The Effect of Consumer Label Use on Diet Quality

This study examines the effect of consumer label use on diet quality of Americans. Using an endogenous switching regression model, results indicate that label use improves diet quality by as much as 14 points on a 100 point Healthy Eating Index Scale. Use of health claims among the information on food labels provides the highest level of improvement in diet quality.

Sung-Yong Kim, Texas A&M University[1]
Rodolfo M. Nayga, Jr., Texas A&M University[2]
Oral Capps, Jr., Texas A&M University[3]
Beverly Tepper, Rutgers University[4]

Introduction

Many Americans are not meeting dietary recommendations. The U.S. Department of Agriculture revealed that only about 12 percent of Americans are eating healthfully. This statistic is a concern because four of the top ten causes of death in the United States – heart disease, cancer, stroke, and diabetes – are associated with poor diets. Diet-related health conditions cost society an estimated $250 billion annually in medical costs and lost productivity (Frazao, 1995).

Concerns about the effect of diet on health have resulted in the legislation of the Nutritional Labeling and Education Act (NLEA) and its implementation in 1994. The NLEA instituted sweeping changes to replace the voluntary system of labeling established by the Food and Drug Administration (FDA) in 1973 (Nayga, 1996). Hence, the NLEA was implemented to provide consistent, understandable, and usable nutritional labels that can help consumers make healthier food choices and, therefore, improve the quality of their diets.

The purpose of this study is to assess the effectiveness of the NLEA in terms of diet quality improvements. Specifically, this study will attempt to determine the characteristics of consumers who use nutritional labels as well as to evaluate the effect of consumer label use on diet quality as measured by the Healthy Eating Index (HEI) developed by the USDA.

The Econometric Model

To assess the effect of consumer label use on diet quality, endogenous switching regression model is employed to control for unobserved heterogeneity in the label use decision. It consists of diet quality equations for label users and non-label users, and an equation for the label use decision. Define HEI_i as the observed diet quality; HEI_1 and HEI_0 as the diet quality of label user and non-label user, respectively; I_i^* as a latent variable that determines label use decision; I_i as an indicator variable that equals one if consumer uses nutritional labels and equals zero otherwise; X as a vector of observed characteristics that affect diet quality and *Z* as vector characteristics that affect label use. The endogenous switching regression model is written as

$$HEI_1 = X'\beta_1 + \varepsilon_1 \quad (1)$$
$$HEI_0 = X'\beta_0 + \varepsilon_0 \quad (2)$$
$$I^* = Z'\gamma + u \quad (3)$$
$$I = 1 \text{ if and only if } I^* > 0, \text{ otherwise } I = 0$$

The observed HEI is defined as $HEI_i=HEI_{1i}$ if and only if I=1; $HEI_i=HEI_{0i}$, if and only if I=0. The error terms of the above equations, ε_1, ε_0 and μ are assumed to have a trivariate normal distribution. Since the choice of using labels or not is endogenous, the error terms in equation (1) and (2), conditional on the sample selection criterion, have a nonzero expected value. Thus OLS estimates of β are biased. Sample selection corrected Healthy Eating Index

equations are specified following Lee (1976):

$$HEI_1 = X'\beta_1 + \sigma_{1u}\phi(Z'\gamma)/\Phi(Z'\gamma) + \upsilon_1 \quad \text{if } I = 1 \qquad (4)$$
$$HEI_0 = X'\beta_0 - \sigma_{0u}\phi(Z'\gamma)/1\text{-}\Phi(Z'\gamma) + \upsilon_0 \quad \text{if } I = 0 \qquad (5)$$

where σ_{1u} (σ_{0u}) indicates the covariance between $\varepsilon_1(\varepsilon_0)$and u, the function $\phi(\bullet)$ indicates normal probability distribution, and the function $\Phi(\bullet)$ indicates normal cumulative density. The new residuals, $\upsilon_1=\varepsilon_1 + \sigma_{1u}$ and $\upsilon_0=\varepsilon_0 + \sigma_{0u}$ are uncorrelated. The two-step procedure for estimation (4) and (5) involves first calculating the Mill's ratio, $\phi(Z'\gamma)/\Phi(Z'\gamma)$ and $-\phi(Z'\gamma)/1-\Phi(Z'\gamma)$, using probit estimates of (3). The ordinary least squares estimation is used next to estimate (4) and (5). Then, the variance-covariance matrix of two-step estimates is adjusted using the procedure described by Maddala (1983).

Data

Besides the use of the 1994-96 Continuing Survey of Food Intakes by Individuals (CSFII) data for the HEI variable, the 1994-96 Diet and Health Knowledge Survey (DHKS) data, the companion data of the CSFII, also are used in this study. The DHKS includes detailed information about the individual's socioeconomic background and questions on label usage. The empirical work uses DHKS respondent files, providing a sample size of 5405.

The name, definitions, and means for principal variables are exhibited in Table 1. The dependent variables include the Healthy Eating Index (HEI) developed by USDA, and a binary label use variable. The HEI provides a summary measure of people's overall diet quality (Bowman et al., 1998). The mean of HEI score is 63.72 for label users and 58.93 for non-label users. The analysis also is disaggregated by type of information contained on food labels. The five types of information that are presented on the food label are (1) the list of ingredients; (2) the short phrases on the label like "low-fat" or "light" or "good source of fiber"(nutrient content claims); (3) the nutrition

Table 1
Definition of Variables.

	Description	Means	Std.Dev
Dependent Variables			
LBUSE	General use of label (yes=1; no=1)	0.8294	0.3762
HEI	Health Eating Index	62.8991	13.7452
Explanatory Variables			
INCOME	Household income(10,000 dollars)	3.5082	2.6401
INCMSQ	Square of household income	19.2765	26.1452
AGE	Age of respondent (in years)	50.9071	17.1429
AGESQ	Square of age of respondent	2885.3582	1808.1358
MALE	Respondent is male (1=yes; 0=no)	0.5034	0.5000
B_RACE	Respondent is black (1=yes; 0=no)	0.1164	0.3207
O_RACE	Respondent is other nonwhite race (1=yes; 0=no)	0.0633	0.2435
EMPLOYED	Respondent is employed (1=yes; 0=no)	0.5819	0.4933
CITY	Respondent resides in the central city (1=yes; 0=no)	0.2949	0.4560
NOMETRO	Respondent resides in the non-metropolitan (1=yes; 0=no)	0.2666	0.4422
EDUCATION	Schooling in years	12.6459	3.0932
NE	Respondent resides in the Northeast (1=yes; 0=no)	0.1919	0.3938
WE	Respondent resides in the West (1=yes; 0=no)	0.2019	0.4014
MW	Respondent resides in the Midwest (1=yes; 0=no)	0.2514	0.4339
FSP	Participant in the food stamps program (1=yes; 0=no)	0.0788	0.2695
EXERCISING	Respondent has regular exercise (1=yes; 0=no)	0.4846	0.4998
BMI_SP	Body-mass ratio of respondent	27.9155	11.3339
SMOKING	Respondent is smoking now (1=yes; 0=no)	0.2564	0.4367

DHA	Diet-health awareness (index)	5.7937	1.6159
PYRAMID	Knowledge on Food Pyramid Guideline(index)	2.3604	1.2312
SHOPPER	Respondent is major food shopper (1=yes; 0=no)	0.6949	0.4605
SPECDIET	Respondent has special diet (1=yes; 0=no)	0.1741	0.3792
NHSP	Respondent is non-Hispanic (1=yes; 0=no)	0.9221	0.2680

panel that tells the amount of calories, protein, fat, and such in a serving of the food; (4) the information about the size of serving; and (5) the statement on the label that describe health benefits of nutrients or foods (health claims). About 76.4% of the sample used the list of ingredients, 74.2% used the nutrient content claims, 75.4% used the nutrition panel, 67.6% used the serving size, and 67.9% used the health claims. About 82.9 % of the sample used at least one of these information on the label, defined as the general use of label. Binary variables (1=use; 0=not use) are used to capture the decision to use each type of information on the food label.

Independent variables consist of personal or household characteristics, demographic factors, participation in government programs such as the Food Stamp Program, and knowledge about the food guide pyramid. Personal or household characteristics include body mass index, age, gender, level of education, ethnicity, race, exercise status, smoking status, employment status, and special diet status. Other demographic factors include region, urbanization, household size, and income. Consumer's knowledge about the Food Pyramid Guidelines (PYRAMID) is constructed as a measure of diet-health knowledge. The variable measures how much consumers know the servings recommended for the five primary food groups (grains, fruits, vegetables, dairy and meat) in the Food Guide Pyramid. Since PYRAMID reflects the answers to 5 questions, the variable has values ranging from 0 to 5.

The other variables in the label use probit equation are a dummy variable indicating whether the individual is a major shopper or not and a variable reflecting consumer's awareness about the linkage between diet and health (DHA). The variable DHA is constructed following Variyam et al. (1996). The eight questions in the DHKS used to construct the DHA variable take the general form: "Have you heard about any health problems that might be related to being overweight and how much of a particular nutrient (such as fat, fiber, salt, calcium, cholesterol, and sugar) a person eats?. Each answer of "Yes" is given a value of one while each answer of "No" is given a value of zero. Since DHA reflects the answers to 7 questions, the variable has a lower limit of zero and an upper limit of seven.

Empirical Results

Probit Label Use Model

The first stage probit model was estimated for the general label use (The estimated results were omitted because of the limitation of space).[5] Based on the results, there is a nonlinear relationship between income and label use. The probability of label use increases with income until an income level of about $59,800 before it declines with subsequent increases in income. Also, a nonlinear relationship exists between age and general label use. The probability of label use increases until age 44 before declining with subsequent increases in age.

Males are less likely to use labels than females. Results also indicate that education is significantly and positively related to label use. Urbanization and regional differences also are evident in the results. Specifically, individuals who reside in nonmetro areas are less likely to use labels than those who reside in suburban areas. In addition, individuals who reside in the South are less likely to use labels than individuals from other regions. Non-hispanics are less likely to use labels than others. Individuals who are on a special diet are more likely to use labels than individuals who are not on a special diet. Individuals who are more informed about the link between diet and health also are more likely to use nutritional labels. This result is consistent with the argument that poorly informed consumers tend to underestimate the marginal benefit of label use. Major food shoppers are more likely than others to use information on the label when shopping than others. This finding is comforting since a household's major food shopper can potentially influence the quality of the diet of individual household members just from the types of foods he or she decides to purchase.

Heathy Eating Index Models

The second-stage estimates of the endogenous switching-regression model for general label use are exhibited in Table 2. The parameter estimates for education, body mass index, exercise, food stamp participation, and knowledge about the food guide pyramid are statistically significant and have the expected signs in the model for label users. In the model for non label-users, these coefficients are insignificant, but the estimates for age, household income, the level of urbanization (i.e., city), and some regions (i.e., midwest) are statistically significant.

Results based on the general label use model indicate that income is positively related to diet quality (i.e., HEI) of label users. Income is not significant in the non-label user equation. Black label users and non-label users have HEIs that are about three points lower than the HEIs of white label users and non-label users, respectively. Label users of other races, however, have higher HEI than white label users.

Employed label users have a lower HEI than unemployed label users. The reason for this result is not clear. However, it is possible that the diet quality of employed label users is lower because they do not have as much time

Table 2
Parameter Estimates of the HEI Equations.

	Label User		Label Non-User	
	Parameters	t-values	Parameters	t-values
Constant	59.2580**	22.592	51.9270**	10.364
INCOME	0.5894*	1.963	1.1210	1.674
INCMSQ	-0.0212	-0.760	-0.0980	-1.514
HHSIZE	-0.1528	-0.927	-0.1614	-1.891
AGE	-0.0399	-0.525	-0.2659	-1.891
AGESQ	0.0019*	2.449	0.0034*	2.533
MALE	0.2982	0.530	-0.9156	-0.709
B_RACE	-3.1097**	-4.639	-3.2132*	-2.287
O_RACE	2.8826**	3.116	2.7028	1.272
EMPLOYED	-1.6043**	3.158	-2.1652	-1.935
CITY	0.1353	0.281	2.3789*	2.087
NOMETRO	-1.4254*	-2.750	-0.3648	-0.346
EDUCATION	0.3093**	3.205	0.1938	1.078
NE	1.6808**	2.934	3.0121*	2.299
WE	1.7406**	3.008	1.4781	1.142
MW	0.6376	1.205	0.3526**	2.979
NHSP	-1.8547*	-2.174	0.4309	0.218
FSP	-2.9135**	-3.535	-1.7525	-1.127
BMI_SP	-0.0517**	-3.026	0.0873*	2.335
SPECDIET	2.8795**	5.318	3.9822**	2.771
EXERCISE	1.4659**	3.799	0.3537	0.400
SMOKING	-4.3461**	-9.464	-4.4899**	-4.780
PYRAMID	0.5986**	3.681	0.0843	0.248
LAMBDA	-7.6420**	-4.300	-3.6971**	-2.073
N	4483		922	
R^2	0.153		0.171	

* indicates significance at 5% level; ** indicates significance at 1% level

as the unemployed to spend on food shopping to make the more appropriate decisions regarding the quality of foods they need to buy. Consistent with prior hypothesis, education is positively related to HEI in all label user equations.

Non-label users from central cities have an HEI that is more than two points higher than non-label users from suburban areas. On the other hand, label users from nonmetro areas have an HEI that is about 1.5 points lower than label users from suburban areas. Regionally, label and non-label users from the northeast have higher HEIs than those from south. Label users from the west and non-label users from midwest also have higher HEI than their counterparts from the south.

Nonhispanic label users have a HEI that is almost two points lower than hispanic label users. More importantly, food stamp participants who are label users have a HEI that is almost three points lower than non-food stamp participants who are label users. This result implies that the food stamp program does not improve the diet quality of participants to the level of non-participants, despite the use of the labels.

Body mass index is negatively related to HEI for label users. As expected, those who are on special diet have higher HEIs than those who are not on a special diet. In addition, label users who regularly exercise have a higher HEI than label users who do not exercise. Smokers, whether label or non-label users, have HEI which are more than four points lower than those of non-smokers. Label users with higher knowledge about the food guide pyramid also have higher HEIs.

Self-selection occurs in both label user and non-label user equations because the Mill's ratios (variable lamda) are all statistically significant. These estimates imply that self-selection bias could have occurred if the endogenous switching model was not employed in the estimation of the equations.

Label Use and Diet Quality Improvements

To evaluate the benefit of label use, we need to consider the total gross benefit for label users. For each label user with characteristics X and Z, we can compare the outcome HEI when using the label, [E (HEI_1 | I=1)] and the expected potential outcome when not using the label, [E (HEI_0 | I=1)]. Thus, their current decisions are compared to what they would have been if they had not used the labels. The expected gross benefit in terms of diet quality due to label use is

$$E(HEI_1 \mid I=1)-E(HEI_0 \mid I=0)= X'(\beta_1 - \beta_0)+(\sigma_{1u} - \sigma_{0u})\phi(Z'\gamma)/\Phi(Z'\gamma) \qquad (6)$$

The difference in the expected HEIs is calculated for all label users. The sample average of differences are reported in Table 3. The effects of consumer label use on diet quality also are estimated for each of the five types of information on the labels. Consumer label use increases the average expected diet quality by a range of 9.06 and 13.96 points, depending on the type of information. Improvement in the diet is highest when consumers use health benefit statements on the labels.

In terms of the distribution, 54 % of the sample get an improvement of between 5 to 10 points in diet quality, while 39 % get an improvement of 10 to 20 points when using labels (see General Label Use column in Table 3). For the types of information, about three-fourths of the sample get an improvement of 10 to 20 points when using ingredient, health claims, and nutrition panel, while 92 % to 98% of the sample get an improvement of 10 to 20 points when using information concerning serving size or health benefit statement in the label.

Table 3
The Effect of Consumer Label Use on the Diet Quality.

Types of information on the label used	General label use	List of Ingredient	Nutrient content claims	Nutrition panel	Serving size	Health claims
Sample average of the difference in the expected HEI	9.06	11.61	11.59	11.51	12.89	13.96
	Distribution of the average of the difference in the expected HEI(%)					
Less than 0	1.43	0.07	0.00	0.12	0.00	0.00
0 to 5	5.44	0.94	0.70	1.33	0.06	0.03
5 to 10	54.34	24.22	21.21	25.02	6.98	1.55
10 to 20	38.79	74.71	78.09	73.51	92.42	98.39
Over 20	0.00	0.05	0.00	0.02	0.00	0.03

Conclusions and Implications

Public concerns about the effect of diet on health have resulted in the legislation of the NLEA. The passage of the NLEA has been expected to provide improvement in the diet quality by encouraging consumers to make healthier food choices. An endogenous switching regression technique is employed to assess the effect of consumer label use on diet quality. The empirical results show that label use, indeed, has a positive effect in improving die quality.

The key findings in this study are of great importance in terms of public policy because of the tremendous benefits that improved diets can provide the society in general in terms of lives saved and reduction of health care costs. For instance, McNutt (1992) estimated that the health care savings from improved and better diets could amount to $3.6 billion to $21 billion. Zarkin et al.(1993) also estimated that the number of discounted life-years that could be gained nationwide during the first 20 years after the implementation of the NLEA ranges from about

40,000 to a high of 1.2 million. USDA also estimates that improved dietary patterns could save $43 billion in medical care costs and lost productivity resulting from disability associated with coronary heart disease, cancer, stroke and diabetes each year, and prevent over 119,900 premature deaths among individuals 55-84 years of age, valued at $28 billion per year (Frazao, 1995). On the other hand, the FDA estimated that the NLEA would cost the food industry $1.4 billion to $2.3 billion and the government $163 million over next 20 years. These estimates, however, are contingent upon the presumption that consumers' diets are improved by their use of food labels.

Appendix

For question about this paper you may contact Rudy Nayga by telephone, fax or electronic mail.

Telephone: (409) 845-8376
FAX: (409) 862-8679
E-Mail: rnayga@tamu.edu

References

Bowman, S.A., Lino, M., Gerrior, S.A., & Basiotis, P.P. (1998). The Healthy Eating Index: 1994-96. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion.

Bulter, J.S., & Raymond, J.E. (1996). The Effect of the Food Stamp Program on Nutrient Intakes. Economic inquiry, 34 (4), 781-98.

Frazao, E. (1995). The American Diet: Health and Economic Consequences. U.S. Department of Agricuture, Economic Research Service.

Guthrie, J., Fox, J., Cleveland, L., & Welsh, S. (1995). Who Uses Nutrition Labeling and What Effects Does Label Use e on Diet Quality? Journal of Nutrition Education, 27 (4), 153-72.

Lee, L. (1978). Unionism and Wage Rates: A Simultaneous Equation Model with Qualitative and Limited Dependent Variables. International Economic Review, 19 (2), 415-33.

Maddala, G.S. (1983). Limited Dependent and Qualitative Variables in Econometrics, New York, NY: Cambridge University Press.

McNutt, K. (1992). 3.6 to $21 Billion Benefit from New Labeling Regulations. Nutrition Today, 27, 39-43.

Nayga Jr., R.M. (1996). Determinants of Consumers' Use of Nutritional Information on Food Packages. Journal of Agricultural and Applied Economics, 28 (2), 303-12.

Variyam, J.N., Blaylock, J., & Smallwood, D. (1996). A Probit Latent Variable Model of Nutrient Information and Dietary Fiber Intake. American Journal of Agricultural Economics, 78 (3), 629-39.

Zarkin, G.A., Dean, N., Mauskopf, J.A., & Williams, R. (1993). Potential Health Benefits of Nutrition Label Changes. American Journal of Public Health, 83 (5), 17-24.

Endnotes

1 Graduate Research Assistant, Department of Agricultural Economics.

2 Associate Professor, Department of Agricultural Economics.

3 Professor, Department of Agricultural Economics.

4 Associate Professor, Department of Food Science

5 More information about the result of estimation is available from the authors upon request.