THE ROLE OF HEALTH KNOWLEDGE IN DETERMINING DIETARY FAT INTAKE

Kurt A. Carlson and Brian W. Gould

An endogenous switching regression model is used to estimate the effect that health information has on the food purchase decisions of meal planners. Specifically, we examine how information concerning the health implications of dietary fat intake influences the meal planner's daily intake of total and saturated fat. This analysis uses the 1989 to 1990 and 1990 to 1991 Continuing Survey of Food Intake by Individuals (CSFII), and companion Diet and Health Knowledge Surveys (DHKS). Our results show that health information regarding dietary fat intake has a significant impact on meal planner food choices.

Coronary Heart Disease (CHD) has long been recognized as a major health concern for adults. The causes of CHD have been studied for decades, and several factors have been found to contribute to its onset, including smoking, obesity, lack of exercise, and diet. For the most part, the role of diet has focused on dietary fat intake. In the mid-1960's, researchers provided the initial link between saturated fat intake and risk of CHD (Keys, Anderson, and Grande).

Recently, the importance of diet in preventing CHD has been re-emphasized. For example, recent recommendations are for less than 30 percent and 10 percent of daily caloric intake to originate from total and saturated fat, respectively (American Heart Association; U.S. Department of Health and Human Services). There is increasing evidence that consumers are becoming more aware of the link between saturated fat intake and CHD, and are incorporating fat intake recommendations into food purchase decisions (Dairy Council Digest; Borra). For example, Putler and Frazao reported that from 1970 to 1988, awareness of adverse health effects associated with a high fat diet increased from 8 percent to 55 percent of surveyed respondents. A 1986 survey conducted by the FDA found that more than 60 percent of the respondents reported changes in eating patterns as the result of health concerns (Mueller).

Identifying factors that affect dietary fat intake is of primary interest for several reasons. The food industry is interested in providing products consistent with consumer demands. Given changing attitudes towards fat, producers and marketers need information about these attitudes in order to provide desired products in a timely fashion. Currently, many food manufacturers are acting to capitalize on the trend toward healthy foods by developing markets for new low-fat/cholesterol products. In 1991, introduction of new reduced-fat/low cholesterol products increased 39 percent from 1990 introductions (Cheese Reporter, p.1).

To be effective at changing diets, nutritionists require accurate information about factors influencing attitudes towards dietary fat intake. Health information is assumed to be linked to diet quality via attitudes about fat and food choices made in the presence of these attitudes. The causality begins with health knowledge which leads to attitude development and/or adjustment about dietary fat which influences food choices, and in turn affects diet quality. If factors influencing these attitudes can be clearly identified, then nutritionists can map strategies for improving diets.

Previous econometric analyses of factors affecting nutrient demand have been primarily concerned with measuring the impact of participation in government programs. Some examples of government programs that have influenced nutrient intake include: food stamps, national school lunch, national school breakfast,
and nutrition education programs (Akin et al.; Butler, Ohls, and Posner; Morgan; Davis and Neenan; Devaney and Fraker; and Long). Capps and Schmitz noted that most of these studies found participation in government food assistance programs to have a positive influence on nutrient intake, ceteris paribus (p. 22).

Previous nutritional science analyses have focused on nutrient intake for individuals with similar socio-economic characteristics. Examples of such studies concerned with fat intake include Hackett et al., Reid et al., and Terry, Oakland, and Ankeny. Hackett et al. investigated dietary sources of fat among English adolescents, Reid et al. examined fat intake changes among individuals diagnosed with CHD, and Terry, Oakland, and Ankeny studied characteristics relating to males' (ages 35 to 55) adoption of reduced total and saturated fat diets. Given that these studies are concerned with fat intake as it relates to a specific group of people, they overlook the impact of socio-economic characteristics on dietary fat intake.

In one of the few attempts to incorporate health information in a food purchase model, Brown and Schrader developed a time-series based "cholesterol index," calculated as the cumulative number of clinical articles published from 1966 to 1987 that examined the link between cholesterol and heart disease. Quarterly shell egg disappearance was negatively related to this index. Subsequent studies that have applied Brown and Schrader's cholesterol index, include Capps and Schmitz, and Yen and Chern.

Jensen, Kesavan, and Johnson created a health awareness index and used it as a determinant of dairy product purchases. The index had a positive effect on dairy product purchases both in terms of the probability of consuming and quantity purchased. In a subsequent study, Jensen and Kesavan found that age and income positively influenced a nutrient attitude/awareness index (NAI), and that NAI had a positive influence on consumption of dairy and milk products.

The present study extends previous work by using cross-sectional data to determine the impact of health knowledge on total and saturated fat intake. Moreover, the current study uses a non-homogeneous sample of U.S. meal planners to estimate fat intake. For each fat type, a switching regression model estimates the meal planner's fat intake, dependent on health knowledge status.

**Description of the Econometric Model**

This study determines whether fat intake varies depending on meal planner knowledge of possible health consequences associated with dietary fat intake. The econometric model is based on the assumption that an individual maximizes utility (U) which is a function of food and other goods consumed. Assume, as in Akin et al., that nutrient contents of each food are known to consumers. Further, hypothesize that the diet/health information search activity of an individual is endogenous to the overall nutrient intake decision. Then, the decision variables are: (1) search activity with respect to the health impacts of food (nutrient) consumption; and (2) level of nutrient intake.

Representing health knowledge status by \( \Omega \) for \( M \) individuals, the utility maximization problem can be represented as:

\[
\text{MAX } U = U(Q_1, \ldots, Q_F, C | D, \Omega) \quad O_i
\]

s.t. \( Y = \sum_{j=1}^{F} P_j Q_j + C_i \) \( (1) \)

where \( U \) is an \((M \times 1)\) utility vector; \( Q_j \) is an \((M \times 1)\) vector of quantities of the \( j \)th commodity consumed; \( C \) is an \((M \times 1)\) vector of a composite good consumed; \( F \) is the number of food commodities; \( D \) is an \((M \times n)\) matrix of \( n \) individual demographic characteristics; \( Y \) is an \((M \times 1)\) vector of income; and \( P_j \) is the price of food \( j \) relative to the normalized price of \( C \).

From equation (1), we obtain the individual's demand functions:

\[
Q_j = Q_j(P_1, \ldots, P_F, Y | D, \Omega) \quad (2)
\]

\((\forall j = 1, \ldots, F)\).

Each unit of food \((Q_j)\) has \( \alpha_{ij} \) units of nutrient \( h \). Total and saturated fat intake is then:
\( N_h = \sum_{j=1}^{F} \alpha_{hj} Q_j \)  
\( (h = T, S), \)

where \( N_T \) and \( N_S \) are \((M \times 1)\) vectors of total and saturated fat intake, respectively.

Substituting equation (2) into equation (3), two nutrient demand equations can be represented as:

\( N_h = N_h(\mathbf{P}, \mathbf{Y} | \Omega, A) \)
\( (h = T, S), \)

where \( \mathbf{A} \) is the set of nutrient coefficients such that \( \alpha_{hj} \in \mathbf{A} \). We can also use equations (2) and (3) to obtain the indirect utility function:

\[
\Gamma(\mathbf{P}, \mathbf{Y}) = \max \{ U[N_T, N_S] \} \]

Following Blaylock and Blisard (1992, 1993), we define latent variable \( I_h' \) to be the net benefits of an individual obtaining health knowledge regarding the health impacts of alternative nutrients:

\( I_h' = \Gamma(\mathbf{P}, \mathbf{Y}) - \Gamma(\mathbf{P}, \mathbf{Y}) \)
\( (h = T, S), \)

where \( \Gamma' \) is utility with optimal levels of search, and \( \Gamma \) is without search, given prices and income. \( I_h' \) relates to a set of individual characteristics via:

\( I_h' = Z\gamma_h - \varepsilon_h, \)

where \( Z \) is an \((M \times R)\) matrix of household characteristics; \( \gamma_h \) is an \((R \times 1)\) vector of parameters; and \( \varepsilon_h \) is an \((M \times 1)\) error term vector which is distributed \( N(0,1) \). \( I_h' \) is not observed, but the binary variable, \( I_h \), is observable and related to \( I_h' \) by:

\[
I_h = \begin{cases} 
1 & \text{iff } Z\gamma_h \geq \varepsilon_h \\
0 & \text{iff } Z\gamma_h < \varepsilon_h.
\end{cases}
\]

An example of \( I_h \) could be whether an individual is aware of the relationship between saturated fat intake and CHD.

This formulation allows individuals without health knowledge to exhibit different nutrient consumption behavior than those with such knowledge. That is:

\[
\begin{align*}
N_{1h} & \iff I_h = 1 \\
N_{2h} & \iff I_h = 0,
\end{align*}
\]

where:

\[
N_{1h} = X_i\beta_{1h} + \nu_{1h} \]
\( (r = 1, 2), \)

and \( X_i \) is an \((M \times K)\) matrix of explanatory variables; \( \beta_{rh} \) is a \((K \times 1)\) vector of parameters; and \( \nu_{rh} \) is an \((M \times 1)\) error term vector for the \( r^{th} \) knowledge regime, identifying total fat \( T \) and saturated fat \( S \).

The assumption that search behavior is endogenous to the individual implies that \( N_{1h}, N_{2h}, \) and \( I_h \) are trivariate normal:

\[
\begin{pmatrix}
N_{1h} \\
N_{2h}
\end{pmatrix}
\]

Following Akin et al., assume that individuals self-select into equation (10) where \( \sigma_{as} \) are non-zero. To see this, differentiate between conditional and unconditional nutrient consumption. Unconditional expected nutrient intake, regardless of health knowledge status, is calculated as the sum of the probability that an individual is in a particular health knowledge regime times expected nutrient intake for the individual in each regime:

\[
E(N) = \Phi(Z\gamma) E[N_1 | I = 1] + (1 - \Phi(Z\gamma)) E[N_2 | I = 0],
\]

\( ^1 \)For notational simplicity, the nutrient subscripts \( (h=T,S) \) have been dropped from this point forward. Thus, the model development will proceed for a specific nutrient (i.e., total or saturated fat).
where $\Phi$ is the standard normal cumulative distribution function (Poier and Ruud; Dolton and Makepeace; Huang, Raunikar, and Misra; Lee and Brown). From this, expected conditional nutrient intakes are:

$$E(N_{l} | I = 1) = X_{l} \beta_{1} - \sigma_{1u}\left(\frac{\phi(Z_{Y})}{\Phi(Z_{Y})}\right)$$

$$E(N_{2} | I = 0) = X_{l} \beta_{2} + \sigma_{2u}\left(\frac{\phi(Z_{Y})}{1-\Phi(Z_{Y})}\right),$$

where $\phi$ denotes the standard normal density function, and the last term in each equation is the expected value of the error term, $E(v_{i}|I)$.

Equation (14) implies that the estimation equations are:

$$N_{1} = X_{l} \beta_{1} - \sigma_{1u}\left(\frac{\phi(Z_{Y})}{\Phi(Z_{Y})}\right) - \epsilon_{1}$$

$$N_{2} = X_{l} \beta_{2} + \sigma_{2u}\left(\frac{\phi(Z_{Y})}{1-\Phi(Z_{Y})}\right) + \epsilon_{2},$$

where $\epsilon_{1}$ and $\epsilon_{2}$ are the new-corrected residuals, with zero conditional means and heteroscedastic variances (Huang, Raunikar, and Misra; Maddala).

Following Heckman, a two-step estimation procedure is used to estimate equation (15), where $\hat{\gamma}$, obtained from Probit maximum likelihood methods, are used in place of $\gamma$. This two-step method provides consistent estimates of $\beta_{1}$, $\beta_{2}$, $\sigma_{1u}$, and $\sigma_{2u}$; however, since $\hat{\gamma}$ are used, estimation of equation (15) results in parameter variances that are biased downwards (Maddala). To calculate correct variances, it is necessary to compute $\sigma_{1u}$ and $\sigma_{2u}$ since they are not directly estimated. The methods for calculating $\sigma_{1u}$, $\sigma_{2u}$, and the correct variances for estimated parameters are taken from Maddala (pp. 225-27).

Similar to McDonald and Moffitt, conditional and unconditional fat intake are differentiated when estimating marginal impacts of a change in explanatory variables. Following Dolton and Makepeace, and Kimhi, the effect of a change in an explanatory variable on expected nutrient demand, conditional on health knowledge status is shown in equation (16) for $(s = 1, \ldots, K)$, where $X_{s}$ is the explanatory variable of concern.

**Data Description and Estimation Procedures**

Data used in this analysis are the 1989 to 1990 and 1990 to 1991 CSFII and companion Diet and Health Knowledge Surveys (DHKS). The DHKS contains information on diet, health, and food safety issues for individuals identified as main meal planner/ preparer in the CSFII. The CSFII contains information on food intakes by individuals over three consecutive days. Unlike earlier versions of the CSFII, dietary information is collected for all household members. Individual intake of food energy, 29 nutrients, and other dietary components are provided by the CSFII. In the present analysis, the focus is on meal planner intake of total and saturated fat.

Only meal planners providing three-day food records are used in this analysis. The 1989 to 1990 and 1990 to 1991 data are pooled creating a data set of 2,901 meal planner observations. Observations were deleted for one or more of the following reasons: (1) less than three days of reported consumption; (2) lack of

$^{3}$Testing stationary preferences was not the focus of this analysis; therefore, we assumed the role of health knowledge was constant across survey years, and therefore were able to pool the 1989 to 1990 and 1990 to 1991 data. As one reviewer suggested, a cross-validation test might provide support for this assumption. A chi-squared test of the null hypothesis that there were no differences in coefficients across years was conducted. The null hypothesis could not be rejected at the 0.01 level, but could be rejected at the 0.05 level. Given the focus of the study and concerns about manuscript length, we maintained the assumption and conducted our analysis on pooled data.

$$\frac{\partial E(N_{l} | I = 1)}{\partial X_{s}} = \beta_{ls} = \gamma_{s} \sigma_{1u}
\left(\frac{\phi(Z_{Y})}{\Phi(Z_{Y})}\left(Z_{Y} - \phi(Z_{Y})\right)\right)$$

$$\frac{\partial E(N_{2} | I = 0)}{\partial X_{s}} = \beta_{2s} = \gamma_{s} \sigma_{2u}
\left(\frac{\phi(Z_{Y})}{1-\Phi(Z_{Y})}\left(Z_{Y} - \phi(Z_{Y})\right)\right).$$
the role of health knowledge in determining dietary fat

Carlson, Gould

377

data on meal planner's education; or (3) missing household income data.

The binary variables used as dependent variables in total (THEART) and saturated fat (SHEART) Probit models are defined in Table 1. Respondents in the DHKS, identified as the main meal planners, were asked if they were aware of any health problems that might be related to the amount of fat a person eats. If the respondents answered “yes,” they were asked to elaborate as to the types of possible health problems. This study focuses on whether respondents were aware of the link between total and saturated fat intake and CHD.

More than 55 percent of meal planners indicated some knowledge of the relationship between total fat intake and heart problems. Mean daily meal planner total and saturated fat intake for the entire sample and conditional on health knowledge status are given in Table 1. Notice that more than one-third of meal planner total fat intake comes from saturated fat.

The relationship between socio-economic characteristics and attitudes towards basic nutrition and health beliefs has been the focus of previous nutritional research (Jensen and Kesavan; Jensen, Kesavan, and Johnson; Feick, Herrmann, and Warland). In addition, previous econometric analyses have considered such variables as seasonality, race, sex, income, region, education of the household head, household head being on a diet, and number of household members as factors affecting nutrient intake (Akin et al.; Butler, Ohls, and Posner; Horton and Campbell). The independent variables used in our analysis of health knowledge status and dietary fat intake are defined in Table 2.

Equations (7) and (15) are used to estimate meal planner total and saturated fat intake. Each fat type requires estimation of a first-stage Probit equation relating health knowledge status to a set of meal planner characteristics, and a pair of second-stage nutrient intake equations, one for each health knowledge regime. For the total fat intake models, the dummy variable used to represent health awareness is THEART and the fat intake variable is TFAT. For the saturated fat model, the dependent variables are SHEART and SFAT. Both Probit and regression models are estimated using the GAUSS software system.

Table 3 reports an evaluation of the overall fit of the econometric models. For the Probit equations, $\chi^2$ statistics test the null hypothesis of zero slope coefficients; percent of observations correctly predicted by each Probit equation, and Maddala's pseudo $R^2$ are also provided. For the fat intake equations, squared correlation coefficients of predicted and actual conditional intakes, $r^2$, and adjusted $R^2$ values are presented.

For each Probit equation, the $\chi^2$ statistic is significant, resulting in a rejection of the null hypothesis. Each Probit equation successfully predicted at least 63 percent of the observations. These percentages are reasonable when compared to those obtained from similar Probit estimations. For example, Feick, Herrmann, and Warland correctly predicted from 51 to 59 percent of the observations in their analysis of nutrition information sources. The Maddala pseudo $R^2$ values from the Probit equations range in value from 0.086 to 0.103. These values are reasonable when compared to Jensen and Kesavan's values of 0.06 and 0.03, in their analysis of demand for dairy products.

The bottom of Table 3 provides adjusted $R^2$ values for the conditional intake equations. These estimates are similar to previous studies. Butler, Ohls, and Posner obtained adjusted $R^2$ values ranging from 0.057 to 0.081 for nutrient intake equations. Similarly, Horton and Campbell obtained adjusted $R^2$ values ranging from 0.044 to 0.078 in their analysis of labor market activity influence on nutrient intake.

Factors Affecting Health Knowledge

Putler and Frazao noted that awareness of the adverse health effects associated with a high-fat diet varies across socio-economic groups (p. 16). Specifically, respondents with higher education and income levels were more aware of

---

No price data is collected in the CSFII/DHKS, so prices are excluded from the analysis. The low mean income level is due to the large sample of low-income households included in the CSFII. Mean income, excluding the low-income households, is greater than $28,000.
the health consequences of dietary fat intake. Similarly, Lutz, Blaylock, and Smallwood, in analysis of food consumption behavior, concluded that low-income households have not followed national trends towards consumption of lower fat foods; also, in addition to food prices and culturally-based eating habits, food choices were determined by household socio-economic characteristics (p. 17). In this study, income and education are hypothesized to contribute to the ease with which individuals obtain and process nutritional information. Consumers with higher income and/or education levels presumably have higher opportunity costs of obtaining and processing information, and therefore implement the most efficient sources and methods.

Table 4 gives Probit parameter estimates associated with explaining health knowledge status of the main meal planner. With dummy variables used in the Probit models, the base observation is one in which the meal planner is not on a low-fat diet, does not have high blood pressure or cholesterol, watches four hours of television or more, belongs to a family with no children, is white, has a high school education, does not use package label nutrient information on a regular basis, and resides in the Pacific region. In both the total and saturated fat Probit equations, nine of 16 parameter estimates are significant at 0.05 level.

During the course of undertaking a low-fat diet, the meal planner is hypothesized to become more aware of nutrition's role in determining health, and adjusts household food purchases. Surprisingly, the variable LOWFDIET does not significantly impact the probability that the meal planner is aware of the health implications of total or saturated fat intake. This may be due to the influence of individuals who use dieting as an exclusive means for managing body weight, and who are unaware of other health risks associated with excessive fat intake.

The meal planner's level of control over fat content of foods consumed is represented by the dummy variables, NUTQDUM and COMPODUM. NUTQDUM takes on a value of unity if nutrition is important or very important to the meal planner while shopping. COMPODUM takes on a value of unity if the meal planner compares nutrients sometimes or always while deciding among different brands of similar foods. Approximately 79 percent and 55 percent of meal planners indicated positive responses to these two variables, respectively. A positive Probit coefficient for either/both of these variables is hypothesized to represent concern for

4The omitted television variable was four hours or more of television per day. This may seem excessive; however, nearly 30 percent of the meal planners in our survey were of this type. The television categories were created with the objective of maintaining a relatively even proportion of meal planners across categories.

Table 3. Summary Statistics for the Econometric Models

<table>
<thead>
<tr>
<th>Equation Statistic</th>
<th>THEART</th>
<th>SHEART</th>
</tr>
</thead>
<tbody>
<tr>
<td>1865.89</td>
<td></td>
<td>1851.19</td>
</tr>
<tr>
<td>260.22 (15)</td>
<td></td>
<td>314.70 (15)</td>
</tr>
<tr>
<td>0.086</td>
<td></td>
<td>0.103</td>
</tr>
<tr>
<td>63.1</td>
<td></td>
<td>63.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equation Statistic</th>
<th>TFAT THEART=1</th>
<th>TFAT THEART=0</th>
<th>SFAT SHEART=1</th>
<th>SFAT SHEART=0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.135</td>
<td>0.146</td>
<td>0.127</td>
<td>0.168</td>
<td></td>
</tr>
<tr>
<td>0.380</td>
<td>0.397</td>
<td>0.372</td>
<td>0.422</td>
<td></td>
</tr>
</tbody>
</table>

Note: R² is defined as the squared correlation coefficient between predicted and actual fat intake.
Table 4. Parameter Estimates of the Probability of Recognizing Health Problems

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total Fat/Health Problem</th>
<th>Saturated Fat/Health Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated Coefficient</td>
<td>Standard Error</td>
</tr>
<tr>
<td>INTERCEPT</td>
<td>-0.822*</td>
<td>0.507</td>
</tr>
<tr>
<td>NUTQDUM</td>
<td>0.386*</td>
<td>0.307*</td>
</tr>
<tr>
<td>COMPDUM</td>
<td>0.210*</td>
<td>0.142*</td>
</tr>
<tr>
<td>ln(INCOME)</td>
<td>0.186*</td>
<td>0.086</td>
</tr>
<tr>
<td>ln(MPAGE)</td>
<td>0.086</td>
<td>-0.017</td>
</tr>
<tr>
<td>NOHIGH</td>
<td>-0.065</td>
<td>-0.150*</td>
</tr>
<tr>
<td>POSTHIGH</td>
<td>0.149*</td>
<td>0.079</td>
</tr>
<tr>
<td>COLLEGE</td>
<td>0.336*</td>
<td>-0.382*</td>
</tr>
<tr>
<td>BLACK</td>
<td>-0.240*</td>
<td>-0.352*</td>
</tr>
<tr>
<td>HISPANIC</td>
<td>-0.170</td>
<td>-0.113</td>
</tr>
<tr>
<td>HIBPCHOL</td>
<td>0.013</td>
<td>0.013</td>
</tr>
<tr>
<td>TV0</td>
<td>0.144</td>
<td>0.253*</td>
</tr>
<tr>
<td>TV0_1</td>
<td>0.149*</td>
<td>0.060</td>
</tr>
<tr>
<td>TV2_3</td>
<td>0.184*</td>
<td>0.188*</td>
</tr>
<tr>
<td>LOWFDIET</td>
<td>0.103</td>
<td>0.146</td>
</tr>
<tr>
<td>CHILDREN</td>
<td>-0.021</td>
<td>-0.064</td>
</tr>
</tbody>
</table>

Note: * denotes significance at the 0.05 level. INCOME used in the estimation is scaled by 10,000.

According to Moorman and Matulich, age reflects the mental and physical ability of consumers to engage in healthy behavior. The authors stated that research on how age influences health maintenance is mixed. They concluded that age had a positive influence on health maintenance behavior and a negative influence on health information acquisition. That is, in general, as people get older they suffer more illnesses and become more concerned with maintaining health, but are less able to obtain and interpret health information due to deteriorating cognitive abilities. Alternatively, younger meal planners have grown up in an era where health information is readily available, making them more likely to have solidly-formed attitudes about nutrition’s role in health maintenance. Jensen and Kesavan hypothesized that older people perceive the risk of health problems differently than younger people, and therefore have different behavior regarding health issues. In this analysis, MPAGE is hypothesized to capture these potentially-offsetting cohort effects. Table 4 shows that MPAGE does not significantly contribute to health knowledge status. In the presence of offsetting age cohort effects, this result is to be expected.

Individuals from households with greater incomes are hypothesized to be better able to incur costs of obtaining health-related information. In fact, INCOME positively impacts health knowledge status in both Probit equations. This is consistent with the findings of Jensen, Kesavan, and Johnson who reported that income positively influences the probability that a household is aware of the benefits of calcium intake.

The positive relationship between education and health knowledge may reflect the increased ability of those with higher educational levels to process health information (Moorman and Matulich). The education variable COLLEGE positively impacts the probability of health awareness for both fat types while the variable NOHIGH negatively impacts the same, for the saturated fat equation only. The relatively large nutritional intake and willingness to incur search activity costs (time) necessary to obtain nutrient content information. Table 4 indicates, as expected, that each of these variables has a positive impact on health knowledge status.

In households with greater incomes are hypothesized to be better able to incur costs of obtaining health-related information. In fact, INCOME positively impacts health knowledge status in both Probit equations. This is consistent with the findings of Jensen, Kesavan, and Johnson who reported that income positively influences the probability that a household is aware of the benefits of calcium intake.
estimated coefficient for COLLEGE indicates that education plays a key role in determining the level of health awareness. The role of ethnicity is consistent across fat Probit equations with BLACK and HISPANIC meal planners having a lower probability of possessing health knowledge than non-minority meal planners.

The amount of television watched might aid the information collection process, and thereby positively influence health knowledge. However, it is hypothesized that individuals who watch four hours of television or more per day (controlling for respondent age) are less active and have poorer diets than people who watch three hours of television or less per day (TV0, TV0_1, and TV2_3). Therefore, these three variables, acting as proxies for activity level, will be positively related to meal planners' health knowledge status. Table 4 indicates that four of six parameter estimates associated with the television variables are positive and significant. This result indicates that active meal planners are more likely to be aware of the relationship between fat intake and CHD than meal planners who watch more than four hours of television per day.

Factors Affecting Total and Saturated Fat Intake

Table 5 presents conditional total and saturated fat intake regression parameter estimates. The last row of Table 5 provides the estimated correlation coefficient, ρ, between the error terms of the first stage (Probit) and the second stage conditional demand equations, for each fat type. The second from the last row of Table 5 shows estimated sample selection correction coefficients and their respective standard errors. The correction factors (σ_u) control for self-selection bias; that is, meal planners choose an awareness regime (search activity level) according to the effect they perceive it will have on fat intake and utility. Both correction coefficients are significant for the health "aware" regime (THEART=1 and SHEART=1). In contrast, neither correction factor in the "unaware" regime is significant. In their study of the impact of health awareness on dairy product demand, Jensen, Kesavan, and Johnson found that all correction factor estimates were insignificant. Therefore, they argued that factors inducing the consumption decision (Probit equation determinants) are not important in determining conditional demand via their impact on the likelihood of consuming.

Our results for the "unaware" regime parallel those of Jensen, Kesavan, and Johnson. For this regime, the factors inducing health knowledge awareness are not important in determining the level of fat intake via their impact on the probability of being "aware." Determining the impact of the correction factors is complicated and cannot be accomplished through comparison of the estimated coefficients alone (Dolton and Makepeace). In his analysis of union/nonunion wage differentials, Lee reported positive correction factor effects which he claimed result from the individual's choice of a work regime (union or nonunion) that pays a higher wage than received by an average worker in that regime with the same characteristics and working circumstances. The correction factor effect, which Lee referred to as the "term truncation effect," is the product of the Mill's ratio and its respective estimate (e.g., σ_u (φ(Zγ)/Φ(Zγ)).

Though consumers in our analysis do not directly choose their regime, they do choose the search activity cost level they are willing to bear in order to attain health knowledge awareness. Calculation of correction factor effects reveals significant negative effects for meal planners in the health aware regime. Following Lee's argument, these negative effects indicate that meal planners in the "aware" regime indirectly choose their awareness regimes (via search activity choice) to minimize fat intake and maximize the utility derived from it, relative to individuals with the same characteristics who are also aware of the link between dietary fat intake and CHD.

Columns [a] and [c] of Table 5 give fat intake parameter estimates conditional on the meal planner being aware of the fat intake/health status link. Inspection of these results indicates

---

1BLACK is defined as black and not of hispanic descent and HISPANIC is defined as white or black and of hispanic descent. Similar procedures are used to construct the sex-specific race dummy variables.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Total Fat</th>
<th>Saturated Fat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \text{THEART} = 1 )</td>
<td>( \text{THEART} = 0 )</td>
</tr>
<tr>
<td></td>
<td>( \text{[a]} )</td>
<td>( \text{[b]} )</td>
</tr>
<tr>
<td>BLACKFEM</td>
<td>-4.044</td>
<td>-0.867</td>
</tr>
<tr>
<td></td>
<td>(2.581)</td>
<td>(2.416)</td>
</tr>
<tr>
<td>BLACKMAL</td>
<td>10.716*</td>
<td>22.572*</td>
</tr>
<tr>
<td></td>
<td>(4.809)</td>
<td>(3.798)</td>
</tr>
<tr>
<td>HISPFEM</td>
<td>-1.243</td>
<td>-6.065*</td>
</tr>
<tr>
<td></td>
<td>(3.013)</td>
<td>(2.925)</td>
</tr>
<tr>
<td>HISPMAI</td>
<td>5.822</td>
<td>18.769*</td>
</tr>
<tr>
<td></td>
<td>(6.793)</td>
<td>(6.388)</td>
</tr>
<tr>
<td>WHITEMAL</td>
<td>21.053*</td>
<td>20.954*</td>
</tr>
<tr>
<td></td>
<td>(1.860)</td>
<td>(2.163)</td>
</tr>
<tr>
<td>LOWFDIET</td>
<td>-8.748*</td>
<td>-11.164*</td>
</tr>
<tr>
<td></td>
<td>(2.260)</td>
<td>(2.922)</td>
</tr>
<tr>
<td>In(MPAGE)</td>
<td>-12.096*</td>
<td>-10.496*</td>
</tr>
<tr>
<td></td>
<td>(1.793)</td>
<td>(1.806)</td>
</tr>
<tr>
<td>In(INCOME)</td>
<td>2.787*</td>
<td>2.477*</td>
</tr>
<tr>
<td></td>
<td>(1.012)</td>
<td>(1.095)</td>
</tr>
<tr>
<td>PREGLACT</td>
<td>7.071</td>
<td>6.131</td>
</tr>
<tr>
<td></td>
<td>(3.728)</td>
<td>(4.196)</td>
</tr>
<tr>
<td>NE</td>
<td>6.604*</td>
<td>5.397</td>
</tr>
<tr>
<td></td>
<td>(3.063)</td>
<td>(3.995)</td>
</tr>
<tr>
<td>MIDATL</td>
<td>4.708*</td>
<td>5.608</td>
</tr>
<tr>
<td></td>
<td>(2.342)</td>
<td>(2.874)</td>
</tr>
<tr>
<td>ENC</td>
<td>5.183*</td>
<td>7.046*</td>
</tr>
<tr>
<td></td>
<td>(2.200)</td>
<td>(2.705)</td>
</tr>
<tr>
<td>ENC</td>
<td>5.129*</td>
<td>4.560</td>
</tr>
<tr>
<td></td>
<td>(2.919)</td>
<td>(3.711)</td>
</tr>
<tr>
<td>SATL</td>
<td>1.492</td>
<td>4.179</td>
</tr>
<tr>
<td></td>
<td>(2.211)</td>
<td>(2.572)</td>
</tr>
<tr>
<td>ESC</td>
<td>2.672</td>
<td>5.222</td>
</tr>
<tr>
<td></td>
<td>(2.869)</td>
<td>(3.239)</td>
</tr>
<tr>
<td>WSC</td>
<td>3.317</td>
<td>3.595</td>
</tr>
<tr>
<td></td>
<td>(2.635)</td>
<td>(2.979)</td>
</tr>
<tr>
<td>MNT</td>
<td>3.214</td>
<td>0.205</td>
</tr>
<tr>
<td></td>
<td>(3.207)</td>
<td>(3.617)</td>
</tr>
<tr>
<td>( \sigma^2 )</td>
<td>-8.981*</td>
<td>-5.863</td>
</tr>
<tr>
<td></td>
<td>(4.370)</td>
<td>(4.288)</td>
</tr>
<tr>
<td>( \rho_\sigma^2 )</td>
<td>-0.358</td>
<td>-0.225</td>
</tr>
</tbody>
</table>

Note: * denotes significance at the 0.05 level. INCOME used in the estimation is scaled by 10,000. Standard deviations are presented in parentheses. The correlation between the first and second stages, \( \rho \), is calculated by \( \rho = \sigma^2 / (\text{sqrt}(SSE/d.f.)) \) (Poirer and Ruud, p. 252).
a significant negative relationship between meal planner age and conditional fat intake. Though meal planner race is a significant factor in determining health knowledge status, race dummy variables were not significant in conditional demand equations, so we coupled them with gender-specific dummy variables (BLACKFEM, BLACKMAL, HISPFEM, HISPMAI, and WHITEMAL). Table 5 shows that for all races, demand for total and saturated fat is positively influenced by the meal planner being male. This disparity in fat intake, on the basis of gender, is consistent with recommended daily allowances of energy intake. Similar to the impact on awareness probability, income positively impacts both total and saturated fat intake.

As expected, being on a low-fat diet (LOWFDIET) has a negative impact on the demand for both total and saturated fat. Age has a negative impact on the intake of both fat types. Presumably, this is due to the increased health awareness that is associated with additional years of planning meals as well as the decreasing need for caloric intake that occurs as an adult ages. For example, the National Research Council of the Food and Nutrition Board reported that the daily energy requirement for adult males decreases from 2,900 kcal to 2,300 kcal after age 50.

Hytten and Chamberlain noted that the caloric cost of pregnancy is about 400 kcal/day. Half of these additional calories are provided by increased food consumption. The other half are accounted for by a reduction in physical activity. Some of the 200 kcal derived from additional food intake are hypothesized to be provided by increased fat intake. PREGLACT positively impacts the demand for saturated fat in the unaware regime only. This is not surprising. One would expect a pregnant meal planner to increase fat intake as she increases caloric intake; and if unaware of the health consequences, to increase saturated fat intake in the process.

Total and saturated fat intakes follow similar patterns with respect to regional variation. Relative to the Pacific Region, meal planners living in the NE, MIDATL, and ENC (see Table 2 for definitions) regions with information concerning the implications of total fat intake consume more total fat. Meal planners living in these regions who do not have information concerning the implications of saturated fat intake consume more saturated fat, relative to those living in the Pacific Region. Other regions do not provide significant impacts on total and saturated fat intake when compared to the Pacific Region. Total fat intake is positively influenced by the ENC region for both regimes; however, the impact of this region, across health knowledge regimes, appears to be different as is evident from the relative magnitudes of these two coefficients.

Conclusions

The objective of this analysis was to investigate the role of health information in determining dietary fat intake. Two periods of the Continuing Food Intake Surveys (CSFII) and the companion Diet and Health Knowledge Surveys (DHKS) were pooled. The sample used for estimation had 2,901 meal planners, all of whom reported three days of food consumption. Two endogenous switching regression models were estimated; one for total fat and one for saturated fat. The sample was partitioned on the basis of the meal planner's awareness level regarding the relationship between fat intake and CHD.

Several factors positively influence the probability that an individual is aware of the link between fat intake and CHD. The level of concern that a meal planner exhibits over fat content while shopping, as indicated by the variables NUTQDUM and COMQDUM, has a positive influence on nutritional awareness. Other factors having a positive influence on the probability of being aware are education, income, and being white. Programs to improve health awareness should be focused on minority groups, people with relatively less education and/or income, and people who watch more hours of television. In addition, results suggest

\[ \text{The National Research Council of the Food and Nutrition Board daily energy recommendations for males and females between the ages of 19 and 50 are 2,900 kcal and 2,200 kcal, respectively.} \]
that encouragement of product nutrient comparisons will increase health awareness. Surprisingly, dieting does not have a significant impact on the probability of being aware. This supports the argument that a segment of society uses dieting as a means for weight loss while being unaware of the relationship between fat intake and health problems other than obesity. Therefore, efforts to increase health awareness about the link between fat intake and CHD should also focus on dieters.

Given the structure of our model, the fat intake results show how socio-economic characteristics influence nutrient intake, conditional on awareness level. Several characteristics have an affect on conditional fat intake including income, meal planner age, whether the meal planner is pregnant and/or lactating, sex, and region of residence. For both awareness regimes and fat types, male meal planners have higher fat intake than white females (base meal planner). Part of this result is due to men's greater daily intake requirements, and part likely results from men being less willing to forgo high fat foods. This argument is consistent with previous findings which indicate that men consider taste to be more important than nutrition when considering consumption options (Schafer).

Using equation (16), conditional total and saturated fat intake elasticities for changes in INCOME and MPAGE are calculated. The positive impact of income on both total and saturated fat intake, despite its positive impact on awareness probability, is due to increased food demand which coincides with higher income levels. As evidence, notice that all income elasticities are positive. The negative impact of meal planner age on both total and saturated fat intake is attributed to cohort effects. Again, in support of our argument, all meal planner age elasticities are negative. By viewing the negative meal planner age elasticities in the context of cohort effects, prospects are that as current young meal planners age, they will continue to be conscious of the health consequences of their food choices.

This research has focused on meal planner fat intake. An extension of our research would be to estimate the impact of health information on consumption of specific disaggregate food commodities (e.g., low-fat and full-fat cheeses). The CSFII is unable to accommodate such an effort as it lacks disaggregate consumption information for foods used as ingredients. For example, consumption of food such as mozzarella cheese, which is used almost exclusively as an ingredient, is dramatically underreported by the CSFII. Currently we are undertaking efforts, similar to those of Jensen and Kesavan, to merge the CSFII with other data.

Table 6. Condition INCOME and MPAGE Fat Intake Elasticities

<table>
<thead>
<tr>
<th>Type of Fat Intake</th>
<th>Independent Variable</th>
<th>INCOME</th>
<th>MPAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Fat</td>
<td>Health Knowledge Status</td>
<td>Aware</td>
<td>Unaware</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.029</td>
<td>0.029</td>
</tr>
<tr>
<td>Saturated Fat</td>
<td></td>
<td>0.027</td>
<td>0.032</td>
</tr>
</tbody>
</table>

[Received December 1993. Final version received March 1994.]

References


