3 (September 1992b): 261-276.

Anderson, Simon P, and Andre De Palma. "The Logit as a Model of Product Differentiation." *Oxford Economic Papers* (Oxford University Press) 44 (January 1992a): 51-67.

Asche, Frank, and Cathy R. Wessels. "On Price Indices in the Almost Ideal Demand System." *American Journal of Agricultural Economics* 79, no. 4 (1997): 1182-1185.

Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." *Econometrica* 63, no. 4 (July 1995): 841-890.

Binkley, James, Abbott Sharon, Christine Wilson, and Kevin McNamara. "Determinants of Functional Food Consumption." *A workshop on the Use of Scanner Data in Policy Analysis.* Washington: USDA - ERS, 2000.

Chan, Tat Y. "Estimating a Continuous Hedonic Choice Model with an Application to Demand for Soft Drinks." *RAND Journal of Economics* 37, no. 2 (2006): 466-482.

Cotterill, Ronald W., and Tirtha Dhar. *Oligopoly Pricing with Differentiated Products: The Boston Fluid Milk Market Channel.* Connecticut: Food Marketing Policy Center, 2003.

Cropper, M, L Deck, and K McConnell. "On the choice of functional form for hedonic functions." *Review of Economics and Statistics*, no. 70 (1988): 668-675.

Gateway to Government Food Safety Information. June 1999. http://www.foodsafety.gov/~dms/flg-7a.html (accessed June 15, 2008).

Harris, Mike. "Properties of Scanner Data." *A Workshop on the Use of Scanner Data in Policy Analysis.* Washington: USDA - ERS, 2003.

Hendel, Igal. "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns." *Review of Economic Studies* 66, no. 2 (1999): 423-446.

Jensen, Helen H, and Huffman K. Sonya. "Demand for Enhanced Foods and the Value of Nutrition Enhancements of Food: The Case of Margarines." *A workshop on the Use of Scanner Data in Policy Analysis.* Washinton: USDA - ERS, 2003.

Ladd, George W., and Veraphol Suvannunt. "A model of Consumer Goods Characteristics." *American Journal of Agricultural Economics* (American Agricultural Economics Association) 58 (August 1976): 504-510.

Lohr, Luanne. "Factors Affecting International Demand and Trade in Organic Food Products." *University of Georgia Faculty Series*, February 2001.

McFadden, Daniel. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Residential Location*, by et.al. Karlgvist, 75-96. Amsterdam: North Holland, 1978.

Moschini, Giancarlo. "Units of Measurement and the Stone Index in Demand System Estimation." *American Journal of Agricultural Economics* (Wiley-Blackwell) 77, no. 1 (February 1995): 63-68.

Nevo, Aviv. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica* 69, no. 2 (2001): 307-342.

Nevo, Aviv. "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry." *RAND Journal of Economics* 31, no. 3 (2000): 395-421.

Nii, Jenifer K. *Milk prices set to rise* . April 28, 2004 . http://deseretnews.com/dn/view/0,1249,595059211,00.html (accessed June 20, 2008).

Nimon, Wesley, and John Beghin. "Are Eco-Label Valauble? Evidence from the Apparel Industry." *American Journal of Agricultural Economics* (American Agricultural Economics Association) 81, no. 4 (November 1999): 801-811.

Palmquist, Raymond B. "Property Value Models." In *Handbook of Environmental Economics*, edited by Goren K Maler and Vincent Jeffrey. Elsevier, 2003.

Pinske, J, M. Slade, and C. Brett. "Spatial Price Competition:A Semiparametric Approach." *Econometrica*, 2002: 1111-1155.

Rojas, Christian, and Peterson B. Everett. "Demand for Differentiated Products: Price and Advertising Evidence from the U.S. Beer Market." *International Journal of Industrial Organization*, 2008: 288-307.

Rosen, R. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, no. 82 (1974): 34-55.

Silver, Mick. "Scanner Data and Price Indices." *A workshop on the Use of Scanner Data in Policy Analysis.* Washington: USDA - ERS, 2003.

Wohlgenant, Micheal, and George C. Davis. "Demand Elasticities from a Discrete Choice Model: The Natural Christmas Tree Market." *American Journal of Agricultural Economics*, no. 75 (1993): 730-738.

Table 2.1 Core Data Household Demographics vs. Entire Data Household Demographics

CORE DATA DEMOGRAPHICS

Household Income

HHIncome	Percent	Cumulative
3	0.51	0.51
4	0.45	0.96
6	0.45	1.4
8	1.4	2.81
10	2.26	5.07
11	4.56	9.63
13	7.43	17.06
15	6.38	23.44
16	7.46	30.9
17	7.46	38.36
18	7.24	45.6
19	7.14	52.74
21	11.42	64.16
23	9.76	73.92
26	15.91	89.83
27	10.17	100

Household Size

HHSize	Percent	Cumulative
1	20.98	20.98
2	42.6	63.58
3	15.78	79.37
4	12.79	92.16
5	5.23	97.39
6	1.63	99.01
7	0.61	99.62
8	0.19	99.81
9	0.19	100

Age and Presence of Children

Children	Percent	Cumulative
1	2.46	2.46
2	7.02	9.47
3	7.14	16.61
4	3.22	19.83
5	0.41	20.25
6	4.5	24.74
7	0.99	25.73
No Children	74.27	100

Household Race

HHRace	Percent	Cumulative
White	82.53	82.53
Black	8	90.53
Asian	2.2	92.73
Others	1.21	93.94
Hispanic	6.06	100

ENTIRE DATA DEMOGRAPHICS

Household Income

HHIncome	Percent	Cumulative
3	0.81	0.81
4	1.32	2.12
6	1.14	3.27
8	1.62	4.89
10	3.05	7.94
11	5.44	13.38
13	7.73	21.11
15	6.89	28
16	7.85	35.84
17	6.97	42.82
18	6.83	49.64
19	6.45	56.09
21	11.05	67.14
23	8.78	75.92
26	14.87	90.79
27	9.21	100

Household Size

HHSize	Percent	Cumulative
1	26.1	26.1
2	37.4	63.5
3	14.91	78.41
4	13.13	91.55
5	5.55	97.1
6	1.92	99.02
7	0.6	99.62
8	0.25	99.87
9	0.13	100

Age and Presence of Children

Children	Percent	Cumulative
1	3.71	3.71
2	6.44	10.15
3	7.95	18.11
4	3.33	21.44
5	0.57	22.01
6	4.3	26.31
7	0.87	27.17
No Children	72.83	100

Household Race

HHRace	Percent	Cumulative
White	78.83	78.83
Black	10.07	88.9
Asian	2.13	91.03
Others	1.82	92.85
Hispanic	7.15	100

Table 2.2 Food Labeling Guide

Food Labeling CFR References

Reference Values for Nutrition Labeling

(Based on a 2000 Calorie Intake; for Adults and Children 4 or More Years of Age)

NUTRIENT	UNIT OF MEASURE	DAILY VALUES
Total Fat	grams (g)	65
Saturated fatty acids	grams (g)	20
Cholesterol	milligrams (mg)	300
Sodium	milligrams (mg)	2400
Potassium	milligrams (mg)	3500
Total carbohydrate	grams (g)	300
Fiber	grams (g)	25
Protein	grams (g)	50
Vitamin A	International Unit (IU)	5000
Vitamin C	milligrams (mg)	60
Calcium	milligrams (mg)	1000
Iron	milligrams (mg)	18
Vitamin D	International Unit (IU)	400
Vitamin E	International Unit (IU)	30
Vitamin K	micrograms (µg)	80
Thiamin	milligrams (mg)	1.5
Riboflavin	milligrams (mg)	1.7
Niacin	milligrams (mg)	20
Vitamin B_6	milligrams (mg)	2
Folate	micrograms (µg)	400
Vitamin B_{12}	micrograms (µg)	6
Biotin	micrograms (µg)	300
Pantothenic acid	milligrams (mg)	10
Phosphorus	milligrams (mg)	1000
Iodine	micrograms (µg)	150
Magnesium	milligrams (mg)	400
Zinc	milligrams (mg)	15
Selenium	micrograms (µg)	70
Copper	milligrams (mg)	2
Manganese	milligrams (mg)	2
Chromium	micrograms (µg)	120
Molybdenum	micrograms (µg)	75
Chloride	milligrams (mg)	3400

Nutrients in this table are listed in the order in which they are required to appear on a label in accordance with 101.9© This list includes only those nutrients for which a Daily Reference Value (DRV) has been established in 101.9(c)(9) or a Reference Daily Intake (RDI) in in 101.9(c)(8)(iv).

Source: http://www.foodsafety.gov/~dms/flg-7a.html REV. Jan 30, 1998

Table 2.3 Mean Attribute Values by Milk Type

Attribute	All	Milk Types 0.00%	1.00%	2.00%	3.25%	SOY
Purchase Frequency	525323	147268	92825	172825	97126	15279
Percentage	100.00%	28.03%	17.67%	32.90%	18.49%	2.91%
Organic Claim (%)	2.44%	0.58%	1.01%	0.48%	0.61%	63.81%
Promotion	8.67%	7.90%	9.29%	9.76%	6.27%	14.16%
CFLF	5.04%	3.87%	1.19%	2.21%	0.94%	98.99%
Vitamin-Mineral Label	96.79%	98.69%	96.88%	97.53%	97.18%	68.99%
Protein per Serving	8.40	8.81	8.41	8.54	8.04	5.07
Carbohydrate per Serving	13.09	12.69	13.54	13.15	12.23	18.88
Fat per Serving	3.68	0.53	2.26	4.77	8.15	2.41
Cholesterol DRI*	4.89	1.68	4.01	6.55	8.44	0.10
Sodium DRI*	4.97	5.29	4.75	5.24	4.30	4.40
Vitamin-Mineral DRI*	11.50	12.02	11.35	11.92	10.54	8.71
Servings per Package	14.71	14.93	14.69	15.63	13.66	8.40

Table 2.4 Price, Quantity and Market Share Summary Statistics by Milk Type

Variable	Mean	Std Dev	Minimum	Maximum
Quantities				
2.00%	12925.71	725.88	11305.65	14727.50
0.00%	10524.45	623.34	8949.75	12707.13
3.25%	6345.95	611.65	4973.75	7567.25
1.00%	6526.62	409.16	5348.00	7639.25
SOY	615.31	142.79	282.90	1018.50
Prices*				
2.00%	17.82	1.33	15.82	21.49
0.00%	17.45	1.14	15.72	20.97
3.25%	19.35	1.59	16.82	22.97
1.00%	18.33	1.26	16.09	21.76
SOY	34.85	2.02	28.56	41.70
Market Shares				
2.00%	34.00%	0.96%	31.42%	36.15%
0.00%	27.13%	0.88%	24.34%	30.08%
3.25%	18.07%	1.02%	15.35%	20.73%
1.00%	17.66%	0.76%	15.71%	19.75%
SOY	3.14%	0.58%	1.61%	4.81%

*Prices are measured in terms of cents per serving. A gallon of milk has 16 servings in a single package

Table 2.5 Rotterdam and LA/AIDS Model Mean Elasticity Estimates

RM MODEL

Hicksian (Compansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.47811*	0.1501	0.270839*	0.1109	0.119638	0.096	0.170366*	0.0845	-0.00518	0.0525
SKIM	0.339535*	0.139	-0.57498*	0.1665	0.09129	0.1128	0.087705	0.0867	0.075706	0.0703
FULLFAT	0.225038	0.1806	0.136974	0.1693	-0.52492*	0.1598	0.074177	0.1375	0.152301	0.0842
PERCENT1	0.328035*	0.1627	0.134706	0.1332	0.075931	0.1407	-0.63591*	0.1727	-0.08251	0.1064
SOY	-0.05627	0.57	0.65582	0.6087	0.879308	0.4864	-0.46536	0.5998	-1.0135*	0.261

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.81794*	0.1485	-0.00023	0.1125	-0.06103	0.0977	-0.00612	0.0856	-0.03647	0.0527
SKIM	-0.04698	0.1408	-0.88329*	0.1703	-0.11419	0.1115	-0.11303	0.0871	0.040115	0.0705
FULLFAT	-0.05415	0.177	-0.08573	0.1729	-0.67334*	0.1607	-0.07082	0.1372	0.126593	0.0843
PERCENT1	0.006093	0.1603	-0.1221	0.138	-0.09522	0.1414	-0.80311*	0.1735	-0.11215	0.1061
SOY	-0.44975	0.5934	0.341951	0.6115	0.670121	0.4937	-0.66971	0.6024	-1.04973*	0.2594

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	0.999341*	0.0688	1.136638*	0.0858	0.82103*	0.0915	0.94675*	0.0923	1.157123*	0.3527

Nonlinear FIML Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	DW
WDLQ200	4.5	203.5	0.0311	0.00015	0.0124	0.6339	0.6276	2.8865
WDLQ000	4.5	203.5	0.034	0.00017	0.0129	0.5643	0.5568	3.1004
WDLQ325	4.5	203.5	0.0196	9.6E-05	0.00981	0.3604	0.3494	2.9245
WDLQ100	4.5	203.5	0.0231	0.00011	0.0107	0.3859	0.3753	2.996

LogL	2738.944
Observations	208
Parameters Estimated	22
AIC	-5433.888
BIC	-5426.8906

(*) Significant at the 95% level.

Table 2.5 Rotterdam and LA/AIDS Model Mean Elasticity Estimates (Cont.)

LA/AIDS MODEL

Hicksian (Compansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.47057*	0.1486	0.273262*	0.1111	0.118547	0.0957	0.173484*	0.0844	-0.00434	0.0524
SKIM	0.342573*	0.1392	-0.58569*	0.1662	0.095962	0.1129	0.090182	0.0867	0.075812	0.07
FULLFAT	0.222985	0.1801	0.143984	0.1694	-0.52289*	0.1604	0.0787	0.138	0.153777	0.0838
PERCENT1	0.334038*	0.1625	0.138511	0.1331	0.080561	0.1412	-0.6461*	0.1723	-0.08461	0.1064
SOY	-0.04717	0.5692	0.656734	0.6067	0.887827	0.4838	-0.4772	0.6002	-1.02019*	0.2495

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.8157*	0.1481	-0.00204	0.1129	-0.06494	0.0971	-0.00576	0.085	-0.03612	0.0526
SKIM	-0.05374	0.1409	-0.90182*	0.1704	-0.11473	0.1112	-0.11564	0.0874	0.039319	0.0703
FULLFAT	-0.05583	0.1768	-0.07842	0.1735	-0.67111*	0.1614	-0.0661	0.1371	0.128103	0.0838
PERCENT1	0.013446	0.1607	-0.11722	0.1374	-0.08988	0.1417	-0.8126*	0.1729	-0.11413	0.1061
SOY	-0.30791	0.5903	0.448751	0.6086	0.749211	0.4902	-0.61262	0.6008	-1.00138*	0.00779

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.014948*	0.0671	1.165451*	0.0865	0.819931*	0.0926	0.94278*	0.0919	0.766761*	0.3254

Nonlinear FIML Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	DW
DW200	5.5	202.5	0.0312	0.000154	0.0124	0.0124	-0.0096	2.8978
DW000	5.5	202.5	0.0339	0.000167	0.0129	0.0402	0.0188	3.0909
DW325	5.5	202.5	0.0195	0.000096	0.00981	0.0469	0.0257	2.9172
DW100	5.5	202.5	0.0231	0.000114	0.0107	0.0196	-0.0022	2.9906

LogL	2741.83
Obs	208
Parameters Estimated	22
AIC	-5439.7
BIC	-5432.7

(*) Significant at the 95% level.

Table 2.6 DM RM Distance Parameter Estimates

One Dimensional			
Distance	Estimate	S.E.	LogL
L1 - W	0.02308*	0.00388	2729
L2 - F	0.014967*	0.00295	2726
L3 - Org	0.031222*	0.00451	2735
L4 - Size	0.019692*	0.00322	2730
Two Dimensional			
Distance	Estimate	S.E.	LogL
L5 - FO	0.033606*	0.00554	2730
L6 - FS	0.023599*	0.00373	2731
L7 - OS	0.033336*	0.00486	2735
Three Dimensional			
Distance	Estimate	S.E.	LogL
L8 - FOS	0.035592*	0.00578	2731

(*) Significant at the 95% level

Table 2.7 DM RM Distance Parameter Estimates and Elasticities

RM MODEL (v1 F/OS/NNFOS)

Parameter	Estimate	S.E
L2	0.000868	0.00868
L7	0.050755*	0.0121
N3	-0.03432**	0.0177
B0	-0.29452*	0.1437
B1	0.422161	0.2952
B2	0.006547	0.0051
B3	8.763477	9.165
LogL	2736.343	
Parameters Estimated	15	
n	208	
AIC	-5442.686	
BIC	-5437.915	

RM MODEL (v2 F/OS/NNFO)

Parameter	Estimate	S.E
L2	-0.00476	0.00824
L7	0.048292*	0.0118
N1	-0.01924	0.0165
B0	-0.28926*	0.1433
B1	0.403694	0.2966
B2	0.006634	0.00511
B3	8.591353	9.1314
LogL	2734.59	
Parameters Estimated	15	
n	208	
AIC	-5439.18	
BIC	-5434.40905	

RM MODEL (v3 F/O/NNFOS)

Parameter	Estimate	S.E
L2	-0.00262	0.00799
L3	0.041876*	0.0103
N1	-0.01939	0.0165
B0	-0.28387*	0.143
B1	0.39592	0.2963
B2	0.005494	0.00512
B3	8.864848	9.1188
LogL	2734.468	
Parameters Estimated	15	
n	208	
AIC	-5438.936	
BIC	-5434.165	

RM MODEL (Full Model)

Parameter	Estimate	S.E
L1	-0.07458	0.1188
L2	-0.07927	0.2599
L3	1.18717	1.3434
L4	-4.23229	5.7564
L5	7.106881	15.3052
L6	5.997615	7.8858
L7	-2.69147	3.1506
L8	-7.09475	15.4787
N1	-0.01427	0.0358
N2	-0.03428	0.0553
N3	-0.0382	0.0268
B0	-0.24206	0.4567
B1	0.104616	0.6489
B2	0.009977	0.0118
B3	7.631876	33.9058
LogL	2740.689	
Parameters Estimated	23	
n	208	
AIC	-5435.378	
BIC	-5428.062543	

(*) Significant at the 95% level.
(**) Significant at the 90% level.

Table 2.7 DM RM Distance Parameter Estimates and Elasticities (Cont.)

RM MODEL (v1 F/OS/NNFOS)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.56713✓*	0.09	-0.13375✓*	0.04	-0.06065✓*	0.03	-0.14331✓*	0.04	0.01764✓	0.02
SKIM	-0.21✓*	0.04	-0.7719✓*	0.11	-0.04218✓	0.04	-0.1428✓*	0.05	0.02565✓	0.02
FULLFAT	-0.25✓*	0.08	0.01462✓	0.05	-0.77109✓*	0.13	0.10112✓	0.05	0.06691✓*	0.02
PERCENT1	-0.06✓	0.07	-0.1649✓*	0.07	0.087856✓	0.06	-0.8204✓*	0.12	0.06614✓*	0.03
SOY	0.14✓	0.26	0.21379✓	0.18	0.328789✓*	0.16	-0.76379✓	0.41	-1.07712✓*	0.25

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.00✓*	0.07	1.12955✓*	0.08	0.842333✓*	0.09	0.92892✓*	0.09	1.16047*	0.35

RM MODEL (v2 F/OS/NNFO)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.57172✓*	0.09	-0.1487✓*	0.03	-0.07594✓*	0.03	-0.11592✓*	0.04	0.00075✓	0.02
SKIM	-0.22927✓*	0.04	-0.7746✓*	0.11	-0.05729✓	0.03	-0.11017✓*	0.05	0.01213✓	0.01
FULLFAT	-0.19211✓*	0.08	-0.0062✓	0.05	-0.76109✓*	0.13	0.07726✓	0.05	0.04652✓*	0.02
PERCENT1	-0.08971✓	0.06	-0.1154✓	0.07	0.06176✓	0.05	-0.81815✓*	0.12	0.04006✓	0.03
SOY	-0.65982✓	0.41	0.09657✓	0.15	0.209719✓	0.13	0.1854✓	0.16	-0.99337✓*	0.25

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.004027✓*	0.07	1.13006✓*	0.08	0.835939✓*	0.09	0.9318✓*	0.09	1.1615✓*	0.35

(*) Significant at the 95% level.
(✓) The estimated parameter is within the 95% confidence interval of original model.

Table 2.7 DM RM Distance Parameter Estimates and Elasticities (Cont.)

RM MODEL (v3 F/O/NNFO)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.57637✓*	0.09	-0.1533✓*	0.03	-0.06319✓*	0.03	-0.11798✓*	0.04	0.00302✓	0.02
SKIM	-0.23471✓*	0.04	-0.7586✓*	0.11	-0.05327✓	0.04	-0.12513✓*	0.05	0.01146✓	0.01
FULLFAT	-0.16925✓*	0.08	-0.0007✓	0.05	-0.78134✓*	0.13	0.07492✓	0.05	0.04395✓*	0.02
PERCENT1	-0.09289✓	0.06	-0.1387✓	0.07	0.059499✓	0.05	-0.79555✓*	0.12	0.04065✓	0.03
SOY	-0.63952✓	0.41	0.0908✓	0.15	0.195251✓	0.13	0.18895✓	0.16	-0.996✓*	0.25

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.004125✓*	0.07	1.1292✓*	0.08	0.836959✓*	0.09	0.93206✓*	0.09	1.16052✓*	0.35

RM MODEL (v4 Full)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.67878✓*	0.13	-0.1602✓	0.11	0.062871✓	0.11	-0.1348✓	0.12	0.00798✓	0.12
SKIM	-0.11222✓	0.14	-0.8869✓*	0.18	-0.05254✓	0.10	-0.08072✓	0.08	0.04163✓	0.07
FULLFAT	-0.12051✓	0.22	-0.0076✓	0.16	-0.65109✓*	0.17	-0.24393✓	0.34	0.31728X	0.31
PERCENT1	0.074593✓	0.18	-0.0643✓	0.13	-0.05958✓	0.16	-0.81577✓	0.48	0.10029X	0.11
SOY	-0.41811✓	0.61	0.35163✓	0.61	0.683634✓	0.49	-0.69889✓	0.61	-1.07435✓*	0.27

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.011953✓*	0.07	1.12306✓*	0.09	0.860618 ✓*	0.09	0.90298 ✓*	0.09	1.15609 ✓*	0.35

(*) Significant at the 95% level.
(✓) The estimated parameter is within the 95% confidence interval of original model.

Table 2.8 DM LA/AIDS Distance Parameter Estimates and Elasticities

LA/AIDS MODEL (v1 F/OS/NNFOS)

Parameter	Estimate	S.E
L2*	0.023669	0.00871
L7	-0.02005	0.0122
N3**	-0.03319	0.0176
B0	-0.11237	0.1423
B1	0.444024	0.2936
B2	0.004629	0.00504
B3	6.101079	9.0321
LogL	2740.78	
Parameters Estimated	15	
n	208	
AIC	-5451.56	
BIC	-5446.789	

LA/AIDS MODEL (v2 F/OS/NNFO)

Parameter	Estimate	S.E
L2*	0.02147	0.00825
L7*	-0.02123	0.0119
N1*	-0.02868	0.0162
B0	-0.10661	0.1415
B1	0.429713	0.2943
B2	0.004522	0.00504
B3	6.087335	8.9711
Log L	2740.051	
Parameters Estimated	15	
n	208	
AIC	-5450.102	
BIC	-5445.331	

LA/AIDS MODEL (v3 F/O/NNFOS)

Parameter	Estimate	S.E
L2*	0.020755	0.00797
L3**	-0.01889	0.0104
N1**	-0.02845	0.0162
B0	-0.10811	0.1416
B1	0.433966	0.2946
B2	0.00503	0.00504
B3	5.933291	8.9804
LogL	2740.125	
Parameters Estimated	15	
n	208	
AIC	-5450.25	
BIC	-5445.479	

LA/AIDS MODEL (Full Model)

Parameter	Estimate	S.E
L1	0.021772	0.1169
L2	-0.03676	0.2574
L3	0.738448	0.8256
L4	-2.16745	3.8772
L5	0.810522	13.7101
L6	3.214653	5.1839
L7	-1.80957	1.8817
L8	-0.66407	13.8467
N1	0.003599	0.0358
N2	-0.0357	0.0317
N3*	-0.03906	0.0225
B0	-0.05725	0.2608
B1	0.297472	0.5041
B2	0.007605	0.01
B3	1.070981	17.1152
LogL	2744.229	
Parameters Estimated	23	
n	208	
AIC	-5442.458	
BIC	-5435.1425	

(*) Significant at the 95% level.
(**) Significant at the 90% level.

Table 2.8 DM LA/AIDS Distance Parameter Estimates and Elasticities (Cont.)

LA/AIDS MODEL (v1 F/OS/NNFOS)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.73872 ✓ *	0.093	-0.02473 ✓	0.037	-0.01077 ✓	0.03	-0.10964 ✓ *	0.038	0.0424 ✓ *	0.02
SKIM	-0.0819 ✓	0.044	-0.87671 ✓ *	0.107	-0.06291 ✓	0.036	-0.16364 ✓ *	0.048	0.0161 ✓	0.02
FULLFAT	-0.14438 ✓	0.078	-0.00638 ✓	0.054	-0.73583 ✓ *	0.129	-0.01512 ✓	0.053	0.0372 ✓	0.02
PERCENT1	0.006821 ✓	0.066	-0.1868 ✓ *	0.068	-0.03113 ✓	0.059	-0.76803 ✓ *	0.121	0.0566 ✓	0.03
SOY	0.54612 ✓ *	0.254	0.24834 ✓	0.171	0.2291 ✓	0.158	-0.71178 ✓	0.41	-1.0735 X*	0.24

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.014519 ✓ *	0.066	1.164164 ✓ *	0.084	0.83971 ✓ *	0.089	0.926241 ✓ *	0.091	0.7617 ✓ *	0.32

LA/AIDS MODEL (v2 F/OS/NNFO)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.73847 ✓ *	0.093	-0.03152 ✓	0.034	-0.01749 ✓	0.026	-0.10382 ✓ *	0.039	0.0355 ✓ *	0.02
SKIM	-0.08974 ✓ *	0.043	-0.87026 ✓ *	0.105	-0.06945 ✓ *	0.035	-0.15673 ✓ *	0.047	0.0105 ✓	0.01
FULLFAT	-0.13083 ✓	0.078	-0.01571 ✓	0.051	-0.72322 ✓ *	0.128	-0.02562 ✓	0.049	0.0289 ✓	0.02
PERCENT1	-0.00627 ✓	0.057	-0.17575 ✓ *	0.074	-0.0419 ✓	0.052	-0.7517 ✓ *	0.124	0.0461 ✓	0.02
SOY	-0.44405 ✓	0.4	0.200045 ✓	0.143	0.18013 ✓	0.128	0.288305 ✓	0.152	-0.9875 ✓ *	0.24

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.016366 ✓ *	0.065	1.164073 ✓ *	0.083	0.8378 ✓	0.089	0.924533 ✓	0.089	0.7631 ✓	0.32

(*) Significant at the 95% level.
(✓) The estimated parameter is within the 95% confidence interval of original model.

Table 2.8 DM LA/AIDS Distance Parameter Estimates and Elasticities (Cont.)

LA/AIDS MODEL (v3 F/O/NNFO)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.73344 ✓*	0.093	-0.03061 ✓	0.033	-0.02415 ✓	0.029	-0.10341 ✓*	0.039	0.0346 ✓*	0.02
SKIM	-0.0887 ✓*	0.042	-0.87394 ✓*	0.103	-0.0727 ✓*	0.036	-0.1507 ✓*	0.046	0.0107 ✓	0.01
FULLFAT	-0.14188 ✓	0.079	-0.02036 ✓	0.052	-0.70954 ✓*	0.132	-0.0267 ✓	0.048	0.0297 ✓	0.02
PERCENT1	-0.0068 ✓	0.056	-0.1664 ✓*	0.073	-0.0431 ✓	0.052	-0.75773 ✓*	0.12	0.0457 ✓	0.02
SOY	-0.44667 ✓	0.397	0.200827 ✓	0.142	0.18487 ✓	0.127	0.286221 ✓	0.152	-0.9896 ✓*	0.24

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.01635 ✓*	0.065	1.164337 ✓*	0.083	0.83728 ✓*	0.089	0.924465 ✓*	0.089	0.7643 ✓*	0.32

LA/AIDS MODEL (v4 Full)

Marshallian (Uncompansated) Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
PERCENT2	-0.75432 ✓*	0.13	-0.08835 ✓	0.105	0.02379 ✓	0.102	-0.08733 ✓	0.11	-0.044 ✓	0.12
SKIM	-0.02272 ✓	0.09	-0.91842 ✓*	0.17	-0.08864 ✓	0.102	-0.12576 ✓	0.081	0.0368 ✓	0.07
FULLFAT	-0.0936 ✓	0.215	-0.05215 ✓	0.16	-0.61529 ✓*	0.163	-0.2082 ✓	0.173	0.329 X	0.18
PERCENT1	0.073386 ✓	0.176	-0.12588 ✓	0.125	-0.02005 ✓	0.152	-0.84451 ✓*	0.178	0.1063 X	0.08
SOY	-0.27494 ✓	0.595	0.423628 ✓	0.605	0.77548 ✓	0.489	-0.62323 ✓	0.606	-1.067 X*	0.25

Expenditure Elasticities

	PERCENT2		SKIM		FULLFAT		PERCENT1		SOY	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Estimate	1.025268 ✓*	0.067	1.153367 ✓*	0.087	0.85533 ✓*	0.093	0.905369 ✓*	0.094	0.766 ✓*	0.32

(*) Significant at the 95% level

(✓) The estimated parameter is within the 95% confidence interval of original model

Table 2.9 Hedonic Attribute Estimates

Parameter Estimates			NonWeighted Linear		NonWeighted Semi-Log		Weighted Linear		Weighted Semi-Log	
Variable	*Label*	*DF*	*Parameter Estimate*	*Std. Error*	*Parameter Estimate*	*Std. Error*	*Parameter Estimate*	*Std. Error*	*Parameter Estimate*	*Std. Error*
Intercept	Intercept	1	-33.032	0.382	1.777	0.015	-20.627	0.367	1.667	0.018
MOrgClaim	Marketing organic Claim	1	8.379	0.095	0.346	0.004	10.962	0.090	0.428	0.004
MSoy	Marketing Soy Dummy	1	-1.458	0.166	-0.261	0.007	-9.351	0.155	-0.367	0.008
MPromo	Marketing Promotion Dummy	1	-2.094	0.038	-0.117	0.002	-1.583	0.026	-0.100	0.001
MCFLF	Marketing Lactose Cholesterol Free	1	21.149	0.074	0.740	0.003	23.537	0.069	0.857	0.003
MVitminIndex	Marketing Vitamin Mineral Index	1	7.122	0.080	0.177	0.003	4.497	0.077	0.140	0.004
NProtein	Nutrient Protein Content (g)	1	4.046	0.043	0.088	0.002	2.734	0.042	0.085	0.002
NCarboHydrt	Nutrient Carb Content (g)	1	1.421	0.007	0.040	0.000	0.991	0.007	0.033	0.000
NLipid_Tot	Nutrient Lipid (Fat) Content (g)	1	0.609	0.018	0.020	0.001	0.861	0.014	0.032	0.001
NPerCholestrl	Nutrient Cholesterol DRI Max	1	0.057	0.018	0.002	0.001	-0.358	0.014	-0.011	0.001
NPerSodium	Nutrient Sodium DRI Max	1	-2.982	0.033	-0.085	0.001	-2.010	0.027	-0.060	0.001
NPerVitMinIndex	Nutrient Vitamin-Mineral Percentage Index	1	0.919	0.016	0.025	0.001	0.789	0.016	0.020	0.001
SQuantity	Purchased Serving Size Quantity	1	-0.315	0.001	-0.014	0.000	-0.106	0.001	-0.005	0.000
Adj R-Sq	Adjusted R-Square Value		0.4146		0.3933		0.3666		0.2872	

All estimates are significant at 99% level.

Table 2.10 HM RM Hedonic Parameter Estimates and Elasticities

RM MODEL (v1 Weighted Semi-Log)

Parameter	Estimate	S.E
L1*	0.045299	0.00934
N1**	-0.02808	0.0143
B0	0.214714	0.8369
B1	0.121864	0.1766
B2	-1.01493	2.4436
LogL	2733.933	
Parameters Estimated	13	
n	208	
AIC	-5441.866	
BIC	-5437.7312	

RM MODEL (v2 Weighted Linear)

Parameter	Estimate	S.E
L1*	0.04888	0.0108
N1**	-0.02903	0.014
B0	-0.12377	0.3359
B1	0.167975	0.1691
B2	-0.03055	0.9529
Log L	2733.662	
Parameters Estimated	13	
n	208	
AIC	-5441.324	
BIC	-5437.1892	

RM MODEL (v3 Semi-Log)

Parameter	Estimate	S.E
L1*	0.048656	0.00931
N1*	-0.03441	0.0144
B0	0.634955	1.7322
B1	0.181281	0.1864
B2	-2.31899	5.0862
LogL	2735.395	
Parameters Estimated	13	
n	208	
AIC	-5444.79	
BIC	-5440.6552	

RM MODEL (Linear)

Parameter	Estimate	S.E
L1*	0.045983	0.00975
N1*	-0.02484	0.0128
B0	-0.21546	0.3766
B1	0.168077	0.1692
B2	0.237685	1.0722
LogL	2733.419	
Parameters Estimated	13	
n	208	
AIC	-5440.838	
BIC	-5436.7032	

(*) Significant at the 95% level.
(**) Significant at the 90% level.

Table 2.10 HM RM Hedonic Parameter Estimates and Elasticities (Cont.)

RM MODEL (v1 Weighted Semi-Log)

Marshallian (Uncompansated) Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.58219	✓*	0.073	-0.1543	✓*	0.032	-0.06221	✓*	0.029	-0.13968	✓*	0.026	0.014624	✓	0.01
SKIM	-0.23758	✓*	0.041	-0.66177	✓*	0.069	-0.06858	✓*	0.031	-0.15166	✓*	0.033	0.021177	✓	0.012
FULLFAT	-0.21598	✓*	0.055	-0.02272	✓	0.045	-0.76388	✓*	0.102	0.067422	✓	0.045	0.057354	✓*	0.017
PERCENT1	-0.08507	✓	0.057	-0.17789	✓*	0.052	0.05226	✓	0.049	-0.74816	✓*	0.099	0.057505	✓*	0.018
SOY	-0.80413	✓*	0.397	0.165978	✓	0.138	0.26605	✓*	0.118	0.277125	✓*	0.118	-1.10084	✓*	0.223

Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	1.00159	✓*	0.066	1.131397	✓*	0.084	0.83581	✓*	0.089	0.928491	✓*	0.09	1.195812	✓*	0.349

RM MODEL (v2 Weighted Linear)

Marshallian (Uncompansated) Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.5627	✓*	0.071	-0.15103	✓*	0.035	-0.06739	✓*	0.03	-0.13774	✓*	0.022	0.017243	✓	0.011
SKIM	-0.24161	✓*	0.039	-0.63681	✓*	0.068	-0.08469	✓*	0.028	-0.15926	✓*	0.033	0.025429	✓	0.013
FULLFAT	-0.23445	✓*	0.054	-0.04318	✓	0.043	-0.72543	✓*	0.089	0.056071	✓	0.044	0.06399	✓*	0.02
PERCENT1	-0.24481	✓*	0.046	-0.02224	✓	0.06	0.04012	✓	0.051	-0.755	✓*	0.099	0.068716	✓*	0.022
SOY	0.11396	✓	0.168	0.203516	✓	0.153	0.30236	✓*	0.133	-0.58834	✓	0.349	-1.2373	✓*	0.21

Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	0.99028	✓*	0.068	1.143999	✓*	0.086	0.8347	✓*	0.09	0.930271	✓*	0.092	1.205815	✓*	0.346

(*) Significant at the 95% level.

(✓) The estimated parameter is within the 95% confidence interval of original model.

Table 2.10 HM RM Hedonic Parameter Estimates and Elasticities (Cont.)

RM MODEL (v3 Semi-Log)

Marshallian (Uncompansated) Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.56365	✓*	0.07	-0.24411	X*	0.029	-0.0533	✓	0.029	-0.05006	✓	0.028	0.018339	✓	0.01
SKIM	-0.34698	X*	0.041	-0.67887	✓*	0.088	-0.05248	✓	0.032	-0.0377	✓	0.034	0.025401	✓*	0.012
FULLFAT	-0.23633	✓*	0.055	-0.00277	✓	0.046	-0.80522	✓*	0.098	0.087546	✓	0.046	0.064075	✓*	0.017
PERCENT1	-0.08305	✓	0.055	-0.20944	✓*	0.054	0.06785	✓	0.048	-0.79527	✓*	0.106	0.06185	✓*	0.018
SOY	-0.96784	✓*	0.404	0.198443	✓	0.136	0.30491	✓*	0.117	0.30658	✓*	0.115	-1.03902	✓*	0.222

Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	0.99685	✓*	0.067	1.117287	✓*	0.084	0.83724	✓*	0.089	0.957616	✓*	0.09	1.196924	✓*	0.348

RM MODEL (v4 Linear)

Marshallian (Uncompansated) Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.56985	✓*	0.071	-0.1526	✓*	0.034	-0.07478	✓*	0.027	-0.13113	✓*	0.022	0.014126	✓	0.01
SKIM	-0.33404	X*	0.041	-0.63571	✓*	0.068	-0.08494	✓*	0.027	-0.05925	✓	0.031	0.020717	✓	0.012
FULLFAT	-0.08683	✓	0.046	-0.04371	✓	0.041	-0.71176	✓*	0.105	-0.08527	✓	0.044	0.060277	✓*	0.018
PERCENT1	-0.23021	✓*	0.046	-0.03243	✓	0.056	0.03662	✓	0.049	-0.76707	✓*	0.102	0.062058	✓*	0.02
SOY	0.07991	✓	0.161	0.161634	✓	0.144	0.28031	✓*	0.126	-0.49314	✓	0.326	-1.23748	✓*	0.21

Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	0.99261	✓*	0.068	1.14303	✓*	0.085	0.83433	✓*	0.09	0.927115	✓*	0.091	1.208766	✓*	0.348

(*) Significant at the 95% level.
(✓) The estimated parameter is within the 95% confidence interval of original model.

Table 2.11 HM LA/AIDS Hedonic Parameter Estimates and Elasticities

LA/AIDS MODEL (v1 Weighted Semi-Log)

Parameter	Estimate	S.E
L1**	-0.01543	0.00864
N1	0.004367	0.0036
B0	0.548367	0.9511
B1	0.222612	0.1743
B2	-1.66567	2.787
LogL	2736.999	
Parameters Estimated	13	
n	208	
AIC	-5447.998	
BIC	-5443.8632	

LA/AIDS MODEL (v2 Weighted Linear)

Parameter	Estimate	S.E
L1	0.000155	0.005
N1*	-0.02441	0.00824
B0	0.147344	0.3285
B1	-0.27978	0.2264
B2	-0.08365	0.9301
Log L	2740.587	
Parameters Estimated	13	
n	208	
AIC	-5455.174	
BIC	-5451.0392	

LA/AIDS MODEL (v3 Semi-Log)

Parameter	Estimate	S.E
L1**	-0.01825	0.00939
N1	0.005588	0.00383
B0	2.171825	2.0703
B1	0.134952	0.1789
B2	-6.43623	6.0902
LogL	2740.125	
Parameters Estimated	13	
n	208	
AIC	-5454.25	
BIC	-5450.1152	

LA/AIDS MODEL (v4 Linear)

Parameter	Estimate	S.E
L1	-0.0055	0.00479
N1	-0.00486	0.00653
B0	0.154432	0.5498
B1	0.141158	0.1779
B2	-0.47111	1.5405
LogL	2736.371	
Parameters Estimated	13	
n	208	
AIC	-5446.742	
BIC	-5442.6072	

(*) Significant at the 95% level.

(**) Significant at the 90% level.

Table 2.11 HM LA/AIDS Hedonic Parameter Estimates and Elasticities (Cont.)

LA/AIDS MODEL (v1 Weighted Semi-Log)

Marshallian (Uncompansated) Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.80146	✓*	0.097	-0.04307	✓	0.031	-0.04254	✓	0.0273	-0.02997	✓	0.02	-0.016	✓	0.009
SKIM	-0.10865	✓*	0.038	-0.88402	✓*	0.081	-0.07742	✓*	0.0289	-0.06596	✓*	0.023	-0.0247	✓*	0.011
FULLFAT	0.01007	✓	0.038	-0.01997	✓	0.043	-0.86597	✓*	0.0999	-0.041	✓	0.042	-0.0227	✓	0.016
PERCENT1	-0.05832	✓	0.053	-0.03847	✓	0.038	-0.0642	✓	0.045	-0.81587	✓*	0.13	-0.0276	✓	0.017
SOY	0.04607	✓	0.13	-0.10582	✓	0.129	-0.12346	✓	0.1093	-0.12648	✓	0.11	-0.4648	X	0.281

Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	1.0114	✓*	0.066	1.172133	✓*	0.083	0.8175	✓*	0.0901	0.940475	✓*	0.088	0.77447	✓*	0.326

LA/AIDS MODEL (v2 Weighted Linear)

Marshallian (Uncompansated) Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.92569	✓*	0.069	0.002672	✓	0.022	0.00189	✓	0.0167	-0.06989	✓*	0.026	0.00042	✓	0.005
SKIM	-0.05618	✓	0.032	-0.88767	✓*	0.06	-0.02973	✓	0.0194	-0.11893	✓*	0.031	-0.005	✓	0.007
FULLFAT	-0.07828	✓	0.051	0.045302	✓	0.029	-0.59896	✓*	0.1476	0.029762	✓	0.025	0.00545	✓	0.009
PERCENT1	-0.12265	✓*	0.055	0.012539	✓	0.034	0.00853	✓	0.027	-0.59567	✓*	0.146	0.00167	✓	0.01
SOY	0.05395	✓	0.128	0.043347	✓	0.108	0.02942	✓	0.0826	-0.75069	✓*	0.255	-0.2224	X	0.269

Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	0.99152	✓*	0.065	1.166589	✓*	0.083	0.83513	✓*	0.0884	0.956464	✓*	0.09	0.84638	✓*	0.329

(*) Significant at the 95% level.
(✓) The estimated parameter is within the 95% confidence interval of original model.

Table 2.11 HM LA/AIDS Hedonic Parameter Estimates and Elasticities (Cont.)

LA/AIDS MODEL (v3 Semi-Log)

Marshallian (Uncompansated) Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.79147	✓*	0.093	-0.03463	✓	0.026	-0.04976	✓	0.029	-0.04937	✓	0.029	-0.019	✓	0.01
SKIM	-0.09683	✓*	0.033	-0.90909	✓*	0.073	-0.08663	✓*	0.0314	-0.08972	✓*	0.033	-0.0279	✓*	0.012
FULLFAT	0.00459	✓	0.039	-0.03371	✓	0.046	-0.88392	✓*	0.0951	-0.05546	✓	0.046	-0.028	✓	0.017
PERCENT1	-0.07206	✓	0.056	-0.04521	✓	0.039	-0.08031	✓	0.0496	-0.77723	✓*	0.13	-0.0327	✓	0.018
SOY	0.04747	✓	0.129	-0.13967	✓	0.137	-0.15784	✓	0.1184	-0.15746	✓	0.117	-0.3842	X	0.303

Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	1.01201	✓*	0.066	1.169079	✓*	0.083	0.81415	✓*	0.0907	0.944367	✓*	0.089	0.79171	✓*	0.326

LA/AIDS MODEL (v4 Linear)

Marshallian (Uncompansated) Elasticities

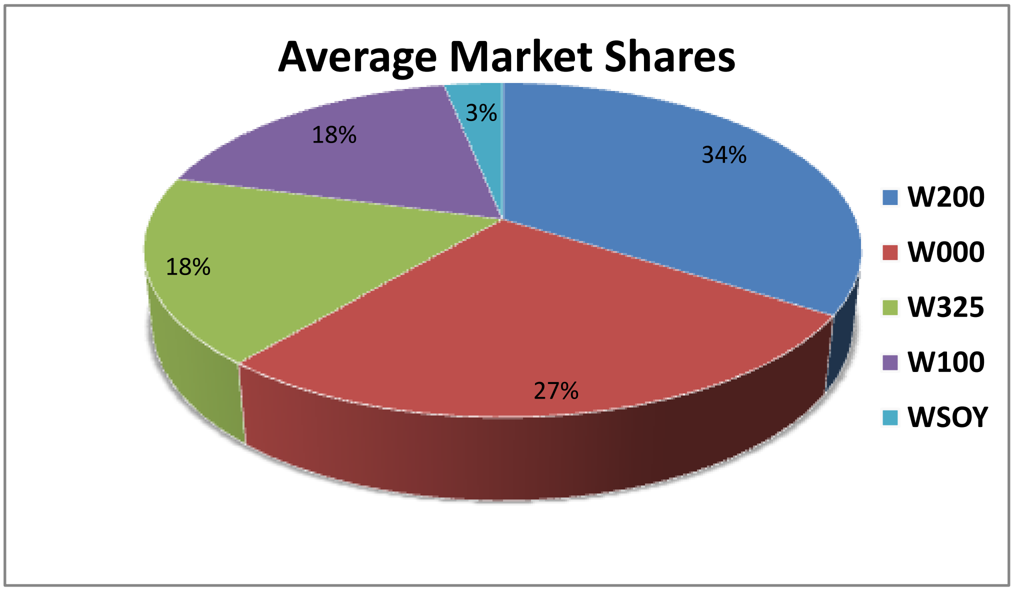
	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
PERCENT2	-0.86962	✓*	0.071	-0.01604	✓	0.022	-0.01391	✓	0.0164	-0.02967	✓	0.023	-0.0057	✓	0.005
SKIM	-0.0949	✓*	0.036	-0.92799	✓*	0.062	-0.04618	✓*	0.0194	-0.04796	✓*	0.021	-0.0122	✓	0.006
FULLFAT	0.03666	✓	0.035	0.026185	✓	0.029	-0.88997	✓*	0.0982	-0.01948	✓	0.038	-0.0048	✓	0.009
PERCENT1	-0.03475	✓	0.048	-0.01043	✓	0.034	-0.01392	✓	0.0269	-0.86332	✓*	0.154	-0.0091	✓	0.01
SOY	0.02583	✓	0.128	0.008896	✓	0.108	-0.01472	✓	0.0832	-0.17276	✓	0.202	-0.5986	X*	0.239

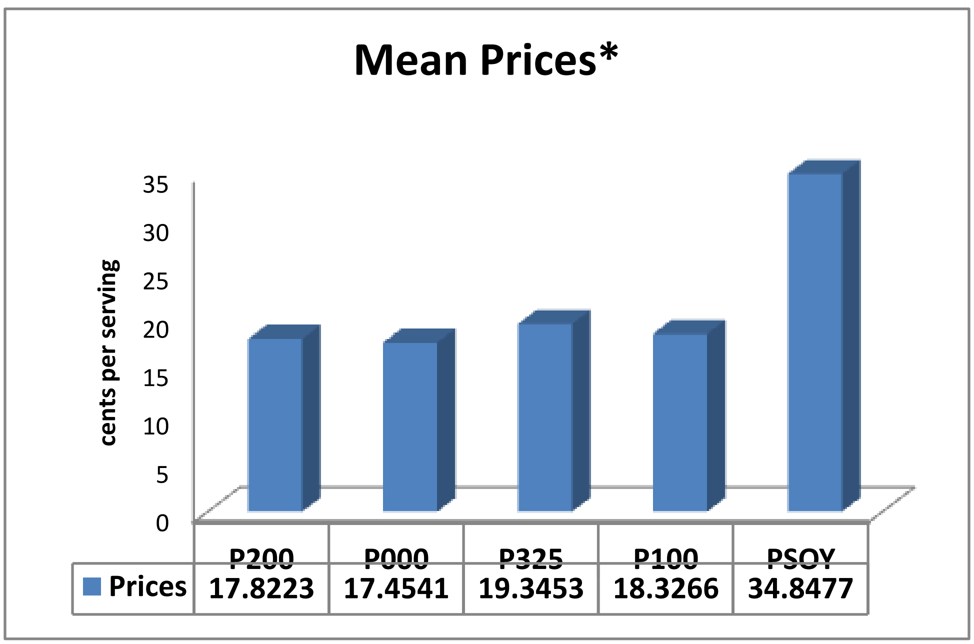
Expenditure Elasticities

	PERCENT2			SKIM			FULLFAT			PERCENT1			SOY		
	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.	Estimate		S.E.
Estimate	1.00768	✓*	0.066	1.174939	✓*	0.083	0.82294	✓*	0.0892	0.941857	✓*	0.088	0.75133	✓*	0.329

(*) Significant at the 95% level.
(✓) The estimated parameter is within the 95% confidence interval of original model.

Figure 2.1 Average Market Shares and Prices





*Prices are measured in terms of cents per serving. A gallon of milk has 16 servings in a single package.

Chapter 3

3. A Discrete Approach to the Demand for Specialty Milk Types Based on Hedonic Prices

Introduction

The USDA Dietary Food Guide suggests consuming 2 to 3 cups of dairy products per day for a 2000 Calorie Diet (2005). Although milk is rich in nutritional value, the presence of lactose in milk might cause uncomfortable digestive issues in some individuals. Digesting lactose requires special enzymes which might not be available in sufficient amounts for some non-European people. When an individual does not have these enzymes, milk consumption can cause gas, cramps and/or diarrhea. Asians are among the racial groups who do not have lactose digesting enzymes. As a result, soymilk has been consumed as substitute for centuries. In recent years, soymilk also has become popular in American households. It is not only lactose free, but it is also cholesterol free which makes it a valuable alternative to those who have cholesterol concerns. Organic milk is also another alternative for health concerned individuals whether they have lactose intolerance or not. However, organic milk, soymilk and CFLF milk products have significantly higher price premiums than other milk types.

In order to identify the factors that derive the demand for these specialty milk types we first estimated hedonic regressions to obtain implicit prices of product attributes. In accordance with hedonic theory, the price of each good can be characterized by a set of its attributes. If we define this attribute set as $x = [x_1, \ldots, x_k]$ then the price of a good can be explained by the relative unit prices and quantities of its characteristics. (Ladd and Suvannunt 1976) In this case, a functional relationship exists between the price of a good and its' characteristic vector x such that $p = f(x) + \mu$ where μ is the error vector. The results of these hedonic regressions give us the implicit prices for not only functional enhancements but also product contents such as protein, fat, and carbohydrate. The estimates from the hedonic regressions indicate positive values for organic and CFLF attributes whereas soy attribute is expected to have a negative value likely due to its unconventional taste.

In the second stage, we applied a probabilistic approach where a consumer chooses to buy or not to buy a specialty milk for a given time period. Through discrete choice modeling, we can determine which types of households are more likely to purchase specialty milk types. According to the binary logistic approach, the probability for occurrence of an event is

$$P[Y = 1\backslash x] = \frac{e^{x'\beta}}{1 + e^{x'\beta}} = \Lambda\,(x'\beta) \qquad (3.1)$$

Both household demographics and the significant hedonic prices are included among the explanatory variables in the binary logistic estimations. The results indicate Asian and Black households have a much higher probability of purchasing specialty milk types. Moreover they are also less responsive to changes in attribute prices and other factors affecting purchase probabilities.

The next section in this paper includes a literature review on hedonic and discrete choice models. The detailed derivation of the data, the sampling, aggregation issues, product attributes and household characteristics used in the estimation are explained in the data section. In the model section we derive the hedonic price theory model and the discrete choice logistic model. The results for the hedonic attribute prices from the first-stage estimation and the logistic demand for specialty milk types from the second-stage estimation are discussed in the results section. Finally, the paper concludes with a summary of findings and directions for future research.

Literature Review

According to hedonic theory, the price of a good can be explained by relative unit prices and quantities of its characteristics. Therefore a functional relationship exists between the price of a good and its' characteristics vector x such that $p = f(x) + \mu$ where μ is the error term. The partial derivatives of the hedonic functions give us the implicit marginal prices of these characteristics.

While analyzing the implicit prices in food market, Ladd and Suvannunt (1976) define the characteristic vector in terms of nutritional attributes. Starting from a large set of attributes and iteratively eliminating the insignificant ones, they reduced the characteristic vector into food energy, protein, carbohydrate, phosphorus, iron, potassium, riboflavin, and ascorbic acid. Their results indicate that some of these characteristics have negative hedonic prices which could arise from unpleasant taste, odor or color.

Rosen's (1974) approach to modeling hedonics is more general than that of Ladd and Suvannunt's because it also allows for inclusion of consumer demographics in the characteristic vector. As Palmquist (2003) indicates, these models allow the econometrician to estimate a schedule of prices for differentiated goods.

While analyzing the factors that determine demand for wine, Nerlove (1995) estimated a hedonic model where the quantity sold depends on price and quality attributes for each wine type. Since wine is almost entirely imported in Sweden, prices can be taken as exogenous, being determined by factors other than Swedish consumers. In Nerlove's model imported quantities depend on the consumer's taste for wine. His results differ significantly from traditional hedonic regressions where product price is determined by attributes.

Discrete choice models are similar to hedonic models in the sense that they also allow for inclusion of both product and consumer characteristics. In these models the consumer is assumed to choose the best product among alternatives. The distribution of the error terms determines the estimation technique in discrete choice models. McFadden's nested multinomial logit model (1978) is among the most popular discrete choice models. Davis and Wohlgenant (1993) applied this model to natural and artificial Christmas tree choices to estimate demand elasticities. They used a two-stage procedure where consumers first decide whether to purchase or not; given the decision to purchase they then decide on the type of the tree to purchase. Their results indicate households' demographic backgrounds affect their first stage decision. Moreover, the price differences between natural and artificial Christmas trees significantly affect households' choices indicating that these tree types are close substitutes. Cotterill and Dhar (2003) employ a similar approach using the nested logit model on the Boston fluid milk market. In their paper they use a two-stage demand where consumers first choose the retail store, and then choose a specific brand of milk. Their analysis indicates that the demand for milk depends on both milk characteristics and price.

In a series of papers, Anderson and Palma apply the logit model in different ways. In their first paper (1992a) the authors develop the multinomial logit model for modeling product differentiation. A taste parameter is introduced to measure consumers' response in utility to changes in product variety. This method provides a simple parameterization of diversity in consumer tastes. In another paper, Anderson and Palma (1992b) apply a nested logit model to analyze the demand facing multiproduct firms. They specify products that are close substitutes within a nest. Two parameters introduced in their model capture the degree of substitution between nests and within a single nest.

In order to identify the effects of advertising, Ackerberg (2001) uses a binomial discrete choice model of demand. In his model the consumer has only two options. The consumer decides on purchasing or not purchasing a recently introduced product. The author suggests that the effect of advertising is much more significant for inexperienced consumers than consumers who are already familiar with the product.

Constantino (2004) applied a nested multinomial logit model to the demand for refrigerated orange juice. Using scanner data on households, the author is able to identify the effect of brand advertising and prices on consumers' choice of juice type. Richards et al. (2004) apply the random coefficients logit model on household scanner data to test for the effect of nutritional addiction on obesity. By analyzing the factors that derive the demand for snacks, the authors find strong evidence of rational addiction to carbohydrates. They propose imposing taxes that target special components such as carbohydrates and fat content in order to reduce excessive calorie intake among households.

An article by Nti and Larweh (2003) survey consumers to explore the sensory characteristics of flavored soymilk samples in Ghana. By using different combinations of flavors, the authors evaluate the consumers' response of sensory characteristics such as color, taste, mouth feel, and aroma. The addition of flavors at different concentration levels provides an estimate of the optimal level of flavor in the soymilk.

Consumers with stronger perceptions of health and diet are more likely to try innovative and nutritionally enhanced food products. Peng et al. (2006) conduct a survey to determine the factors that affect consumer's acceptability of a CLA enriched dairy product. Their findings indicate that those who perceive a significant connection between health and diet are friendlier to CLA enrichment. Moreover, the health concerned consumers are willing to pay a higher premium for functional enhancements in dairy products. Using the AC Nielsen panel data, Huffman and Jensen (2004) analyze the demand for nutritionally enhanced margarines. They calculate significant premiums on functional enhancements. Moreover, their results indicate that acceptance of functional enhancements increases with higher income levels and older couples are more likely to try functionally enhanced margarines. Chema et al. (2006) study consumer views about soy-based dairy foods with functional attributes. They find that consumers are willing to pay more for higher protein, increased calcium, and lower cholesterol. Similar to other studies they find that consumers who are familiar with soybean products are willing to pay more for functional attributes. Consumers who are unfamiliar with soybean products consider soy inferior on the basis of taste. As the authors point out overcoming the perception of poor taste can increase the market share of soymilk. The majority of the consumers are not familiar with the taste of soymilk. Using a two-stage decision model, Moon et al. (2005) analyze the demand for soy-based food products. Their results indicate that the unappetizing taste has a negative effect on soy consumption and health factors derive demand for soy products. Moreover demographic factors also affect households' decisions; Non-white, more educated, and older families are more likely to purchase soy-based products.

The acceptance of recently introduced products depends on both price and consumer specific characteristics as well as product attributes. As Garretson and Burton (2000) indicate, product labeling can be a valuable source of information to explain product attributes. Many characteristics such as nutritional values and as health/environment related effects can be found on the labels.

Data

We derived our data by matching AC Nielsen Homescan Panel with the USDA-ERS Nutrient Database. The Homescan panel tracks purchases made by thousands of household across United States. Once households purchase items, they record the data using scanners and subsequently the data are uploaded to the AC Nielsen database. We obtained a portion of these data related to dairy purchases between years 1998 and 2005. However, not all participants in the panel regularly report their purchases. In fact many of them exit/enter during these periods. Therefore we selected a core set of households who regularly participated in the panel from 2002 until 2005. These households report purchasing fluid milk products at least 12 times a year. The demographic information on the core data set closely resembles that of the entire sample.

The scanner data possess information on prices, discounts, volumes, expenditures and purchase dates of all dairy products purchased by consumers along with relevant demographic information. Although the AC Nielsen panel includes extensive information about household demographics, the product attributes are limited to discrete variables based on information shown on the labels. Therefore information about product attributes is enriched using the USDA-ERS Nutrient Database. Once milk products are classified according to their attributes, the two databases were matched to complete the information on product attributes. For each milk type, continuous nutritional contents include protein, carbohydrate, fat contents, along with sodium and cholesterol contents which are obtained from the Nutrient Database.

The USDA National Nutrient Database is based on an extensive set of research that is compiled from many resources and provides a reference to most food composition databases. Even in abbreviated form, the food attributes for each product are quite exhaustive and include details on each major nutrient component including protein, carbohydrate and fat content. Although these major components are measured in grams per serving, measurement units for vitamins and mineral are defined in different units. We transformed all values for vitamins and minerals into daily recommended intake percentages provided in a serving size defined by the U.S. Federal Drug Administration.

The issue of almost perfect correlation between minerals and vitamins is solved by combining all vitamin and mineral contents into a single vitamin-mineral index. Sodium and cholesterol are two important components listed on the product labels that may raise health concerns among consumers. In fact, they have a negative effect on the value of the product. Therefore, we excluded sodium from this index and included it as a separate component due to its differential effect.

The first-stage hedonic regressions give us marginal implicit prices for product attributes. Prior to estimation, we deflated the milk prices using a Divisia price index of all milk and soy products. In the logit analysis we used these hedonic prices along with households' demographic variables to estimate the probabilities of purchasing specialty milk types. **Table 3.1** gives us the definitions, means, and standard deviations of the second-stage variables. The mean values for hedonic prices indicate positive prices for CFLF and organic claim attributes whereas the soymilk attribute has a negative hedonic price. Nutritional enhancements that affect mineral and vitamin content are appreciated by consumers whereas sodium and cholesterol content are not desirable. The hedonic prices for main nutritional components are all positive with protein having the highest hedonic price.

The average household annual income is approximately $55.000 and the average size of the household is 2.57. Approximately 25% of the households have kids under the age of 18 and 71% of the panel participants are married. Approximately 29% of the household leaders are not fully employed. Although that number might be seen as a very high percentage, it includes all households where adults might be retired, doing social work for non-profit, employed temporarily or working less than full-time. The education level of households is based on the highest education level attained by any of the adult members in the house. We used both continuous and categorical variables for the education level. The continuous education level is scaled as an index from 1 to 6 and calculated as the average education of adult members in the household.

Table 3.2 shows the demand for specialty milk types by each demographic classification, and **Table 3.3** presents household profiles by demand for specialty milk types. On average, specialty milk consumers are smaller in size, older than rest, and have a lower ratio of kids. The ratio of single households is higher among specialty milk consumers whereas the ratio of married households is lower. Moreover the percentage of Asian and Black consumers, compared to other households, is more than twofold among specialty milk type consumers. Income affects soymilk purchases differently than organic milk and CFLF milk. As can be seen from **Table 3.2** the average income of organic milk and CFLF milk consumers is higher than the average income for all consumers whereas the average income of soymilk consumers is lower.

Table 3.3 presents the purchase behavior for each demographic group. In order to see the effect of continuous demographic variables, we transformed household income, size and age into dummies based on whether they are above or below average values. The results in **Table 3.3** conform to **Table 3.2**. The households with above average income have a higher preference for organic and CFLF milk and a lower preference for soymilk. Not having children under 18, having a smaller family size, and being older than average increases the percentage of specialty milk consumption. We can also observe that households with higher education levels have a higher preference for specialty milk types than households with lower education levels. Another striking result is the preference of Asian and Black households. Approximately 12% of these households purchase CFLF milk whereas only 4.38% of white households prefer CFLF milk. The percentage of organic milk and soymilk shoppers is also higher among minority groups compared to white households.

Model

Rosen (1974) developed the structural model to calculate the unobservable attribute prices. His model also allows for the inclusion of consumer characteristics. If we let $x =$

$(x_1, \ldots, x_k)$ denote the attributes of a given product, the price would be related to these attributes such that:

$$P = P(x) = P(x_1, \ldots, x_k) \tag{3.2}$$

Each consumer would be maximizing their utility subject to their budget constraint:

$$\max_{\{x\}} U(x_1, \ldots, x_k, x_0; \alpha^n) \quad s.t. \;\; I^n = P(x) + x_0 \tag{3.3}$$

Here x_0 refers to the numerairé index of all other goods and I^n is the normalized income of consumer n. The first-order conditions (FOC) for utility maximization are

$$\begin{gathered} \frac{\partial U^n}{\partial x_i} = \lambda^n . p_i \qquad i = 1 \ldots k \\ \frac{\partial U^n}{\partial x_0} = \lambda^n \\ I^n = P(x) + x_0 \end{gathered} \tag{3.4}$$

The ratio of the 1st and 2nd equations above gives the marginal rate of substitution between the attributes and the numerairé good. For each consumer that expression gives $p_i = \partial U^n_{x_i} / \partial U^n_{x_0}$. Thus the price of an attribute x_i is equal to the MRS between this attribute and the numerairé. For each individual n if we define the bid function as $\theta^n = \theta^n(x, I^n, U^n, \alpha^n)$ the utility is implicitly given as $U^n(x, I^n - \theta^n, \alpha^n) = U^n$ .The FOCs for the redefined utility function suggests that the consumer's marginal bid function for an attribute would be equal to the marginal price of that attribute

$$\theta_{x_i}(x, u) = p_i \quad for\ all\ i = 1, \ldots, k \tag{3.5}$$

By conceptualizing the individual bid function in this way, Rosen's model can explain the differences in the effects of marginal characteristic prices in consumer decisions. In the first stage implicit prices are calculated by estimating the parameters that explain the relationship between product attributes and prices of products. If we define this relationship as linear

$$p_i = \sum_{j=1}^{k} x_{ji}\beta_j + \varepsilon = f_i(x) + \ \varepsilon \qquad \left(\frac{\partial f_i(x)}{x_{ij}} = \beta_j\right) \tag{3.6}$$

where x_{ji} is the amount of attribute j in product i, then the partial derivative of the hedonic regression, β_j, indicates how much the price of product i changes with respect to an additional unit of change in characteristic x_{ji}. However, since our data are based on weekly time series, the price of an attribute might differ between regions and can change over time. Therefore, we specified our first stage hedonic model as a semi-log model where prices of products are measured in log form

$$p_i = \beta_a \prod_{j=1}^{k} \exp\,(\beta_j . x_{ji}) \ \ or \ln p_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ji} \ where \ \beta_0 = \ln \beta_a \tag{3.7}$$

In this form β_j would be equal to the derivative of $\log p_i$ with respect to x_{ji} such that

$$\beta_j = \frac{\partial lnp_i}{\partial x_{ji}} = \frac{\partial p_i / p_i}{\partial x_{ji}} \ and \ \frac{\partial p_i}{\partial x_{ji}} = \beta_j p_i \tag{3.8}$$

The coefficient β_j shows the marginal percentage change rate in prices due to a unit change in an attribute. This method is convenient for the purpose of calculating the price premiums in percentage forms. It also allows us to differentiate the hedonic prices for each product such that the price of the product also affects the prices of characteristics within the product. Once we derive the hedonic prices of each attribute over time as $p_{x_{ji,t}} = \beta_j p_{i,t}$ we can use these prices along with the background information on consumers to estimate their demand for attributes in the second stage.

In traditional hedonic models the quantities of attributes are related to the hedonic prices and consumer demographics in the second stage to estimate the demand for attributes. However, because we want to identify consumer profiles for each type of milk, we decided to estimate a discrete demand model in the second stage of estimation. Using the discrete choice approach, we can focus on how household demographics affect purchase decision. Moreover, this approach also enables us to use household level data. For multiple choice sets, varieties of logistic models are the most commonly applied distributions. We used the logit model to explain systematic taste variation among households based on demographic information.

The logit model is based on the assumption that the consumer chooses the product yields the highest utility among the feasible set of alternatives. Assuming that the unobserved factors are independent across time and they are independent and identically distributed, the logit model can be used. The model is derived by defining the utility received from a product as a sum of both deterministic and random components. If a decision maker labeled n faces j alternatives then the utility can be decomposed as

$$U_{nj} = V_{nj} + \varepsilon_{nj} \ \ \forall j \tag{3.9}$$

Assuming that the random part of the utility is distributed as i.i.d. type I extreme value, the density function for the error terms can be stated as

$$f\left(e_{nj}\right) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}} \tag{3.10}$$

Following Train’s (2003) notation the probability that the consumer n chooses product i is

$$\begin{aligned} P\left(U_{ni} > U_{nj}, \forall i \neq j\right) &= P\left(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}, \forall\ i \neq j\right) \\ &= P(\varepsilon_{nj} < \varepsilon_{ni} + V_{ni} - V_{nj}, \forall i \neq j) \end{aligned} \tag{3.11}$$

Some algebraic manipulation gives the following logit choice probability for a decision maker n choosing product i over other alternatives:

$$P\left(U_{ni} > U_{nj}, \forall i \neq j\right) = \frac{\exp\,(V_{ni})}{\sum_j \exp\,(V_{nj})} \tag{3.12}$$

The estimation of the parameters usually utilizes Maximum Likelihood Estimation. If we define the utility as linear in parameters such that $V_{nj} = \beta^{'} x_{nj}$ the logit probability can be rewritten as

$$P_{ni} = \frac{e^{\beta^{'} x_{ni}}}{\sum_j e^{\beta^{'} x_{nj}}} \tag{3.13}$$

In our binary logit model a household has the option to buy or not to buy a specific type of milk for a given week. Since the difference between two extreme value variables is also distributed as logit, the logit probability with only two alternatives can be written as

$$P_n = \frac{e^{\beta' x_n}}{1 + e^{\beta' x_n}} \tag{3.14}$$

In this model x_n refers to location and time dependent price variables faced by households along with household specific demographic information. The vector β refers to the corresponding coefficients that measure the effects of these variables on specific probabilities. Since we aim to estimate the systematic variation among household tastes for specialty milk types, we can assume that the unobserved factors that affect the decision makers are independent not only between households but also across time. This assumption enables us to treat our panel data as pooled cross-sectional data. In this case, adding the time (week) subscripts gives us the following utility obtained from product j at period t by household n:

$$U_{njt} = V_{njt} + \varepsilon_{njt}\ \forall j, t \tag{3.15}$$

Therefore the choice probabilities for each period can be stated as

$$P_{nit} = \frac{e^{\beta_1' x^1{}_{nt} + \beta_2' x^2{}_n}}{1 + e^{\beta_1' x^1{}_{nt} + \beta_2' x^2{}_n}} \tag{3.16}$$

In this notation x^1 refers to time and household specific hedonic prices faced by households and x^2 refers to the background information on households.

Logit Elasticities

In order to calculate the elasticities we first estimated the derivatives of the choice probabilities. The derivative of a choice probability for commodity i with respect to an observed factor x_n can be stated as

$$\frac{\partial P_{nit}}{\partial x_{nit}} = \partial \left(\frac{e^{\beta' x_{nit}}}{\sum_j e^{\beta' x_{njt}}} \right) \Big/ \partial x_{nit} \tag{3.17}$$

Assume that the utility is linear in parameters with a coefficient of β_n on factor x_n. If x_n affects the utility derived from commodity i then the derivative above can be calculated as

$$\frac{\partial P_{nit}}{\partial x_{nit}} = \frac{\partial V_{nit}}{\partial x_{nit}} P_{nit}(1 - P_{nit}) = \beta_n P_{nit}(1 - P_{nit}) \tag{3.18}$$

The elasticity of P_{nit} with respect to x_{nit} is

$$E_{x_{nit}}^{nit} = \frac{\partial P_{nit}}{\partial x_{nit}} \frac{x_{ni}}{P_{ni}} = \beta_n P_{nit}(1 - P_{nit}) \frac{x_{nit}}{P_{nit}} = \beta_n x_{nit}(1 - P_{nit}) \tag{3.19}$$

Results

In the first stage, we estimated both linear and semi-log hedonic regressions. The linear hedonic regression can be stated as

$$p_{it} = \sum_{j=1}^{k} \beta_j x_{jit} \tag{3.20}$$

where p_{it} is the price per serving size of i^{th} item at period *t*, x_{jit} is the amount of j^{th} attribute in product *i* at period *t*. In this form the coefficient on attribute quantities, β_j gives us the implicit price of an attribute. In the semi-log form the product price is defined in log terms such that

$$\log p_{it} = \sum_{j=1}^{k} \beta_j x_{jit} \tag{3.21}$$

The variables in semi-log form are defined in the same way with linear form. However, the coefficients on attribute quantities show the percentage changes in prices with respect to changes in attribute quantities. In order to calculate the implicit attribute prices we adjust for the prices by multiplying the price of the product with attribute coefficients. In this case, the implicit prices depend on product prices where the term $\beta_j * p_{it}$ equals the hedonic price. The semi-log form is not only convenient but it also adds variety to the attribute prices such that the same attribute adds a higher value to a product with higher price.

We used two methods to adjust for inflation. First we included a weekly trend variable to minimize the effect of time on the results for both linear and semi-log regressions. Second we deflated the prices using a Divisia price index formula. When we compare the semi-log forms, we observe that there is not much difference between using deflated prices and using a trend variable. However, if we do not deflate the prices or if we remove the trend variable, not only the estimation power (R-square) reduces significantly but also the model overestimates the parameters. Therefore in the first stage, we used a semi-log hedonic regression where prices are deflated using the Divisia price index.

The results for hedonic estimations are given in **Table 3.4**. We observe that the soy attribute has a negative self attribute, possibly due to its unconventional taste. However soymilk has many other desirable attributes; it is naturally LCFC and mostly organic which increases the final value of soymilk. Being LFCF adds a significant 66% price premium and being organic increases the prices by 33%. We also observe that households behave rationally by assigning positive values on desirable or healthy attributes and negative values on unhealthy attributes. Vitamin enhancement labels and nutritional values are both desirable attributes with positive hedonic prices. However cholesterol and sodium content are undesirable characteristics which consumers are willing to pay to avoid their presence. Among the major nutrients, protein content is significantly more valued than fat or carbohydrate content. Products on promotion are on average 18% cheaper than non-discounted products. Each additional serving in a package reduces the price by 3.5%. Glass, carton, box packages are more expensive compared to plastic package, whereas purchasing in a pouch (bundled) reduces the price per serving. Also, compared to the grocery stores, convenience stores are slightly more expensive, whereas drug stores, mass merchandisers, supercenters, and club stores are relatively cheaper. In fact, club stores are 19% cheaper per serving than grocery stores. Since grocery and convenience stores offer the convenience of location and better customer service we observe higher prices in these stores.

We used the results derived from the first-stage hedonic estimation along with the background information on the households to estimate demand functions for specialty milk types. Since there are too many variables based hedonic prices and demographic information on households, we applied a stepwise elimination method to select and use only the most significant variables in the binary logit stage.

Table 3.5 shows the logit parameter estimates and odds ratios for three types of specialty milk products which are soymilk, cholesterol free lactose free (CFLF) milk and organic milk type. The odds ratio is an indicator of how likely a household within a certain class will purchase the special type of milk compared to other household classes. A lower odds ratio indicates a smaller likelihood whereas a higher odds ratio indicates a larger likelihood of purchasing specialty milk. If the odds ratio between two classes is equal to 1 then there is no significant difference in consumer groups regarding their purchase behavior.

For all specialty milk types the coefficient on household size is significantly negative. As can be seen by the odds ratios, larger households have lower probabilities of purchasing specialty milks. On the other hand, the coefficient on household age is significantly positive with an odds ratio greater than one. This indicates elderly households are more health concerned and prefer specialty milk types more than other households.

Most of the demographic factors affect the probabilities in the same direction for all specialty milk types. One striking result in all logit estimations is the way race affects purchase behavior. Minority households have a much higher probability of choosing specialty milk types than white households. In fact, the odds ratio for Asians vs. White and Black vs. White households is greater than three. As suggested by Dr. Steinman (2002), approximately 90-95% of black individuals are deficient in the enzymes that digest lactose. This ratio might even be higher among Asians. The outcomes of logistic estimation support the scientific view that Asians and Blacks suffer most from lactose intolerance (Press 2005). In addition, if the household is employed, the chances of purchasing specialty milk are higher compared to the unemployed households. Education level is also another factor that affects consumers' decision. The households' attitude towards functionally enhanced specialty milk types gets more favorable with a higher education level. Households with an education level of college degree or higher have a greater chance of purchasing specialty milks compared to those with lower education levels. To get a better understanding of the effect of education, we also estimated the logit model using a linear education index that is scaled from 1 to 6. The results suggested a significantly positive coefficient for the effect of education level.

When we analyze the logit demand for soymilk and CFLF milk, we observe that the price of soymilk attribute has a positive effect on probabilities of purchasing. Although it might seem contrary to demand theory to observe positive own-price effects we need to recall that the soy attribute has a negative hedonic price. The hedonic regressions confirm that soy taste is an undesirable attribute which consumers are willing to pay to avoid. A positive increase in this attribute implies a reduction in willingness to avoid soy taste, thus an improvement in consumers' attitude towards soy taste—consistent with the theory of demand. As consumers become familiar with the soy taste, the chances that they will buy soymilk increases and logit estimation reveals that relationship. Another interesting result is the positive coefficient on organic dummy. If the price of organic attribute increases, then the probability of purchasing soymilk or CFLF milk increases. This result indicates organic milk is a substitute of soymilk and CFLF milk. Given the high price premiums for specialty milk types, organic milk can be a feasible alternative to soymilk for households concerned about health/environmental issues.

The effect of income on logit demand for soymilk is insignificant. The frequency analysis in **Table 3.2** suggests that the average income of households who purchase soymilk is slightly lower than that of the entire sample. The same frequency analysis also reveals that the purchasers of organic and CFLF milk have significantly higher incomes than average households. The logit model supports these findings, indicating that a higher income increases the probability of purchasing organic milk and soymilk products.

The presence of children under 18 in the household increases the probability for purchasing CFLF milk but it has a negative effect on soymilk and organic milk purchase probabilities. The gender of the head of the household also affects the probabilities of purchasing different specialty milk types. If the household is male, then the probability of purchasing soymilk and organic milk decreases whereas households with a female head have a lower probability of purchasing CFLF milk. That might be because males prefer CFLF milk and females prefer soymilk and/or organic milk. Another interesting result is the effect of marriage on household's preference. If the couple is married than the chances of organic milk purchase are higher and marriage reduces the probabilities of purchasing soymilk and CFLF milk.

We also estimated the logit demand elasticities for each type of milk. **Table 3.6** gives the outcome for mean and median elasticity estimates. The logit elasticities are calculated as percentage changes in purchase probabilities for percentage changes in explanatory variables. However, household size and age change in single unit increments, so calculated the elasticities for these variables based on unit changes instead of percentage changes. Thus the terms EUnitHHSize and EUnitHHAge are percentage changes in probabilities for a unit change in household size and age respectively.

Own-price elasticities are negative as expected. Since the soy attribute price is negative, the elasticities with respect to the soy attribute price measures how the probabilities change when consumers' attitudes towards soy taste changes. For example the logit demand elasticity of soymilk with respect to soy attribute price is -0.6859. That means a one percentage increase in consumers' willingness to avoid soy taste reduces their chances of purchasing soymilk by 0.6859 percent.

Other than the income elasticity and household size elasticity, the elasticities for soymilk and CFLF milk are very close to each other. Income elasticity of soymilk is almost zero. As the parameter estimates in **Table 3.6** suggest, income does not have much effect on soymilk consumption behavior. However, income effects the household's decision to purchase CFLF and organic milk. Among all milk types, the income elasticity of organic milk is the highest. The household size elasticity is negative in all models. The household size elasticity of CFLF milk is almost twice that of organic milk and even higher than twice the household size elasticity of soymilk. The mean elasticity with respect to household age is positive as expected, however the effect of household age is much higher for soymilk and CFLF milk compared to organic milk.

It is also of interest to see how different consumer types respond to changes in hedonic prices, household income, size and age. Therefore we also calculated logit elasticities for each consumer type when consumers are classified according to their marital status, education level, employment and household leader type, race, and presence of kids. **Tables 3.7, 3.8 and 3.9** give the results of elasticity estimations by different household groups for soymilk, CFLF milk and organic milk respectively. The elasticities do not indicate significant differences between households when they are classified according to education level or employment status. However, we observe significant differences in elasticities based on household's marital status, leadership, racial profile and children status.

For all specialty milk types, married households and households with children are more responsive to changes in hedonic prices, household income, size and age. Changes in hedonic prices or continuous household attributes affect married households' and households with childrens' decisions more than non-married households. In addition female head of households are less responsive when it comes to soymilk and organic milk purchase decisions yet they are more responsive to changes in factors that affect CFLF milk purchases. Another striking result is the way that elasticities differ among households' racial profile. Asian and Black households not only have higher probabilities of purchasing specialty milk types but they also have lower elasticities than other households. White households are the most hedonic price elastic and income elastic group among different races. Since lactose intolerance is highest among the minority households, they prefer specialty milk types without lactose more than white households. They are also less responsive to changes in factors that might affect their decisions.

Summary

In this paper we applied a discrete choice model of specialty milk consumption based on hedonic prices. The first-stage results from hedonic estimation indicated that consumers derive positive utilities from healthy nutritional attributes. Although soymilk has a significant price premium over traditional milk types due to its desirable attributes such as being naturally CFLF and mostly organic, the soy attribute itself has a negative marginal implicit price. The consumers are willing to pay extra to avoid soymilk's unconventional and possibly unappealing taste. It is of soymilk producers' interest to make consumers familiar with the soy taste through promotions to avoid the negative attitude towards soymilk. Our results also reveal important suggestions for dairy producers. In order to have a higher share in the premium priced niche market of functionally enhanced milk, it is in dairy producers' interests to remove cholesterol and lactose from their products. Being organic adds a 33% price premium whereas being LFCF adds 66% price premium on milk products.

In the second stage we applied a logit model of specialty milk type choice for soymilk, CFLF milk, and organic milk. The explanatory variables are based on the implicit hedonic prices from the first stage and the household demographics. Our results indicate that smaller and elderly households have a higher probability of participating in the specialty milk market. The most striking result is the effect of race on households' decision. Non-white households have a higher probability of purchasing specialty milks. In particular, the likelihood of purchasing specialty milk types is almost 2-3 times more for Asian and Black households than white households. Our results support the scientific view that race is a significant factor in lactose intolerance such that Asian and Black households are more sensitive to the presence of lactose in their diet. Moreover, these households have the smallest price and income elasticities, suggesting that they are loyal customers of specialty milk types. For many households, cholesterol free and/or lactose milk is a necessity rather than luxury. In fact, the income elasticity of soymilk is insignificant. These results suggest important policy implications. Policy makers might reconsider the current food pyramid which suggests 2-3 servings of dairy products each day since that is not possible for many households. For lactose intolerant households, alternative dairy products might also be suggested. The scientific research on processes to remove cholesterol and lactose will have the greatest welfare effects on minority households.

The two-stage hedonic-logistic model applied in this paper can also be expanded in several directions. While analyzing the logit price elasticities, we used hedonic prices which are hypothetical prices derived from the relationship between product prices and attributes. It is also of our interest to see how changes in product prices will affect the logit demand elasticities. Another direction of research might relax the assumption of single choices among exclusive alternatives. Such a model would enable us to estimate the demand for specialty milk types simultaneously in a single model instead of separate equations.

References

Ackerberg, Daniel A. "Empirically distinguishing informative and prestige effects of advertising." *RAND Journal of Economics* 32, no. 2 (2001): 216-333.

Anderson, Simon P, and Andre de Palma. "Multiproduct Firms: A Nested Logit Approach." *The Journal of Industrial Economics* 40, no. 3 (September 1992b): 261-276.

Anderson, Simon P, and Andre De Palma. "The Logit as a Model of Product Differentiation." *Oxford Economic Papers* (Oxford University Press) 44 (January 1992a): 51-67.

Chema, S. Kambua, Leonie A. Marks, Josephy L. Parcell, and Maury Bredahl. "Marketing Biotech Soybeans with Functional Health Attributes." *Canadian Journal of Agricultural Economics* 54 (2006): 685-703.

Costantino, Cesar. *Three Essays on Vertical Product Differentiation: Exclusivity, Non-Exclusivity and Advertising.* PhD Thesis, University of Maryland, College Park, 2004.

Cotterill, Ronald W., and Tirtha Dhar. *Oligopoly Pricing with Differentiated Products: The Boston Fluid Milk Market Channel.* Connecticut: Food Marketing Policy Center, 2003.

Garretson, J., and S. Burton. "Effects of nutrition facts panel values, nutrition claims, and health claims on consumer attitudes, perceptions of disease-related risks, and trust." *Journal of Public Policy and Marketing* 19, no. 2 (2000): 213-227.

Huffman, Sonya K, and Jensen Helen H. "Demand for Enhanced Foods and the Value of Nutritional Enhancements of Food: The Case of Margarines." Denver: American Agricultural Economics Association, 2004.

Ladd, George W., and Veraphol Suvannunt. "A model of Consumer Goods Characteristics." *American Journal of Agricultural Economics* (American Agricultural Economics Association) 58 (August 1976): 504-510.

McFadden, Daniel. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Residential Location*, by et.al. Karlgvist, 75-96. Amsterdam: North Holland, 1978.

Moon, Wanki, Siva K Balasubramanian, and Rimal Arbindra. "Perceived Health Benefits and Soy Consumption Behavior:Two-Stage Decision Model Approach." *Journal of Agricultural and Resource Economics* (Western Agricultural Economics Association) 30, no. 2 (2005): 315-332.

Nerlove, Mark. "Hedonic Price Functions and the Measurement of Preferences: The Case of Swedish Wine Consumers." *European Economic Review* 39 (1995): 1697-1716.

Nti, Christina A., and Patience M. Larweh. "Production and Sensory Sharacteristics of Flavoured Soymilk Samples." *International Journal of Consumer Studies* (Blackwell Publishing Ltd) 27, no. 3 (June 2003): 181-184.

Palmquist, Raymond B. "Property Value Models." In *Handbook of Environmental Economics*, edited by Goren K Maler and Vincent Jeffrey. Elsevier, 2003.

Peng, Yanning, Gale E. West, and Cindy Wang. "Consumer Attitudes and Acceptance of CLA - Enriched Dairy Products." *Canadian Journal of Agricultural Economics* 54 (2006): 633-684.

Press, Associated. *News Max.* October 1, 2005. http://archive.newsmax.com/archives/articles/2005/10/30/215232.shtml (accessed Sep 15, 2008).

Richards, Timothy J, Paul M Patterson, and Abe Tegene. "Faculty Working Paper Series." *Morrison School of Agribusiness and Resource Management.* JULY 9, 2004. http://agb.east.asu.edu/workingpapers/0407.pdf (accessed September 12, 2008).

Rosen, R. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, no. 82 (1974): 34-55.

Steinman, Harris. "Milk Allergy and Lactose Intolerance ." *Science in Africa.* May 2002. http://www.scienceinafrica.co.za/2002/may/milk.htm (accessed September 15, 2008).

Train, Kenneth E. *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press, 2003.

U.S. Department of Health and Human Services and U.S. Department of Agriculture. *Dietary Guidelines for Americans.* 2005. http://www.health.gov/dietaryguidelines/dga2005/document/pdf/DGA2005.pdf (accessed September 28, 2008).

Wohlgenant, Micheal, and George C. Davis. "Demand Elasticities from a Discrete Choice Model: The Natural Christmas Tree Market." *American Journal of Agricultural Economics*, no. 75 (1993): 730-738.

Table 3.1 Second Stage Explanatory Variables

Variable	Type	Explanation	Mean	Std Dev	Minimum	Maximum
PMSoy	Continous	Hedonic Price of Soy Attribute	-9.730	1.636	-20.576	-5.084
PMCFLF	Continous	Hedonic Price of CFLF Attribute	23.562	3.428	9.369	36.119
PMOrgClaim	Continous	Hedonic Price of Organic Attribute	11.737	1.873	6.586	32.527
PMVitMinIndex	Continous	Hedonic Price of VitMin Attribute (Marketing)	2.481	0.251	1.897	3.015
PNProtein	Continous	Hedonic Price of Protein Content	1.097	0.108	0.848	1.333
PNCarboHydrt	Continous	Hedonic Price of Carb Content	0.401	0.040	0.311	0.505
PNLipid_Tot	Continous	Hedonic Price of Lipid Content	0.373	0.036	0.287	0.497
PNPerCholestrl	Continous	Hedonic Price of Cholesterol Content	-0.039	0.004	-0.048	-0.031
PNPerSodium	Continous	Hedonic Price of Sodium Content	-0.971	0.096	-1.192	-0.754
PNPerVitMinIndex	Continous	Hedonic Price of VitMin Content (Nutrition)	0.346	0.034	0.269	0.422
HHIncome	Continous	Household Annual Income	55368.97	30299.52	2500	115000
HHSize	Continous	Household Size	2.569	1.325	1	9
HHKids	Dummy	Dummy for Kids in Household	0.256	0.436	0	1
HHAge	Continous	Household Leader's Age	55.886	11.458	27	70
HHEducation	Continous	Household Education Level (Continous)	4.156	0.982	1	6
HHEducation2	Dummy	Upto 12th Grade without Diploma	0.026	0.160	0	1
HHEducation3	Dummy	High School Diploma	0.193	0.395	0	1
HHEducation4	Dummy	Some College without Degree	0.337	0.473	0	1
HHEducation5	Dummy	College Degree	0.323	0.468	0	1
HHEducation6	Dummy	Graduate or Professional Degree	0.121	0.326	0	1
HHCompMarried	Dummy	Married Household	0.716	0.451	0	1
HHCompNonRelated	Dummy	Household Living with NonRelated	0.028	0.166	0	1
HHCompRelated	Dummy	Household Living with Related	0.083	0.276	0	1
HHCompAlone	Dummy	Household Living Alone	0.172	0.377	0	1
HHMaritalMarried	Dummy	Household Married	0.718	0.450	0	1
HHMaritalWidowed	Dummy	Household Widowed	0.082	0.274	0	1
HHMaritalDivorced	Dummy	Household Divorced	0.094	0.292	0	1
HHMaritalSingle	Dummy	Household Single	0.106	0.308	0	1
HHEmpBoth	Dummy	Both Parents Employed	0.317	0.465	0	1
HHEmpFemale	Dummy	Only Female is Employed	0.165	0.371	0	1
HHEmpMale	Dummy	Only Male is Employed	0.229	0.420	0	1
HHEmpNone	Dummy	No Adult is Currently Employed	0.289	0.453	0	1
HHLeaderFemale	Dummy	Household Leader is Female	0.081	0.273	0	1
HHLeaderMale	Dummy	Household Leader is Male	0.173	0.378	0	1
HHLeaderBoth	Dummy	Household Leader is Both	0.746	0.435	0	1
HHRaceAsian	Dummy	Asian Household	0.023	0.149	0	1
HHRaceBlack	Dummy	Black Household	0.064	0.244	0	1
HHRaceHispanic	Dummy	Hispanic Household	0.060	0.237	0	1
HHRaceOthers	Dummy	Other Households	0.013	0.113	0	1
HHRaceWhite	Dummy	White Household	0.841	0.366	0	1

Table 3.2 Analysis of Demand for Specialty Milk Types by Household Demographics

Analysis Variable	Total Sample	MOrgClaim		MCFLF		MSoy	
		0	1	0	1	0	1
Frequency	525292	512366	12926	498666	26626	510044	15248
HHIncome	$ 55,369.28	$ 55,259.34	$ 59,727.14	$ 55,325.03	$ 56,198.00	$ 55,377.55	$ 55,092.64
HHKids	25.58%	25.75%	18.93%	26.00%	17.68%	25.88%	15.48%
HHSize	2.57	2.57	2.35	2.59	2.26	2.58	2.25
HHAge	55.89	55.87	56.43	55.75	58.47	55.81	58.56
HHCompAlone	17.20%	17.10%	20.93%	16.84%	23.80%	17.00%	23.71%
HHCompMarried	71.65%	71.74%	68.17%	72.02%	64.76%	71.91%	63.05%
HHCompNonRelated	2.85%	2.85%	2.72%	2.81%	3.57%	2.82%	3.89%
HHCompRelated	8.31%	8.31%	8.18%	8.33%	7.86%	8.28%	9.35%
HHEducation	4.16	4.15	4.26	4.15	4.28	4.15	4.25
HHEducation2	0.03	0.03	0.02	0.03	0.02	0.03	0.02
HHEducation3	0.19	0.19	0.16	0.20	0.16	0.19	0.17
HHEducation4	0.34	0.34	0.33	0.34	0.31	0.34	0.33
HHEducation5	0.32	0.32	0.35	0.32	0.39	0.32	0.35
HHEducation6	0.12	0.12	0.13	0.12	0.12	0.12	0.13
HHEmpBoth	31.69%	31.67%	32.52%	32.06%	24.85%	31.88%	25.37%
HHEmpFemale	16.52%	16.47%	18.37%	16.31%	20.36%	16.42%	19.86%
HHEmpMale	22.85%	22.87%	22.41%	22.82%	23.48%	22.79%	25.07%
HHEmpNone	28.94%	28.99%	26.70%	28.81%	31.31%	28.91%	29.69%
HHLeaderBoth	74.58%	74.70%	70.02%	74.92%	68.19%	74.82%	66.57%
HHLeaderFemale	8.13%	8.06%	10.90%	8.06%	9.45%	8.02%	11.82%
HHLeaderMale	17.29%	17.24%	19.08%	17.02%	22.37%	17.16%	21.61%
HHMaritalDivorced	9.40%	9.37%	10.74%	9.26%	12.18%	9.32%	12.15%
HHMaritalMarried	71.81%	71.91%	68.17%	72.19%	64.76%	72.08%	63.05%
HHMaritalSingle	10.62%	10.61%	11.25%	10.57%	11.59%	10.59%	11.67%
HHMaritalWidowed	8.16%	8.12%	9.84%	7.98%	11.47%	8.01%	13.13%
HHRaceAsian	2.28%	2.23%	4.50%	2.13%	5.16%	2.23%	4.07%
HHRaceBlack	6.36%	6.16%	14.17%	5.89%	15.20%	6.06%	16.19%
HHRaceHispanic	5.96%	5.98%	5.04%	5.95%	6.04%	6.02%	3.84%
HHRaceOthers	1.30%	1.30%	1.27%	1.32%	0.99%	1.31%	1.01%
HHRaceWhite	84.10%	84.33%	75.01%	84.72%	72.62%	84.38%	74.90%
MSOY	2.90%	1.08%	75.27%	0.03%	56.74%	0.00%	100.00%
MOrgClaim	2.46%	0.00%	100.00%	0.64%	36.64%	0.63%	63.81%
MCFLF	5.07%	3.29%	75.47%	0.00%	100.00%	2.26%	98.99%

This frequency analysis above indicates the percentage of right hand side variables in terms of left hand variables. For example 16.19% of soymilk is purchased by black households whereas only 6.06% of all non soymilk products are purchased by black households.

Table 3.3 Analysis of Household Profiles by their Demand for Specialty Milk Types

Analysis Variable	MOrgClaim	MCFLF	MSoy		MOrgClaim	MCFLF	MSoy
All Households	2.46%	5.07%	2.91%	ALL OTHERS	97.54%	94.93%	97.09%
HHIncomeAbAv	2.97%	5.39%	3.09%	ALL OTHERS	2.14%	4.88%	2.79%
HHKids	1.82%	3.50%	1.76%	ALL OTHERS	2.68%	5.61%	3.30%
HHSizeAbAv	2.35%	3.93%	2.23%	ALL OTHERS	2.53%	5.76%	3.32%
HHAgeAbAv	2.56%	6.08%	3.56%	ALL OTHERS	2.33%	3.77%	2.06%
HHCompAlone	3.00%	7.03%	4.01%	ALL OTHERS	2.35%	4.67%	2.68%
HHCompMarried	2.34%	4.59%	2.56%	ALL OTHERS	2.76%	6.31%	3.79%
HHCompNonRelated	2.35%	6.36%	3.97%	ALL OTHERS	2.46%	5.04%	2.88%
HHCompRelated	2.42%	4.81%	3.28%	ALL OTHERS	2.46%	5.10%	2.88%
HHEducation2	1.80%	3.19%	2.08%	ALL OTHERS	2.48%	5.13%	2.93%
HHEducation3	2.08%	4.23%	2.52%	ALL OTHERS	2.55%	5.28%	3.00%
HHEducation4	2.44%	4.65%	2.89%	ALL OTHERS	2.47%	5.29%	2.92%
HHEducation5	2.67%	6.18%	3.17%	ALL OTHERS	2.36%	4.55%	2.78%
HHEducation6	2.72%	5.07%	3.08%	ALL OTHERS	2.43%	5.07%	2.88%
HHEmpBoth	2.52%	3.98%	2.33%	ALL OTHERS	2.43%	5.58%	3.18%
HHEmpFemale	2.74%	6.26%	3.51%	ALL OTHERS	2.41%	4.84%	2.79%
HHEmpMale	2.41%	5.21%	3.19%	ALL OTHERS	2.47%	5.03%	2.83%
HHEmpNone	2.27%	5.49%	2.98%	ALL OTHERS	2.54%	4.91%	2.88%
HHLeaderBoth	2.31%	4.64%	2.60%	ALL OTHERS	2.90%	6.35%	3.83%
HHLeaderFemale	3.30%	5.90%	4.23%	ALL OTHERS	2.39%	5.00%	2.79%
HHLeaderMale	2.72%	6.57%	3.64%	ALL OTHERS	2.41%	4.76%	2.76%
HHMaritalDivorced	2.81%	6.59%	3.77%	ALL OTHERS	2.42%	4.92%	2.82%
HHMaritalMarried	2.34%	4.57%	2.55%	ALL OTHERS	2.78%	6.35%	3.81%
HHMaritalSingle	2.61%	5.54%	3.19%	ALL OTHERS	2.44%	5.02%	2.87%
HHMaritalWidowed	2.97%	7.12%	4.67%	ALL OTHERS	2.42%	4.89%	2.75%
HHRaceAsian	4.86%	11.48%	5.20%	ALL OTHERS	2.40%	4.92%	2.85%
HHRaceBlack	5.49%	12.14%	7.41%	ALL OTHERS	2.26%	4.59%	2.60%
HHRaceHispanic	2.08%	5.15%	1.88%	ALL OTHERS	2.48%	5.07%	2.97%
HHRaceOthers	2.40%	3.85%	2.25%	ALL OTHERS	2.46%	5.09%	2.92%
HHRaceWhite	2.19%	4.38%	2.59%	ALL OTHERS	3.87%	8.75%	4.60%

The frequency analysis above indicates the percentage of left hand side variables in terms of right hand variables. For example 12.14% of Black households have purchased CFLF milk where this ratio is 4.59% among all other households.

Table 3.4 Results of Stage 1 Hedonic Estimation

	Time Adjustment			
	Week Adjustment		No week parameter	
Dependent Variable	**PPPCup**	**LogPPPCup**	**LogDefPPPCup**	**LogPPPCup**
	Week Trend	Week Trend	Deflated	No adjustment
	Estimate	Estimate	Estimate	Estimate
Intercept	-7.50685	2.2544	2.31653	2.32372
Marketing Soy Dummy	-7.76391	-0.29258	-0.29127	-0.28696
Marketing Lactose Cholesterol Free	19.35535	0.65619	0.66074	0.67561
Marketing organic Claim	9.10183	0.33252	0.33766	0.35081
Marketing Vitamin Mineral Index	4.79728	0.14614	0.14536	0.14458
Marketing Promotion Dummy	-3.23968	-0.19571	-0.18385	-0.12169
Nutrient Protein Content (g)	2.29756	0.06148	0.0642	0.07108
Nutrient Carb Content (g)	0.7441	0.02239	0.0228	0.02313
Nutrient Lipid (Fat) Content (g)	0.62913	0.02175	0.02142	0.02131
Nutrient Cholesterol DRI Percentage of Max	-0.17633	-0.00253	-0.0023	-0.00233
Nutrient Sodium DRI Percentage of Max	-1.80266	-0.05351	-0.05639	-0.06222
Nutrient Vitamin-Mineral Percentage Index	0.74765	0.01988	0.02013	0.02059
# of servings in a package	-0.79602	-0.03559	-0.03572	-0.03638
Week	0.02098	0.0012		
Marketing Container Box	4.91994	0.0921	0.08098*	0.09094
Marketing Container Carton	0.011'	0.01476	0.01179	0.003*
Marketing Container Glass	11.33017	0.35409	0.35696	0.36018
Marketing Container Pouch	-10.05194	-0.51911	-0.52486	-0.51375
Location Channel Drug	-1.22728	-0.06076	-0.05947	-0.0561
Location Channel MassMerchant	-0.51239	-0.0245	-0.02527	-0.02065
Location Channel SuperCenter	-0.71043	-0.02363	-0.02069	-0.00454
Location Channel Club	-3.09501	-0.19031	-0.18677	-0.16988
Location Channel Convenience	0.19588	0.0134	0.01573	0.0149
Location Channel Other	-1.89019	-0.10809	-0.10821	-0.10147
Adj R-Sq	0.5387	0.4664	0.4611	0.4246

(*) Significant at 5% level.

(**) Significant at 10% level.

(') Insignificant.

All variables are significant at 1% level unless indicated otherwise.

Table 3.5 Estimated Logit Coefficients and Odds Ratios

Parameter	Soymilk Estimate	Soymilk Odds Ratios	CFLF Milk Estimate	CFLF Milk Odds Ratios	Organic Milk Estimate	Organic Milk Odds Ratios
Intercept	-3.9509	_	-3.0317	_	-3.1576	_
PMSoy	0.0725	1.075	0.063	1.065		
PMCFLF	-0.00877*	0.991	-0.00889	0.991		
PMOrgClaim	0.0192	1.019	0.0118	1.012	-0.0336	0.967
HHSize	-0.1138	0.892	-0.2444	0.783	-0.144	0.866
HHAge	0.0266	1.027	0.0257	1.026	0.00411	1.004
HHIncome	0.0000001385'	1'	2.10E-06	1'	4.78E-06	1'
HHKids						
None	0.1169	1.263	-0.0548	0.896	0.1232	1.279
HHrace						
Asian	0.4854	2.458	0.6241	3.368	0.4142	2.323
Black	0.7917	3.339	0.6385	3.416	0.6007	2.799
Hispanic	-0.4685	0.947'	-0.1171	1.605	-0.3218	1.113
Other	-0.3946	1.02'	-0.5555	1.035'	-0.2644	1.179
HHLeader						
Both	0.0364'	1.125	0.2008	1.247	-0.3129	0.684
Female	0.0452**	1.135	-0.1806	0.852	0.2454	1.195
HHEmp						
Both	0.0615	1.424	-0.0547	1.097	0.1126	1.271
Female	0.0282'	1.377	0.0456	1.213	0.0172'	1.155
Male	0.2019	1.638	0.1565	1.355	-0.00249'	1.133
HHEducation						
Upto High School	-0.3672	0.61	-0.4526	0.587	-0.162	0.829
High School Degree	-0.0837	0.81	-0.0727	0.858	-0.043'	0.933'
Some College	0.11	0.983'	0.0675	0.987'	0.0803	1.056'
College Degree	0.2141	1.091	0.3773	1.346	0.0985	1.075
HHMarital						
Married	-0.117	0.791	-0.1198	0.787	0.183	1.442

The Odds Ratio for HHKids compares the odds for households with no kids to those with kids under age of 18
The Odds Ratio for HHRace compares the odds for non-white households to white households for each race
The Odds Ratio for HHLeader compares the odds for nonmale leaded households to male leaded households
The Odds Ratio for HHEmp compares the odds for employed households to unemployed households
The Odds Ratio for HHEducation compares the odds for households with different education levels to those with graduate education
The Odds Ratio for HHMarital compares the odds for married households to nonmarried households

(**) Significant at 10% level.
(') Insignificant.
All variables are significant at 1% level unless indicated otherwise.

Table 3.6 Logit Elasticities*

	Soymilk		
Variable	Mean	Std Dev	Median
EPMSOY	-0.6859008	0.118323	-0.6665012
EPMCFLF	-0.2012689	0.029788	-0.1981196
EPMOrgclaim	0.2188849	0.035487	0.2150536
EHHIncome	0.0074439	0.00408	0.0064768
EUnitHHSize	-11.049581	0.204873	-11.082052
EUnitHHAge	2.5827666	0.047888	2.5903567

	CFLF Milk		
Variable	Mean	Std Dev	Median
EPMSOY	-0.5829898	0.103635	-0.5663044
EPMCFLF	-0.1995361	0.030548	-0.1964032
EPMOrgclaim	0.1315594	0.021901	0.129308
EHHIncome	0.1100298	0.060153	0.0966722
EUnitHHSize	-23.199548	0.768345	-23.343036
EUnitHHAge	2.4395596	0.080796	2.4546482

	Organic Milk		
	Mean	Std Dev	Median
EPMOrgclaim	-0.3848311	0.062508	-0.3780441
EHHIncome	0.2577687	0.140206	0.2237411
EUnitHHSize	-14.047284	0.170925	-14.08107
EUnitHHAge	0.4009329	0.004879	0.4018972

*The hedonic price and income elasticities are calculated as the percentage changes in probabilities for percentage changes in hedonic prices and income. The elasticities for household size and age are measured in terms of percentage changes in probabilities for unit changes in size or age.

Table 3.7 Soymilk Elasticities by Selected Household Demographics

		Mean	*Std Dev*	*Median*
HHMarital				
Married	EPMSOY	-0.6892	0.1185	-0.6694
377247	EPMCFLF	-0.2022	0.0298	-0.1990
	EPMOrgclaim	0.2198	0.0355	0.2160
	EHHIncome	0.0082	0.0040	0.0075
	EUnitHHSize	-11.0899	0.1652	-11.1125
	EUnitHHAge	2.5922	0.0386	2.5975
NonMarried	EPMSOY	-0.6776	0.1175	-0.6587
148076	EPMCFLF	-0.1990	0.0297	-0.1959
	EPMOrgclaim	0.2165	0.0353	0.2127
	EHHIncome	0.0056	0.0036	0.0049
	EUnitHHSize	-10.9470	0.2543	-11.0064
	EUnitHHAge	2.5588	0.0594	2.5727
HHKids				
0	EPMSOY	-0.6846	0.1189	-0.6652
390961	EPMCFLF	-0.2008	0.0298	-0.1977
	EPMOrgclaim	0.2183	0.0355	0.2145
	EHHIncome	0.0070	0.0039	0.0063
	EUnitHHSize	-11.0051	0.2039	-11.0503
	EUnitHHAge	2.5724	0.0477	2.5829
1	EPMSOY	-0.6897	0.1164	-0.6700
134362	EPMCFLF	-0.2026	0.0296	-0.1994
	EPMOrgclaim	0.2205	0.0354	0.2167
	EHHIncome	0.0087	0.0042	0.0088
	EUnitHHSize	-11.1789	0.1437	-11.2240
	EUnitHHAge	2.6130	0.0336	2.6235
HHLeader				
Both	EPMSOY	-0.6883	0.1183	-0.6687
391780	EPMCFLF	-0.2019	0.0298	-0.1988
	EPMOrgclaim	0.2196	0.0355	0.2157
	EHHIncome	0.0081	0.0040	0.0074
	EUnitHHSize	-11.0852	0.1716	-11.1091
	EUnitHHAge	2.5911	0.0401	2.5967
Female	EPMSOY	-0.6750	0.1166	-0.6556
42717	EPMCFLF	-0.1983	0.0296	-0.1951
	EPMOrgclaim	0.2156	0.0349	0.2115
	EHHIncome	0.0066	0.0039	0.0057
	EUnitHHSize	-10.8989	0.2431	-10.9549
	EUnitHHAge	2.5475	0.0568	2.5606
Male	EPMSOY	-0.6807	0.1186	-0.6619
90826	EPMCFLF	-0.1997	0.0298	-0.1968
	EPMOrgclaim	0.2173	0.0354	0.2138
	EHHIncome	0.0049	0.0032	0.0043
	EUnitHHSize	-10.9670	0.2549	-11.0264
	EUnitHHAge	2.5635	0.0596	2.5774

		Mean	*Std Dev*	*Median*
HHrace				
Asian	EPMSOY	-0.6328	0.0855	-0.6194
11990	EPMCFLF	-0.1891	0.0239	-0.1864
	EPMOrgclaim	0.2047	0.0302	0.2015
	EHHIncome	0.0103	0.0042	0.0112
	EUnitHHSize	-10.7859	0.2196	-10.8270
	EUnitHHAge	2.5211	0.0513	2.5307
Black	EPMSOY	-0.6331	0.1044	-0.6166
33403	EPMCFLF	-0.1875	0.0276	-0.1843
	EPMOrgclaim	0.2047	0.0340	0.2007
	EHHIncome	0.0071	0.0038	0.0067
	EUnitHHSize	-10.5358	0.3453	-10.5672
	EUnitHHAge	2.4627	0.0807	2.4700
Hispanic	EPMSOY	-0.6673	0.0983	-0.6500
31294	EPMCFLF	-0.1982	0.0262	-0.1945
	EPMOrgclaim	0.2145	0.0326	0.2108
	EHHIncome	0.0077	0.0042	0.0074
	EUnitHHSize	-11.1641	0.1013	-11.1907
	EUnitHHAge	2.6095	0.0237	2.6158
Other	EPMSOY	-0.6551	0.0924	-0.6400
6832	EPMCFLF	-0.1951	0.0253	-0.1927
	EPMOrgclaim	0.2124	0.0325	0.2086
	EHHIncome	0.0069	0.0040	0.0058
	EUnitHHSize	-11.1209	0.1094	-11.1217
	EUnitHHAge	2.5994	0.0256	2.5996
White	EPMSOY	-0.6931	0.1201	-0.6737
441804	EPMCFLF	-0.2030	0.0300	-0.1999
	EPMOrgclaim	0.2208	0.0356	0.2171
	EHHIncome	0.0074	0.0041	0.0064
	EUnitHHSize	-11.0863	0.1204	-11.0902
	EUnitHHAge	2.5914	0.0281	2.5923

Table 3.8 CFLF Milk Elasticities by Selected Household Demographics

		Mean	Std Dev	Median
HHMarital				
Married	EPMSOY	-0.5867	0.1036	-0.5698
377247	EPMCFLF	-0.2007	0.0305	-0.1975
	EPMOrgclaim	0.1323	0.0219	0.1301
	EHHIncome	0.1210	0.0592	0.1110
	EUnitHHSize	-23.3215	0.6655	-23.4392
	EUnitHHAge	2.4524	0.0700	2.4648
NonMarried	EPMSOY	-0.5735	0.1030	-0.5577
148076	EPMCFLF	-0.1964	0.0305	-0.1935
	EPMOrgclaim	0.1296	0.0218	0.1274
	EHHIncome	0.0821	0.0531	0.0710
	EUnitHHSize	-22.8891	0.9118	-23.1032
	EUnitHHAge	2.4069	0.0959	2.4294
HHKids				
0	EPMSOY	-0.5809	0.1040	-0.5643
390961	EPMCFLF	-0.1988	0.0306	-0.1956
	EPMOrgclaim	0.1310	0.0219	0.1288
	EHHIncome	0.1034	0.0581	0.0927
	EUnitHHSize	-23.0687	0.7614	-23.2350
	EUnitHHAge	2.4258	0.0801	2.4433
1	EPMSOY	-0.5889	0.1024	-0.5722
134362	EPMCFLF	-0.2018	0.0304	-0.1985
	EPMOrgclaim	0.1331	0.0219	0.1308
	EHHIncome	0.1292	0.0619	0.1295
	EUnitHHSize	-23.5800	0.6532	-23.7844
	EUnitHHAge	2.4796	0.0687	2.5011
HHLeader				
Both	EPMSOY	-0.5858	0.1035	-0.5689
391780	EPMCFLF	-0.2005	0.0305	-0.1973
	EPMOrgclaim	0.1322	0.0219	0.1299
	EHHIncome	0.1201	0.0593	0.1107
	EUnitHHSize	-23.3064	0.6802	-23.4282
	EUnitHHAge	2.4508	0.0715	2.4636
Female	EPMSOY	-0.5767	0.1023	-0.5600
42717	EPMCFLF	-0.1976	0.0304	-0.1945
	EPMOrgclaim	0.1302	0.0216	0.1279
	EHHIncome	0.0982	0.0583	0.0854
	EUnitHHSize	-22.9993	0.7727	-23.1751
	EUnitHHAge	2.4185	0.0813	2.4370
Male	EPMSOY	-0.5738	0.1039	-0.5584
90826	EPMCFLF	-0.1964	0.0307	-0.1937
	EPMOrgclaim	0.1295	0.0219	0.1276
	EHHIncome	0.0722	0.0472	0.0635
	EUnitHHSize	-22.8335	0.9669	-23.0653
	EUnitHHAge	2.4011	0.1017	2.4254

		Mean	Std Dev	Median
HHrace				
Asian	EPMSOY	-0.5138	0.0756	-0.5022
11990	EPMCFLF	-0.1791	0.0246	-0.1762
	EPMOrgclaim	0.1175	0.0183	0.1154
	EHHIncome	0.1449	0.0587	0.1572
	EUnitHHSize	-21.6274	0.9536	-21.6763
	EUnitHHAge	2.2742	0.1003	2.2794
Black	EPMSOY	-0.5222	0.0922	-0.5098
33403	EPMCFLF	-0.1804	0.0287	-0.1775
	EPMOrgclaim	0.1194	0.0210	0.1172
	EHHIncome	0.1026	0.0549	0.0931
	EUnitHHSize	-21.4632	1.2662	-21.6290
	EUnitHHAge	2.2570	0.1331	2.2744
Hispanic	EPMSOY	-0.5606	0.0856	-0.5466
31294	EPMCFLF	-0.1942	0.0266	-0.1909
	EPMOrgclaim	0.1274	0.0198	0.1253
	EHHIncome	0.1131	0.0615	0.1067
	EUnitHHSize	-23.1679	0.6008	-23.2628
	EUnitHHAge	2.4362	0.0632	2.4462
Other	EPMSOY	-0.5600	0.0798	-0.5474
6832	EPMCFLF	-0.1945	0.0255	-0.1919
	EPMOrgclaim	0.1284	0.0198	0.1260
	EHHIncome	0.1028	0.0589	0.0860
	EUnitHHSize	-23.4904	0.4015	-23.4586
	EUnitHHAge	2.4701	0.0422	2.4668
White	EPMSOY	-0.5914	0.1041	-0.5747
441804	EPMCFLF	-0.2020	0.0304	-0.1989
	EPMOrgclaim	0.1332	0.0218	0.1310
	EHHIncome	0.1095	0.0602	0.0963
	EUnitHHSize	-23.3710	0.4503	-23.4001
	EUnitHHAge	2.4576	0.0473	2.4606

Table 3.9 Organic Milk Elasticities by Selected Household Demographics

		Mean	*Std Dev*	*Median*
HHMarital				
Married	EPMOrgclaim	-0.3856	0.0626	-0.3788
377247	EHHIncome	0.2827	0.1379	0.2575
	EUnitHHSize	-14.0654	0.1567	-14.0932
	EUnitHHAge	0.4015	0.0045	0.4022
NonMarried	EPMOrgclaim	-0.3829	0.0623	-0.3761
148076	EHHIncome	0.1943	0.1252	0.1703
	EUnitHHSize	-14.0011	0.1953	-14.0466
	EUnitHHAge	0.3996	0.0056	0.4009
HHKids				
0	EPMOrgclaim	-0.3845	0.0626	-0.3777
390961	EHHIncome	0.2432	0.1358	0.2213
	EUnitHHSize	-14.0160	0.1684	-14.0566
	EUnitHHAge	0.4000	0.0048	0.4012
1	EPMOrgclaim	-0.3857	0.0622	-0.3790
134362	EHHIncome	0.3002	0.1441	0.3047
	EUnitHHSize	-14.1383	0.1434	-14.1770
	EUnitHHAge	0.4035	0.0041	0.4046
HHLeader				
Both	EPMOrgclaim	-0.3854	0.0626	-0.3787
391780	EHHIncome	0.2808	0.1383	0.2574
	EUnitHHSize	-14.0690	0.1556	-14.0963
	EUnitHHAge	0.4016	0.0044	0.4023
Female	EPMOrgclaim	-0.3810	0.0616	-0.3737
42717	EHHIncome	0.2302	0.1363	0.1978
	EUnitHHSize	-13.9251	0.1812	-13.9697
	EUnitHHAge	0.3974	0.0052	0.3987
Male	EPMOrgclaim	-0.3840	0.0625	-0.3774
90826	EHHIncome	0.1717	0.1115	0.1516
	EUnitHHSize	-14.0109	0.1966	-14.0668
	EUnitHHAge	0.3999	0.0056	0.4015

		Mean	*Std Dev*	*Median*
HHrace				
Asian	EPMOrgclaim	-0.3597	0.0542	-0.3537
11990	EHHIncome	0.3550	0.1433	0.3874
	EUnitHHSize	-13.7031	0.1998	-13.7038
	EUnitHHAge	0.3911	0.0057	0.3911
Black	EPMOrgclaim	-0.3657	0.0611	-0.3583
33403	EHHIncome	0.2507	0.1310	0.2426
	EUnitHHSize	-13.6126	0.2470	-13.6031
	EUnitHHAge	0.3885	0.0070	0.3883
Hispanic	EPMOrgclaim	-0.3746	0.0574	-0.3683
31294	EHHIncome	0.2660	0.1434	0.2560
	EUnitHHSize	-14.1011	0.1048	-14.1092
	EUnitHHAge	0.4025	0.0030	0.4027
Other	EPMOrgclaim	-0.3712	0.0574	-0.3642
6832	EHHIncome	0.2381	0.1355	0.1990
	EUnitHHSize	-14.0504	0.1380	-14.0237
	EUnitHHAge	0.4010	0.0039	0.4003
White	EPMOrgclaim	-0.3879	0.0627	-0.3814
441804	EHHIncome	0.2554	0.1397	0.2229
	EUnitHHSize	-14.0856	0.0940	-14.0905
	EUnitHHAge	0.4020	0.0027	0.4022