United States
Department of
Agriculture

ERS

Technical
Bulletin
Number 1842

An Economic Research Service Report

Modeling Nutrient Intake

The Role of Dietary Information

Jayachandran N. Variyam
James Blaylock
David Smallwood

It's Easy To Order Another Copy!

Just dial 1-800-999-6779. Toll free in the United States and Canada. Other areas, please call 1-703-834-0125.

Ask for *Modeling Nutrient Intake: The Role of Dietary Information* (TB-1842).

The cost is $9.00 per copy. For non-U.S. addresses (including Canada), add 25 percent. Charge your purchase to your Visa or MasterCard. Or send a check (made payable to ERS-NASS) to:

ERS-NASS
341 Victory Drive
Herndon, VA 22070

We'll fill your order by first-class mail.

Modeling Nutrient Intake: The Role of Dietary Information. By Jayachandran N. Variyam, James Blaylock, and David Smallwood. Food and Consumer Economics Division, Economic Research Service, U.S. Department of Agriculture. Technical Bulletin No. 1842.

Abstract

Measuring the effect of nutrition information on an individual's intake of dietary fiber is crucial for understanding and quantifying the linkage between nutrition education programs and actual consumption behavior. Dietary information and intake data for a sample of U.S. household meal planners are used to estimate the effect of fiber-specific information on dietary fiber intake. The information variables are measured using survey responses to a series of questions on fiber content of foods, attitude toward consuming fiber-rich foods, and awareness of fiber-health links. Results indicate that attitudes about eating high-fiber foods and awareness of the linkage between dietary fiber intake and some diseases are important determinants of an individual's fiber intake.

Keywords: Dietary fiber, dietary knowledge, attitudes, fiber-disease awareness, latent variables, simultaneous model.

Washington, D.C. 20005-4788 May 1995

Contents

	Page
Summary	v
Introduction	1
The Conceptual Framework	2
The Modeling Framework	3
Data and Variables	6
Empirical Results	9
Fiber Information	10
Fiber Consumption	13
Some Final Observations	20
References	21
Appendix: Model Estimation	23

Summary

Understanding the relationship between an individual's socio-economic/demographic characteristics and dietary knowledge and effects of these on nutrient intake is crucial for designing, targeting, and evaluating nutrition education programs and monitoring the Nation's progress toward dietary goals. In this study, we focus on a nutrient that has received widespread publicity in the past few years: dietary fiber. Using data from the U.S. Department of Agriculture's Continuing Survey of Food Intake by Individuals and the companion Diet and Health Knowledge Survey, we tested whether fiber information, as measured by fiber-specific nutrition knowledge, attitudes toward consuming fiber-rich foods, and fiber-disease awareness, translates into increased intake of dietary fiber. We estimate the direct effects of household, socio-demographic, health, and diet-related characteristics of the respondents on their fiber intake, as well as the indirect effects of these characteristics on fiber intake, through their influence on respondents' knowledge of the fiber content of foods, their attitude toward consuming fiber-rich food, and their awareness of fiber-health linkages.

Our results show that fiber information plays a key role in determining dietary fiber intake and describes direct and indirect channels through which various consumer characteristics influence consumption. We find that attitudes toward consuming fiber-rich foods and awareness of fiber-disease linkages are more important influences on fiber intake than specific knowledge about the fiber content of individual foods.

Our results, if verified in other studies, have widespread implications for nutrition education programs. We conclude that, among the fiber information variables, diet-disease awareness has the largest effect on fiber consumption. This is not surprising given that avoiding health problems has the most immediate and transparent economic benefits to the consumer. Therefore, consumers possessing this type of information may be expected to actively change their dietary patterns compared with consumers who do not have this information, or with those who have only general notions about the benefits of dietary fiber. This suggests that nutrition education programs aimed at increasing the general awareness of fiber-health-disease linkages will likely have the highest payoffs in terms of raising dietary fiber-intake levels. We also find a distinct attitudinal dimension beyond fiber-disease awareness that influences fiber consumption. Our results imply that nutrition education strategies emphasizing general attitudinal messages such as five-a-day intake are likely to have greater effect in modifying dietary patterns than strategies emphasizing specialized knowledge about the nutrient content of foods.

Years of formal education, a vegetarian diet, and the respondent's age are found to be significantly associated with higher dietary fiber intake. Smokers and black respondents have significantly lower dietary fiber intake than average. Income is positively associated with awareness of fiber-disease relationships, as well as knowledge of the fiber content of specific foods. However, the direct effect of income on fiber intake is negative, because many higher income individuals tend to regard some fiber-rich foods as inferior to other kinds of foods. This causes the net effect of income on fiber intake to be negative. Therefore, strategies expecting improved nutrition through higher purchasing power alone may be flawed.

The estimated influence of respondent gender and Hispanic ethnicity provides vivid testimony to the importance of isolating a variable's direct and indirect effects on dietary fiber intake. First, the higher level of fiber consumption (per 1,000 kilo calories) by female respondents compared with males may be chiefly due to females' higher level of fiber-related disease awareness and attitude. Second, in the case of Hispanic respondents, higher fiber consumption from ethnically related diet choices is negated by lower fiber consumption from diet choices associated with lower levels of disease awareness and concern about eating fiber-rich foods.

Modeling Nutrient Intake
The Role of Dietary Information

Jayachandran N. Variyam, James Blaylock, and David Smallwood

"For the two out of three adult Americans who do not smoke and do not drink excessively, one personal choice seems to influence long-term health prospects more than any other: what we eat." (*The Surgeon General's Report on Nutrition and Health*, 1988)

Introduction

Dietary excesses are associated with the three leading causes of death in the United States: coronary heart disease, some types of cancer, and stroke. In addition, deaths from noninsulin-dependent diabetes mellitus (also called type II diabetes, or adult-onset diabetes), the seventh leading cause of death are also linked to dietary excesses (National Center for Health Statistics, 1993). Together, deaths from diseases associated with dietary excesses account for nearly two-thirds of the deaths occurring each year in the United States.

The costs of poor diets to both society and individuals are high and continue to increase. The American Heart Association estimates that about 1.5 million heart attacks occur each year, and that coronary heart disease costs Americans an estimated $52 billion in direct health care spending and lost productivity (Frazao, 1994). Associated costs for cancer are even higher, $104 billion. Experts agree that our diets are an important factor in deaths from these diseases although not the only factors (others being genetics, the environment, and aging). A recent study by McGinnis and Foege (1993) suggests that dietary factors and sedentary activity patterns together account for at least 300,000 U.S. deaths per year. Some experts estimate that 35 percent of all cases of cancer could be prevented through dietary changes alone (Doll and Peto, 1981).

After an extensive literature review, Willett (1994) remarks: "One clear conclusion from existing epidemiologic evidence is that many individuals in the United States have suboptimal diets and that the potential for disease prevention by improved nutrition is substantial." In a similar vein, Nobel laureate and economist, Robert Fogel (1994) notes that "Improved ... nutrition explains roughly 30 percent of the growth of per capita income in Britain between 1790 and 1980."

One nutrient that has received widespread publicity in the past few years is dietary fiber. Scientific evidence shows a strong link between diets high in food containing dietary fiber and lower incidence of some cancers (Block, Patterson, and Subar, 1992; U.S. Dept. of Health and Human Services, 1988, 1991). A number of studies have linked diets high in soluble-fiber with reduced blood cholesterol levels (U.S. Dept. of Health and Human Services, 1988, 1991). Elevated blood cholesterol levels are one of the chief risk factors in heart disease (Browner, Westenhouse and Tice, 1991; Kushi and Kottke, 1990; Oliver, 1990; U.S. Dept. of Health and Human Services, 1988; Willett, 1994). Also "Eating foods with fiber is important for proper bowel function and can reduce symptoms of chronic constipation, diverticular disease, and hemorrhoids," according to the *Dietary Guidelines for Americans*. The guidelines note that diets low in dietary fiber may increase the risk of developing certain types of cancer.

Despite the importance of sound dietary practices and the press coverage given to diet-related topics, results from a recent U.S. Department of Agriculture (USDA) survey lead to the conclusion that misconceptions, confusion, and lack of motivation are barriers to dietary change. Almost 90 percent of meal planners agreed with the statement: "What you eat can make a big difference in your chance of getting a disease, like heart disease or cancer." However, only 50 percent were aware of health problems associated with fiber and about 25 percent of survey participants said that it was of low or very low importance to them to eat at least six servings a day of bread, cereal, and other grain products. Lack of knowledge about the nutrient content of foods was also a problem: 23 percent of the respondents did not know that popcorn contained more fiber than a comparable amount of pretzels. These statistics suggest that those charged with educating the public on nutrition issues face a complex task with a number of difficult challenges.

The principal objective of most nutrition education programs is to impart additional knowledge of food composition, instill better attitudes about healthy eating, and heighten awareness of diet-disease relationships. Presumably, better knowledge, attitudes, and awareness lead to better food choice. But do they, and if yes, how? Previous evidence on these questions is limited and mixed (Capps and Schmitz, 1991; Kushi and others, 1988; Shepherd and Stockley, 1985). We take a fresh look at this complex issue by focusing on a particular nutrient and developing a rigorous statistical model to identify and quantify the association between an individual's socio-economic/demographic characteristics and the current nutrition information level and its effect on dietary fiber intake. In other words, we want to know if better nutrition knowledge and attitudes translate into increased intake of a particular nutrient (dietary fiber), while controlling for other factors. Population characteristics may explain consumption directly by capturing the variations in economic costs or tastes, and indirectly by capturing the variations in dietary knowledge, attitude, and awareness. Knowledge of the relative strengths of these direct and indirect effects is critical for the design and implementation of effective nutrition education programs.

While the linkages among knowledge, attitude, and consumption are conceptually straightforward, the statistical estimation of these relationships poses some challenges. Unlike consumption, information variables such as knowledge, attitude, and awareness are abstract theoretical constructs with no directly observable counterparts. Therefore, their measurement and modeling raise special problems. The approach usually used to measure these variables is by using multiple observable indicators. For example, knowledge of the nutrient content of foods may be indicated by a battery of questions where consumers are presented with pairs of food items and are required to identify from each pair the item with the higher nutrient content. Given such multiple indicators, a statistically meaningful method for combining them into a knowledge variable is required.

A second modeling issue arises from the fact that indicators are usually discrete variables (for example, 1 if Yes or 0 if No; 1 if Not at all Important to 6 if Very Important). Treating discrete indicators as continuous variables raises various econometric problems that lead to statistically biased and inefficient estimates (Greene, 1990). A satisfactory specification should account for the discrete nature of the indicators.

We apply, for the first time in the nutrition literature, a statistical framework that unifies the measurement of unobserved information variables with multiple discrete indicators and the estimation of their influence on intake in a simultaneous model. We hope to contribute to better evaluation and targeting of the Nation's educational needs regarding dietary fiber consumption by identifying the socio-demographic profile of individuals with least favorable dietary fiber knowledge, attitude, and awareness, as well as the association between knowledge, attitude and awareness, and actual dietary fiber intake. We also hope to develop and illustrate an analytical framework that can be expanded and applied to examine other diet-health behavior issues and their interrelations.

The Conceptual Framework

Prompted by the increasing economic significance of health issues and the growing recognition of dietary factors in the cause and prevention of important diseases, economists have extended traditional food demand analysis to examine the intake of nutrients and other dietary components (Behrman and Deolalikar, 1988; Capps and Schmitz, 1991). The primary focus of this research has been on a better understanding of nutrient choices across the population, both independently and within the context of various nutrition enhancement and food assistance programs (for example, Behrman and Deolalikar, 1987; Pitt and Rosenzweig, 1985). The theoretical underpinning for nutrient-intake analysis is provided by integrating Becker's (1965) theory of household production with the characteristics model of consumer demand developed by Lancaster (1971). In this framework, households combine various inputs to produce "commodities," including the health of family members, so as to maximize a joint utility function. Some of the inputs (for example, food) derive their value by supplying characteristics (nutrients) necessary for the production of some outputs (health). Subject to the constraints of household technology and resources, this maximization process generates individual and household demand functions for the inputs and input characteristics.

To formalize this notion, assume that a representative household with T members has a joint utility function:

$$U = U(F, \underline{z}, \underline{h}, \underline{l}), \quad U' > 0, \; U'' < 0, \tag{1}$$

where **F** is a matrix of foods consumed, and $\underline{z}$, $\underline{h}$, and $\underline{l}$ are vectors of non-foods, health status, and leisure for each family member (Behrman and Deolalikar, 1988; Pitt and Rosenzweig, 1985). Health and food intakes enter directly into the utility function because good health is valued in itself and because foods are consumed for reasons other than their nutritional value, such as for their taste.

Given household assets and market prices, the preference function expressed in equation 1 is maximized, subject to two sets of constraints. The first is a set of technology constraints determining the production of health and nutrient intakes of the household members. The second set of constraints binds time and income into a full-income constraint.

The health production function for each household member t is given by:

$$h_t = h(\underline{c}_t, \underline{g}_t, \underline{x}_t, u_t),\ t = 1,\ldots,T, \tag{2}$$

where $\underline{c}_t$ is a vector of nutrients consumed, $\underline{g}_t$ is a vector of non-food health inputs such as medical services, $\underline{x}_t$ is a vector of exogenous health-related personal characteristics such as age and sex, as well as household characteristics and health-related environmental factors common to all members, and u_t is an exogenous health endowment beyond the individual's or household's control. Elements of the nutrient intake vector are constrained to the maximum levels attained by linear combinations of foods consumed.

Under the assumption that the relevant functions have desirable properties to ensure unique interior solutions, the first-order conditions for the maximization of equation 1 give, among other relations, member-specific nutrient demand equations as a function of prices, income, and socio-demographic and household characteristics, and u_t.

Introducing dietary knowledge and attitudes explicitly into this model is simply a recognition of these factors as exogenous components similar to $\underline{x}_t$ that influence the production of health and nutrients. Food for household consumption is bought and prepared by main meal planners/preparers. Thus, their nutrition information level influences members' health through the channels of food choice, nutrient production, and the use of nutrients with other factors in the health production function. Hence, the reduced-form nutrient demand function may be written as:

$$\underline{c}_t = c(\underline{p}, I, \underline{x}_t, \underline{\eta}_t, u_t), \tag{3}$$

where $\underline{p}$ is a vector of prices, I is the household income, and $\underline{\eta}_t$ is a vector of dietary information variables such as nutrition knowledge and attitudes.

The Modeling Framework

In specifying an empirical version of equation 3 for modeling dietary fiber intake, we make two abstractions. First, since our analysis is concerned with intake choices made at a given time by a cross-section of households, we assume that there is limited price variation across households. The extant price variation is assumed to be captured by variables for geographic region and location of the household. Second, we focus on modeling the fiber intake of one individual, the main meal planner/preparer, per household.

In equation 4, suppose c_i represents the amount of dietary fiber consumed by the ith individual, i = 1,...,N, where N is the number of households. Suppose there are K fiber information variables such as fiber knowledge, attitude, and disease awareness, and let $\underline{\eta}_i = (\eta_{i1} \ldots \eta_{iK})'$ be a (K×1) vector of the ith individual's measurements on these variables. Our basic fiber consumption model is a linear equation relating c_i to $\underline{\eta}_i$ and a (P×1) vector $\underline{x}_i$ consisting of the individual's economic, socio-demographic, and dietary characteristics:

$$c_i = \beta_0 + \underline{\beta}_1' \underline{\eta}_i + \underline{\beta}_2' \underline{x}_i + \xi_i^*, \tag{4}$$

where β_0 is a scalar intercept, $\underline{\beta}_1$ and $\underline{\beta}_2$ are conformable vectors of structural coefficients, and ξ_i^* is a random error term that captures the effect of u_t in equation 3 and other idiosyncratic variation, and is assumed to be distributed independently over the individuals with variance ρ^2.

The fiber information variables $\underline{\eta}_i$ in equation 4 assumed to be predetermined by the vector $\underline{x}_i$:

$$\eta_{ik} = \gamma_{0k} + \underline{\gamma}_{1k}' \underline{x}_i + \zeta_{ik},\ k = 1,\ldots,K, \tag{5}$$

where γ_{0k} is a scalar intercept, γ_{1k} is a (P×1) vector of structural coefficients and ζ_{ik} is a term capturing residual variation in the kth information variable assumed to be randomly distributed over the individuals with the covariance structure:

$$cov(\zeta_{ik},\zeta_{jl}) = \begin{cases} \psi_k^2 & for \quad i = j \ and \ k = l \\ \psi_{kl} & for \quad i = j \ and \ k \neq l, \ k,l = 1,\dots,K \\ 0 & for \quad i \neq j. \end{cases} \tag{6}$$

Thus, the fiber information variables are allowed to be correlated for the same individual and uncorrelated otherwise.

Additional restrictions are necessary for identifying the parameters in equations 4 and 5. If it is assumed that:

$$cov(\xi_i^*,\zeta_{jk}) = 0 \ for \ all \ i,j,and \ k, \tag{7}$$

the model is fully recursive and no further restrictions are necessary on the structural parameter vectors $\underline{\beta}_2$ and $\underline{\gamma}_{1k}$ for identification. However, if $cov(\xi_i^*,\zeta_{ik})$ is unrestricted for all $i = j$ and $k = 1,\dots,K$, $\underline{\beta}_2$ and $\underline{\gamma}_{1k}$ must contain zero restrictions, so that some variables in $\underline{x}_i$ affect only c_i , some variables in $\underline{x}_i$ affect only η_{ik}, and the rest of the variables in $\underline{x}_i$ affect both c_i and η_{ik}.

If η_{ik} are observed and measured on a continuous scale, the regression parameters in equations 4 and 5 can be readily estimated by two-stage least squares. However, as noted in the introduction, our data source does not provide directly observed measurements of an individual's stock of fiber-related information. Instead, the data contain an individual's responses to fiber-related questions that indirectly provide an indication of that person's underlying information levels. A further complication in measuring η_{ik} arises from the fact that the responses to the indicator questions are not continuous, but binary or ordinal. Ignoring for the moment this discrete nature of the indicator responses, suppose each of the fiber information variables has M_k indicators y^*_{ikm}, $m = 1, \dots, M_k$. We formalize the notion that the indicators or indirect measurements y^*_{ikm} are generated by an underlying unobserved or latent fiber information variable, η_{ik}, by the following measurement model:

$$y^*_{ikm} = \lambda_{0km} + \lambda_{1km}\eta_{ik} + \varepsilon^*_{ikm}, \tag{8}$$

where λ_{0km} is a scalar intercept, λ_{1km} is a scalar "loading" parameter, and ε^*_{ikm} is an independently distributed error term with the covariance structure:

$$cov(\varepsilon^*_{ikm},\varepsilon^*_{jln}) = \begin{cases} \theta^2_{km} \ for \quad i = j, \ k = l, \ and \ m = n \\ 0 \quad otherwise. \end{cases} \tag{9}$$

Thus, each of the latent fiber information variables is assumed to account for all the interdependencies among its respective indicators, so that y^*_{ikm} are conditionally independent, given η_{ik}. In addition, it is also assumed that:

$$cov(\zeta_{ik},\varepsilon^*_{jlm}) = 0 \ for \ all \ i,j,k,l, \ and \ m, \tag{10}$$

and

$$cov(\xi_i^*,\varepsilon^*_{jkm}) = 0 \ for \ all \ i,j,k, \ and \ m. \tag{11}$$

Equations 4, 5, and 8 constitute a structural equations model with latent variables (for example, Bollen, 1989). Identification of this model requires additional restrictions beyond the covariance restrictions in equations 6, 7, and 9-11. In general, identification requires that each latent variable has two or more indicators and that the origin and scale of each latent variable are fixed (Bollen, 1989, pp. 238-254). We set the origin and scale of η_{ik} by letting $\lambda_{0km} = 0$ and $\lambda_{1km} = 1$ for one of its indicators y^*_{ikm}. In some instances, a latent variable may have only one indicator. Identification can be achieved in such instances by imposing the additional restriction that $\varepsilon^*_{ikm} = 0$ so that $y^*_{ikm} = \eta_{ik}$. Thus, for a fiber information variable with only one indicator, the indicator would enter directly on the right hand side of equation 4.

If y^*_{ikm} are observed and continuous, our fiber consumption model would be complete and the estimation of the system equations 4, 5, and 8 can proceed by methods such as maximum likelihood (Bollen, 1989). In the present case, however, y^*_{ikm} are observed on a binary or ordinal scale. In this discrete variables case, y^*_{ikm} may be viewed as underlying indexes that generate binary or ordinal outcomes as they cross some unknown thresholds. We assume that for indicators y^*_{ikm} with binary outcomes, the observed response y_{ikm} taking a value of 0 or 1, is generated by the relation:

$$y_{ikm} = 1 \ \ if \ y^*_{ikm} > \mu_{km}, \ 0 \ otherwise, \tag{12}$$

and for indicators y^*_{ilm} ($l \neq k$) with ordinal outcomes, the observed response y_{ilm} taking a value of r, r = 1,...,R_m, is generated by the relation:

$$y_{ilm} = r \ \ if \ \ \mu_{lm,r-1} \leq y^*_{ilm} < \mu_{lm,r}, \tag{13}$$

where μ_{km} and $\mu_{lm,r}$ are unknown threshold values with normalizations $\mu_{lm,0} = -\infty$, $\mu_{lm,R_m} = +\infty$, and R_m is the number of categories of the mth ordinal indicator.

Assuming the error terms are normally distributed, the structural equation 5, the measurement equation 8, and the threshold relations in equations 12 and 13 constitute a probit latent variable model, first proposed by Muthen (1979). Substituting for η_{ik} in equations 4 and 8 from equation 5, we have the reduced forms:

$$C_i = \alpha_0 + \underline{\alpha}_1'\underline{x}_i + \xi_i, \tag{14}$$

where $\alpha_0 = \beta_0 + \underline{\beta}_1'\underline{\gamma}_0$, $\underline{\gamma}_0 = (\gamma_1 \ldots \gamma_K)'$, $\underline{\alpha}_1 = \Gamma_1'\underline{\beta}_1 + \underline{\beta}_2$, $\Gamma_1' = (\underline{\gamma}_1 \ldots \underline{\gamma}_K)$, and $\xi_i = \underline{\beta}_1'\underline{\zeta}_i + \xi^*_i$, $\underline{\zeta}_i = (\zeta_{i1} \ldots \zeta_{iK})'$, and

$$y^*_{ikm} = \pi_{0km} + \underline{\pi}_{1km}'\underline{x}_i + \varepsilon_{ikm}, \tag{15}$$

where $\pi_{0km} = \lambda_{0km} + \lambda_{1km}\gamma_{1k}$, $\underline{\pi}_{1km} = \lambda_{1km}\underline{\gamma}_{1k}$ and $\varepsilon_{ikm} = \lambda_{1km}\zeta_{ik} + \varepsilon^*_{ikm}$. Thus, the complete model in terms of observables consists of the fiber consumption equation and the binary and ordinal probit equations for the observed discrete indicators of the fiber information variables. Although the error terms ε^*_{ikm} are uncorrelated among themselves and with ξ^*_i, the reduced-form errors in equations 14 and 15 are correlated since:

$$cov(\varepsilon_{ikm},\varepsilon_{iln}) = \begin{cases} \lambda^2_{1km}\psi^2_k + \theta^2_{km} & for \ \ k = l \ and \ m = n \\ \lambda_{1km}\lambda_{1kn}\psi^2_k & for \ \ k = l \ and \ m \neq n \\ \lambda_{1km}\lambda_{1ln}\psi_{kl} & for \ \ k \neq l \ and \ all \ m,n, \end{cases} \tag{16}$$

and

$$cov(\xi_i,\varepsilon_{ikm}) = \underline{\beta}_1'\underline{\psi}_k\lambda_{1km}, \tag{17}$$

where $\underline{\psi}_k = (\psi_{1k} \ldots \psi_{Kk})$. Thus, the structural parameters of interest in equations 4, 5, and 8 are embedded nonlinearly in M correlated equations for the observables, where $M = 1 + M_1 + \ldots + M_K$. Application of the maximum likelihood estimation procedure to this model involves evaluation of the joint probabilities of observing different configurations of the indicator responses for each individual. Since such evaluation requires multi-dimensional integration, the maximum likelihood procedure is computationally complex except for small values of M. Models of this nature involving structural equations with latent variable and discrete indicators have been previously studied by Lee (1982), Muthen (1983), and Sobel and Arminger (1992). To estimate the model parameters, we adopt a computationally tractable three-stage procedure discussed in this literature.[1]

[1] See the appendix for a complete discussion of the estimation procedures we used.

Data and Variables

This research is based on analysis of the 1989 and 1990 USDA Continuing Survey of Food Intake of Individuals (CSFII) and the companion Diet and Health Knowledge Survey (DHKS) conducted by the Human Nutrition Information Service. The CSFII gathers information on the dietary intake of individuals and personal health-related data in households over a 3-day period. The first day's data were collected in a personal in-home interview using a 1-day dietary recall. The second and third days' data are from a 2-day dietary record kept by all respondents. Personal and household characteristics data such as income, age, weight, height, race, and education were also collected.

The 1989 and 1990 DHKS were conducted as a detailed 30-minute telephone follow-up survey to the CSFII. Individuals identified in the CSFII as the main meal planner/preparer for the household were contacted about 6 weeks after the CSFII and asked a series of questions about their diet and health knowledge, awareness, and attitudes.

The DHKS was designed so that information from it could be linked to information on food consumption from the CSFII. These surveys are the first nationally representative data to provide a combined and detailed look at both the intake of foods eaten as well as the individual's knowledge, attitudes, and awareness of diet and nutrition and their relationship to health and behavior.

To maintain a strong and consistent linkage between these two surveys, the analysis presented in this report is restricted to include only the main meal planner/preparer who reported 3 days of complete intake data. After merging DHKS data for 1989 and 1990 and eliminating cases with missing values, our final sample consists of 2,466 observations.

The DHKS data provide indicators for measuring respondents' knowledge of the fiber content of foods, their attitude toward consuming fiber-rich foods, and their awareness of health problems due to insufficient fiber intake. Table 1 presents these indicator questions and the percentage of correct and incorrect responses. Use of the terms "knowledge," "attitude," and "awareness" to describe groups of indicators is somewhat arbitrary and mainly done for ease of exposition. However, we believe these terms are the best, but probably not the only, describers of the underlying information content of the questions asked on the DHKS.

There are six questions on the DHKS that ask respondents to identify which of two food items has the higher fiber content. The pairs are meat or fruit, cornflakes or oatmeal, white or whole wheat bread, orange juice or apple, pretzels or popcorn, and lettuce or kidney beans. We use these questions to gauge respondents' knowledge of the fiber content of specific foods. In all cases, a relatively high percentage of respondents identified the high-fiber choice correctly (given in the table as the second item). For example, 93 percent of the respondents correctly identified whole wheat bread as a better source of fiber than white bread. The lettuce or kidney beans pair drew the lowest number of correct responses at 59 percent.

Attitude toward consuming fiber-rich food is measured by three indicator questions with ordinal response categories ranging from 1 for not at all important to 6 for very important. These questions are a general question about eating foods with adequate fiber, and two specific questions about consuming five servings a day of fruits and vegetables and six servings a day of breads, cereals, and grains. Note that the label "attitude toward consuming fiber-rich foods" could be changed to "general nutrition knowledge about fiber" without any loss of generality. More meal planners/preparers felt that eating at least six servings of breads, cereals, and other grain products was important or very important (37 percent) than those who felt it was slightly or not at all important (22 percent). Responses to the importance of eating at least five servings of fruits and vegetables follow a similar pattern: 45 percent consider it important or very important and 24 percent consider it slightly or not at all important. A large majority, 64 percent, said it is important for them to eat foods with adequate fiber; only 8 percent said it is not important.

The disease-awareness variable has only a single indicator (having heard about health problems related to how much fiber a person eats). We call this variable "disease-awareness," but we could have just as easily called it "knowledge about health problems." The sample was equally split between yes and no responses.

Average daily dietary fiber intake and fiber density (grams of dietary fiber consumed per 1,000 kilo calories of energy) from 3 days of data are reported in the last row of table 1. On average, the respondents consumed 12.1 grams of dietary fiber per day. The National Cancer Institute recommends daily intake of 20 to 30 grams for a 2,000-kilo calorie diet.

Table 1—Fiber responses, percentage of response, and fiber intake

Knowledge of the most fiber in food:						
			Incorrect	Correct		
			Percent			
Meat/fruit			20.4	79.6		
			(7.2)	(8.0)		
Cornflakes/oatmeal			18.3	81.7		
			(7.7)	(7.9)		
White/whole wheat bread			6.7	93.3		
			(6.8)	(7.9)		
Orange juice/apple			23.9	76.1		
			(7.7)	(7.9)		
Pretzels/popcorn			23.6	76.4		
			(7.2)	(8.0)		
Lettuce/kidney beans			41.3	58.7		
			(7.6)	(8.0)		
Attitude about importance of fiber:						
	Not at all important 1	2	3	4	5	Very important 6
			Percent			
Eat foods with adequate fiber	3.9	3.8	11.8	16.7	20.9	42.9
	(7.7)	(6.4)	(6.9)	(7.5)	(8.0)	(8.4)
Eat at least five servings per day of fruits and vegetables	15.7	8.4	13.1	17.6	14.6	30.7
	(7.7)	(6.9)	(7.1)	(7.3)	(8.1)	(8.9)
Eat at least six servings per day of breads, cereals, and other grain products	10.0	12.3	18.3	21.9	13.7	23.7
	(7.4)	(7.1)	(7.7)	(7.8)	(8.1)	(8.6)
Fiber-disease awareness:						
			No	Yes		
			Percent			
Heard about health problems related to fiber			49.8	50.2		
			(7.6)	(8.1)		
Dietary fiber intake			Mean	Standard deviation		
			Grams			
			12.1	6.3		
			(7.9)	(3.5)		

Note: Computed from 2,466 observations. The mean fiber density (grams of fiber per 1,000 kilo calories) is reported in parentheses.

Table 1 also reports the mean fiber densities for the categories of each fiber question. It is readily apparent that, on average, respondents with more fiber information, that is, those choosing the correct high-fiber food, those aware of health problems related to fiber, and those who consider it important or very important to eat fiber-rich food, also tend to have higher fiber consumption.

The explanatory variables used in our analysis are listed in table 2. Annual household income, household size, and the main meal planners' age and body mass index (BMI) are continuous variables, and the others are dummy variables. BMI is a ratio of the body weight (in kilograms) divided by the square of height (in meters). The region and location of the household and the main meal planner's education have multiple categories and are represented by groups of dummy variables with the Northeast, city, and less-than-high school as their respective omitted base categories.

The sample is heavily weighted toward females (80 percent), because in U.S. households a female is usually the main meal planner/preparer. This is important to remember when generalizing or interpreting our empirical findings. Blacks are more heavily represented in our sample than their population percentage would indicate, while Hispanics are slightly under-represented. Our sample also contains more high school and college graduates than the general population.

The explanatory variables fall into three broad categories. Income, program participation, household size, presence of children, region, and location of residence are household characteristics. Gender, race, education, employment status, and age are socio-demographic characteristics of the main meal planner. Smoking, vegetarianism, whether or not on a special diet, whether or not taking fiber supplements, and BMI are health-related or dietary characteristics of the main meal planner.

Except for income, we do not have *a priori* expectations about effects of the various household characteristics on knowledge, attitude, awareness, or consumption. The effect of income is complicated. On the one hand, a higher income may be expected to give better access to dietary information and thus affect knowledge, attitude, awareness, and

Table 2--Explanatory variables

Variable	Mean	Standard deviation
Income (1,000 dollars)	23.18	21.70
Program[1]	0.07	0.26
Household size	2.63	1.47
Children	0.41	0.49
Midwest[2]	0.25	0.43
South	0.34	0.47
West	0.20	0.40
Suburban[3]	0.47	0.50
Nonmetro	0.23	0.42
Female	0.80	0.40
Black	0.15	0.35
Hispanic	0.06	0.23
High school[4]	0.51	0.50
College	0.34	0.47
Post-graduate	0.10	0.31
Age (years)	47.56	18.33
Not employed	0.43	0.50
Smoker	0.26	0.44
Vegetarian	0.03	0.18
Special diet	0.16	0.37
Fiber supplement	0.06	0.24
BMI	25.70	5.51

Note: Household income and size, and the main meal planners' age and body mass index (BMI) are continuous variables, and the others are dummy variables.

[1] Participated in the food stamp or Women, Infants, and Children (WIC) program.

[2] Northeast omitted.

[3] City omitted.

[4] Less-than-high school omitted.

consumption positively. On the other hand, fiber-rich foods such as breads, grains, and cereals may be inferior goods whose consumption falls as income increases. *A priori* information is lacking on which of these effects will dominate. Thus, we must rely on the data and the model to determine the outcome.

Among socio-demographic variables, we expect the education variables to have positive effects since higher education enables better acquisition and processing of nutrition information. The traditional role of females in food preparation and shopping leads us to expect they have higher stocks of nutrition information than males. The race and age variables are expected to capture variations in fiber information, and food preferences and consumption induced by cultural backgrounds and dietary habits. If the main meal planner is not employed outside the home, this is likely to affect time allocation for food preparation versus other activities and hence may affect consumption.

Since grains, fruits, and vegetables are the major sources of dietary fiber, we expect vegetarians to be more familiar and informed about these fiber-rich foods and have relatively higher dietary fiber consumption. Conversely, smokers are probably less concerned about health issues and hence possess less dietary information. To the extent that dietary habits of smokers differ from non-smokers, fiber consumption would also be affected.

The variables "special diet" and "fiber supplement" are included to control for the likely higher fiber densities of respondents on a diet or the effects of taking fiber supplements. The BMI is included to control for the effects of variations in the amount of food consumed due to weight and height. We expect the BMI to be negatively correlated with fiber intake, because complex carbohydrates are bulky and less energy-dense than fats. Thus, individuals with higher BMI's may receive more of their kilo calories from foods high in fats and protein and fewer kilo calories from foods rich in complex carbohydrates (Canadian Paediatric Society, 1993).

Empirical Results

In empirical estimation, the consumption equation is specified in loglinear form with the dependent variable fiber density and the continuous independent variables income, age, and BMI converted to their logarithms. Thus, coefficients of the latter variables can be interpreted as elasticities.

An examination of unconditional tetrachoric correlations among the six pairs of knowledge indicators revealed that the lettuce/kidney beans pair may have confused the respondents. The correlations between the first five pairs listed in table 1 are all positive and relatively high, ranging from 0.252 to 0.579. However, the lettuce/kidney beans pair was negatively correlated with four of the other pairs, the highest correlation being -0.096. The only positive correlation was a relatively small 0.067 with meat/fruit. Therefore, we dropped the lettuce/kidney beans pair from further analysis.

In general, an explanatory variable may be expected to have similar directional effects on both fiber information and fiber consumption. However, some variables, such as diet-related variables, may act mainly through their direct effect on consumption, while the effect of other variables may be mostly through their influence on information. As suggested earlier, we do not have much theoretical guidance on which explanatory variable, if any, to exclude from either the consumption equation or the information equations. Therefore, in all our specifications, we include most variables in both the consumption and information equations and allow the estimated models to reveal the relative strengths of direct and indirect effects. The only restriction common across estimated models is the exclusion of BMI from the information equations. Thus, BMI is assumed to have only a direct effect on consumption.

However, if we allow the error terms between the consumption and the information equations to be freely estimated, some exclusionary restrictions on the consumption equation are necessary for identification. For this purpose, we restrict the effects of education, gender, and program participation to be entirely indirect, that is, these variables are excluded from the consumption equation, and their effect on consumption is exclusively through their effects on fiber information levels. Education is the most likely candidate for exclusion since it is the variable most closely related to the process of acquiring and using dietary information. Program participation is excluded since these programs have an educational component. Since our dependent variable is fiber intake relative to energy intake (that is, fiber density), it seems reasonable to assume that the remaining female-male difference is purely due to difference in information levels.

We begin by estimating a fully recursive model (labeled Model I), where the error terms of the consumption and the information equations are assumed to be uncorrelated. All variables except BMI enter both the information and consumption equations. BMI is excluded from the information equations. Model II relaxes the assumption on the error covariances and allows them to be free parameters. However, besides excluding BMI from the information equations as in Model I, Model II excludes education, gender, and program participation from the consumption equation.

Besides these two models, we consider some additional specifications to assess the relative influence of fiber information variables, as well as the relative role of direct and indirect effects of the explanatory variables on fiber consumption. All additional models have the same specification of the measurement and structural equations for the information variables and only the specification of the fiber consumption equation is changed. The parameter estimates of the measurement and structural equations for the information variables are nearly identical for Models I and II.

Fiber Information

Table 3 reports estimates of the measurement model for the fiber information variables for Model I. The last two columns of the table present two pseudo-R^2 measures for assessing the goodness-of-fit of the binary and ordinal indicators. $R^2_{\varepsilon^*}$ is the pseudo-R^2 for the measurement equation 8 and R^2_{ε} is the pseudo-R^2 for the reduced-form equation 15. They are computed (Laitila, 1993) as:

$$R^2_{\varepsilon^*} = 1 - \frac{var(\varepsilon^*_{km})}{var(y^*_{km})}, \tag{18}$$

and

$$R^2_{\varepsilon} = 1 - \frac{var(\varepsilon_{km})}{var(y^*_{km})}, \tag{19}$$

where

$$var(y^*_{km}) = \lambda^2_{1km}\underline{\gamma}'_{1k}var(\underline{x})\underline{\gamma}_{1k} + \lambda^2_{1km}\psi^2_k + \theta^2_{km}. \tag{20}$$

The two pseudo-R^2 measures differ with respect to the error terms involved in the calculation. $R^2_{\varepsilon^*}$ is based on the measurement error of the indicators whereas R^2_{ε} includes the measurement error of the latent information variables as well. Since the disease-awareness variable has only one indicator, its pseudo-R^2 values are identical, 0.17. For the knowledge and attitude indicators, the two pseudo-R^2 measures show that the latent variable formulation captures a large part of the indicator variation. The difference between the two pseudo-R^2 measures gives the amount of indicator variation explained by the error term of the information variable that the indicator is measuring. Therefore, the explained part of the latent information variables appears to account for a lesser extent of the indicator variation as compared with the unexplained part, particularly for the attitude indicators.

The estimated loading coefficients for all the knowledge and attitude indicators are highly significant (absolute t-values are given in parentheses). The estimated coefficients and the pseudo-R^2 values together show the relative quality of indicators as indirect measurements of the latent knowledge or attitude variables. Thus, the fiber content questions for the food pairs orange juice/apple, meat/fruit, and white/whole wheat bread are relatively better indicators of fiber knowledge than the cornflakes/oatmeal and pretzels/popcorn pairs. The importance to the respondents of eating at least five servings of fruits and vegetables is the weakest among attitude indicators.

Table 4 reports the structural estimates (γ_{0k} and $\underline{\gamma}_{1k}$) for fiber knowledge, attitude, and disease awareness for Model I. The last row of the table reports the pseudo-R^2 for the estimated equations, computed as:

$$R^2_{\zeta} = 1 - \frac{var(\zeta_k)}{var(\eta_k)}. \tag{21}$$

Table 3--Model I measurement estimates for fiber knowledge, attitude, and disease awareness

Indicator	λ_{0km}	λ_{1km}	$R^2_{\varepsilon^*}$	R^2_ε
Knowledge (η_1):				
Meat/fruit	-0.53	0.86	0.36	0.14
	(5.00)	(12.62)		
Cornflakes/oatmeal	-0.23	0.71	0.25	0.10
	(2.15)	(10.70)		
White/whole wheat bread[1]	0.00	1.00	0.46	0.18
Orange juice/apple	-0.59	0.88	0.37	0.15
	(5.13)	(11.79)		
Pretzels/popcorn	-0.45	0.72	0.26	0.10
	(3.93)	(10.77)		
Attitude (η_2):				
Foods with adequate fiber[1]	0.00	1.00	0.57	0.10
Fruits and vegetables	-0.10	0.60	0.22	0.04
	(1.49)	(16.33)		
Breads, cereals, grains	-0.11	0.71	0.30	0.05
	(1.35)	(17.00)		
Disease awareness (η_3):				
Fiber-related health problems[1]	0.00	1.00	0.17	0.17

Note: Absolute t-values are in parentheses.
[1] The coefficients are normalized to 0 and 1 for identification.

The equations fit reasonably well. The regressors explain 40 percent of the estimated variation in the latent knowledge variable and 17 percent each of the attitude and disease awareness variables. All three equations have a large number of statistically significant coefficients.

Given the normalization in the measurement model, a positive coefficient in the knowledge equation indicates the corresponding variable increases the odds that a person is more knowledgeable about the fiber content of foods. Positive coefficients in the attitude or awareness equations indicate that the corresponding variable increases the odds that a person thinks it is important to eat fiber-rich foods or is aware of health problems associated with fiber. We used a 10-percent level to test the statistical significance of parameter estimates. All explanatory variables have a significant effect on at least one fiber information variable. We focused on some prominent results. Income has a significant influence on all three fiber information variables. The result for income is interesting in one important respect: it is the only variable that has significant but opposite effects on knowledge and attitude. While a higher income level is associated with increased knowledge of the dietary fiber content of foods and fiber-related disease awareness, the effect on attitude toward eating fiber-rich foods is negative. This result supports the argument that higher income individuals tend to view fiber-rich foods, such as many grain products, as inferior goods.

Participation in either the food stamp or Women, Infants, and Children (WIC) program is associated with lower odds that the respondent is knowledgeable about the fiber content of foods and aware of fiber-disease relationships. This is somewhat surprising since some nutrition guidance is often provided in these programs. However, it is difficult to attribute causality because these programs target individuals in need of nutrition assistance.

Among other household characteristics, region and locality of residence are statistically significant in several instances. Residence in the West or Midwest increases the odds of disease-awareness compared with residence in the Northeast. Westerners are more likely than residents of the Northeast to believe it is important to eat fiber-rich foods, all else held constant. This finding lends some credence to the common perception that westerners are more health-conscious consumers than their peers living in other regions. The South has a significant negative coefficient on knowledge, indicating that those who score low on the fiber-content test are more likely to live in the South than in the Northeast.

Not surprisingly, female meal planners/preparers are more knowledgeable about fiber sources, more likely to believe it is important to eat fiber-rich foods, and more aware of fiber/health relationships than men. Likely explanations for this finding

Table 4--Model I structural estimates for fiber knowledge, attitude, and disease awareness

Explanatory variable	Knowledge (η_1)		Attitude (η_2)		Disease awareness (η_3)	
Constant	0.152	(0.39)	0.137	(0.36)	-1.724	(3.66)
Log income	0.129	(4.16)	-0.069	(2.48)	0.080	(2.07)
Program	-0.235	(3.00)	-0.098	(1.25)	-0.288	(2.59)
Household size	0.026	(1.22)	0.022	(1.05)	0.004	(0.15)
Children	-0.056	(0.85)	0.080	(1.33)	0.129	(1.62)
Midwest	-0.056	(0.86)	-0.061	(1.10)	0.285	(3.88)
South	-0.205	(3.45)	-0.081	(1.57)	0.066	(0.94)
West	-0.021	(0.33)	0.274	(4.58)	0.479	(6.04)
Suburb	0.085	(1.75)	-0.002	(0.05)	0.035	(0.59)
Nonmetro	0.130	(2.29)	-0.007	(0.14)	0.139	(1.98)
Female	0.385	(7.31)	0.227	(4.53)	0.123	(1.91)
Black	-0.371	(6.37)	-0.026	(0.52)	-0.154	(2.08)
Hispanic	-0.500	(6.07)	-0.191	(2.43)	-0.345	(3.34)
High school	-0.045	(0.50)	0.661	(7.08)	0.071	(0.55)
College	0.303	(3.06)	0.588	(5.87)	0.389	(2.86)
Post-graduate	0.443	(3.67)	0.687	(6.12)	0.951	(6.07)
Log age	-0.018	(0.29)	0.419	(6.39)	0.113	(1.39)
Not employed	-0.049	(1.07)	0.007	(0.17)	0.019	(0.32)
Smoker	-0.118	(2.59)	-0.156	(3.75)	-0.178	(3.10)
Vegetarian	0.423	(3.79)	0.340	(4.42)	0.262	(1.77)
Diet	0.255	(4.30)	0.242	(4.95)	0.323	(4.56)
Supplement	0.032	(0.36)	0.226	(2.90)	0.220	(2.00)
R^2_ζ	0.395		0.168		0.173	

Note: Absolute t-values are in parentheses.

are that females receive more formal and informal training in nutrition than males.

Ethnicity and race are influential variables for explaining the variation in knowledge, attitude, and awareness. Knowledge and awareness coefficients indicate that those with lower fiber information levels are more likely to be black or Hispanic. The negative effect on information levels is greater for Hispanics than for blacks. These findings probably relate to cultural barriers and limited access to print and electronic media sources for information. However, no black-white racial differences were found in the attitude toward eating fiber-rich foods.

Years of formal education weigh heavily in favor of increased knowledge levels, increased awareness, and better attitudes. Both college and post-graduate dummy variables are consistently significant at the 1-percent level. This evidence is consistent with the findings of Kushi and others, who found that a larger proportion of respondents with high educational attainment than those with lower educational attainment were able to answer dietary knowledge questions correctly.

One interesting result is that a high school education *vis-a-vis* a lower education level does not increase the odds of higher knowledge of fiber content or disease awareness. However, having at least a high school education has a positive effect on the attitude variable. A likely explanation for this difference is that the specificity of information associated with knowledge and awareness questions requires additional education whereas just a high school education can make a significant difference in accumulating the more general information associated with the attitude questions. A similar result appears for age: while an increase in age does not significantly increase knowledge of the fiber content of foods, it does increase the probability of a respondent's stressing the importance of eating fiber-rich foods as well as being aware of fiber-health linkages. This finding may result from differing dietary needs of the elderly and diet modification prompted by physiological changes that occur with aging. Controlling for others factors, there are no significant differences in the information levels of employed versus stay-at-home meal planners.

As expected, vegetarianism and smoking are two key personal habits that exert significant and opposing effects on all fiber information variables. The result for smoking is consistent with the notion that smokers are less health conscious than nonsmokers and, hence, less likely to gather or retain dietary information.

While the above results confirm several *a priori* expectations regarding the influence of individual characteristics, they also provide new insights into the differential effects of these characteristics on fiber knowledge, attitude, and disease awareness. As discussed above, race, high school education, age, and income have differing effects on attitude *vis-a-vis* knowledge. These distinctions are important in designing and targeting knowledge-related and attitude-related information for nutrition education.

The residual correlations ψ_{kl} give information on the strength of linkages between fiber knowledge, attitude, and disease awareness, after accounting for the influence of the explanatory variables. The estimated residual correlation (standard error) between knowledge and disease awareness is a relatively large 0.309 (0.029). Attitude, however, has a relatively low correlation of 0.076 (0.019) with knowledge and an insignificant correlation of 0.018 (0.024) with disease awareness. Thus, it appears that the attitude variable is capturing a distinct dimension of the consumer's fiber-related information set compared with fiber knowledge and disease awareness. The latter two variables, on the other hand, probably reflect much of the same aspects of fiber-related information.

Fiber Consumption

Table 5 reports the coefficient estimates of the fiber consumption equation for Models I and II, as well as three additional specifications labeled Models III through V. The three rows prior to the last row of table 5 present R^2 measures: $R^2_{\xi *}$ and R^2_ξ are two pseudo-R^2 measures for the consumption equation computed analogous to $R^2_{\epsilon *}$ and R^2_ϵ in equations 18 and 19. R^2_m is a pseudo-R^2 for the entire model (that is, the consumption equation plus the measurement and structural equations for the fiber information variables) computed as:

$$R^2_m = \frac{Q(\hat{\delta}_b) - Q(\hat{\delta}_m)}{Q(\hat{\delta}_b)}, \tag{22}$$

where $Q(\hat{\delta}_b)$ and $Q(\hat{\delta}_m)$ are the minimized values of the minimum distance function (equation 3 in the appendix) for a base model b and the hypothesized model m (Bollen, 1989). The last row of table 5 reports the Akaike Information Criterion (AIC) for each model.

The reported R^2_m values are calculated from a base model of which the structural and measurement equations have been restricted to include only the intercepts. Models I and II have identical overall fits with an R^2_m of 0.78. The R^2_ξ are also similar at 0.21. However, based on AIC, Model II is preferable over Model I. This suggests that freeing the error covariances between the consumption and information equations while excluding education, gender, and program participation from the consumption equation is a more valid specification than the recursive specification that restricts error covariances to zero while including education, gender, and program participation in the consumption equation. We tried alternative restrictions to identify Model II, such as excluding smoking from the consumption equation, but none gave as satisfactory results as Model II.

Turning first to coefficient estimates of the fiber information variables, it can be seen that these variables have significant positive effects on fiber consumption. Under Model I, coefficients of knowledge and attitude variables are significant, and under Model II, attitude and awareness have significant positive effects. However, the awareness coefficient is insignificant under Model I and the knowledge coefficient is insignificant under Model II. This result could be due to the knowledge and disease awareness variables capturing the same dimension of fiber information, as suggested earlier by an examination of their residual correlations. Further support for this interpretation is provided when Model II is re-estimated after dropping the knowledge variable from the consumption equation. The resulting model, Model III in table 5, shows a large increase in the size of awareness and attitude coefficients over Model II. The sum of the fiber information coefficients, however, remains approximately the same at 0.264 under Model II and 0.266 under Model III. Model III has the same R^2_ξ and R^2_m as Model II, but has a lower AIC than Model II. Therefore, we choose Model III as our preferred model. In the discussion below, we focus on Model III results, and refer to Model I results for drawing attention to the differences induced by the different specifications.

Table 5--Structural estimates for fiber consumption

Explanatory variable	Model I		Model II		Model III		Model IV		Model V	
Constant	-5.463	(32.77)	-5.259	(26.31)	-5.179	(26.27)	-5.445	(32.86)	-4.976	(25.72)
η_1	0.078	(2.48)	0.067	(1.13)	--		--		--	
η_2	0.084	(6.38)	0.071	(1.44)	0.093	(1.90)	--		0.129	(2.48)
η_3	-0.007	(0.49)	0.126	(2.53)	0.173	(5.35)	--		0.163	(4.98)
Log income	-0.038	(3.42)	-0.046	(3.41)	-0.040	(3.10)	-0.033	(3.16)	-0.035	(2.62)
Program	-0.057	(2.18)	--		--		-0.085	(3.29)	--	
Household size	-0.000	(0.02)	-0.001	(0.06)	0.001	(0.09)	0.004	(0.48)	-0.005	(0.45)
Children	-0.026	(1.08)	-0.044	(1.60)	-0.053	(1.90)	-0.019	(0.80)	-0.052	(1.91)
Midwest	0.038	(1.63)	-0.001	(0.02)	-0.017	(0.62)	0.026	(1.18)	-0.003	(0.10)
South	0.021	(0.99)	0.011	(0.41)	-0.005	(0.21)	-0.005	(0.23)	0.001	(0.03)
West	0.110	(4.41)	0.050	(1.33)	0.022	(0.69)	0.144	(6.41)	0.029	(0.99)
Suburban	0.022	(1.34)	0.020	(1.09)	0.024	(1.28)	0.031	(1.95)	0.028	(1.52)
Nonmetro	0.000	(0.02)	-0.016	(0.74)	-0.013	(0.59)	0.014	(0.74)	-0.016	(0.71)
Female	0.010	(0.48)	--		--		0.073	(3.98)	--	
Black	-0.101	(4.47)	-0.083	(3.07)	-0.100	(4.25)	-0.126	(6.39)	-0.096	(4.10)
Hispanic	0.076	(2.48)	0.110	(2.89)	0.094	(2.53)	0.032	(1.12)	0.094	(2.56)
High school	-0.007	(0.18)	--		--		0.049	(1.37)	--	
College	0.029	(0.72)	--		--		0.087	(2.22)	--	
Post-graduate	0.099	(2.18)	--		--		0.180	(4.15)	--	
Log age	0.248	(9.84)	0.241	(7.01)	0.227	(6.63)	0.270	(11.12)	0.210	(6.12)
Not employed	0.033	(2.07)	0.029	(1.64)	0.026	(1.40)	0.035	(2.25)	0.031	(1.68)
Smoker	-0.097	(6.24)	-0.078	(4.05)	-0.076	(3.68)	-0.120	(7.88)	-0.088	(4.42)
Vegetarian	0.158	(4.70)	0.128	(3.30)	0.133	(3.16)	0.222	(6.99)	0.126	(2.76)
Diet	0.008	(0.46)	-0.029	(1.21)	-0.033	(1.31)	0.042	(2.47)	-0.028	(1.17)
Supplement	0.044	(1.53)	0.019	(0.58)	0.011	(0.33)	0.053	(1.99)	0.004	(0.13)
Log BMI	-0.110	(3.15)	-0.114	(3.31)	-0.113	(3.28)	-0.100	(2.89)	-0.128	(3.81)
$R^2_{\xi^*}$	0.252		0.149		0.095		0.206		0.140	
R^2_{ξ}	0.213		0.211		0.211		0.214		0.206	
R^2_m	0.777		0.777		0.777		0.765		0.744	
AIC	747.954		744.710		743.842		818.690		--	

Note: The dependent variable is log fiber density. The sample size is 2,466. Absolute t-values are in parentheses. The -- indicates that the variable is not included in that model.

The above results, if verified in other studies, have widespread implications for nutrition education programs. The overall conclusion from the first three models is that, among the fiber information variables, disease awareness has the greatest effect on fiber consumption. This is not surprising given that avoiding health problems has the most immediate and transparent economic benefits to the consumer. Therefore, consumers possessing this type of information may be expected to actively change their dietary patterns compared with consumers who do not have this information, or with those who have only general notions about the benefits of dietary fiber. This suggests that nutrition education programs aimed at increasing the general awareness of fiber-health-disease linkages will likely have the highest payoffs in terms of raising dietary fiber intake levels. The results also suggest that there is a distinct attitudinal dimension beyond disease awareness that influences fiber consumption. Nutrition education strategies emphasizing general attitudinal messages such as the five-a-day intake are likely to have greater effect in modifying dietary patterns than strategies emphasizing specialized knowledge about the nutrient content of foods.

The above conclusions about the effects of fiber information on consumption apply without reference to any particular population subgroup. The relative importance of these effects can yet vary when specific subgroups are considered. These differences become clear when the direct and indirect effects of variables identifying the subgroups are discussed below.

The coefficient estimates of the observed explanatory variables reported in table 5 capture the variables' direct effect on fiber consumption. These variables also affect fiber consumption indirectly through their effects on the fiber information variables. These separate effects can be seen by rewriting the reduced-form consumption equation 14 as follows:

$$C_i = \alpha_0 + \beta_{11}\underline{\gamma}_1'\underline{x}_i + \beta_{12}\underline{\gamma}_2'\underline{x}_i + \beta_{13}\underline{\gamma}_3'\underline{x}_i + \underline{\beta}_2\underline{x}_i + \xi_i, \tag{23}$$

where β_{11}, β_{12}, and β_{13} are the coefficients of knowledge, attitude, and disease awareness variables. The second to fourth terms on the right hand side represent the indirect effect of the observed explanatory variables on consumption and the fifth term represents the direct effect. The sum of indirect and direct effects gives the total effect of observed explanatory variables on consumption. The indirect and total effects for Models I and III are reported in tables 6 and 7. For convenience, the respective direct effects are repeated in the first column of each table.

The coefficients of the dummy variables in tables 6 and 7 can be interpreted as the approximate percentage change in fiber density for the respective category compared with the base category (the exact percentage change is given by $100\{\exp(\beta)-1\}$, where β is the coefficient estimate). The coefficients for income, age, and BMI are elasticities. In tables 8 and 9, we translate selected coefficient estimates to their quantitative effects by computing the changes in predicted fiber density due to changes in an explanatory variable, holding other variables constant.

Household income, the primary economic variable, has a significant negative total effect on dietary fiber intake. Despite the different specifications, the direct and total income effects remain similar under Models I and III. The income elasticity is -0.03, which implies a reduction in fiber density by about 0.03 gram for a 10-percent rise in income. The direct effect of income dominates, confirming that as income rises, consumption might be shifting to foods low in fiber. Our result implies that an earlier finding in the context of developing countries' nutrition, suggesting that an increase in income by itself may not be enough to improve nutrition (Behrman and Deolalikar, 1987), may be equally applicable to developed countries.

Under Model III, fiber density of meal planners from households participating in the food stamp or WIC programs is about a half gram lower than respondents from nonparticipating households. Most of the lower consumption occurs through lower fiber-health awareness. Attitude effect is insignificant. This is consistent with the earlier result that attitude toward consuming fiber-rich food is not significantly different between program participants and non-participants. Overall, the results suggest that if participants are targeted for nutrition education, greater emphasis on increasing disease awareness and fiber knowledge would be desirable.

Among other household variables, household size and presence of children do not have much effect on fiber consumption. Among regions, only the West has significant indirect and total effects under Model III, and both effects are positive. The paths through which residency in the West affects consumption is different under Model III as compared with Model I. Under Model I, 86 percent of the total effect is direct. The direct effect becomes insignificant under Model III and the total effect is dominated by the indirect effect through greater awareness. Despite this difference, however, the total effect is quite similar under the two models. The Model III result seems more plausible given the large effect that residency in the West has on awareness as shown in table 4. Based on Model III, a change from the Northeast to the West increases fiber consumption by 14 percent, which translates, on average, to 1 gram of dietary fiber per 1,000 kilo calories. About 64

Table 6--Direct, indirect, and total effects of explanatory variables on fiber consumption for Model I

Explanatory variable	Direct effect		Indirect effect: Knowledge		Indirect effect: Attitude		Total effect	
Log income	-0.038	(3.42)	0.010	(2.07)	-0.006	(2.34)	-0.034	(3.26)
Program	-0.057	(2.18)	-0.018	(1.92)	-0.008	(1.23)	-0.082	(3.17)
Household size	-0.000	(0.02)	0.002	(1.10)	0.002	(1.05)	0.004	(0.45)
Children	-0.026	(1.08)	-0.004	(0.81)	0.007	(1.30)	-0.024	(1.03)
Midwest	0.038	(1.63)	-0.004	(0.82)	-0.005	(1.09)	0.026	(1.17)
South	0.021	(0.99)	-0.016	(2.05)	-0.007	(1.53)	-0.002	(0.09)
West	0.110	(4.41)	-0.002	(0.33)	0.023	(3.60)	0.128	(5.67)
Suburban	0.022	(1.34)	0.007	(1.42)	-0.000	(0.05)	0.028	(1.76)
Nonmetro	0.000	(0.02)	0.010	(1.74)	-0.001	(0.15)	0.009	(0.47)
Female	0.010	(0.48)	0.030	(2.32)	0.019	(3.74)	0.058	(3.15)
Black	-0.101	(4.47)	-0.029	(2.39)	-0.002	(0.52)	-0.131	(6.60)
Hispanic	0.076	(2.48)	-0.039	(2.28)	-0.016	(2.31)	0.024	(0.82)
High school	-0.007	(0.18)	-0.003	(0.49)	0.056	(4.73)	0.045	(1.24)
College	0.029	(0.72)	0.024	(1.99)	0.050	(4.33)	0.099	(2.52)
Post-graduate	0.099	(2.19)	0.034	(2.12)	0.058	(4.45)	0.185	(4.26)
Log age	0.248	(9.84)	-0.001	(0.28)	0.035	(4.62)	0.281	(11.53)
Not employed	0.033	(2.07)	-0.004	(0.99)	0.001	(0.17)	0.029	(1.88)
Smoker	-0.097	(6.24)	-0.009	(1.75)	-0.013	(3.25)	-0.118	(7.74)
Vegetarian	0.158	(4.70)	0.033	(2.10)	0.029	(3.57)	0.218	(6.85)
Diet	0.008	(0.46)	0.020	(2.10)	0.020	(4.00)	0.046	(2.67)
Supplement	0.044	(1.53)	0.003	(0.36)	0.019	(2.69)	0.064	(2.38)
Log BMI	-0.110	(3.15)	--		--		-0.110	(3.15)

Note: Absolute t-values are in parentheses. Direct and indirect effects may not add up to the total effect, because indirect effect through disease awareness is not reported. The -- indicates that the variable is not included in that model.

percent (0.63 gram) of this additional fiber consumption is due to greater awareness of health problems related to fiber. The midwestern and the southern respondents do not have significantly different fiber consumption than the northeasterners. Suburban residents have about a 3-percent higher fiber density than city residents, while fiber densities of nonmetro residents and city residents are not significantly different.

Socio-demographic variables have considerable influence on fiber consumption. The direct effect of gender on fiber density is statistically insignificant under Model I. This implies that, holding knowledge and attitude constant, female respondents do not have significantly different fiber intakes than males. However, the indirect effects are positive and significant, contributing to a significant total effect. Under Model III, we restrict gender to have only indirect effects. Model III estimates imply that females have 5 percent (a third of a gram) higher fiber density than males, contributed almost evenly by better awareness and attitude.

Black respondents have significantly lower fiber consumption compared with other races; 12 percent lower (about 0.9 gram) dietary fiber per 1,000 kilo calories, under Model III. While most of this is a direct effect, 21 percent of this lower consumption is due to the lower awareness of black respondents regarding health problems associated with low-fiber intake. It is interesting to note that the attitude effect is insignificant. Thus, nutrition education programs targeted toward blacks may need to emphasize fiber-health linkages.

The effects due to the Hispanic ethnicity of respondents provide perhaps the best illustration of the importance of isolating direct and indirect effects on consumption. Under Model III, for a given level of awareness and attitude, the direct effect of Hispanic ethnicity is to increase fiber density by about 10 percent (0.7 gram) over respondents of other ethnic groups. However, Hispanics' lower levels of disease awareness, and to a lesser extent, lower levels of concern about eating fiber-rich foods, exert indirect negative effects on their fiber intake, rendering their total fiber consumption not significantly different from that of other ethnic groups.

Table 7--Direct, indirect, and total effects of explanatory variables on fiber consumption for Model III

Explanatory variable	Direct effect		Indirect effect: Awareness		Indirect effect: Attitude		Total effect	
Log income	-0.040	(3.10)	0.013	(1.80)	-0.006	(1.51)	-0.033	(3.29)
Program	--		-0.058	(3.21)	-0.010	(1.11)	-0.068	(3.62)
Household size	0.001	(0.09)	0.001	(0.17)	0.002	(0.93)	0.004	(0.47)
Children	-0.053	(1.90)	0.023	(1.55)	0.007	(1.09)	-0.023	(0.98)
Midwest	-0.017	(0.62)	0.050	(3.18)	-0.006	(0.99)	0.027	(1.21)
South	-0.005	(0.21)	0.011	(0.91)	-0.008	(1.20)	-0.001	(0.10)
West	0.022	(0.69)	0.082	(3.95)	0.025	(1.75)	0.129	(5.72)
Suburban	0.024	(1.28)	0.006	(0.60)	-0.000	(0.04)	0.030	(1.87)
Nonmetro	-0.013	(0.59)	0.024	(1.83)	-0.001	(0.13)	0.010	(0.52)
Female	--		0.025	(2.33)	0.021	(1.79)	0.046	(3.49)
Black	-0.100	(4.25)	-0.027	(1.97)	-0.002	(0.45)	-0.129	(6.54)
Hispanic	0.094	(2.53)	-0.058	(2.79)	-0.018	(1.48)	0.019	(0.65)
High school	--		0.004	(0.20)	0.061	(1.92)	0.065	(2.23)
College	--		0.061	(2.61)	0.054	(1.90)	0.115	(3.98)
Post-graduate	--		0.152	(4.65)	0.062	(1.89)	0.215	(5.94)
Log age	0.227	(6.63)	0.019	(1.31)	0.039	(1.82)	0.285	(11.89)
Not employed	0.026	(1.40)	0.004	(0.37)	0.000	(0.11)	0.030	(1.92)
Smoker	-0.076	(3.68)	-0.030	(2.52)	-0.014	(1.71)	-0.120	(7.93)
Vegetarian	0.133	(3.16)	0.045	(1.73)	0.031	(1.74)	0.209	(6.72)
Diet	-0.033	(1.31)	0.057	(3.63)	0.023	(1.81)	0.047	(2.73)
Supplement	0.011	(0.33)	0.037	(1.79)	0.021	(1.53)	0.068	(2.58)
Log BMI	-0.113	(3.28)	--			--	-0.113	(3.28)

Note: Absolute t-values are in parentheses. The -- indicates that the variable is not included in that model.

Indirect effects dominate the total effect of the education variables on consumption, especially for the college-educated, under Model I. In Model III, we impose the restriction that education has no direct effect on consumption. Aside from test criteria cited earlier, this restriction is supported by the relative sizes of the education effects under Models I and III. Despite the restriction, the total effect of the education variables rise considerably under Model III. High school education adds about 7 percent (half a gram) of fiber density and a college education adds about 12 percent (0.83 gram) higher fiber density over the fiber density of those with less than a high school education. The effect of post-graduate education is even more substantial: the indirect effects through awareness and attitude contribute over 1 gram and about half a gram of higher fiber density. Adding the direct and indirect effects indicates that post-graduate respondents have 1.64 grams of higher fiber density than respondents with less than a high school education.

Dietary fiber intake increases significantly with increases in the respondent's age. The age elasticity is 0.3, which translates to slightly over two-tenths of a gram of higher fiber density for a 10-percent increase in age. Like income, the age effect is also principally direct. Unlike most previous variables, however, the contribution of the indirect effect to the total effect is dominated by attitude. This is in line with the earlier finding that age affects attitude much more than it affects disease awareness (see table 4).

Meal planners who are not employed full time outside the home consume about 0.2 gram more fiber per 1,000 kilo calories than those with outside employment. Interestingly, the entire positive effect of not working outside the home is direct. The implied difference in dietary patterns may be related to the different time allocations for food preparation and shopping by the employed and not employed groups.

As expected, smokers consume significantly lower amounts of dietary fiber than nonsmokers: a total of 11 percent lower (0.9 gram) in terms of fiber density under Model III. Sixty-three percent (0.7 gram) of this lower consumption is through the direct effect, indicating that smokers and nonsmokers have different dietary patterns even after adjusting for fiber information levels. The lower level of smokers' awareness of fiber-health linkages is also quantitatively important, reducing

Table 8--Average changes in fiber consumption (in grams per 1,000 kilo calories) due to changes in selected explanatory variables: Direct, indirect and total effects for Model I

Explanatory variable	Direct effect	Indirect effect		Total effect
		Knowledge	Attitude	
Income	-0.028	0.007	-0.004	-0.025
Program	-0.414	-0.132	-0.059	-0.591
West	0.847	-0.012	0.177	0.986
Female	0.074	0.218	0.139	0.425
Black	-0.714	-0.205	-0.016	-0.927
Hispanic	0.571	-0.292	-0.120	0.177
High school	-0.049	-0.025	0.396	0.319
College	0.211	0.172	0.362	0.725
Post-graduate	0.759	0.263	0.442	1.412
Age	0.184	-0.001	0.026	0.209
Not employed	0.242	-0.028	0.005	0.217
Smoker	-0.700	-0.066	-0.095	-0.852
Vegetarian	1.302	0.270	0.236	1.793
Diet	0.062	0.149	0.154	0.348
Supplement	0.334	0.019	0.146	0.487
BMI	-0.081	--	--	-0.081

Note: Figures for income, age, and BMI are for a 10-percent increase in the respective variable. Direct and indirect effects may not add up to the total effect, because indirect effect through disease awareness is not reported. The -- indicated that the variable is not included in that model.

Table 9--Average changes in fiber consumption (in grams per 1,000 kilo calories) due to changes in selected explanatory variables: Direct, indirect and total effects for Model III

Explanatory variable	Direct effect	Indirect effect		Total effect
		Awareness	Attitude	
Income	-0.030	0.010	-0.005	-0.025
Program	--	-0.422	-0.070	-0.492
West	0.165	0.632	0.194	0.991
Female	--	0.183	0.153	0.336
Black	-0.708	-0.190	-0.015	-0.914
Hispanic	0.704	-0.434	-0.131	0.139
High school	--	0.029	0.427	0.456
College	--	0.441	0.390	0.831
Post-graduate	--	1.160	0.475	1.635
Age	0.168	0.014	0.029	0.211
Not employed	0.190	0.028	0.003	0.221
Smoker	-0.548	-0.214	-0.104	-0.866
Vegetarian	1.086	0.372	0.255	1.713
Diet	-0.248	0.431	0.171	0.354
Supplement	0.083	0.280	0.157	0.519
BMI	-0.084	--	--	-0.084

Note: Figures for income, age, and BMI are for a 10-percent increase in the respective variable. The -- indicates that the variable is not included in that model.

fiber density by two-tenths of a gram. This result supports the previous finding by McPhillips and others (1994) that smokers eat a less healthful diet, and specifically, consume less dietary fiber than nonsmokers. Such dietary behavior on the part of smokers may be related to evidence that smokers underestimate health risks relative to nonsmokers (Viscusi, 1990).

Although vegetarians constitute only 3 percent of the sample, they form the most distinctive population subgroup in the dietary fiber profile. Controlling for other factors, vegetarians, on average, consume about 23 percent (1.7 gram) per 1,000 kilo calories more dietary fiber than non-vegetarians. Not surprisingly, having better awareness and attitude plays a lesser role in this higher consumption than the fact that vegetarian diet in itself is rich in dietary fiber. Thus, even after controlling for awareness and attitude levels, vegetarians have 14 percent (1.1 grams) higher fiber density than non-vegetarians.

The effects of being on a special diet and use of fiber supplement provide further evidence on the importance of isolating the direct and indirect paths to dietary differences. Since both variables are diet related, one may expect direct effects to dominate. However, the results show that the significant positive effects of these variables are almost entirely due to indirect effects. Those on a special diet tend to possess a higher level of disease awareness, as well as a better attitude toward consuming fiber-rich food. This higher information level causes these individuals to have about 5 percent (0.35 gram) higher fiber density than those not on a special diet. The indirect effects of awareness and attitude are even larger for those using fiber supplements, contributing to about half a gram of higher fiber density for this group as compared with those not using supplements.

Finally, as expected, the elasticity of dietary fiber consumption with respect to body mass is negative and significant. For a 1-percent increase in body mass, the amount of dietary fiber consumed falls by 0.1 percent. For a 10-percent increase in BMI, this translates to a decrease of 0.08 gram of dietary fiber per 1,000 kilo calories.

The above results affirm the key role of fiber information variables in determining dietary fiber intake and describes direct and indirect channels through which various consumer characteristics influence consumption. Several results regarding the relative influence of various characteristics are notable. While the large effect of vegetarianism is expected, the relative strength of college and post-graduate education is somewhat surprising. Our results are in agreement with the findings of Kushi and others (1988) on educational attainment and nutrient intake. They too found a significant positive association between fiber intake and level of education. Our results support their conjecture that the level of nutrition information may mediate the relationship between educational attainment and nutrient intake.

A notable aspect of our results with respect to the effect of education is the shift in relative importance from attitude to disease awareness as education rises from high school to the post-graduate level. Almost the entire effect of a high school education occurs through a higher level of concern toward consuming fiber-rich foods. The high school-educated tend to be no more aware of fiber-health linkages than those with a lesser education, and therefore, the contribution of disease awareness to fiber density is insignificant. However, the contribution of awareness rises considerably for the college-educated and even more substantially for those with post-graduate education whereas the contribution of attitude does not differ much from that for the high school-educated. This points to the likely high pay-off for a strategy targeting fiber-health awareness campaigns to those with less than a college education.

Two results providing clear testimony to the importance of isolating direct and indirect effects emerge from the coefficients for the female and Hispanic variables. Earlier studies have shown that females have higher dietary fiber intake than males (Kushi and others, 1988). The present results show that the higher level of fiber consumption by female respondents as compared with males may be chiefly due to females' higher level of knowledge and attitude. In the case of Hispanics, higher fiber consumption from ethnically related diet choices is negated by lower fiber consumption from diet choices associated with lower knowledge and attitude levels. Thus, targeting nutrition information toward male meal planners and Hispanic meal planners might be a useful strategy.

To further test the overall effect of the fiber information on fiber intake, we estimated Model IV, which excludes all three fiber information variables from the consumption equation. Both R_ξ^2 and R_m^2 fall for this model and the AIC is higher, all indicating poorer fit.

Some Final Observations

This report has used a new econometric framework for modeling unobserved effects on nutrient consumption and applied it to explain variations in dietary fiber intakes from the 1989-90 CSFII data. Researchers often add socio-demographic variables in nutrient intake equations to capture variation unexplained by purely economic variables, such as prices and income. However, the exact role of these variables in the consumers' decisionmaking process is left unclear. Our results provide strong evidence that the socio-demographic variables may be mediating the relationship between consumers' awareness of diet-health linkages, their attitudes toward consuming healthier foods, and their actual nutrient consumption. Identifying these relationships is useful for designing, targeting, and evaluating nutrition education programs for specific population subgroups.

However, the strength of the relationships we have identified should be interpreted subject to an important caveat. The ability of our latent variable measurement model to capture the true variation in fiber information levels is only as good as the observable indicators of these variables. To the extent that the survey questions underestimate the true awareness, attitude, and knowledge levels, our results will underestimate the effects of these variables on fiber consumption. Such mismeasurement can result from such factors as the question format and wording. As an example of the effect this may have on the results, consider the attitude questions in table 1. Although the mean fiber density increases from the second to the sixth category for each of the three attitude questions, the first category has a higher mean fiber density than the second, third, and fourth categories for the first two attitude questions. This could be because some respondents are choosing the first category (the option that they consider it not at all important to eat foods with adequate fiber, for example) since they feel they already consume enough dietary fiber. To get an idea of the quantitative effect of this problem, we converted the six-point attitude scale to a two-point scale by combining categories 1 to 4 to one category and categories 5 and 6 to a second category. After the recoding, the attitude variables, therefore, have three binary indicators. We re-estimated Model III with these binary attitude indicators and the results are reported under Model V, table 5. The estimated attitude coefficient increases to 0.129 from 0.093 under Model III. The attitude variable increases in importance relative to the awareness variable.

A second important caveat relates to the generalization of our results. First, our sample consists of main meal planners. To generalize the results to a wider population, we need information on the cross effects of meal planners' knowledge and attitude levels on the intakes of other household members. Second, we have focused on a single nutrient: dietary fiber. It would be useful to investigate knowledge and attitude effects on other key dietary components such as fat and cholesterol. We hope to address these issues in our future research.

References

Becker, G. S. "A Theory of the Allocation of Time," Economic Journal. Vol. 75 (1965): 493-517.

Behrman, J. R., and A. B. Deolalikar. "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India," Journal of Political Economy. Vol. 95 (1987): 492-507.

Behrman, J. R., and A. B. Deolalikar. "Health and Nutrition," Handbook of Development Economics. H. Chenery and T. N. Srinivasan (eds.), North-Holland, Amsterdam, Vol. I (1988): 631-771.

Block, G., B. Patterson, and A. Subar. "Fruits, Vegetables, and Cancer Prevention: A Review of the Epidemiological Evidence," Nutrition and Cancer. Vol. 18 (1992): 1-29.

Bollen, K. A. Structural Equations with Latent Variables. Wiley, New York, 1989.

Browner, W. S., J. Westenhouse, and J. A. Tice. "What if Americans Ate Less Fat? A Quantitative Estimate of the Effect on Mortality," Journal of the American Medical Association. Vol. 265 (1991): 3285-3291.

Canadian Paediatric Society. "Nutrition Recommendations Update... Dietary Fat and Children." Report of the Joint Working Group of the Canadian Paediatric Society and Health Canada, Ottawa, 1993.

Capps, O., Jr., and J. D. Schmitz. "A Recognition of Health and Nutrition Factors in Food Demand Analysis," Western Journal of Agricultural Economics. Vol. 16 (1991): 21-35.

Chamberlain, G. "Panel Data," Handbook of Econometrics, Volume II. Z. Griliches and M. D. Intriligator (eds.), North-Holland, Amsterdam (1984): 1247-1318.

Doll, R., and R. Peto. "The Causes of Cancer: Quantitative Estimates of Available Risks of Cancer in the United States Today," Journal of the National Cancer Institute. Vol. 66 (1981): 1191-1308.

Fogel, R. W. "Economic Growth, Populations Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." Working Paper No. 4638, National Bureau of Economic Research, Cambridge, MA (1994).

Frazao, B. "The High Costs of Poor Diets," Food Review. 17 (1994): 2-3.

Greene, W. H. Econometric Analysis. Macmillan, New York, 1990.

Kushi, L. H., A. R. Folsom, D. R. Jacobs, R. V. Luepker, P. J. Elmer, and H. Blackburn. "Educational Attainment and Nutrient Consumption Patterns: The Minnesota Heart Survey," Journal of the American Dietetic Association. Vol. 88 (1988): 1230-1236.

Kushi, L. H., and T. E. Kottke. "Dietary Fat and Coronary Heart Disease: Evidence of a Causal Relation," Preventing Disease: Beyond the Rhetoric. R. B. Goldbloom and R. S. Lawrence (eds.), Springer-Verlag, New York (1990): 385-400.

Laitila, T. "A Pseudo-R^2 Measure for Limited and Qualitative Dependent Variable Models," Journal of Econometrics. Vol. 56 (1993): 341-356.

Lancaster, K. Consumer Demand: A New Approach. Columbia University Press, New York, 1971.

Lee, L-F. "Health and Wage: A Simultaneous Equation Model with Multiple Discrete Indicators," International Economics Review. Vol. 23 (1982): 199-221.

McGinnis, J., and W. Foege. "Actual Causes of Death in the United States," Journal of the American Medical Association. Vol. 18 (1993): 2207-2212 .

McPhillips, J. B., C. B. Eaton, K. M. Gans, and others. "Dietary Differences in Smokers and Nonsmokers from Two Southeastern New England Communities," Journal of the American Dietetic Association. Vol. 94 (1994): 287-292.

Muthen, B. "A Structural Probit Model with Latent Variables," Journal of the American Statistical Association. Vol. 74 (1979): 807-811.

Muthen, B. "Latent Variable Structural Equation Modeling with Categorical Data," Journal of Econometrics. Vol. 22 (1983) 43-65.

National Center for Health Statistics. "Advance Report of Final Mortality Statistics, 1991," Monthly Vital Statistics Report. Vol. 42, No. 2, suppl. Hyattsville, MD: Public Health Service, 1993.

Oliver, M. F. "Introduction: Lipids and Coronary Disease - Resolved and Unresolved Problems," British Medical Bulletin. Vol. 46 (1990): 865-872.

Pitt, M. M., and M. R. Rosenzweig. "Health and Nutrient Consumption Across and within Farm Households," The Review of Economics and Statistics. Vol. 67 (1985): 212-223.

Shepherd, R., and L. Stockley. "Fat Consumption and Attitude Towards Food with a High Fat Content," Human Nutrition: Applied Nutrition. Vol. 39A (1985): 431-442.

Sobel, M. E., and G. Arminger. "Modeling Household Fertility Decisions: A Nonlinear Simultaneous Probit Model," Journal of the American Statistical Association. Vol. 87 (1992): 38-47.

U.S. Department of Agriculture and U.S. Department of Health and Human Services. Nutrition and Your Health: Dietary Guidelines for Americans. 3rd edition, HGB-232, 1990.

U.S. Department of Health and Human Services, Public Health Service. The Surgeon General's Report on Nutrition and Health. DHHS (PHS) Publication No. 88-50210 (1988).

U.S. Department of Health and Human Services, Public Health Service. Healthy People 2000. DHHS (PHS) Publication No. 91-50212, (1991).

Viscusi, K. W. "Do Smokers Underestimate Risks?" Journal of Political Economy. Vol. 98 (1990): 1253-1269.

Willett, W. C. "Diet and Health: What Should We Eat?" Science. Vol. 264 (1994): 532-537.

Appendix: Model Estimation

The estimation and inference framework for our model is based on the minimum distance method (for example, Chamberlain, 1984). Let $\underline{y}_i = (c_i\ y^*_{i11}\ \ldots\ y^*_{iKM_K})'$ be an $(M\times1)$ vector of the reduced-form dependent variables and let $\underline{\delta}$ be a $(Q\times1)$ vector of all the unknown model parameters. Then, our model can be expressed in the form:

$$\underline{y}_i = \Omega(\underline{\delta})\underline{\tilde{x}}_i + \underline{\upsilon}_i, \tag{1}$$

where $\underline{\tilde{x}}_i = (1\ \underline{x}_i')'$, $\Omega(\underline{\delta})$ is an $(M\times\tilde{P})$ matrix, $\tilde{P} = P+1$, and $\underline{\upsilon}_i = (\xi_i\ \varepsilon_{i11}\ \ldots\ \varepsilon_{iKM_K})'$ is an independently distributed vector of error terms with the covariance structure:

$$cov(\underline{\upsilon}_i, \underline{\upsilon}_i) = \Sigma(\underline{\delta}). \tag{2}$$

The elements of Ω and Σ are twice continuously differentiable functions of the unknown parameter vector $\underline{\delta}$. Let f be a vector of the elements of Ω, the nonredundant elements of Σ, and the elements of $\underline{\mu}$, a vector of the threshold parameters mapping y^*_{ikm} to y_{ikm}. Let $\underline{\omega}$ be an estimator of f. Then, our objective is to estimate $\underline{\delta}$ from an estimate of $\underline{\omega}$, under the restriction $\underline{\omega} = f(\underline{\delta})$. Suppose $\hat{\underline{\omega}}$ is an estimate of $\underline{\omega}$ obtained from appendix equation 1. Then, the minimum distance estimator for $\underline{\delta}$ is the solution to:

$$\min[\hat{\underline{\omega}} - f(\underline{\delta})]' \hat{W}^{-1} [\hat{\underline{\omega}} - f(\underline{\delta})], \tag{3}$$

where $\mathbf{W}$ is a positive definite matrix. For the optimal choice of $\hat{\mathbf{W}}$, it has been shown that:

$$\sqrt{N}(\hat{\underline{\delta}} - \underline{\delta}^0) \xrightarrow{D} N(\underline{0}, [F'W^{-1}F]^{-1}), \tag{4}$$

where $\underline{\delta}^0$ is the true value of $\underline{\delta}$ and $\mathbf{F} = \partial f(\underline{\delta}^0)/\partial\underline{\delta}'$ (Chamberlain, 1984). The optimal $\hat{\mathbf{W}}$ for this minimum distance estimator is provided by the consistent estimate of the variance-covariance matrix of $\underline{\omega}$. The optimal minimum distance estimator has several attractive properties. Chamberlain has shown that it is more efficient than the quasi-maximum likelihood estimator for non-normal distributions. It is also robust against general forms of heteroscedasticity.

The parameter vector $\underline{\delta}$ is estimated in two stages. In the first stage, the reduced-form consumption equation 14 is estimated by ordinary least squares and the reduced-form probit equation 15 is estimated by maximizing their marginal maximum likelihood functions. The marginal probit equation for a binary indicator is:

$$Pr[y_{ikm} = 1 | \underline{x}_i] = \Phi(\pi_{0km} + \underline{\pi}_{1km}'\underline{x}_i - \mu_{km}), \tag{5}$$

and for an ordinal indicator is:

$$Pr[y_{ilm} = r | \underline{x}_i] = \Phi(\mu_{lm,r} - \underline{\pi}_{0lm} - \underline{\pi}_{1lm}'\underline{x}_i) - \Phi(\mu_{lm,r-1} - \underline{\pi}_{0lm} - \underline{\pi}_{1lm}'\underline{x}_i),\ l \neq k, \tag{6}$$

where Φ is the normal cumulative distribution function. For identifying appendix equations 5 and 6, the threshold parameters μ_{km} for the binary indicators and $\mu_{lm,1}$ for the ordinal indicators are normalized to zero (since equation 15 includes an intercept), and the error variances of ε_{ikm} are normalized to 1 (for example, Greene, 1990). These M single-equation estimations yield $\hat{\Omega}$, $\hat{\underline{\mu}}$, and $\hat{\sigma}_1^2$, the error variance of the reduced-form fiber consumption equation. Note that the normalization of the error variances of ε_{ikm} to identify the probit equations implies that the rest of the diagonal elements of Σ are equal to one.

In the second stage, the reduced-form error covariances, or the off diagonal elements of Σ, are estimated by maximizing the bivariate likelihood functions for all pairs of observables, conditional on the first-stage parameter estimates and normalizations. For example, the conditional joint probability of the reduced-form fiber consumption equation and the sth binary indicator in appendix equation 1, is given by:

$$Pr[y_{ikm}=1,\ C_i | \underline{x}_i] = \int_{-\infty}^{z_s} \phi_2(u, z_1, \hat{\sigma}_1^2, \sigma_{1s})du, \tag{7}$$

where $z_1 = \hat{\alpha}_0 + \hat{\underline{\alpha}}_1'\underline{x}_i$, $z_s = \hat{\pi}_{0km} + \hat{\underline{\pi}}_{1km}'\underline{x}_i$, and ϕ_2 is the bivariate normal density function. The conditional joint probability of the sth binary indicator and the tth ordinal indicator is given by:

$$Pr[y_{ikm}=1,\ y_{ilm}=r\,|\underline{x}_i] = \int_{-\infty}^{z_s}\int_{z_{t,r-1}}^{z_{t,r}} \phi_2(u,v,\sigma_{st})\,dv\,du, \tag{8}$$

where $z_{t,r-1} = \hat{\mu}_{lm,r-1} - \hat{\pi}_{0lm} + \hat{\underline{\pi}}_{1lm}'\underline{x}_i$, $z_{t,r} = \hat{\mu}_{lm,r} - \hat{\pi}_{0lm} + \hat{\underline{\pi}}_{1lm}'\underline{x}_i$, and $s, t = 2, ..., M$, $s \neq t$, $k \neq l$. The joint probability functions must be modified appropriately for bivariate combinations of continuous, binary, and ordinal variables.

The first-stage estimates $\hat{\Omega}$, $\hat{\underline{\mu}}$, and $\hat{\sigma}_1^2$, as well as the covariances $\hat{\sigma}_{1s}$ and $\hat{\sigma}_{st}$ obtained at the end of the evaluation of all the conditional bivariate likelihoods, complete the estimation of $\hat{\underline{\omega}}$. These estimates are consistent and:

$$\sqrt{N}(\hat{\underline{\omega}} - \underline{\omega}) \xrightarrow{D} N(\underline{0}, \boldsymbol{W}). \tag{9}$$

A consistent estimate of **W** is required to evaluate appendix equation 3. This matrix is computed at the end of stage two (Lee, 1982; Sobel and Arminger, 1992). In stage three, the structural parameter estimates and their standard errors are obtained by minimizing appendix equation 3.

ERS
U.S. Department of Agricu
Economic Research Servic
SUMMARY OF REPORT #AIB-705

Consumer Interest in Nutrition Grows, and the Food Sector Responds

October 1994

Contact: Elizabeth Frazao, 202/219-0911

As evidence of the link between diet and health grows in the United States, many consumers are changing their diets. Food consumption patterns have changed dramatically in the last 20 years. Eating patterns are slowly shifting toward healthier diets, although there is still considerable room for improvement in meeting Federal food guidance recommendations. The food sector is clearly aware that nutrition is important to many consumers, and has been active in responding to consumer demand for foods with improved nutrient profiles. Meats, for example, are much leaner now than even 10 years ago, due to improved breeding practices and changes in meat trimming practices.

Consumer Concerns About Nutrition: Opportunities for the Food Sector, a recent report from USDA's Economic Research Service, reveals that many consumers want to improve their diets, but claim they lack the information to do so. Research has shown that many of the changes Americans have made in their food choices end up canceling each other out. To assist consumers in choosing a healthier diet, the Federal Government has overhauled its nutrition labeling regulations. In mid-1994, new nutrition labels became mandatory for most processed foods. Although nutrition labeling remains voluntary for fresh produce, meats, and seafood, the regulations contain strong incentives for the information to be made available to consumers.

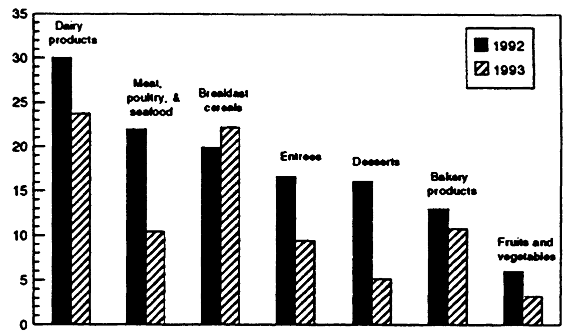
New or reformulated products have also abounded. More than 4,500 claims were made about the high nutrient content of new foods in 1992--nearly four times the number made in 1988, and a 5-percent increase above the number filed in 1991. The number of nutrient content claims on new products fell significantly in 1993, possibly related to the new mandatory nutrition labeling regulations. These regulations may push manufacturers to reformulate their products to further improve their nutrient content to meet the new definitions and requirements for health claims and nutrient descriptors.

Although Americans are making some dietary changes, they enjoy the taste of high-fat foods and do not seem willing to give them all up. If food companies can develop lower-fat products that taste like traditional high-fat foods, and provide consumers with acceptable low-fat substitutes, the food industry can help consumers to eat less fat without having to greatly change their eating habits.

1992 and 1993 product introductions that are low, reduced, or nonfat, by product category

Dairy products and breakfast cereals were the main food groups introducing lower fat products in 1993.

Percent of new introductions that are lower fat



Source: *Prepared Foods.*

To Order This Report...

The information presented here is excerpted from ***Consumer Concerns About Nutrition: Opportunities for the Food Sector,*** AIB-705, by Elizabeth Frazao. The cost is $9.00.

To order, dial **1-800-999-6779** (toll free in the United States and Canada) and ask for the report by title. Please add 25 percent for shipment to foreign addresses (including Canada). Charge to your VISA or MasterCard, or send a check (made payable to ERS-NASS) to:

ERS-NASS
341 Victory Drive
Herndon, VA 22070.

We'll fill your order via first class mail.