The Welfare Effect of Organic Milk

Hee-Jung Choi *

and

Michael Wohlgenant **

Department of Agricultural and Resource Economics

North Carolina State University

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2010 AAEA,CAES, & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010

* PhD. Candidate, Department of Agricultural and Resource Economics, North Carolina State University.
Contact: Hee-Jung Choi; hjchoi@ncsu.edu.

** William Neal Reynolds Distinguished Professor, Department of Agricultural and Resource Economics, North Carolina State University

I. Introduction

Recent innovations in agricultural industry make new products, such as Genetically Modified (GM) food and Organic food, available to consumers. Although various opinions on the effect of GM food products are not in agreement among nutrition experts, consumers' concern on health and environment increases demand for organic food products. This study analyses the structural changes driven by organic food introduction into the U.S. food sector in terms of its welfare effect. According to the Organic Trade Association (OTA), organic food sales in the U.S. were $13.8 billion in 2005, which is 2.5% of total food sales. This is an increase from 1.9% in 2003 and from 0.8% in 1997. Increasing trends of demand are expected to continue and estimated to rise to $23.8 billion by 2010 (Nutrition Business Journal, 2004). Public policy also played a significant role in the expansion of organic food sector. The National Organic Standards, which is implemented by the U.S. Department of Agriculture (USDA) in 2002, specify the production process for processing, distributing, and growing organic food. The policy also restricts the use of Organic logo by allowing it only to the products whose profile meet the standards. Consumers are not only exposed to more information on organic standard by this policy adoption, but the logo also provides an easy way for consumers to recognize qualified organic products.

As part of the organic food market, organic milk market has also been growing. According to the USDA, organic cow milk and soy milk drinks are the top two categories among processed organic products other than fresh products. Organic milk first appeared in conventional supermarkets in 1993 and 8 conventional supermarkets were selling organic milk in 1996 (Glaser and Thompson, 2000). After the introduction of organic milk in the market, sales of organic milk have been growing; organic milk and cream sales increased from $15.8 million to $104 million, from 1996 to 2000. The industry shows a dramatic increase in sales during the early 2000s, which coincides with the implementation of National

Organic Standards[1] in 2002 and a price increase of conventional milk in 2004. As of 2005, organic milk and cream sales were over $1billion, which is 25% up from 2004 sales. A noticeable fact is that overall sales of milk have remained constant since the mid-1980s, which indicates that organic milk sales not only increased, but also expanded in its market share in overall milk industry (Miller and Blayhey, 2006). This also implies that not only new firms entered the organic milk industry during the past years, but also existing firms increased the supply of organic milk by recruiting and assisting conventional milk producers converting their product to organic milk[2]. (USDA, Retail and Consumer Aspects of the Organic Milk Market)

As interest in organic market grow, agricultural researchers have conducted some studies on organic milk. An earlier study by Glaser and Thompson (2000) considers demand for branded milk, private label milk and organic milk based on supermarket scanned data collected from 1988 to 1999 by AC Nielsen and Information Resources, Inc (IRI). According to the study, demand elasticities computed from nonlinear Almost Ideal Demand System (AIDS) framework indicate that organic milk demand is more elastic than private-label and branded milk. Although this analysis well describes the sensitivity of organic milk demand in the early stages of introduction, it cannot account for the current market analysis because the organic market has grown competitive to the extent that private label milk suppliers also produce organic milk.

The type of consumer more likely to purchase organic products has also been of interest to marketing researchers. Lohr (2001) characterizes organic consumers as White, affluent and well-

[1] Organic dairy products, as defined by USDA, are made from the milk of animals raised organic management. The animals are raised separately from the herd of conventional dairy animals. The animals are not given hormones or antibiotics. The animals receive preventive medical care, such as vaccines, and dietary supplements of vitamins and minerals. (Recent Growth Patterns in the U.S. Organic Foods Market, USDA)

[2] To convert from conventional to organic production, the cow must be fed a diet consisting of at least 80 percent organic feed for 9 months and then 100 percent organic feed for 3 additional months, or must be grazed on land that is managed under a certified organic plan. (Recent Growth Patterns in the U.S. Organic Foods Market, USDA)

educated. Lohr and Semali (2000) conclude that parents of young children or infants are more likely than those without children to purchase organic food. Dimitri and Venezia (2007) provide descriptive statistics on the socio demographic characteristics of organic milk consumers using 2004 Nielsen Homescan panel. They also conclude that the typical organic milk consumer is white, well-educated and living in a household headed by someone younger than 50 years old, which is not different from the description on general organic consumers studied by others. Alviola and Capps (2008) provide a more formal statistical analysis with the same data as Dimitri et al. implementing Heckman two-step procedure. Their conclusion largely agrees with Dimitri et al.

However, there are very few studies on the welfare effect of the introduction of organic products. Dhar and Foltz (2005) estimate a demand system for recombinant bovine growth hormone (rBST) free labeled milk, organic labeled milk, and unlabeled (conventional) milk utilizing the quadratic AIDS framework and full information maximum likelihood estimation techniques. The data used for the analysis consist of weekly sales and prices from 1997 to 2002 in twelve cities over the U.S. They find that organic milk and rBST-free milk are complements to each other, while unlabeled milk is a substitute for both rBST-free and organic milk. In addition, the amount of consumers' benefit from the specialty milks is analyzed through the competitive effect (CE) and variety effect (VE), where the former implies the amount of price reduction by existing competitors after the introduction of a new competitor and the latter is the willingness-to-pay changes for having more options to choose. They claim that consumers will benefit 2 cents per gallon by the price reduction in unlabeled milk (CE) and 17 cents from the option of having rBST-free and organic milk. However, this research has some limitations. First, the data and categorization used in the analysis is likely outdated because of the rapidly changing organic milk market. As mentioned above, the National Organic Standards, as well as many media reports on health issues, contributed to changing consumers' perception of organic milk.

Hence, it is very likely that consumption behavioral patterns on organic and conventional milk changed after this policy. This idea can be supported by the report that more consumers have bought the higher premium especially after 2002 (Dimitri et al, 2005). In addition, consumers' concern on health risk changed the structure of milk market. The nation's largest diary process, Dean Foods, no longer sells rBST treated milk, and the top 3 grocery retailers, Wal-Mart, Kroger, and Costco, claimed not to sell such milk in their stores. Therefore, rBST treated milk is not in the product space for most of consumers after the early 2000s; hence, the categorization used in their study will not be valid any longer. Plus, rBST-free and organic variables are not from the data, created by authors. The authors state that the rBST-free and organic labels are recognized after interviewing with manufacturers. Second, the model does not control other relevant factors such as fat contents or flavor. Based on the fact that consumers' preference on fat contents has been changed, it is possible that this model over states the welfare effect of organic milk. Finally, the demand elasticities estimated are conditional on the expenditure of milk.

In the early stage of introduction of organic milk, there existed only two organic milk suppliers in the nation, Organic Valley and Horizon Organic, and they were available in some limited areas. Thus, the structure of competition in milk market is rather between organic milk and non organic milk. As the organic milk market has expanded, however, the competitive structure has also changed to the extent that there are several national and local brand organic milk plus private-labeled organic milk carried by conventional supermarkets. Therefore, in order to establish an appropriate analysis of current milk market, the competition at the brand level rather than commodity group level should be taken account into the model.

In this light, the objective of this study is to analyze the demand for organic and conventional milk at both brand level and commodity group level. A brand level milk product is defined by its fat contents, organic claim, flavor and the name of the supplier. A group commodity milk is defined by its fat contents, organic claim and flavor. Detailed explanations on categorization will be provided in the data section. AC Nielsen Homescan data from 2004 to 2005 is used for the study[3]. The multi-stage demand approach is used to estimate the demand for milk products at the brand level following Hausman (1997). The Linear Approximate AIDS (LA/AIDS) model is adopted for the functional form of demand equations. Unconditional (on the expenditure) elasticities among brands and group commodities are estimated using the methodology suggested by Carpentier et al.(2001). In addition, the welfare effect of the introduction of organic milk will be analyzed at both the brand and commodity group level.

Previous studies on brand level demand analysis are reviewed in section II and descriptive statistics from Nielsen Homescan data are presented in section III. Section IV explains model specification and estimation techniques for demand analysis and the results are shown in section V, and the welfare analysis is provided in section VI.

II. Demand Estimation for Differential Goods

1) Classical Approach

Researchers have been interested in developing methodologies to estimate demands for differentiated goods as disaggregated data and advanced computational devices have become available. For example, Hausman et al. (1994) proposed a multi-stage demand system with an application to beer market and Berry et al. (1995) introduced mixed logit approach with an application to automobile demand. Although the applications to disaggregated data are recent innovations, their basic ideas are from the existing demand approaches.

[3] AC Nielsen established organic variable since 2002, but the data before 2004 imply that organic cow milk is not introduced or the consumer perceptions of organic products are lacking in the market this study focuses on. The recorded organic purchases are occurred in soy milk category. Therefore, the data from 2004 to 2005 are used for demand estimation and price values before 2004 are used to calculate the virtual price.

Since Stone (1954) derived the very first demand system, Linear Expenditure System (LES), by imposing theoretical restrictions on a simple linear demand system, researchers have developed various functional forms of demand equations to reflect the reality better so that be able to test whether the theoretical restrictions are true. Theil (1965) and Barton (1966) derived Rotterdam model (RM) by substituting Slutsky decomposition into a differentiated double log demand function. Homogeneity, symmetry and negativity are tested, but Barton finds that the empirical results of RM are consistent with theory only with the application to highly aggregated data while disaggregated applications conflict with theory. Different approaches that give more functional flexibility to the model were suggested by a great number of researchers in order to find a model consistent with theory. The basic idea of flexible functional forms is to approximate direct utility function, indirect utility function or cost function by some specific functional forms and give it enough parameters so that it is flexible enough to approximate an arbitrary utility or cost function. The demand functions are derived through duality. Christensen, Jorgenson, and Lau (1975)'s translog model approximates the indirect utility function by a quadratic function of logs of normalized prices, and derive Marshallian demand by Roy's identity. The authors also test theoretical restrictions, but homogeneity does not hold for this model. Another famous approach known as Almost Ideal Demand System (AIDS) is introduced by Deaton and Muellbauer (1980). The model is derived from the cost function of generalized Gorman polar form using Shephard's lemma. The test with this model is also inconsistent with theoretical restrictions.

As summarized above, a list of conventional demand systems has been developed in the past in efforts to find flexible functional forms that are close to reality, and those proposed models were tested to examine whether the models are consistent with consumer theory. Empirical evidence implies that currently available demand models and data fail to support the theoretical restrictions. Researchers appear to have different interpretation of the results. Some researchers, such as Christensen et al., conclude that the theory of demand does not hold. But most researchers, such as Deaton and Muellbauer, carefully conclude that none of the existing models perfectly define demands and measure elasticies, and estimate the best approximation imposing the theoretical restriction. Although the imposed restrictions reduce the number of parameters to be estimated, its application is very limited depending on the number of equations in the system.

2) Logit Approach

Another approach to estimate demand will be discrete choice models. McFadden (1974) argues that the conventional demand approach assumes all individuals in a population have a common behavior rule. The logit model starts from the indirect utility functions of individuals instead of a “representative” utility, taking account heterogeneity of individual tastes. The indirect utility function consists of common utility and random utility. Based on the revealed preference theory, probability of choice can be presented as an integral of cumulated joint density functions. This probability directly can be interpreted as the share of demand, but it has to be transformed into a closed form of a function for estimation purpose. Luce (1956) derived the probability of choice formula from the conditions satisfies Independence of Irrelevant Alternative (IIA). McFadden (1974) derived the same probability of choice formula under the assumption of Gumbel distribution for the error terms.

A nice feature of logit models is that the tastes that vary systematically with respect to observed variables can be captured while the tastes that vary with unobserved variables cannot be handled. Also, the logit models solve the problem of having large number of parameters which conventional demand system suffers from because the indirect utility functions in discrete choice model are not defined with the prices of all the products in the system. However, the substitutability in logit models is very restrictive. Since logit models exhibit the independence of irrelevant alternatives (IIA) property, it constrains the cross price elasticities. The IIA claims that the ratio of choice probabilities between two goods does not change even if the third irrelevant good is introduced. This might be a plausible assumption in some cases, it is not behaviorally accurate in many cases.

Nested logit models have been used to overcome the limitations of IIA (Ben Akiva 1973, Train et al. 1987). The approach is based on the assumption that consumers make decisions in sequence. Consumers would choose whether to participate in the economic activity in question, then select a specific choice. Within each data step, the IIA assumption holds. However, across different steps, the ratio of probabilities can depend on the attributes of other alternatives in those nests and IIA does not hold. Nested logit approach is still limited in its ability to account for unobserved preferences. To relax the IIA assumption and account for heterogeneity, the mixed logit model is used (Train et al. 1987 and Berry 1995). It is an extension of standard logit that allows the coefficients to vary across individuals by assuming the coefficients have distributions rather than fixed numbers.

3) Multi-stage Demand

Hausman et al.(1994, 2002) apply Gorman's multi-stage budgeting approach into the demand for differentiated products. Strotz (1957) discussed that consumers allocate expenditure among broad groups of commodities in the first stage of budgeting, and then allocate individual commodities within each group if the utility function is separable. Gorman (1959, 1971) developed Strotz's discussion in detail. He argues that 'separability' is not enough to explain the consumer's multi-stage budgeting behavior. He shows that, under the assumption of 'weak separability', consumers allocate their income into broad groups of commodities at higher stage of budgeting and more detailed within-group allocation happens at lower stage. Weakly separable preferences allow the last stage demand functions to be presented only with the group expenditure and the prices of products within that group. However, in order for the higher stage demand functions to be expressed with total expenditure and the price indices of each group, additive separability and Gorman generalized polar form of indirect utility functions, or homothetic preference is required.

Hausman et al. apply this approach to estimate the brand level demand in beer market, whereas Gorman's original approach is conducted at the aggregate level of economy. The basic idea of this model is to let the top level demand corresponds to the overall demand of beer, and the middle level demand corresponds to the demands for different segments of beers. The lowest stage of the demand system corresponds to each brand of beer. The underlying assumption of separability eases the problem of dimensionality where the system of demand equations suffered from its numerous coefficients to estimate. Weak separability is assumed at the lowest level of utility maximization problem so that the demand for each brand can be presented as a function of group expenditure and the prices of own and other brands in the same group. Additive separability is required at the higher stage of utility maximization in order for the higher stage demand to be presented as a function of total beer expenditure and price indices of segments. Although additive separability has nice features which reduces the number of coefficients allowing the demand function to be written with group price and quantity indices instead of commodity prices, it is not a realistic nor plausible assumption. Hausman does not explicitly discuss the assumptions for the higher stage demand, but it seems that he adopts Carpentier and Guyomard's (2001) approximation of first-stage allocation process instead of imposing additive separability. Carpentier states that, if preferences are weakly separable and the group price indices being used do not vary too greatly with the utility level, allocation between groups of commodities by two stage budgeting will be consistent with unconditional demand analysis, thus the first stage demand function with price index can be approximately rationalized without a strong assumption.

III. Data

As mentioned above, the data used in this study are the Nielsen Homescan panel data. The sample is selected among volunteers based on both demographic and geographic targets. Stratification is done by AC Nielsen to ensure that the sample matches the U.S. Census. The panelist members are required to scan the items purchased with handheld scanner and transfer the information to AC Nielsen each week. Thus, the data are recorded on a weekly basis. Unobserved data should be interpreted as infrequency of sales rather than infrequency of records since it is mandatory for the members to transfer data every week. If a member fails to comply with the rule and does not report more than a month, then the panelist membership is terminated.

The nationally representative sample consists of purchase histories of milk products by 49,114 households from 2002 to 2005. 8,866 households participated in 2002, 18,539 households in 2003, 40,327 and 37,338 households in 2004 and 2005 respectively. The sample contains information on demographics such as income, household size, age of head, number of child, employment, education and race. Demographic distributions are presented in <Table1>. Half of the sample is from under $45,000 income class and the other half is from above $45,000 income class. More than half of the households consist of single or two members, and 75 percent of the sample have no children under 18. 72 percent of male or female household heads are employed more than 30 hours a week and 70 percent of them have at least college degree. The shares of organic milk purchase by different demographic characteristics are provided in <Table 2>. Households with small number of members tend to purchase more organic milk than large families, middle income class is less likely to purchase organic milk than low income and high income classes. Also, the data show that the households only with under-6-year-old children are relatively more likely to purchase organic milk than any other households.

A number of physical product characteristics, weekly market level prices and quantities purchased are also included in the data. Important characteristics to differentiate milk products are fat contents, flavor and organic claim[4]. The fat contents are categorized into five types; non-fat, 1% low fat, 2% reduced fat, whole milk and soy&lactose-free milk. Flavor is categorized into flavored and not flavored. <Table 3> provides the market shares of products distinguished by these characteristics each year. 2% reduced fat milk brings the largest share of milk sales up to 35% during the period, and the market shares of 2% milk and whole milk have decreasing trends while the shares of non-fat, low fat and soy&lactose-free milk have moderately increasing trends. The share of organic milk vs. non-organic milk also shows an increasing trend in this sample.

In this study, a product is defined at the brand level with three different characteristics of products; fat contents, organic claim, flavored or not. Many kinds of flavors are consolidated into flavored for simplicity. Different fat contents of a brand are treated as different products, and organic milk and non-organic milk of a same brand are treated as different products as well. Different brands with same fat contents and flavor and the same organic claim are, of course, regarded as different products. But different sizes and different types of containers are not distinguished in the products defined in this study. The commodity groups are aggregated across different brands with the same characteristics. For example, the 2% reduced fat-organic-unflavored group commodity milk is an aggregation of different products within the group of 2% reduced fat-organic-unflavored milk. Hence, there are 20 group commodities with the categorization mentioned above. The quantities of group commodities are the aggregation across brand level products with same characteristics and their prices are the price indices of each group. In terms of time frequency, weekly purchase data are aggregated into monthly records in order to minimize infrequency problem. According to the definition of product above, there exist 1,902 products in the nation. However, it is notable that specific brands of milk appear only in specific areas and only a few brands dominate the local markets while a large number of

[4] Some of variables such as fat contents and flavors are not precisely recorded so that those variables are created from the UPC (Universal Product Code) description.

residuals take only 1~5% of market share. Hence, it is concluded that the brand-level milk market is highly localized and dominated by a few brands so this study needs to focus on some specific market. Raleigh-Durham-Chapel hill and Charlotte markets are chosen and brands with market share larger than 1% are considered.

1,634 households participated in the survey from 2002 to 2005 in RDU (Raleigh-Durham-Chapel hill and Charlotte) area. 103 households participated in 2002, 481 households in 2003, 1440 households and 1319 households in 2004 and 2005 respectively. Among the panelists, 471 households participated for one year and 1004 households participated for two years. There are 80 households and 47 households who participated for three and four years, respectively. The demographics in this area show similar features as the national demographics. However, although AC Nielsen established organic variable since 2002, the data before 2004 imply that organic cow milk is not introduced or the consumer perceptions of organic products are lacking in this market. The organic purchases are occurred only in soy milk category according to the data during 2002 and 2003. Therefore, the data from 2004 to 2005 are used to estimate consumer demand and the welfare effects are analyzed under the assumption that organic cow milk is introduced in this area since 2004. The price values of conventional milk prior to 2004 are used to calculate the virtual prices in the welfare analysis. The shares of each type of milk sales are described in <Table 4>. The figures are similar to the national sample. The organic milk takes about 2.5 percent of the milk market and the 2% reduced fat milk takes the largest share.

There exist 249 products in the area, but only 58 products take more than 97% of the milk market. Hence, only the 58 products are included in this study. The products can be categorized into 20 groups according to the characteristics mentioned above, which are fat contents, flavor and organic claim. Market shares and average prices of products in each group, and the number of brands with larger than 1% of market share within each group are shown in <Table 5>. Conventional non-flavored non-organic milk dominates the market with 92% market share. Soy and lactose free milks are priced higher than cow milk among non organic milk. Organic cow milk has higher per unit prices than conventional cow milk as expected, but soy and lactose free milk are not priced differently between organic and non organic.

IV. Model

1) Multi Stage Demand System

Hausman's three stage demand systems approach is adopted to estimate the demands of milk. The first stage demand is defined as total demand of milk; the second stage is defined as demands for group commodities; the third (lower) stage estimates brand level demands within groups. It is assumed that the direct utility function is weakly separable into sub-utilities and the current weighted true cost of living price indices for each groups vary only slightly with corresponding sub-utility levels so that the empirical variation of price index with sub-utilities can be neglected. The latter assumption allows to avoid strong assumptions, such as strong separability or homothetic preference, in the upper stage of demand system (Carpentier and Guyomard, 2001). The econometric functional form of brand level demand equation is specified as Linear Approximate Almost Ideal Demand System (LA/AIDS):

$$(1) \quad w_{it} = \alpha_{it} + \sum_{j=1}^{n} \gamma_{ij} \ln(p_{jt}) + \beta_i \ln\left(\frac{m_t^G}{P_t^G}\right) + \varepsilon_{it}, i, j \in G, i, j = 1, 2, ..., n$$

where $i = 1, 2, ..., n$ denotes the brands of milk in group G and t denotes time period. p_{jt} are the price of product j consumers face in time period t. m_t^G are total group expenditure on group G in period t, that is, $m_t^G = \sum_{i=1}^{n} p_{it} q_{it}$ and ε_{it} is an error term. P_t^G is the Linear Approximate AIDS price index of brands in group G period t.

In order to estimate the group commodity demand in the second stage, Stone Index is computed for the price indices of each segment using mean values of market shares of each brand. LA/AIDS is used to specify the middle level equation. (Hausman states in his paper that the difference in functional form does not make difference in outcomes.)

$$(2) \qquad q_{mt} = \beta_m \log \frac{y_{Bt}}{P_t^B} + \sum_{k=1}^{M} \delta_{ij} \log \pi_{kt} + e_{mt}$$

m = 1, …, M, t = 1, …, T

where q_{mt} is the share of segment m in period t, y_{Bt} is total milk expenditure, and π_{kt} is segment price indices in the period of t.

The first level equation, which explains the overall demand for milk, can be specified as

$$(3) \qquad \log u_t = \beta_0 + \beta_1 \log y_t + \beta_2 \log \prod_t + Z_t \delta + e_t$$

where ut is overall consumption of milk, yt is disposable income, πt is price index for milk, and Zt are the variables that account for time trends.

2) Specification and Estimation

As I mentioned above, data used in this study are micro-level survey data. When it comes to demand analysis using this type of data, one cannot avoid the issue that some products are not consumed by at least some economic agents in some periods. Even though the data used in this study for the lower level of multistage demand equation are not disaggregated as to the household level, the data are still disaggregated to some degree of brand level and indicate zero purchases for some brands in some periods.

Setting aside the difficulties of estimating latent dependent variable models, missing regressor difficulties are first encountered because prices are not observed for non purchased products. Three simple solutions for this problem are 1) to discard all incomplete observations and estimate population parameters using the remaining observations, 2) to use zero-order methods which substitute sample means for the missing values, and 3) to use first-order methods which substitute predicted values from simple regression for the missing values. However, these methods are criticized because of sample selection bias. Many researchers suggest various missing value procedures mostly utilizing demographic or product characteristics. For example, Heckman procedure and Amemiya's principle require both regressands and regressors in demand systems to be endogenous so that the variability of regressors can be explained with other exogenous variables. However, in multi-stage demand approach, it is impossible to incorporate quality adjusting price equations because the assumption of separability does not allow volatilities in the exogenous variables that explain price variation, such as characteristics of products. Therefore, a simple regression method seems to be the only feasible approach to treat the missing price problems. The unobserved unit prices are predicted following Perali and Chavas (2000). The unit prices at UPC level were regressed on characteristics variables, time variables, regional dummies, and interaction terms between characteristics and time variables. The least square results show 0.54 of R-square, but statistically significant coefficients.

Another issue with regard to using micro-level purchasing data is to take the zero purchasing behavior of consumers into account in analysis. This indicates corner solution outcomes of consumer utility maximization problem, which are rational decisions of economic agents. Thus, a Tobit model is suggested to explain the corner solutions.

$$w_{it} = \left\{ \begin{array}{ll} w_{it}^{*} & if \;\; w_{it}^{*} > 0 \\ 0 & if \;\; w_{it}^{*} \leq 0 \end{array} \right\}$$

Assuming random utility hypothesis (RUH) and PIGLOG class utility function, the Marshallian uncompensated demand functions at the household level can be specified as follows:

(4) $$w_{iht}^{*} = \alpha_{iht} + \sum_{j=1}^{n} \gamma_{ij} \ln\left(p_{jht}\right) + \beta_i \ln\left(\frac{m_{ht}^{G}}{P_{ht}^{G}}\right) + \tilde{\varepsilon}_{iht}, i, j \in G, i, j = 1, 2, ..., n$$

where $i = 1, 2, ..., n$ denotes the i's milk product in the demand system, h denotes the household, t denotes the time period. p_{jht} is the price of product j household h faces in time period t. m_{ht} is household h's total group expenditure on milk products in period t, that is, $m_{ht}^{G} = \sum_{i=1}^{n} p_{iht} q_{iht}$ and P_{ht}^{G} is the Linear Approximate AIDS price index for household h in period t. $\tilde{\varepsilon}_{iht}$ is an error term that is heteroscedastic within the share equation for one good and correlated across the share equations for different goods. $\tilde{\varepsilon}_{iht} = \varepsilon_{iht} - \beta_i \sum_j \ln p_{jht} \varepsilon_{jht}$. ε_{jht} is mean zero homoskedastic error term from utility function.

(5) $$\log P_{ht}^{G} = \sum_{i=1}^{n} w_{ih}^{0} \log(p_{iht}) \text{ where } w_{ih}^{0} = \frac{1}{T} \sum_{t=1}^{T} w_{iht}.$$

Household heterogeneity α_{iht} might be specified as

(6) $$\alpha_{iht} = \rho_{i0} + \sum_{k=1}^{s} \rho_{ik} d_{kht} + \rho_{i(s+1)} t + \rho_{i(s+2)} t^2 + c_h$$

Since I have aggregate data, however, the demand function above should be aggregated over households. Aggregating (4) over household yields

(7) $$w_{it}^{*} = \alpha_{it} + \sum_{j=1}^{n} \gamma_{ij} \ln\left(p_{jt}\right) + \beta_i \ln\left(\frac{m_t^{G}}{P_t^{G}}\right) + \tilde{\varepsilon}_{it}, i, j \in G, i, j = 1, 2, ..., n$$

where

$$w_{it}^{*} = \frac{\sum_h m_{ht} w_{iht}}{\sum_h m_{ht}}, \qquad \tilde{\varepsilon}_{it} = \frac{\sum_h m_{ht}^{G} \tilde{\varepsilon}_{iht}}{\sum_h m_{ht}^{G}}, \qquad \alpha_{it} = \rho_{i0} + \rho_{i1} t + \rho_{i2} t^2$$

Therefore, heteroskedastic Tobit model with the two-step estimation approach is adopted for the lower stage demand estimation.

The first and the second stage demand do not require Tobit approach because the aggregated data used in the higher stage do not show the evidence of corner solution outcomes. However, the error terms might not be homoskedastic any longer. Based on the assumption of Random Utility Hypothesis, disturbances of uncompensated demand functions will be heteroskedastic according to the same logic provided above. Hence, the conventional demand systems given in equation (2) and (3) with SUR approach are adopted for the higher stage demand estimation.

Two step estimation

Estimating censored demand system is not an easy task because it involves multiple probability integrals. In the early applications (Wales and Woodland (1983), Lee and Pitt (1986, 1987)), researchers were only able to analyze small systems by taking multiple integrals. Because of recent development in simulation techniques, researchers can numerically evaluate multiple probability integrals and some alternative methods with large system applications are suggested. An application of the simulated maximum likelihood (SML) approach is seen in Kao, Lee and Pitt, and the quasi maximum likelihood (QML) approach which approximates the multivariate likelihood function with a sequence of bivariate function can be seen in Yen, Lin and Smallwood (2003). An alternative that does not involves complicated computational tasks, which is known as two-step estimation, is proposed by Perali and Chavas (2000) and later extended to the panel data framework by Meyerhoefer, Ranney and Sahn (2005). This study adopts Meyerhoefer's two-step estimation because the approach is generalized to the application of panel data while others are applied only with cross-section data and its computational procedure is relatively simple comparing to other approaches.

The basic idea of the two-stage procedure is to estimate an unrestricted heteroskedastic Tobit model equation by equation and find the error correlations, and then recover restricted parameters using the minimum distance method which falls into the GMM framework.

In the first step, the share equation for ith product (7) can be rewritten as follow

$$(8) \qquad w_{it}^{*} = \rho_{i0} + \sum_{j=1}^{n} \gamma_{ij} \ln\left(p_{jt}\right) + \beta_i \ln\left(\frac{m_t^G}{P_t^G}\right) + \tilde{\varepsilon}_{it}, i, j \in G, i, j = 1, 2, ..., n$$

$$\text{where } \tilde{\varepsilon}_{it} = \frac{\sum_h m_{ht}^G \tilde{\varepsilon}_{iht}}{\sum_h m_{ht}^G}, \qquad \tilde{\varepsilon}_{iht} = \varepsilon_{iht} - \beta_i \sum_j \ln p_{jht} \varepsilon_{jht}.$$

ρ_{i0} is product specific fixed effect. As $\tilde{\varepsilon}_{iht}$ is heteroscedastic within each equation and correlated across equations, so does $\tilde{\varepsilon}_{it}$. To get consistent first-step estimates, a heteroscedastic Tobit econometric model is employed for each equation. The variance of the error term is specified using a fairly flexible and general form

$$(9) \qquad Var(\tilde{\varepsilon}_{it}) = \sigma_i^2 \exp\left(z_{it}^{'} \xi_i\right)$$

where z_{it} is a s_z-dimensional vector of variables for product i in period t. Variables of t, t^2, log p_{it} and $\log \frac{m_t^G}{P_t^G}$ were included in z_{it} at first, but only $\log \frac{m_t^G}{P_t^G}$ is included because the other variables seem to hamper the optimization procedure without improving the goodness of fit of the model. As the estimation is conducted for each share equation separately without imposing cross-equation parameter restrictions implied by demand theory, the estimates I obtain from this step are reduced form estimates. In order to recover restricted estimates, reduced form parameter estimates are collected in the vector $\hat{\pi} = \left(\hat{\pi}_1^{'}, ..., \hat{\pi}_n^{'}\right)^{'}$, where $\hat{\pi}_i$ is a $\left(n + s_z + 2\right) \times 1$ vector of reduced form parameter estimates from the i's equation. In the second step of estimation, the cross equation restrictions implied from demand theory are imposed on the reduced form parameters estimated in the first step, and the structural parameters that are consistent with demand theory are calculated. Denote a q-dimensional vector of structural parameters as ψ, then the structural parameters are obtained from the following GMM estimation procedure

$$(10) \qquad \min_{\psi}\left[\hat{\pi}-h(\psi)\right]'\hat{\Omega}^{-1}\left[\hat{\pi}-h(\psi)\right]$$

where $h(\psi)$ is a nonlinear mapping ψ into π that is used to impose the theoretical restrictions on the reduced form parameters. The number of restrictions imposed is $(n+s_z+2)\times n-q$, which is equal to (n-1)*n/2+n+2. Under the null hypothesis that these restrictions are correct, the minimized value of objective function (10) is a chi-square distributed random variable with degree of freedom equals to the number of observation minus the number of restrictions.

The difficulty arises in finding a consistent estimate of Ω. Meyerhoefer et al. (2005) states that the covariance-variance matrix for $\hat{\pi}$ takes the form $\Omega = D_1^{-1}D_2D_1^{-1}$ and the proof is provided in his unpublished dissertation (2002). If $g_t = (g'_{1t},...,g'_{nt})'$ denotes the vector of univariate scores from all of the n equations corresponding to the observation in period t, and H_{it} the univariate Hessian from the i 's equation for the same observation, then $D_1^{-1} = diag\{E(H_{1t})^{-1},...,E(H_{nt})^{-1}\}$ and $D_2 = E(g_t g'_t)$. A consistent estimator for Ω can be obtained by replacing the population moments by their sample counterparts. However, this might not work for this study because the data used in this study do not meet with the condition for large sample theory. These are four years' monthly data so that the number of observations for each brand is at most 48. The data for specific types of milk such as organic milk are established recently, thus very short strings of data are available for special types of milk.

The finite sample properties of GMM estimator seem to be an interesting topic among the econometricians in mid 90s. The July 1996 issue of Journal of Business and Economic Statistics is full of papers on the small sample properties of GMM estimator proposing alternatives for consistent estimator of weighting matrix. Although they are looking at slightly different issues of small sample properties, their conclusions converge to one that the equally weighted matrix, which is equivalent to identity matrix, dominates covariance matrix (or the proposed matrix) in terms of the bias of estimator and over identification test statistics. (i.e. Monte Carlo studies show that the estimates are biased downward and Wald test statistics often exceed asymptotic size which means relying on asymptotic distribution theory leads one to reject the null hypothesis too often.) Therefore, the identity matrix is used in this study.

Elasticities

The unconditional expectation for the budget shares including all the observations is

$$(11) \qquad E(w_{it}) = \Phi_i(\bullet)\times\left[\rho_{i0} + \sum_{j=1}^{n}\gamma_{ij}\ln\left(p_{jt}\right) + \beta_i \ln\left(\frac{m_i^G}{P_i^G}\right)\right] + \sigma_i^2 \exp\left(z_{it}^{'}\xi_i\right)\times\phi_i(\bullet)$$

$$\text{where} \quad \bullet = \frac{\left[\rho_{i0} + \sum_{j=1}^{n}\gamma_{ij}\ln\left(p_{jt}\right) + \beta_i \ln\left(\frac{m_i^G}{P_i^G}\right)\right]}{\sqrt{\sigma_i^2 \exp\left(z_t^{'}\xi_i\right)}}$$

The uncompensated own price, cross price and expenditure elasticities that are conditional on the group expenditure but unconditional on whether the observed budget share is zero or positive can be derived as

(12) $$e_{iit} = -1 + \frac{\partial E(w_{it})}{\partial p_{it}} \frac{p_{it}}{E(w_{it})} = -1 + \frac{\Phi_i(\bullet) \times \left[\gamma_{ii} - \beta_i w_i^0\right] - \frac{1}{2}\left[\phi_i(\bullet) \times \sqrt{\sigma_i^2 \exp\left(z_{it}^{'}\xi_i\right)} \times \xi_i \times w_i^0\right]}{E(w_{it})}$$

(13) $$e_{ijt} = \frac{\partial E(w_{it})}{\partial p_{it}} \frac{p_{it}}{E(w_{it})} = \frac{\Phi_i(\bullet) \times \left[\gamma_{ij} - \beta_i w_j^0\right] - \frac{1}{2}\left[\phi_i(\bullet) \times \sqrt{\sigma_i^2 \exp\left(z_{ht}^{'}\xi_i\right)} \times \xi_i \times w_j^0\right]}{E(w_{it})}$$

$$\text{where} \quad E_{it} = 1 + \frac{\left[\beta_i \times \Phi_i(\bullet)\right] + \frac{1}{2}\left[\phi_i(\bullet)\xi_i \sqrt{\sigma_i^2 \exp\left(z_{ht}^{'}\xi_i\right)}\right]}{E(w_{it})}$$

The mean of elasticities over time are provided in the results section.

$$e_{ij} = \frac{1}{T}\sum_{t=1}^{T} e_{ijt}$$

Unconditional (on expenditure) elasticities are computed following Carpentier and Guyomard (2001). The relationships between second-stage (i.e., conditional) and first-stage (i.e., unconditional) expenditure and price elasticities are established under the assumptions of weakly separable direct utility function and the approximate independence of the true cost of living indices with respect to sub-utility levels. Carpentier and Guyomard provide formulas with two-stage budgeting application, but the results are generalized to the three-stage budgeting application following Edgerton (1997).

V. Welfare Analysis

The total effect on consumers' welfare can be evaluated through compensating variation, which is the difference in consumers' expenditure function before and after the introduction of organic products holding utility constant at the post-introduction level:

(14) $$CV = e(p_1, p_N, r, u_1) - e(p_0, p_N^*(p_0), r, u_1)$$

where p_1 is the vector of post-introduction prices of existing products, p_N is the post-introduction price of the new product, r is a vector of prices of outside industry, and u_1 is the post-introduction utility level.

The function $p_N^*(p)$ defines the 'virtual' price for the new product, which is the reservation price where demand for the new product would be zero given the prices of other products. Following Hausman and Leonard (2002), this total benefit can be broken into two parts:

(15) $$CV = [e(p_1, p_N, r, u_1) - e(p_1, p_N^*(p_1), r, u_1)] + [e(p_1, p_N^*(p_1), r, u_1) - e(p_0, p_N^*(p_0), r, u_1)]$$

and written as $\mathrm{CV} = -(\mathrm{VE} + \mathrm{PE})$. The first term, variety effect (VE), represents the increase in consumer welfare due to the availability of the new products, holding the existing product prices constant at the post-introduction level. This effect not only captures the benefits from having more options but also the benefits from the new characteristics of new products. The second term, price effect (PE), represents the change of consumer welfare due to the change in the prices of existing products. The introduction of new products can lead the prices of existing products to increase or decrease depending on the competitive structure of the industry. If the new products closely compete with the existing products produced by the same manufacturer, the prices of existing products may rise. However, if the products compete closely with the existing products from different manufacturers, the prices of existing products are likely to decrease.

Price changes from the organic milk introduction are predicted at the group commodity level and the whole benefits in dollar values are measured at the aggregate level of milk industry using the price index of whole market. First, virtual prices can be evaluated at the group commodity level by solving a system of second level demand equations that would set the organic group commodities' shares to zero. There are 6 organic commodity groups among the total 16 commodity groups in this application. Given the virtual prices, variety effect can be calculated as follows:

(16) $$VE = [e(p_1, p_N, r, u_1) - e(p_1, p_N^*(p_1), r, u_1)]$$

Hausman (1981) derives straightforward expressions for expenditure functions from uncompensated demand estimates in some special cases such as linear demand or log linear demand equations. First, he simplifies Roy's identity using the implicit function theorem. The simplified expression can be integrated out and the indirect utility function can be obtained in a closed form by solving an ordinary differential equation. Finally, the corresponding expenditure function is obtained through inverting the indirect utility function. An explicit expression for the variety effect can be derived in case of a double log demand equation at the top level equation:

$$(17) \qquad VE = [\frac{1-\beta_1}{(1+\beta_2)y_1^{\beta 1}}\left(P\left(p_1, p_N^*(p_1)\right) exp\left(\delta_0 + \beta_2 lnP\left(p_1, p_N^*(p_1)\right)\right) - X_1\right) + y_1^{1-\beta_1}]^{\frac{1}{1-\beta 1}} - y_1$$

where p_1 and p_N^* represent the post-introduction price indices for non organic commodity groups and the virtual price indices for organic commodity groups. Function P(·) defines the virtual price index for the milk industry in this region. β_1 is the coefficient on log personal disposable income, β_2 is the coefficient on the milk price index from the top level equation, and δ_0 captures the remainder of the variables in the top level equation. y_1 is post introduction personal disposable income and X_1 is actual milk expenditure.

Hausman provides two different methodologies to estimate price effects. First, one can estimate the price effect directly from the data using OLS methodology if both the pre- and post-introduction consumption data for the existing goods are available. Second, one can estimate the price effects indirectly by solving the equilibrium conditions for the assumed model of competition for the post-introduction world. The price effects in this study are estimated with the direct estimation method because both the pre- and post-introduction data are available and the interests of this study are on measuring the realized benefits without imposing any restrictions. The type of competition in this market is not known and assuming certain type of competition restricts the estimates. Therefore, the price effects are estimated by estimating the following equation:

(18) $$logp_{it} = \alpha_i + W_t + I_{it}\delta + \varepsilon_{it}$$

The dependant variable is the log price of the existing conventional milk of type i in time t. The variables α_i and W_t are fixed effects for group i and time t. The regression is conducted separately for log of price of each group and the variable α_i is defined as intercept. W_t are defined with dummy variables for each month. There are 48 months in the data and 47 dummies are included in the regression. I_{it} is a post-introduction indicator. I_{it} equal one if the organic milk is introduced. The coefficient δ measures the amount of price change of existing milk after the organic milk introduction[5].

The overall effect of the organic milk introduction on consumer welfare is the sum of the variety effect and the price effect. Hausman (1981) derived the Compensating Variation in the same way he derived the Variety Effect:

(19) $$CV = [\frac{1-\beta_1}{(1+\beta_2)y_1^{\beta_1}}\left(P(p_0, p_N^*)\, exp\left(\delta_0 + \beta_2 lnP(p_0, p_N^*)\right) - X_1\right) + y_1^{1-\beta_1}]^{\frac{1}{1-\beta_1}} - y_1$$

where $\mathrm{P}(\mathrm{p}_0, \mathrm{p}_\mathrm{N}^*)$ is the milk price index evaluated at the pre-introduction prices for the existing (conventional) milks and the virtual prices for organic types of milk.

VI. Results

Elasticity Estimates

I applied the econometric approach outlined above to the A.C. Nielsen Homescan data to estimate the system of milk demand equations. The estimates of equation (3), top level demand function, directly

[5] The indicator variable equals one after 2004 because consumers in this region began to purchase organic milk since 2004 according to the data. Although organic milk was available nationwide before 2004 and surely in this region, consumers' perception on organic milk seems to start from 2004 in this region. Thus, we assume the starting point of organic milk introduction as January 2004.

give the own price elasticity and the income elasticity, which are -0.2 and 0.88, respectively, in the RDU market. The milk price index in this market is calculated with the given data, the regional disposable income is indirectly obtained from Bureau of Economic Analysis (BEA) and Bureau of Labor Statistics (BLS).

Elasticity estimates for the second stage demand system are provided from <Table 6> to <Table 7>. The second stage demand equations are estimated both with and without the variables that account for time trend. The results partly conflict, but overall implications are not different between two models. Thus, the results with time trends are discussed in this section because the model shows better fits. The value of minimization objective function is smaller and the number of significant estimates at 10% level is larger for the model with time trends. <Table 6> and <Table 7> show conditional and unconditional elasticity estimates, respectively. The elasticities are estimated at the mean of variables. Statistical significances are tested for the conditional elasticity and we find 108 estimates out of 272 are statistically significant at 10% level. All types of milk, except the organic-flavored soy/lactose free milk, show negative own price elasticity. Although the organic-flavored soy/lactose free milk has positive own price elasticity, it cannot be considered as Giffen goods because the estimate is not statistically significant. This might be caused by the imposition of homogeneity by which the coefficients of this group of commodity are computed.

Cross price elasticities do not show a general pattern, but some implications can be drawn from the results. First, cross price elasticities between organic and conventional milk with same fat contents and flavor do not show evidence of substitution patterns between organic and conventional milk. 1% fat unflavored organic and conventional milk have positive cross price elasticities (17.73 and 0.22) implying they are substitutes for each other, whereas organic and conventional unflavored whole milk have negative cross price elasticities (-9.79 and -0.11) suggesting that they are complements to each other. 2% fat milk and soy/lactose-free milk also have negative cross price elasticities between organic and conventional although their conditional elasticities are not significant. It is notable that the magnitude of substitution is not symmetric implying that the amount of organic milk consumption change when the conventional milk price changes is larger than the amount of conventional milk change when the organic milk price changes. Second, cross price elasticities show possible substitution patterns between fat contents although it is hard to conclude that similar fat contents are always substitutes with one another. Within the group of conventional unflavored milk, cross price elasticities show pretty clear substitutability between similar fat contents. Fat free and low fat milk have significant positive cross price elasticities and reduced fat and whole milk also have positive cross price elasticities suggesting that they are substitutes. Low fat and reduced fat milk also show substitutability although their cross price elasticities are not significant and their magnitudes are very small. The results also imply that soy/lactose free milk is substitutable with fat free milk while it is not substitutable with other types of cow milk. Within the group of organic unflavored milk, however, elasticities imply that low fat, reduced fat and whole milk are substitutes for one another although they are not statistically significant according to the conditional elasticity esimates. Soy/lactose free milk and 2% fat milk are substitutes with each other within this group. Cross price elasticities within the group of flavored conventional milk show the similar substitution pattern in the unflavored organic milk group. Therefore, based on the analysis above, this study carefully concludes that organic milk and conventional milk are neither substitute nor complements to each other, but this study rather concludes that milk with similar fat contents are more substitutable to each other. Further, substitution patterns between conventional and organic milk found in the previous aggregate level studies might be driven by the substitutability among fat contents rather than by organic factor.

Elasticity estimates at the brand level are partly provided in <Table 8>. Most of brand milks are cross price elastic to private labeled milk, but not vice versa. In other words, the demand for brand milks varies a lot as the price of private label milk changes, but the demand for private labeled milk is not affected a lot by branded milk prices.

Variety Effects

Virtual prices of organic milk products are computed solving a system of equations by setting quantities of organic milk to zero as described in section V and the results are summarized in <Table 9>. The changes in the virtual prices of each type of organic milk are not consistent with the own price elasticity estimates because the virtual prices resulted from simultaneous effect of organic milk introduction. Virtual price index of milk industry is computed at the original equilibrium. In other words, the index is computed with the mean values of market shares from post-introduction data. Although the differences between actual prices and virtual prices of each type of milk are large, the difference in the milk market index is small because the market share of organic milk is as small as 2.5% on average.

The variety effect is calculated with the formula given in section V which takes the curvatures of demand equations into account. The results imply that consumers in this data set obtain 838.8 dollars per month in total by having organic options in their choice sets and this is 8.2 percent of the milk expenditure. The benefit a representative consumer receives whenever he/she purchases a gallon of milk is 31 cents.

Price Effects

The coefficient estimates of equation (11) are presented in <Table 10>. The estimates show that the majority of prices of existing conventional milks have increased after the introduction of organic milk even though the magnitudes of increments are very small, ranging from 0.05% to 0.17%.

However, the prices have decreased in the cases of 2% and whole fat unflavored milk and soy/lactose free flavored milk. It is hard to draw a generalized conclusion on the price effect of the competition structure of this market from the results because 1) the estimated price effects are very small and the standard errors are relatively large; 2) the directions of price effects conflict with each other; and 3) this model does not capture the supply shock that occurred in 2004. There was a price increase of conventional milk in the middle of 2004 due to the shortage of raw milk production in the U.S., but this model cannot capture the supply shock because the periods of two events coincide. However, ignoring some of the imperfect aspects of the model, this study carefully reaches two possible conclusions on the competition effect, equivalently the price effect. First, competition is structured between the new and the existing products from the same manufacturers. The positive estimates of competitive effects imply cannibalization which is the case that the new product is manufactured by the incumbents so that the prices of existing products are increased. This is also supported by the fact that conventional supermarkets began manufacturing organic milk with their own labels around this time and one of two leading organic milk distributors, Horizon Organic, merged with a conventional dairy distributor, Dean Food, in 2004. Therefore, the coefficient of indicator valued one after 2004 might represent the competition between conventional and organic milk within the same manufacturers. Second, some of the negative price effects for the milk with higher fat contents can be explained by the demand shocks caused by consumers' health concerns on fat contents. Health related concerns on fat contents would cause positive demand shock in lower fat milk market and negative shock in higher fat milk market. The positive shock in lower fat market would lead to the price increase in the lower fat market and the negative shock would result in a lower price in the higher fat milk market.

Finally, the overall effect on consumer welfare of the organic milk introduction is calculated with the formula shown in section V. The compensating variation estimated with the data is 707 dollars per month, which accounts for about 6 percent of the milk expenditure. This number is smaller than the variety effects because the price effect is negative. Therefore, it can be concluded that consumers in this region are willing to spend as much as 8 percent of the overall milk expenditure to have additional choices with different characteristics, i.e. organic milk. However, this benefit is depreciated by competition effects. In other words, the organic milk introduction leads the prices of existing products to rise due to the competition structure of this market, thus the overall benefit will be less than the variety effect but still positive since the negative price effect is smaller than the positive variety effect.

References

Berry, S., Levinsohn, J, and Pakes, A., 1995, "Automobile Prices in Market Equilibrium," Econometrica.

Deaton, A., Muellbauer, J., 1980, "Economics and Consumer Behavior," Chaper 5 & 6, Cambridge University Press.

Dhar, T., and Foltz, J., 2005, "Milk By No Other Name… Consumer Benefits From Labeled Milk," American Journal of Agricultural Economics.

Dimitri, C., and Greene, C., 2002, "Recent Growth Patterns in the U.S. Organic Foods Market," Agricultural Information Bulletin, Economic Research Service/USDA.

Dimitri, C., and Venezia, K., M., 2007, Retail and Consumer Aspects of the Organic Milk Market," Outlook Report, Economic Research Service/USDA.

Garmon, J., S., Huang, C., L., and Lin, B., H., 2007, "Organic Demand: A Profile of Consumers in the Fresh Produce Market," Choices, American Agricultural Economics Association.

Glaser, L. K., and Thompson, G., 2000, "The Demand for Organic and Conventional Milk," Presented at the Western Agricultural Economics Association meeting, Vancouver, British Columbia.

Hausman J., 1981, "Exact Consumer's Surplus and Deadweight Loss," The American Economic Review.

Hausman J., and Newey W., 1995, "Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss," Econometrica.

Hausman, J., Leonard, G., and Zona, J. D., 1994, "The Competitive Analysis with Differentiated Products," Annales D'economie Et De Statistique.

Hausman, J., and Leonard, G., "The Competitive Analysis Using A Flexible Demand Specification."

Hausman, J., Leonard, G., and Zona, J. D., 2002, "The Competitive Effect Of A New Product Introduction: A Case Study," The Journal of Industrial Economics.

Train, K., E., 2003, "Discrete Choice Methods with Simulation," Chapter 3, Cambridge University Press.

<Table 1> National Demographic Distribution

Income	Percent	Size	Percent	Age	Percent	Age Of Child	Percent	Emp	Percent	Edu	Percent
Under $5000	0.82	Single Member	25.35	Under 25 Years	0.48	Under 6 only	3.78	0	37.78	0	30.2
5000-7999	1.26	Two Members	37.99	25-29 Years	2.9	6-12 only	6.32	1	62.22	1	69.8
8000-9999	1.12	Three Members	15.03	30-34 Years	6.33	13-17 only	8.11				
10,000-11,999	1.65	Four Members	13.17	35-39 Years	8.95	Under 6 & 6-12	3.32				
12,000-14,999	2.95	Five Members	5.55	40-44 Years	12.13	Under 6 & 13-17	0.55				
15,000-19,999	5.39	Six Members	1.93	45-49 Years	13.73	6-12 & 13-17	4.3				
20,000-24,999	7.67	Seven Members	0.6	50-54 Years	13.18	Under 6 & 6-12 & 13-17	0.88				
25,000-29,999	6.74	Eight Members	0.26	55-64 Years	21.48	No Children Under 18	72.74				
30,000-34,999	7.89	Nine+ Members	0.13	65+ Years	20.83						
35,000-39,999	6.92										
40,000-44,999	6.83										
45,000-49,999	6.39										
50,000-59,999	11.03										
60,000-69,999	8.86										
70,000-99,999	15.17										
100,000 & Over	9.29										

<Table 2> Organic vs. Non-organic Shares by Demographics

org	Household Size								
	1	2	3	4	5	6	7	8	9
0	95.43	97.05	98.06	97.66	98.29	99.38	99.89	100	100
1	4.57	2.95	1.94	2.34	1.71	0.62	0.11	0	0

org	Household Income															
	3	4	6	8	10	11	13	15	16	17	18	19	21	23	26	27
0	99.39	97.45	95.37	95.42	99.39	97.22	97.38	98.27	99.06	96.93	96.67	97.68	96.38	98.32	97.42	94.45
1	0.61	2.55	4.63	4.58	0.61	2.78	2.62	1.73	0.94	3.07	3.33	2.32	3.62	1.68	2.58	5.55

Org	Age of Children								Employment		Education		Marital Status	
	1	2	3	4	5	6	7	9	0	1	0	1	0	1
0	97.27	98.07	98.98	96.29	84.8	97.81	99.63	97.05	97.24	97.29	98.65	96.56	95.73	97.79
1	2.73	1.93	1.02	3.71	15.2	2.19	0.37	2.95	2.76	2.71	1.35	3.44	4.27	2.21

Org	Age of Head									Race			
	2.5	2.7	3.2	3.7	4.2	4.7	5.2	5.7	6	1	2	3	4
0	99.75	96.36	98.11	97.02	97.23	98.08	96.54	97.01	97.37	97.54	94.74	96.34	99.41
1	0.25	3.64	1.89	2.98	2.77	1.92	3.46	2.99	2.63	2.46	5.26	3.66	0.59

<Table 3> Market Shares by Fat Contents and Organic Claim

	2002	2003	2004	2005
non-fat	23.89	24.03	24.51	27.63
1% low fat	16.92	17.05	18.17	20.06
2% reduced	35.19	34.68	35.80	29.51
whole	21.41	20.78	18.28	19.00
soy & lactose free	2.59	3.46	3.25	3.80
non-organic	98.38	97.75	97.81	97.53
organic	1.62	2.25	2.19	2.47

<Table 4> Market Share by Fat Contents and Organic Claim in RDU

	2004	2005
non-fat	24.7	27.19
1% low fat	15.2	15.38
2% reduced	30.62	28.38
Whole	25.4	24.44
soy & lactose free	4.08	4.61
Non-organic	97.5	97.23
Organic	2.5	2.77

<Table 5> Description of Group in RDU

group	Fat contents	Organic	Flavor	Share	average price per fluid oz	Number of br ands
1	No fat	Nonorganic	No flavor	24.59	0.028274	3
2	1% low fat	Nonorganic	No flavor	14.66	0.028336	3
3	2% reduced	Nonorganic	No flavor	28.59	0.029892	5
4	Whole	Nonorganic	No flavor	24.82	0.031272	7
5	Soy & Lactose free	Nonorganic	No flavor	1.67	0.048348	5
6	No fat	Organic	No flavor	0.18	0.051564	3
7	1% low fat	Organic	No flavor	0.12	0.048173	2
8	2% reduced	Organic	No flavor	0.14	0.052778	3
9	Whole	Organic	No flavor	0.13	0.052442	4
10	Soy & Lactose free	Organic	No flavor	0.84	0.046248	3
11	No fat	Nonorganic	Flavored	0.26	0.050497	2
12	1% low fat	Nonorganic	Flavored	0.3	0.037342	2
13	2% reduced	Nonorganic	Flavored	0.6	0.054402	4
14	Whole	Nonorganic	Flavored	1.21	0.047216	5
15	Soy & Lactose free	Nonorganic	Flavored	0.56	0.041551	4
16	No fat	Organic	Flavored	0	n.a.	0
17	1% low fat	Organic	Flavored	0	n.a.	0
18	2% reduced	Organic	Flavored	0	n.a.	0
19	Whole	Organic	Flavored	0	n.a.	0
20	Soy & Lactose free	Organic	Flavored	1.32	0.045765	3

<Table 6> Conditional Elasticities at the Group Level with Time Trend

Group	expenditure	Fat-free	1% fat	2% fat	whole	Soy lactaid	F-F orgnc	1% orgnc
Fat-free	1.13*	-1.81*	1.52*	-0.50	-0.51	0.57*	0.00	-0.14*
1% fat	0.98*	2.50*	-2.24*	0.66	-1.38*	-0.53*	0.03	0.22*
2% fat	1.11*	-0.41	0.32	-2.71*	1.71*	-0.24	-0.05	-0.19*
Whole	0.75*	-0.41	-0.79*	2.07*	-1.36*	-0.08	-0.09	0.13*
Soy LF	1.03*	4.42*	-2.52*	-2.23	-0.66	-1.43	1.29*	0.50*
F-F orgnc	0.98*	0.04	1.10	-3.74	-6.30	11.61*	-4.08*	-2.67*
1% orgnc	1.16	-18.74*	17.60*	-30.28*	17.24*	8.36*	-4.98*	-7.34*
2% orgnc	1.13	2.15	9.49*	-2.00	-3.45	-4.44*	-0.24	0.93
Whole orgnc	0.60	1.14	-13.97*	19.22*	-9.92*	5.65*	1.27	1.03
Soy LF orgnc	1.14	-0.44	2.41	-2.98	-0.10	-0.95	-0.75*	-0.48
F-F flavor	1.14*	8.08*	0.33	4.69	-8.28*	-3.77*	1.22	1.26*
1% fla	-0.13*	-1.61	3.47	-2.54	0.11	2.55	3.46*	0.57
2% fla	1.10*	-2.15	-1.96	2.90	-0.53	0.89	-0.03	-0.70*
Whole flavor	1.49*	1.44	-4.26*	4.34*	-2.98*	0.42	0.46	1.35*
Soy LF flv	0.43*	-1.93	4.30*	0.04	-2.43	1.09	0.05	0.40
Soy LF org fl	0.78*	-4.60*	0.04	-0.03	5.18*	-0.95	-0.20	0.22

2% orgnc	Whole orgnc	Soy LF orgnc	F-F flavor	1% flavor	2% flavor	Whole flavor	Soy LF flavor	Soy LF org fl
0.03	0.01	-0.03	0.14*	-0.02	-0.10	0.13	-0.07	-0.36*
0.19*	-0.27*	0.24*	0.01	0.05	-0.14	-0.56*	0.23*	0.00
-0.02	0.19*	-0.15	0.07	-0.02	0.11	0.30*	0.00	-0.01
-0.04	-0.11*	0.00	-0.14*	0.00	-0.02	-0.22*	-0.08	0.38*
-0.42*	0.52*	-0.45	-0.52*	0.19	0.31	0.28	0.28	-0.57
-0.20	1.04	-3.22*	1.51	2.32*	-0.10	2.65	0.12	-1.07
1.46	1.58	-3.86	2.91*	0.71	-4.04*	14.35*	1.72	2.16
-1.00*	0.67	4.07*	-1.17	-1.24*	0.82	-4.43*	-1.88	0.58
0.69	-1.91	-0.26	-1.50	0.93	-0.43	0.73	-0.30	-2.98
0.80*	-0.05	-1.11*	1.09*	0.38*	-0.13	0.63	0.53*	0.01
-0.79	-0.99	3.74*	-3.26*	-0.51	0.36	-1.23	-1.79*	-0.20
-1.54*	1.13	2.42*	-0.93	-1.38*	0.78	-4.74*	-0.07	-1.54
0.22	-0.11	-0.17	0.14	0.17	-1.00*	1.60*	-0.21	-0.14
-0.65*	0.10	0.47	-0.27	-0.56*	0.86*	-2.53*	0.42	-0.11
-0.68	-0.10	0.99*	-0.95*	-0.02	-0.28	1.05	-1.78*	-0.19
0.09	-0.46	0.01	-0.05	-0.20	-0.08	-0.10	-0.08	0.43

<Table 7> Unconditional Elasticities with time trend

Group	expenditure	Fat-free	1% fat	2% fat	whole	Soy lactaid	F-F orgnc	1% orgnc
Fat-free	0.99	-1.61	1.65	-0.25	-0.28	0.59	0.00	-0.14
1% fat	0.86	2.68	-2.12	0.87	-1.19	-0.50	0.03	0.22
2% fat	0.98	-0.21	0.45	-2.47	1.93	-0.22	-0.04	-0.19
Whole	0.66	-0.27	-0.70	2.23	-1.21	-0.06	-0.08	0.13
Soy LF	0.91	4.60	-2.40	-2.01	-0.45	-1.41	1.29	0.50
F-F orgnc	0.86	0.21	1.21	-3.53	-6.10	11.63	-4.08	-2.67
1% orgnc	1.02	-18.53	17.73	-30.03	17.47	8.39	-4.97	-7.34
2% orgnc	0.99	2.35	9.62	-1.76	-3.22	-4.41	-0.24	0.93
Whole orgnc	0.53	1.24	-13.90	19.34	-9.79	5.67	1.27	1.03
Soy LF orgnc	1.00	-0.24	2.54	-2.73	0.14	-0.92	-0.75	-0.48
F-F flavor	1.00	8.28	0.46	4.94	-8.05	-3.74	1.22	1.26
1% fla	-0.11	-1.63	3.46	-2.57	0.08	2.54	3.46	0.57
2% fla	0.97	-1.96	-1.83	3.14	-0.31	0.91	-0.03	-0.70
Whole flavor	1.31	1.71	-4.09	4.66	-2.68	0.45	0.47	1.35
Soy LF flv	0.38	-1.85	4.35	0.13	-2.35	1.10	0.06	0.40
Soy LF org fl	0.69	-4.46	0.13	0.14	5.33	-0.94	-0.20	0.22

2% orgnc	Whole orgnc	Soy LF orgnc	F-F flavor	1% flavor	2% flavor	Whole flavor	Soy LF flavor	Soy LF org fl
0.03	0.01	-0.01	0.15	-0.02	-0.09	0.14	-0.06	-0.34
0.19	-0.27	0.25	0.01	0.05	-0.13	-0.54	0.23	0.02
-0.02	0.19	-0.14	0.07	-0.02	0.12	0.32	0.00	0.01
-0.04	-0.11	0.01	-0.14	0.00	-0.01	-0.21	-0.08	0.39
-0.41	0.52	-0.44	-0.52	0.19	0.32	0.29	0.28	-0.55
-0.20	1.05	-3.21	1.52	2.32	-0.09	2.66	0.13	-1.05
1.46	1.58	-3.85	2.91	0.72	-4.03	14.36	1.73	2.18
-1.00	0.67	4.08	-1.16	-1.23	0.83	-4.42	-1.87	0.60
0.69	-1.91	-0.25	-1.50	0.93	-0.42	0.74	-0.29	-2.97
0.80	-0.05	-1.10	1.09	0.38	-0.12	0.65	0.54	0.03
-0.79	-0.99	3.75	-3.26	-0.50	0.36	-1.21	-1.78	-0.18
-1.54	1.13	2.41	-0.93	-1.38	0.78	-4.74	-0.07	-1.54
0.23	-0.11	-0.16	0.15	0.17	-0.99	1.62	-0.21	-0.13
-0.65	0.11	0.48	-0.26	-0.56	0.88	-2.51	0.43	-0.09
-0.68	-0.10	0.99	-0.95	-0.02	-0.28	1.06	-1.78	-0.18
0.09	-0.46	0.02	-0.04	-0.20	-0.07	-0.09	-0.08	0.44

<Table 8> Unconditional Elasticities at the Brand Level

	Income	1_A	1_B	1_C	2_A	2_B	2_C	3_A	3_B	3_D	3_C	3_E
1_A	0.94	-2.02	0.92	-0.14	0.09	1.44	0.03	-0.01	-0.21	0.00	-0.01	0.00
1_B	0.76	0.66	-1.86	0.21	0.07	1.16	0.03	-0.01	-0.17	0.00	-0.01	0.00
1_C	0.99	-0.07	-0.16	-1.07	0.09	1.52	0.04	-0.01	-0.23	0.00	-0.01	0.00
2_A	2.02	0.40	4.23	0.18	-1.24	-3.65	-0.09	0.09	1.82	0.04	0.08	0.02
2_B	0.84	0.17	1.75	0.07	-0.06	-2.13	0.13	0.04	0.75	0.02	0.03	0.01
2_C	0.87	0.17	1.81	0.08	-0.10	-0.92	-1.12	0.04	0.78	0.02	0.03	0.01
3_A	0.98	-0.01	-0.20	-0.01	0.03	0.41	0.01	-1.74	-0.57	-0.03	-0.10	-0.05
3_B	0.98	-0.01	-0.20	-0.01	0.03	0.41	0.01	0.78	-3.98	0.31	0.27	0.13
3_D	1.38	-0.02	-0.28	-0.01	0.04	0.58	0.01	-0.12	-1.84	-1.41	-0.22	0.07
3_C	1.01	-0.01	-0.20	-0.01	0.03	0.43	0.01	0.00	-1.98	0.14	-0.82	-0.94
3_E	0.97	-0.01	-0.19	-0.01	0.03	0.41	0.01	-0.08	-1.27	-0.01	-0.07	-1.05

	Income	4_A	2_B	4_F	4_D	4_G	4_C	4_E	5_B	5_H	5_I	5_J	5_K
4_A	0.64	-1.40	0.44	0.04	0.01	-0.14	-0.12	-0.05	-0.02	-0.04	-0.01	-0.01	-0.01
2_B	0.70	0.37	-2.12	-0.05	0.18	0.01	0.28	-0.03	-0.02	-0.05	-0.01	-0.01	-0.01
4_F	0.66	-0.04	-0.15	-0.68	-0.14	-0.13	-0.07	-0.06	-0.02	-0.04	-0.01	-0.01	-0.01
4_D	0.68	-0.02	-0.56	-0.13	-0.78	0.14	0.05	-0.01	-0.02	-0.04	-0.01	-0.01	-0.01
4_G	0.71	0.03	-2.11	-0.13	0.08	-1.15	-0.18	-0.10	-0.02	-0.05	-0.01	-0.01	-0.01
4_C	0.69	-0.05	-0.09	0.02	-0.04	-0.02	-1.16	-0.02	-0.02	-0.04	-0.01	-0.01	-0.01
4_E	0.66	-0.02	-0.24	0.01	0.00	0.01	-0.03	-1.00	-0.02	-0.04	-0.01	-0.01	-0.01
5_B	2.44	-0.04	-1.04	-0.01	-0.03	-0.05	-0.09	-0.02	-2.52	-2.50	-0.47	-0.94	-0.82
5_H	1.29	-0.02	-0.55	-0.01	-0.01	-0.03	-0.05	-0.01	-0.48	-2.26	-0.36	-0.45	-0.28
5_I	2.45	-0.04	-1.05	-0.01	-0.03	-0.05	-0.09	-0.02	-1.53	-2.51	-1.48	-0.94	-0.82
5_J	2.46	-0.04	-1.05	-0.01	-0.03	-0.05	-0.09	-0.02	-1.54	-2.53	-0.48	-1.95	-1.82
5_K	2.42	-0.04	-1.04	-0.01	-0.03	-0.05	-0.09	-0.02	-1.51	-2.48	-0.47	-0.93	-1.81

*** Numbers represent groups and Alphabets represent brands; Alphabet B represents supermarket labels.

(continue Table 8)

	Income	6_B	6_L	6_M	7_L	7_M	8_B	8_L	8_M	9_B	9_L	9_N	9_M	10_B	10_O	10_M
6_B	0.81	-1.39	-4.11	-0.51	-1.74	-0.71	-0.03	-0.10	-0.07	0.14	0.77	0.10	0.07	-0.17	-2.68	-0.42
6_L	1.61	-3.78	-5.42	-2.83	-3.47	-1.41	-0.06	-0.20	-0.14	0.29	1.55	0.20	0.15	-0.34	-5.35	-0.84
6_M	0.99	-0.56	-4.71	-2.13	-2.14	-0.87	-0.03	-0.12	-0.08	0.18	0.95	0.12	0.09	-0.21	-3.30	-0.52
7_L	1.00	-0.96	-5.92	-0.79	-5.17	-1.85	0.21	0.76	0.53	0.23	1.22	0.16	0.12	-0.21	-3.35	-0.52
7_M	0.99	-0.95	-5.90	-0.78	-4.17	-2.82	0.21	0.75	0.52	0.23	1.22	0.16	0.12	-0.21	-3.34	-0.52
8_B	0.98	-0.05	-0.29	-0.04	0.64	0.26	-0.96	-0.23	0.16	0.10	0.53	0.07	0.05	0.23	3.60	0.56
8_L	1.01	-0.05	-0.29	-0.04	0.66	0.27	0.07	-1.05	-0.08	0.10	0.54	0.07	0.05	0.24	3.70	0.58
8_M	1.10	-0.05	-0.32	-0.04	0.72	0.29	0.14	0.04	-1.33	0.11	0.59	0.08	0.06	0.26	4.04	0.63
9_B	0.61	0.29	1.78	0.24	0.82	0.33	0.12	0.42	0.29	-3.23	-1.28	2.68	2.78	-0.02	-0.26	-0.04
9_L	0.55	0.26	1.60	0.21	0.74	0.30	0.11	0.38	0.26	-0.13	-1.72	-0.15	-0.19	-0.01	-0.23	-0.04
9_N	1.70	0.80	4.97	0.66	2.30	0.93	0.33	1.17	0.81	-0.70	-4.05	-1.63	-0.45	-0.05	-0.71	-0.11
9_M	0.55	0.26	1.62	0.21	0.75	0.30	0.11	0.38	0.26	-0.06	-1.01	0.16	-1.31	-0.01	-0.23	-0.04
10_B	1.27	-0.19	-1.16	-0.15	-0.42	-0.17	0.15	0.54	0.37	-0.01	-0.05	-0.01	0.00	-0.69	-1.01	0.19
10_O	1.06	-0.16	-0.96	-0.13	-0.35	-0.14	0.13	0.45	0.31	-0.01	-0.04	-0.01	0.00	0.05	-1.37	0.05
10_M	1.21	-0.18	-1.10	-0.15	-0.40	-0.16	0.14	0.51	0.36	-0.01	-0.05	-0.01	0.00	-0.80	2.44	-2.50

	Income	11_P	11_Q	12_B	12_D	13_B	13_P	13_R	13_Q	14_B	14_C	14_E	14_S	14_T
11_P	0.11	-0.61	0.37	-0.03	-0.01	0.00	0.01	0.00	0.03	-0.07	-0.03	-0.01	-0.01	-0.01
11_Q	1.01	0.21	-2.49	-0.25	-0.14	0.04	0.05	0.03	0.25	-0.62	-0.28	-0.05	-0.13	-0.07
12_B	-0.07	-0.02	-0.40	-2.47	1.78	0.06	0.07	0.04	0.35	-1.56	-0.70	-0.12	-0.34	-0.19
12_D	-0.13	-0.04	-0.71	0.69	-1.92	0.10	0.12	0.07	0.61	-2.77	-1.24	-0.22	-0.60	-0.33
13_B	0.93	0.00	0.09	0.08	0.05	-0.99	0.14	-0.02	-0.01	0.79	0.35	0.06	0.17	0.09
13_P	0.91	0.00	0.09	0.08	0.04	-0.13	-1.44	0.46	0.17	0.77	0.34	0.06	0.17	0.09
13_R	1.02	0.00	0.10	0.09	0.05	0.01	0.16	-1.04	-0.18	0.86	0.39	0.07	0.19	0.10
13_Q	0.99	0.00	0.10	0.08	0.05	-0.04	-0.04	0.14	-1.07	0.83	0.37	0.07	0.18	0.10
14_B	1.17	-0.01	-0.15	-0.25	-0.14	0.09	0.10	0.06	0.53	-1.82	-0.07	0.06	-0.13	-0.15
14_C	1.30	-0.01	-0.17	-0.27	-0.15	0.10	0.11	0.06	0.59	-0.65	-1.38	-0.07	-0.15	-0.08
14_E	1.30	-0.01	-0.17	-0.27	-0.15	0.10	0.11	0.06	0.59	-0.66	-0.36	-1.07	-0.17	-0.09
14_S	1.45	-0.01	-0.19	-0.31	-0.17	0.11	0.13	0.07	0.66	-0.83	-0.33	-0.07	-1.26	-1.14
14_T	1.32	-0.01	-0.17	-0.28	-0.16	0.10	0.12	0.06	0.60	-0.72	-0.33	-0.06	-0.19	-1.09

<Table 9> Variety Effects

	Unflavored Fat Free	Unflavored Low Fat	Unflavored Reduced Fat	Unflavored Whole	Unflavored Soy/Lacotose Free	Flavored Soy/Lactose Free
Actual Price per floz	0.0509	0.0455	0.0519	0.0531	0.0457	0.0431
Virtual Price per floz	0.627	29.225	11.309	113001.714	0.000	0.745
Virtual Price Index	0.031					
Additional Consumer Welfare (VE, per month in this region)	838.8					
% of Milk Expenditure	8.2					
Benefits per gallon purchase	0.31					

<Table 10> Price Effects

group	post-introduction	intercept	r-squared
fat free unflavored non-organic	0.08 %	-3.74 %	0.30
	(0.20%)	(0.14%)	
1% fat unflavored non-organic	0.05 %	-3.75 %	0.23
	(0.14%)	(0.10%)	
2% fat unflavored non-organic	-0.17 %	-3.47 %	0.10
	(0.16%)	(0.12%)	
whole fat unflavored non-organic	-0.09 %	-3.38 %	0.13
	(0.14%)	(0.10%)	
soy/lactose free unflavored non-organic	0.10 %	-3.29 %	0.08
	(0.11%)	(0.08%)	
fat free flavored non-organic	0.13 %	-3.12 %	0.57
	(0.18%)	(0.13%)	
1% fat flavored non-organic	0.17 %	-3.40 %	0.15
	(0.42%)	(0.30%)	
2% fat flavored non-organic	0.06 %	-3.08 %	0.20
	(0.21%)	(0.15%)	
whole fat flavored non-organic	0.10 %	-3.05 %	0.21
	(0.13%)	(0.09%)	
soy/lactose free flavored non-organic	-0.08 %	-3.07 %	0.09
	(0.29%)	(0.20%)	