Exogenous vs. Endogenous Consumer Preferences and Knowledge about Nutrition

Eliza Mojduska, Rachel Everett, and Aparna Nemana

Paper prepared for presentation at the XIth Congress of the EAAE
(European Association of Agricultural Economists),
'The Future of Rural Europe in the Global Agri-Food System',
Copenhagen, Denmark, August 24-27, 2005

Exogenous vs. Endogenous Consumer Preferences and Knowledge about Nutrition

Contributed Paper
The EAAE Congress-Copenhagen, Denmark
23rd-27th of August 2005

Exogenous vs. Endogenous Consumer Preferences and Knowledge about Nutrition

Introduction

The 1990s and the early 2000s present a striking paradox for consumer food choice in the United States. Awareness of the health impacts of diet increased in the U.S. population, while the implementation of mandatory nutritional labeling in 1994 made it easier for consumers to obtain information on the nutritional composition of food products. However, several measures of health outcomes suggest that nutritional concerns are frequently not important in consumer food choices. Many Americans, adults and children, have poor diets and are becoming overweight at far greater numbers than ever before. According to the American Heart Association, the levels of obesity in the United States increased from 25 percent in 1976-1980 to 33 percent in 1988-1991. In the late 1990s and early 2000s, one in two American adults and one in four American children were overweight or obese (The Centers for Disease Control and Prevention web site—news). The effects are not only cosmetic; the problems of nutrition and obesity foster many deadly ills, from hypertension and heart disease to diabetes and cancer. The estimated cost of this epidemic to the general public health budget by 2020 is estimated to be in the hundreds of billions of dollars (The Centers for Disease Control and Prevention web site—news).

Thus nutritional considerations do not appear to have been strongly influencing consumer demand. What are the important determinants of consumer demand and how have they been changing over time? Understanding these factors is crucial for evaluating the benefits and costs of government regulations, such as labeling, intended to influence consumer food choice and, ultimately, public health; for assessing the competitiveness of U.S. agricultural producers and food processors as they choose product designs and marketing strategies; and for assessing the impact of changing consumer demand for food on the agricultural and food sectors of the U.S. economy.

The purpose of this paper is to propose new theoretical/modeling techniques that evaluate demand on the product level in order to build a comprehensive understanding of the relative importance of different determinants of consumer food choices over time. In particular, we integrate a discrete product choice model with a latent variable model of nutrition

knowledge/concerns. We assume that consumer preferences are endogenous and dependent on the level of knowledge. We focus our attention on modeling the relationship between preferences and knowledge, however, an empirical application of the proposed techniques will allow us to estimate the relative importance of knowledge/concerns, product nutritional quality, incomes, advertising, as well as other company marketing strategies in determining consumer demand for food products. As a result, we will be able to evaluate what is driving consumer food choices and the implications of these drivers for food and nutrition policy in the United State.

The paper most closely related to this one is Variyam et al. (1996). The authors show that when consumer preferences are treated as endogenous and dependent on nutrition knowledge/concerns, then, it is possible to separate the effects of exogenous variables into direct and indirect effects. This gives more precise estimates of the parameters of interest. The main difference between this paper and theirs is that Variyam et al. base their theoretical framework on household production theory and Lancaster's characteristics model of consumer demand for nutrients. Their empirical analysis utilizes disaggregate consumer level survey data. This paper, on the other hand, is concerned with a broad range of determinants of consumer food choices including prices, advertising, and other marketing strategies. We base our theoretical framework on discrete choice demand theory. This theory provides an effective methodology for theoretical modeling and empirical estimation of consumer demand and producer supply parameters in differentiated product markets. Our empirical analysis utilizes aggregate product level data on purchases and disaggregate consumer level survey data on consumer characteristics.

In the following section of the paper, we discuss existing work in the area of consumer food choices and the impact of nutrition knowledge/concerns, particularly work on the relationship between nutrition knowledge and demand for foods or nutrient intakes. Then we present a discrete choice model of consumer demand and a latent variable model of nutrition knowledge/concerns as well as the data sources for our future empirical analysis. In the last section, we provide a discussion of the proposed methods and conclusions.

Supporting Research

Significant effort has gone into modeling, tracking and evaluating the determinants of consumer demand for food over recent years. In these years, new information about the linkages between diet and health and the communication of this information to the general population through private and public campaigns has led to increased awareness of diet-health linkages. This is thought to have led to an increase in the demand for foods of higher nutritional quality and in the availability of such foods in the at-home and away-from-home food markets (Canning et al. 2000, Lin and Frazao 1999, Frazao and Allshouse 1996). However, other evidence (Mojduszka et al. 1999) suggests that the average nutritional quality of foods offered for sale did not improve during this period. The business press has been highlighting sales failures among nutritionally improved food products (*The Food Marketing Institute Report* 1999, 2000).

In this paper, we build on and extend already completed work on the determinants of consumer demand for food products (Mojduszka et al. 2001). Our research approach moves beyond existing work on the determinants of food choices and the impact of nutrition knowledge/concerns, particularly work on the relationship between nutrition knowledge and demand for foods or nutrient intakes. The current literature is primarily based on analysis of aggregate product level data or disaggregate consumer level survey data. Examples of these studies include Brown and Schrader (1990), Capps and Schmitz (1991), Gould and Lin (1994), Chern et al. (1995), Variyam et al. (1996), Adelaja et al. (1997), Chern and Zuo (1997), Kim and Chern (1999), and Chern (2000). These studies advanced our knowledge of the relationship between nutrition knowledge/concerns and consumer food choices but are limited by their focus on theoretical and empirical modeling that utilizes aggregate or disaggregate data sets.

The aggregate analysis approach incorporates consumer health knowledge/concerns into Lancasterian commodity demand models via indexes of information and uses total national sales data for particular commodities. These indexes quantify consumer health knowledge/concerns and their impacts on consumption patterns. Brown and Schrader (1990) and Chern et al. (1995), for example, explore the effects of information by analyzing aggregate national consumption and price data for cholesterol and the fats and oils to determine the impact of increased information on demand. Using their index of cholesterol information, Brown and Schrader show that an increase

in information about cholesterol decreased per capita egg consumption. Chern et al. apply the Brown and Schrader index in a study of the consumption of fats and oil. The authors show that an increase in cholesterol information reduced consumption of butter and lard, but not necessarily of all fats and oil. The advantage of this approach is that it evaluates the impacts of health concerns on actual consumer behavior based on observed consumption data. However, it has major disadvantages for purposes of understanding the determinants of consumer food choices. Chief among them is that analysis based on aggregate product level data does not allow precise estimates of the distribution of consumer utilities. This approach does not allow us to link individual consumer characteristics to product characteristics and does not allow us to obtain preference parameters for each consumer as well as demand elasticities with regard to product characteristics for each product considered. In all of the studies, consumer preferences are treated as exogenous.

On the other hand, analysis based on disaggregate consumer-level survey data (e.g., the Continuing Survey of Food Intakes by Individuals (CSFII) and the Diet and Health Knowledge Survey (DHKS)) uses Lancasterian demand for nutrients models or Marshallian demand for goods models where demands for specific nutrients are then approximated using Engel curves as functions of prices of goods, incomes, nutrition information, and demographic variables. The approach has the advantage of being very useful in handling the unobserved variables of consumption perception and consumer valuation of non-market goods. However, there has been skepticism about the correlation between the knowledge/concerns stated in the survey and actual consumption behavior. Also, consumer preferences are treated as exogenously given in all of the studies but one (Variyam et al. 1996). Variyam et al. (1996) use the CSFII and DHKS to examine the relationship between income, knowledge about diet/health relationships, and the nutrient intake of individuals who plan household meals. The study finds that although higher income was associated with greater knowledge about the fiber content of foods, as people's incomes increased, they reduced fiber consumption, despite its health benefits. While Variyam et al.'s approach raises interesting and important questions about the relationship between nutrition information and the income effect, their approach cannot yield a comprehensive analysis of the determinants of consumer demand because the consumer-level surveys do not provide

information on several important market variables (e.g., prices, advertising, brand strategies) that also influence consumer food choices.

The decade of the 1990s is particularly interesting for analysis of the determinants of consumer food choices because mandatory nutrition labeling was implemented in 1994 under the Nutrition Labeling and Education Act of 1991 (NLEA). The benefits estimated in the regulatory impact analysis for NLEA by the U.S. Food and Drug Administration were based on a forecasted impact of increased consumer information, improved consumer food choices, and ultimately, decreases in chronic disease rates in the United States (*Federal Register* 1991, Zarkin et al. 1991). There have been relatively few empirical studies of the actual effects of mandatory nutrition labeling and only one effectively controls for other important determinants of consumer demand. Existing studies use aggregate or disaggregate data in their analyses.

Using aggregate data, Moorman (1998) investigates the impact of the availability of market information related to the NLEA on the nutritional quality of food product offerings, the nature of competitive rivalry among manufacturers, and consumer activism in using information. She finds that changes in information may confer benefits on the market but that these benefits might be limited in scope. More specifically, the results of her study show that no significant changes occurred in the nutritional quality of food product offerings despite increased consumer activism in using information. Mojduszka et al. (1999) examine nutritional quality change in product offerings in five selected food categories using brand-level data. They find that no significant changes occurred in the average nutritional quality of food products offered for sale by manufacturers after the implementation of the NLEA. In this sense, Mojduszka et al. (1999) confirm Moorman's finding that changes in information availability may have more limited market impacts than was previously theorized (Caswell and Padberg 1992, Moorman 1998). Research using disaggregate consumer-level data offers complementary results. For example, Finke (2000) investigates what contributes to good nutritional outcomes. He utilizes data from the CSFII and the DHKS to analyze the relationship between education and nutrient intake, finding a strong relationship between education level and fat intake. The major shortcoming with these approaches is their comparative static analysis. In other words, they do not expressly include and control all the important determinants of consumer food choice so that the impact of nutrition knowledge/concerns can be accurately assessed. This paper addresses this gap in the literature by

developing new methods that can be used to evaluate whether there was a change in the relative importance of different determinants of food demand over time.

The Methodology

Our paper expands and extends a line of research using discrete choice demand and latent variable models developed to address shortcomings in current approaches to analyzing the determinants of consumer food choices (see Mojduszka et al. 2001). This approach focuses on combining both aggregate store-level product data and disaggregate individual consumer-level survey data. It is based on important developments in the theoretical literature regarding approaches to analysis of consumer and producer behavior in differentiated product markets and on incorporating latent variables into modeling of consumer choice (Ben-Akiva et al., 2002). Our approach draws from and expands on discrete choice models developed by McFadden (1978), Berry (1994), and Berry et al. (1995 and 2000). These models provide an effective approach for the theoretical modeling and empirical estimation of consumer demand and producer supply parameters in differentiated product markets and are consistent with a structural model of equilibrium in oligopolistic industries. In these models, products are bundles of attributes and consumers have preferences defined on this attribute space. Consumer preferences are functions of product attributes (some of which are unobserved) and consumer characteristics. Each consumer chooses the product that maximizes his/her utility and interactions between product and consumer characteristics generate reasonable estimates of individual consumer preference parameters and cross-price and own-price elasticities at the product level. The most recent literature in this area emphasizes gains from using alternative data sources (e.g., aggregate and disaggregate such as significant improvements in the precision of the estimated demand parameters). Increasing precision of the estimates by adding more data is a common strategy in econometrics. This strategy eliminates the need for often very restrictive theoretical and empirical assumptions (Berry et al. 2000).

Initial work by Mojduszka et al. (2001) uses the discrete choice approach to analyze the determinants of demand for frozen entrees (single serving entrees, family pack entrees, frozen dinners, and pizza) for the years 1993-1998. As has been the case in other theoretical and

empirical work in this area, the models used assume that consumer preferences are exogenous. In other words, it assumes that these preferences are determined by forces outside the model and not determined within the model by endogenous influences. For example, the model assumes that consumers choose products based on their preferences, product characteristics, and other factors, with consumer preferences being given. Thus the nutrition knowledge level may attract consumers' attention to a particular nutritionally improved brand and increase their likelihood of buying it but does not alter the consumers' underlying preferences across different brands. Mojduszka et al. (2001) empirically tested the effect of a broad range of determinants on consumer demand for particular brands of frozen entrees. The determinants included nutrient content variables, prices, advertising, socioeconomic and demographic variables, nutrition knowledge and label use variables, and in-store marketing efforts. Results for this product category for the years covered show that prices, advertising, price reductions, and consumer preferences for taste have a significant effect on the demand for frozen entrees, whereas knowledge about nutrition and nutrition label use do not.

Modeling with Exogenous Consumer Preferences

Our paper concentrates on developing a suitable methodology to empirically evaluate the determinants of consumer food choices over time. To date, we have applied a model of this choice assuming exogenous consumer preferences to one product category, frozen entrees, for the years 1993-1998.

Therefore, we first develop and specify a discrete choice demand model using the standard economic assumption that consumer preferences are exogenous. We can analyze how consumer tastes, consumer characteristics (including income, education, nutrition knowledge/concerns and nutrition label use), product characteristics (including prices and the nutritional content of foods), and manufacturers' marketing strategies (including advertising and brand strategies) influence individual food choices as measured by purchases of specific food products.

Estimating the model empirically, we can obtain own- and cross-price elasticities as well as elasticities of demand with respect to product-level nutritional characteristics for all products. The discrete choice demand method is the only one suitable to this purpose because it allows us to

link individual consumer demand for food products to underlying characteristics of these products. The resulting estimations will allow us to evaluate the implications of the demand elasticities for government policy. For example, a positive, large magnitude elasticity of demand with respect to cholesterol would imply that consumers prefer products with higher cholesterol and that government efforts to educate consumers about health problems related to the consumption of high cholesterol foods have not been effective.

To obtain the demand system for heterogeneous consumers and products, we use a discrete choice model of individual consumer behavior with exogenous consumer preference formation (see McFadden, 1978; Berry, 1994; Berry et al. 1995; Nevo, 1997; as well as the literature on product differentiation by Shaked and Sutton, 1982; Perloff and Salop, 1985; Bresnahan, 1987). This demand system serves as a benchmark for the more sophisticated model we develop where consumer preferences are endogenous. We can apply the estimated parameters of the demand system to evaluate the determinants of consumer food choices over time, including the role of publicly and privately provided information about nutrition and the role of food firms' marketing strategies.

Discrete choice models utilize indirect utility functions and assume that the level of utility that a consumer derives from a given product (brand) depends on both product and consumer characteristics. Therefore, we specify the maximum utility derived by consumer i from consuming product j in time period t as:

$$u_{ijt} = \sum_k x_{jkt}\beta_{ik} + \xi_j + \Delta\xi_{jt} + \varepsilon_{ijt}$$

where

$$\beta_{ik} = \overline{\beta}_k + \sum_r D_{irt}\beta_{kr}^m + \beta_k^{um} v_{ik}.$$

The products competing in the market are indexed as j=0, 1,..., J. Product j=0 is the outside good, so that u_{i0} is the utility the consumer derives if she does not purchase any of the J brands and allocates her income to other purchases. The x_{jkt}'s are observed product characteristics,

including price. The ξ_j is the national mean of the unobserved product characteristics and the $\Delta\xi_{jt}$ is a quarter specific deviation from this mean. The β_{ik}'s are the preference parameters of consumer i for product characteristic k. The D_{irt}'s are measured consumer characteristics, where r is a consumer characteristic, including knowledge about nutrition and use of nutrition labels, and v_{ik}'s are unmeasured consumer characteristics from a multi-variate normal distribution. Therefore, the β_{ik}'s are made up of a first component that captures the average preferences (tastes) of all consumers for an attribute and a second component that represents the deviation of individuals from the average preference based on their own characteristics. This latter component is made up of deviations based on both measured (m) and unmeasured (um) consumer characteristics. Finally, the ε_{ijt}'s represent error terms in individual preferences. In this section of the paper, we assume that the consumer nutrition knowledge/concerns are exogenous to our demand system.

We find the consumer level choice model by substituting the second equation into the first equation to obtain:

$$u_{ijt} = \delta_{jt} + \mu_{ijt}, \; for \; j = 0, 1, ..., J,$$

where

$$\delta_{jt} = \sum_k x_{jkt} \overline{\beta}_k + \xi_j + \Delta\xi_{jt},$$

and

$$\mu_{ijt} = \sum_{kr} x_{jkt} D_{irt} \beta^m_{kr} + \sum_k x_{jkt} v_{ik} \beta^{um}_k + \varepsilon_{ijt}.$$

The indirect utility of consumer i from product j in time period t is expressed as the mean utility, referred to as δ_{jt}'s, and the mean zero heteroscedastic deviation from that mean, μ_{ijt}, that captures the effects of the random coefficients, which reflect individual consumer characteristics. In this

case, the contribution of x_k units of the k^{th} product characteristic to the utility of consumer i varies across consumers and is given by:

$$(\overline{\beta}_k + \beta^m_{kr} D_{irt} + \beta^{um}_k v_{ik}) x_{jkt}$$

The mean of the utility from good j, δ_{jt}, is entirely determined by the product characteristics and thus represents a product specific component that does not vary with consumer characteristics. On the other hand, a deviation from that mean, μ_{ijt}, depends on the interaction between consumer and product specific characteristics. As a result, consumers who have a preference for fat, for example, will tend to attach high utility to all fatty products, and this will induce large substitution effects between fatty products. The parameters of the model are $\theta=(\delta, \beta^m, \beta^{um})$. The vector δ includes the linear parameters and the vectors β^m and β^{um} contain the non-linear parameters.

We obtain the aggregate demand system by summing the choices implied by the individual utility model over the distribution of consumer characteristics in the population. We denote the vector of measured and unmeasured individual characteristics by w, therefore,

$$w = (D, v, \varepsilon)$$

We denote its distribution in the population by P_w. Each consumer chooses one unit of the good that maximizes its utility therefore aggregate demand for good j is given by the integral of the density of consumer characteristics over the set of product characteristics that imply a preference for good j:

$$s_{jt}(\delta, \beta^m, \beta^{um}, x) = \int_{A_{jt}} P_w(dw) = \int_{A_{jt}} P_\varepsilon(d\varepsilon) * P_D(dD) * P_v(dv)$$

where

$$A_{jt}(\delta, \beta^m, \beta^{um}; x) = \{w : \max_{r=0,1,\ldots,J} [u_{irt}(w; \delta, \beta^m, \beta^{um}, x)] = u_{irt}\}.$$

By multiplying the market share equation by the number of consumers in the market, M, we obtain the J-vector of demands as $M*s(\delta, \beta^m, \beta^{um}, x)$. We model consumer heterogeneity as a function of the empirical non-parametric distribution of consumer characteristics without

imposing any arbitrary functional forms on this distribution. Thus, given the assumptions on the distribution of the unobserved variables (v and ε), we are able to compute the integral in the market share equation analytically or numerically.

A simulation technique introduced by Pakes (1986) can be applied in order to estimate consistently a random coefficients discrete choice model of consumer demand for all of the food products considered.

Modeling with Endogenous Consumer Preferences

In this section of the paper, our work expands on previously used discrete choice models by treating consumer preferences as endogenous. Doing so is important because preferences are likely endogenous (correlated with knowledge/concerns) and treating them as exogenously given results in inaccurate estimates of the relative importance of various determinants of consumer food choices. Thus, the proposed modeling and empirical analysis can make a significant contribution to the design of more effective information programs and marketing strategies.

To address the problem of estimation, we build on our previously completed work on discrete choice modeling of consumer demand (Mojduszka et al. 2001). This work provides a model of individual consumer utility and demand that is explicitly aggregated to obtain product level demands. It therefore already contains a framework for analyzing aggregate and disaggregate data sources. However, consumer choice of food products may be further conditioned by nutrition knowledge/concerns. To account for this possibility, we assume that consumer choice of food products and nutrition knowledge/concerns are correlated, implying a simultaneous system of equations. We incorporate nutrition information measures in an integrated discrete choice model system of product choice and nutrition knowledge (Ben-Akiva and Bowman, 1998; Ben-Akiva et al., 2003). In this new model, the distribution of consumer utilities depends on both measured and unmeasured individual characteristics. These determine preferences for product attributes (some of which are unobserved) and hence determine demand.

The proposed methodology integrates latent variables in the choice model and incorporates indicators of the latent variables provided by responses to survey questions to aid in estimating the model. Simultaneous estimation techniques can be used, which results in latent

variables that provide the best fit to both the choice and the latent variable indicators. The integrated model framework consists of two main components, the choice model (described in the previous section) and the latent variable model (described below).

For the latent variable model, we need the distribution of the latent variables given the observed variables:

$$D^*_{irt} = h(X_{jkt}; \gamma) + \eta, \text{ for } j = 0, 1, ..$$

This gives us one equation for each latent variable. In addition, for the latent variable model we need the distribution of the indicators conditional on the values of the latent variables.

$$D_{irt} = g(D^*_{irt}, X_{jkt}; \alpha) + \nu$$

This gives us one equation for each indicator (i.e., each survey question). The integrated model consists of equations 1, 3, 9, and 10. Equations 1 and 3 represent the choice model and equations 9 and 10 represent the latent variable model. From equations 1 and 3 and an assumption about the distribution of the disturbances we derive the choice probability conditional on both observable and latent variables. The likelihood function includes complex multidimensional integrals, with dimensionality equal to that of the integral of the underlying choice model plus the number of latent variables. We can estimate the integrated model using a simulation approach.

The changes incorporated into the new model allow us to estimate three sets of parameters using a nested method of moments algorithm. The first quantifies the effect of measured individual characteristics on tastes for product attributes. The second set measures the importance of unmeasured individual characteristics in determining preferences for product attributes. The third set allows us to estimate the effect of product attributes on the mean utility of a product. In other words, the first two sets give direct evidence on the extent to which the demand parameters can be explained by individual characteristics. The aggregate data can then be used to estimate the additional parameters that determine the relationship between product attributes and the mean utility levels of the products.

By integrating the product choice model and the latent variable model of nutrition knowledge/concerns, we can obtain precise estimates of the demand parameters that are crucial in determining consumer food choices over time. Understanding these determinants is central to the design of effective nutrition information programs and to the design of marketing and promotion strategies by producers, manufacturers, and distributors. The results of the study can contribute to precise answers to the question of how consumer information about nutrition affects individual food choices in the market place.

Data

In the near future (by the end of May 2005), we will use the previously described methodology to empirically analyze how consumer tastes, consumer characteristics (including income, education, nutrition knowledge/concerns and nutrition label use), product characteristics (including prices and nutritional content of foods), and manufacturers' marketing strategies (including advertising and brand strategies) influence individual food choices. The final copy of this paper will contain the estimated results of the models with exogenous and endogenous consumer preferences. Producing this comprehensive analysis will require the development and integration of multiple data sources that provide information relevant to the determinants of consumer food choice. These include IRI Info-scan^{TM} Data for quantities, prices, and in-store promotion levels; Nutritional Quality Change Data at the University of Massachusetts and the USDA National Nutrient Database for nutrient content information; National Leading Advertisers Data for advertising expenditures; and USDA Diet and Health Knowledge Survey Data and Consumer Demographics Data for consumer characteristics.

We will obtain data on market shares, prices, brand offerings, and in-store marketing efforts for all the products in selected food categories from the IRI Info-scan database. These data are collected continuously by IRI using scanning devices in a national random sample of supermarkets located in 64 metropolitan and rural areas of the United States.

We will match the IRI Info-scan quarterly market shares, prices, and other data for each product with five other data sources. First, we will match it to the nutrition composition data included in the Nutritional Quality Change Data developed at the University of Massachusetts and to the USDA National Nutrient Database available at the USDA's web site. The IRI Info-scan

data do not provide information on the amounts of nutrients contained in food products. Thus, information on market shares and prices has to be matched with information on the nutritional content of the respective products from the other two data sources.

In order to obtain accurate information on the nutritional content of the products that will be included in our empirical analysis, we will compare, evaluate, and complement two data sources. The USDA National Nutrient Database will be used to complement the Nutritional Quality Change Data, which provides exact information on the nutritional composition of all products offered in a large super-store but does not contain information on all products offered at the national level. Nonetheless, the latter set includes a complete census of all products in the most popular package size offered in 33 food product categories in a representative super-store in New England for the years 1992 through 1999. We will create average nutrient content values based on similar products for products that appear in the scanner data but are missing from the two nutritional quality data sets.

Second, we will obtain information on the distribution of consumer knowledge about nutrition and nutrition label use by sampling individuals from the Diet and Health Knowledge Survey (DHKS) conducted by the U.S. Department of Agriculture. The DHKS surveys 1,966 individuals, 20 years of age or older, who are the main meal planners in their households. The survey includes their answers to questions concerning attitudes toward and knowledge of nutrition, food safety, and diet and health, as well as their use of nutrition labels. We will use those questions from the DHKS that relate to knowledge about nutrition related to calories, fat, cholesterol, and sugar, and to nutrition panel use. We hypothesize that these factors play an important role in consumer food choices. Nutrition panel use can allow consumers to precisely evaluate the nutritional quality of foods they choose. All packaged foods have been required to carry nutrition panels since May 1994. By incorporating this information into our empirical modeling, we will be able to estimate how consumer knowledge of nutrition and use of nutrition panels affect consumer choices of products in the selected food categories over time.

Third, we will obtain information on the distribution of consumer demographic variables by sampling individuals from the Current Population Survey (CPS) carried out each year by the U.S. Bureau of the Census. Consumer per capita income will be constructed by dividing household income by the size of the household. We will also include other important economic

and demographic variables such as: income, education, age, percentage of elderly people, and percentage of women working. The CPS data are representative of the national population statistics from the Bureau of the Census.

Fourth, we will match the IRI Info-scan data with the quarterly expenditures on advertising for each of the products in the four product categories taken from the Leading National Advertisers database for 1993-2002. These data are collected for 11 different types of mass media (e.g., network television, spot television, cable networks, national spot radio, network radio, newspapers, magazines). We will use the total average advertising expenditures on all 11 types of mass media.

At this time, the data phase of our work has been successfully completed. We have compiled one comprehensive data set that consists of all the data information described above. We have also entered the estimation phase of our work that we expect to complete by the end of May 2005.

Discussion and Conclusions

Consumer choices of food products take place in the market place and are influenced by many factors. Sims (1998) stressed the following two sets of factors: those external to the individual (and thus applicable to groups of people) and those that are idiosyncratic and specific to the individual. As the most important external factors Sims identified food, agricultural, and trade policies; available technology; food marketing; and group influence. On the other hand, the most important internal factors were one's genetic nutrient requirements, tastes, attitudes, health and nutrition knowledge, education, occupation, and income. Sims pointed out that those two sets of factors are interconnected. In this paper, we provide new methods that allow for integrated analyses of the external (e.g., government information policy, food firms' marketing strategies) and the internal (e.g., tastes, nutrition knowledge and attitudes, education, income) factors that determine consumer choices of foods, as well as the links between them. Over the years, as the government and nutrition educators have stressed increasingly the relationship between diet and health, average per capita income in the United States has increased, making food more affordable for the average consumer. As people's incomes rise, they place greater value on time and demand

more convenient foods that are often less healthful (Aldrich 1999). In addition, Mojduszka et al. (2001) and Harris (1997) show that taste dominates nutrition in determining which frozen meal or frankfurter to consume. Overall, it appears that the factors of taste, food marketing, rising incomes, and convenience are outweighing nutrition and health concerns in determining demand for individual food products. These trends, however, may not be inevitable. More studies are needed to provide a more comprehensive and systematic assessment of consumer food choices. In this sense, our study provides critical methods (and in the future, critical empirical results) to inform ongoing policy debates relating to nutrition, incomes, advertising, consumer knowledge, and government regulation of information provision.

The rationale for this paper is to provide a suitable methodology for the in-depth analysis of current determinants of consumer food choices that can inform discussion of current consumption trends and their causes. Our motivation is to be able to explain the apparent paradox discussed above that while consumers appear to be more knowledgeable about diet/health relationships, this knowledge is not translating very directly or in some cases at all into improved dietary choices, as well as other apparent paradoxes in patterns of consumer demand. Piecemeal analysis of aggregate trends or of consumer-level survey data can help to inform the discussion but are inherently limited in that they focus on one or only a few factors that influence demand at any given time. In reality, however, several determinants of demand are working simultaneously and what we see as market demand is the outcome of the interaction of these factors.

Our expansion and extension of prior work focuses on developing and (by the end of May 2005) estimating discrete choice models incorporating endogenous formation of consumer preferences. The motivation for this is that to date models have failed to account for the relationship between consumer preferences and knowledge about nutrition. These preferences are unlikely to be exogenous so treating them as exogenously formed results in inaccurate estimates of the relative importance of different determinants of consumer food choices. Estimating consumer preferences as endogenous can correct for this problem.

The results of the empirical application of the integrated model proposed in this paper can be significant in several ways. The results can give a comprehensive empirical picture of the relative importance and strength of factors influencing the actual brand-level food choices of consumers. This information is useful to food producers and processors in analyzing the

importance of different marketing strategies and in forecasting future shifts in demand. The information is critically important for policy makers as they make decisions related to information/labeling, public health, and education programs. For example, expected results from these programs may not materialize, because while the factors they are targeting (e.g., consumer knowledge about diet/health linkages), do have some influence on consumer demand, this influence is currently swamped by other market forces. Clear understanding of current determinants of consumer food choice can inform policy choices that rely on enhancement of these factors. Finally, the results can contribute new, more comprehensive modeling and empirical approaches to brand-level demand estimation.

The most innovative and unique aspect of the proposed work is that it moves beyond the studies described in the introduction to integrate several behavioral models and data sources. As a result, our conclusions can provide important insights into the economic forces that tend to limit the efficient provision and use of nutrition information in consumer choices of foods. This approach differs from other research in this area in several respects. First, the dependant variable focused on is the market share of a particular food product in its product category. This allows a direct analysis of the relative competitive effects of different product attributes on demand as well as of the effect of consumer characteristics on demand. Second, the development by the authors of product-level nutritional profiles allows much more specific and reliable analysis of the nutritional quality of competing products. Third, this approach makes full use of and expands the discrete choice approach to modeling demand for differentiated products.

References

Aldrich, L. (1999) *Consumer Use of Information: Implications for Food Policy*. An Economic Research Service Report, USDA, Agricultural Handbook, Report No. 715.

Anderson, S., De Palma A., and Thisse F. (1989) Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach, *Rev. Econ. Stud.* 56:21-35.

Ben-Akiva M. and J. L. Bowman. (1998) Integration of an Activity-based Model System and a Residential Location Model. *Urban Stud.*, Vol. 35, No. 7:1131-1153.

Ben-Akiva M., J. Walker, A. T. Bernardino, D.A. Gopinath, T. Morikawa, and A. Polydoropoulou. (2002). Integration of Choice and Latent Variable Models. In Perpetual Motion, H.S. Mahmassani (Ed.). Elsevier Science Ltd.

Berry, S., J. Levinshon, and A. Pakes (2000) Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market, Working Paper No. 536, University of Michigan.

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

Berry, S. (1994) Estimating Discrete Choice Models of Product Differentiation, *Rand J. Econ*. 25:242-262.

Bresnahan, T. (1987) Competition and Collusion in the American Automobile Oligopoly: the 1955 Price War, *J. Indust. Econ.* 35:457-482.

Brown, D. J. and L. F. Schrader. (1990) Cholesterol Information and Shell Egg Consumption. Amer. J. Ag. Econ., 72:548-555.

Canning, P., M. Tsigas, C. Lee, G. Schuter, B. O'Roark, S. W. Martinez, L. Aldrich, and N. Blisard (2000) Current Issues in Economics of Food Markets, AIB No. 747, ERS/USDA, September

Cardell, N. S. (1989) *Extensions of the Multinomial Logit: the Hedonic Demand Model, the Non-Independent Logit Model, and the Ranked Logit Model*, Ph.D. Dissertation, Harvard University.

Capps, O. Jr. and J. D. Schmitz (1991) A Recognition of Health and Nutrition Factors in Food Demand Analysis, Western J. Ag. Econ., 16(1): 21-35.

Caswell, J. A. and D. I. Padberg. (1992) Towards More Comprehensive Theory of Food Labels. *Amer. J. Ag. Econ*., 74(2):460-468.

Chamberlain, G. (1982) Multi-Variate Regression Models for Panel Data, *J. Econometrics*, 18(1):5-46.

Chern, W. S. (1997) Impacts of Changing Health Information of Fat and Cholesterol on Consumer Demand: Application of New Indexes, Discussion Paper No. 443, The Ohio State University

Chern, W. E., E. T. Loehman, and S. T. Yen. (1995) Information, Health Risk Beliefs, and the Demand for Fats and Oils. *Rev. Econ. Stat.*, 77:555-564.

Federal Register (1991) Regulatory Impact Analysis of the Proposed Rules to Amend the Food Labeling Regulations, 56 FR 229: 60856-60878.

Finke, M. S. (2000) Did the Nutrition Labeling and Education Act Affect Food Choice in the United States, paper presented at the conference The American Consumer and the Changing Structure of the Food System, Washington D.C., May 4-5.

Frazao, E. and J. E. Allshouse (1996) *Size and Growth of the Nutritionally Improved Foods Market*, AIB-723, USDA/ERS.

Harris, J. M. (1997) The Impact of Food Product Characteristics on Consumer Purchasing Behavior: The Case of Frankfurters. *J. Food Distrib. Res.*, Feb.:92-97.

Hausman, J. (1996) Valuation of New Goods Under Perfect and Imperfect Competition, in T. Bresnahan and R. Gordon, Eds., *The Economics of New Goods*, Studies in Income and Wealth, Vol.58, Chicago: National Bureau of Economic Research.

Hausman, J., Leonard G., and Zona J. D. (1994) Competitive Analysis with Differentiated Products, *Annales D'Economie et de Statistique*, 34:159-80.

Ippolito, P. M. and A. D. Mathios. (1990) Information, Advertising, and Health: A Study of the Cereal Market. *Rand J. Econ*. 21:459-480.

Ippolito, P. M. and A. D. Mathios. (1996) *Information and Advertising Policy: A Study of Fat and Cholesterol Consumption in the United States, 1977-1990*. Bureau of Economics Staff Report, Federal Trade Commission, Washington DC.

Kim, Sam-Ryang and W. S. Chern (1999) Alternative Measures of Health Information and Demand for Fats and Oils in Japan. *J. Cons. Aff*., Vol. 33, No. 1:92-109.

Lin, Bing-Hwan and E. Frazao (1999) Away-From-Home Foods Increasingly Important to Quality of American Diet, AIB No. 749, ERS/USDA, January.

McFadden, D. (1978) Modeling the Choice of Residential Location, in A. Karlgvist, et al., Eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland.

Mojduszka, E. M., J. A. Caswell, and J. M. Harris. (2001) Consumer Choice of Food Products and the Implications for Price Competition and Government Policy. *Agribus.: An Internat. J.*, Vol. 17 (1):81-104.

Nevo, A. (1997) *Demand for Ready-to-Eat Cereal and Its Implications for Price Competition, Merger Analysis, and Valuation of New Goods*. Ph.D. Dissertation, Harvard University.

Pakes, A. (1986) Patents as Options: Some Estimates of the Value of Holding European Patent Stocks, *Econometrica*, 54:755-784.

Pakes, A. and McGuire. (1994) Computation of Markow Perfect Nash Equilibria I: Numerical Implications of a Dynamic Product Model, *Rand J. Econ*. 25:555-589.

Perloff, J. M. and M. B. Ward. (2000) Welfare, Market Power, and Price Effects of Product Diversity: Canned Juices. Paper presented at the Conference: *Industrial Organization and the Food Processing Industry,* June, Toulouse, France.

Perloff, J. M. and Salop S. (1985) Equilibrium with Product Differentiation, *Rev. Econ. Stud.* 52:107-120.

Shaked, A. and Sutton J. (1982) Relaxing Price Competition Through Product Differentiation, *Rev. Econ. Stud.* 49:3-13.

Spence, M. (1976) Product Selection, Fixed Costs, and Monopolistic Competition, *Rev. Econ. Stud.* 43:217-235.

Variyam, J. N., Blaylock J., and D. Smallwood (1996) A Probit Latent Variable Model of Nutrition Information and Dietary Fiber Intake, *Amer. J. Agr. Econ.* 78(August):628-63

Zarkin, G. N. Dean, J. Mauskopf, and D. Neighbors (1991) Estimating Health Benefits of Nutrition Labeling Changes, prepared for Economics Section, Office of Compliance, Center for Food Safety and Nutrition, U.S. Food and Drug Administration, HF