Food Labels and Weight Loss:

Evidence from the National Longitudinal Survey of Youth

Bidisha Mandal [*]

Contact Information:

School of Economic Sciences
PO Box 646210
Hulbert 101
Pullman, WA 99164
Ph: 509-335-7553
bmandal@wsu.edu

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Orlando, FL, July 27-29, 2008.

[*] Bidisha Mandal is an Assistant Professor at the School of Economic Sciences, Washington State University.

Food Labels and Weight Loss:

Evidence from the National Longitudinal Survey of Youth

Abstract

This study investigates the role of nutrition and ingredients information, included in the food labels, as useful tools for individuals who are trying to lose weight. This research has three objectives – examine personal characteristics as predictors of willingness to lose weight conditional on individual's current body mass index, investigate whether those who are trying to lose weight are more likely to read food labels to gather nutritional and ingredients information, and, analyze whether those who want to lose weight and read food labels have a greater propensity to lose weight. Estimates from random effects logistic regressions indicate higher usage of food labels by those who are trying to lose weight, irrespective of their current body mass index. There is also greater likelihood of weight loss in the user group. Future research entails use of more sophisticated econometric techniques to control for self-selection and endogeneity.

Keywords: Nutritional information, ingredient information, body mass index, panel data.

Almost all packaged food, today, carry labels providing essential nutrient and ingredient information. This is a result of the Nutrition Labeling and Education Act (NLEA) passed by the U.S. Congress in 1990 and enacted in 1994, which required all food manufacturers to present standardized nutrition facts on the package. These 'facts' include nutrient-related information such as highlighting percentage of fat and cholesterol, serving size and percent daily value. This mandating has since led to a series of economic, dietetic and behavioral research that investigate the various aspects of food labels and the impact on America's dietary intakes (e.g. Kim et al., 2000; Kristal et al., 2001; Neuhouser et al., 1999; Teisl and Levy, 1997; Variyam and Cawley, 2006; Zarkin et al., 1993).

There is a substantial literature dedicated to exploring the economic perspective of food information (not just food labels) and consumer behavior, e.g. Brown and Schrader (1990), Capps and Schmitz (1991), Chang and Kinnucan (1991), Ippolito and Mathios (1990, 1996, 1999), and Putler (1987). Some of the important findings are mentioned here. Chang and Kinnucan (1991) study the roles of cholesterol information and advertising in explaining consumption trends for fats and oils, focusing on butter. They find decrease in butter consumption in Canada due to increased awareness of the health effects, although consumers' responses to negative information seem to outweigh their responses to positive information. Brown and Schrader (1990) find similar effects with shell egg consumption, such that cholesterol information changing shell eggs' own price and income elasticities. Ippolito and Mathios (1990) find that the increase in advertising health benefits of read-to-eat cereals causes consumers to change their behavior positively and that advertising is an important source of information even after controlling for education. More recent studies in the packaged foods market show

relaxation of health-claim regulations has a positive impact on voluntary information provision, leading to a decrease in fat, saturated fat, and cholesterol consumption (Mathios and Ippolito, 1999). Variyam (2005) provides an economic assessment of a food-away-from-home nutrition labeling policy, including justifications for policy intervention and potential costs and benefits of the policy.

Zarkin et al. (1993) investigate the potential health benefits associated with changes in food consumption since the implementation of NLEA, and conclude that relatively small changes in nutrient intake may generate large public health benefits, such as gain in life expectancy and decrease in number of cases of coronary heart disease and three types of cancer. Although, they also note that not all consumers are likely to respond to the nutrition label changes. Neuhouser et al. (1999) and Kristal et al. (2001) also study the impact of NLEA, and emphasize the usefulness of food labels in limiting fat intakes. Kim et al. (2000) use an endogenous switching regression analysis to measure the impact of food labels on selected nutrient intakes, and find decreased average daily intakes of calories from total fat, saturated fat, cholesterol and sodium with the use of food labels. Finally, Teisl and Levy (1997) find nutrient labeling to significantly affect consumer purchase behavior; though the direction of the switch is ambiguous. They note that providing nutrient information may allow consumers to more easily switch consumption away from 'unhealthy' products in those food categories where taste differences are relatively small between the more and less 'healthy' products; and, consumers might switch towards 'unhealthy' products in categories where differences may be relatively large.

Meanwhile, obesity is fast approaching tobacco as the leading cause of preventable death in the U.S. (McGinnis and Foege, 1993). Annual costs of direct health care and lost productivity resulting from obesity and its consequences were estimated at $99 billion in 1995 dollars, and in 2000, these costs increased to $117 billion (Wolf and Colditz 1998; http://www.fda.gov/fdac/features/2004/304_fat.html). Science journalist Gary Taubes (2002) reports that the obesity epidemic started during the late 1970's when obesity rates shot up from 12-14% to about 22-25%. Kuchler and Golan (2004) investigate whether failure in food markets may help explain the growth of overweight and obesity in the United States. Given the constant onslaught of media coverage devoted to diet and weight, along with information from physicians, government education programs, nutrition labels, and product health claims, it is difficult to believe that Americans are not conscious of the relationship between a healthful diet and obesity. Nevertheless, the authors do find existence of two important information blackout zones – public perceptions of appropriate weight, and information on food sold at restaurants and fast-food establishments. They find that among individuals whom professionals would classify as obese, 13% said that their weight is about right or even too low. Although the NLEA require that manufacturers disclose nutritional information on the label of almost all packaged food, it does not require the same for food purchased at restaurants. This information gap is vital since the nutritional content of food from restaurants tends to be less healthy than food prepared at home (Guthrie et al., 2002).

A recent study, using repeated cross-sectional data, examines whether the nutrition labeling changes introduced by the NLEA impacted body weight among American adults between 1991 and 1998 (Variyam and Cawley, 2006). The results are

significant – the total monetary benefit due to lower mortality, reduced medical expenditures, declining absenteeism, and increased productivity associated with the reduction in body weight since the enactment of NLEA is estimated to be about $166 billion (1991 dollars) over a 20-year period compared to the $1.4 billion – $2.3 billion estimated cost of implementing NLEA. However, little is known about the use of food labels by obese and overweight individuals who are trying to lose weight. The aim of this research is to analyze the efficacy of nutritional and ingredients information included in food labels as useful tools for those who are motivated to lose weight.

Statistical Analysis

Estimation

The objective of this study is threefold. First, the relationship between willingness to lose weight and various personal characteristics is explored. Second, it is investigated whether those who reported trying to lose weight in the 2002 and 2004 NLSY79 surveys were more likely to read food labels to gather nutritional and ingredient information. Finally, the panel nature of the surveys allows analysis of whether those who wanted to lose weight and read food labels had a greater propensity to lose weight. Random effects models are constructed for each body mass index category (BMI) to answer the first two questions. A first difference model is used to study the third problem.

Let W denote whether an individual is trying to lose weight, X be a vector of demographic variables and the highest grade completed by the respondent, and Z be a vector of other personal characteristics, such as income and number of children. Then,

(1) $W = f(X, Z)$

Logistic regression is conducted for the random effects models for each BMI category. Two sets of regression are computed, first includes only the demographic variables and education. The second set includes other covariates too.

Next, let W^s and W^n be dummy variables encoding whether an individual is trying to maintain the same weight and whether an individual is not trying anything, respectively. Let N denote whether an individual reads nutritional information and I denote whether an individual reads ingredients list always, often or sometimes (as opposed to rarely or never). Then,

(2) $N = f(W, W^s, W^n, X, Z)$, and

(3) $I = f(W, W^s, W^n, X, Z)$

In the random effects logistic regressions, W^s is used as the base category and dropped. For all random effects models (equations (1) – (3)), Hausman's specification test is conducted to test for unobserved heterogeneity.

Finally, a first difference model tests whether those who reported trying to lose weight in 2002 and were reading food labels information were more likely to report a lesser body weight in 2004.

(4) $Weight\ loss = f(\Delta N, \Delta Z)$

(5) $Weight\ loss = f(\Delta I, \Delta Z)$

All time-invariant variables such as demographics, education and unobservable effects are dropped. Δ denotes change in the time-varying variables between 2002 and 2004 survey years.

Data

The empirical analysis is implemented using data from the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of 12,686 individuals, who were 14-22 years of age when first surveyed in 1979. The surveys were administered yearly till 1994, and since then the sample has been interviewed biennially. Retention rates was about 90% for the first 16 rounds of the survey, approximately 85% for rounds 17 and 18, and about 77% for latter rounds (detailed NLSY79 documentation is available at http://www.bls.gov/nls/handbook/2005/nlshc3.pdf). Since its inception, the NLSY79 consists of three sub-samples – a cross-sectional sample representative of non-institutionalized civilian youths, a supplemental group to oversample civilian Hispanic, black, and economically disadvantaged non-black, non-Hispanic youths, and a military sample of youths. This study uses survey data from the years 2002 and 2004. Table 1 presents the demographic configuration of the respondents in these two years. Women who reported being pregnant or with a biological child less than 2 years old and individuals in the military are excluded from this study. Among the common respondents in this sample, only 3% reported a different level of education between the two years – 2% reported a higher level in 2004 and 1% reported a lower level (this could be due to miscoding or misreporting). To obtain a single measure of education for each individual in both years, a predicted value is

calculated using three exogenous variables – mother's education, father's education and if any house member received newspaper regularly when the respondent was 14 years old. NLSY79 collected these three measures in the very first round of the survey. The predicted value is used throughout the analysis.

Each round of the survey consists of a core set of questions on labor market experience, education, household income, health, residence and marital status. Wording of these questions, though, might differ from year to year. Some of these factors are used as covariates in this study, and their distribution in years 2002 and 2004 is given in Table 2. The covariates are total household income (measured in $100,000), whether respondent resides in an urban area, number of weeks the respondent was employed since the last interview, if the respondent has any health limitation (although the type of limitation is not recorded), if it is a couple household, and number of biological and adopted children in the household. In selected years, NLSY79 administered additional sets of questions to gather detailed information on job search methods, migration, childcare, fertility decisions, drug and alcohol use, health behaviors, etc. Specifically, in the 2002 and 2004 rounds of survey, respondents were asked about their use of nutritional and ingredients information included in the food labels. The wording of the question is as follows:

> When you buy a food item for the first time, how often would you say you read the nutritional information about calories, fat and cholesterol sometimes listed on the label – would you say always, often, sometimes, rarely or never?

The next question asked:

> When you buy a food item for the first time, how often would you say you read the ingredient list on the package – would you say always, often, sometimes, rarely or never?

The responses to these two questions are presented in Tables 3 and 4 respectively and cross-tabulated across gender-race cohorts. For brevity, two categories have been formulated for the response type. The original survey instrument measures respondent's use of food labels on an ordinal scale corresponding to the five categories – always, often, sometimes, rarely or never. Instead, a two category scale is used in the analysis – high and low. Individuals who reported using the food label information always, often or sometimes were categorized as high end users, while those using the information rarely or never were termed as low end users. Overall, women are more likely to read the information, with white non-Hispanic women paying most attention to the nutritional information on the food labels. In general, individuals are more likely to read the nutritional information than the ingredients list, with the exception of black non-Hispanics. There is also a small increment in the number of users between 2002 and 2004.

One of the main goals of this study is to establish a relationship between demographic variables and the willingness to lose weight. Respondents to the NLSY79 report their weight and height which are then used to calculate their body mass index (BMI), ratio of weight in kilograms to the square of height in meters. Accordingly, the men and women are classified as underweight (BMI < 18.5), healthy (BMI between 18.5 and 24.9), overweight (BMI between 25 and 29.9) or obese (BMI greater than or equal to

30). These self-reported weight and height include some degree of reporting error, which may bias coefficient estimates (Judge et al., 1985). Cawley's (2004) method is used to correct this bias. True height and weight in the NLSY79 are predicted using true and reported values in the Third National Health and Nutrition Examination Survey by regressing actual weight (height) on reported weight (height) and its square for each race-gender cohort (Cawley, 2004). Corrected BMI is calculated as before, and individuals are re-categorized accordingly. The new categories are used throughout the paper.

Distribution of corrected BMI with respect to race and gender is presented in Table 5. Men are more likely to be overweight, while black non-Hispanic and Hispanic women are more likely to be obese. Women are also more likely to be underweight. Underweight individuals are dropped from this analysis given the small sample size. In both the 2002 and 2004 survey years, individuals were asked:

> Are you now trying to lose weight, gain weight, stay about the same, or are you not trying to do anything about your weight?

The responses are presented in Table 6. Two type of information is exhibited in this table. The figures represent percentage of individuals trying to lose weight, gain weight, stay the same or not do anything within each BMI category, as well as percentage of obese, overweight, healthy and underweight individuals within each weight preference category. Of those who reported trying to lose weight in each of the survey years, about 50% were obese, 36% were overweight, 13% were healthy and a very small percentage were underweight (less than 0.3%). On the other hand, approximately 4% obese, 21% overweight, 64% healthy and 10% underweight individuals reported trying to gain weight in both waves. On an average, around 3% individuals reported trying to gain weight.

This category of individuals is excluded from further analysis. The categories 'trying to stay the same' and 'not trying anything' are maintained as separate categories in the analysis to distinguish individuals based on effort. Of all those obese, about two-thirds reported trying to lose weight.

Results

Weight preference and current action

From the 2002 and 2004 survey responses it is evident that a substantial percentage of obese, overweight as well as healthy individuals were trying to lose to weight. The first aim is to relate this group to demographic variables, education and other covariates. Random effects logistic regression is conducted twice, first excluding and then including the covariates. Whether a respondent is trying to lose weight or not is the dependent variable; each of the BMI categories is analyzed separately. The marginal effects are shown in Table 7. The last row in this table presents the results from a Hausman specification test for each model. A large χ^2 value implies significant presence of latent effects, such as level of motivation.

The first set of regression which uses demographic variables and education only shows that irrespective of the BMI category, men are less likely to try to lose weight. Hispanics who are either obese or overweight are more likely to try to lose weight than white non-Hispanics, while black non-Hispanics who are overweight are less likely to try

to lose weight. Finally, more educated obese or overweight individuals and older obese individuals are more likely to try to lose weight.

Next, covariates such as income and urban residency are included. Addition of the covariates does not distort the previous results. Income has a positive marginal effect on obese and overweight individuals. Obese individuals residing in urban areas or with health limitations are more likely to try to lose weight. Overweight individuals with higher number of children (higher than the sample average) are less likely to try to lose weight.

Lastly, from the specification tests, unobserved heterogeneity does not appear to be a major issue in this simplistic analysis. Thus, a predicted probability of whether an individual is trying to lose weight conditional on his or her observed BMI category is calculated, this is to be used in the next section. Although the results are not shown here, predicted probabilities are also calculated for whether an individual is trying to maintain the same weight and whether an individual is not trying anything regarding body weight.

Use of nutritional and ingredients information

The next task involves measuring how likely individuals are to use nutritional and ingredients information on the food labels based upon their weight preference and current action. Again, random effects models are created for each weight category and logistic regression is applied. Marginal effects are presented in Tables 8 and 9 corresponding to the two types of information. Results from the specification tests are given in the last row. The dependent variable is whether an individual reads the information on the food

labels always, sometimes or often, i.e., whether an individual is a high user of such information. Current action regarding weight is the independent variable.

Regression analysis is first conducted using the observed survey responses. For the overweight and healthy subgroups, this results in a large χ^2 value implying significant presence of latent effects. Thus, analysis is revised by replacing raw data with the predicted probabilities (details in the previous section). No covariates are used in this case since the predicted values were obtained using the demographic data, education and other variables. The marginal effects are larger with the predicted values as independent variables, which might be misleading. More sophisticated econometric tools may be applied such as instrumental variables. Overall, it may be concluded that those who are trying to lose weight read nutritional information on the food labels irrespective of their current BMI category, but only obese individuals who are trying to lose weight are more likely to read the ingredients list.

Food labels and propensity to lose weight

The panel data allows one to study whether those who wanted to lose weight and read the information on food labels (in 2002) were in fact more likely to lose weight (as reported in 2004). A first difference estimation is conducted to analyze this issue, and the marginal effects are presented in Table 10. The binary dependent variable measures whether an individual lost more than 0.5 pounds between 2002 and 2004. This particular cut-off is the 60^{th} percentile of weight change between the two survey years, and is used to distinguish between loss of weight and zero gain in weight.

The first difference estimation method includes only time-varying variables, and controls for all time-invariant unobservable effects. Only those who were trying to lose weight in 2002 are eligible, and the analysis was conducted separately for two groups – those who reported not trying to lose any more weight in 2004 and those who reported still trying to lose weight in 2004. For the latter category, weight loss is significantly more likely for those who started to read nutritional and ingredients information on the food labels. For the former category, weight loss is more likely among those who continued to read the nutritional information.

Conclusion

This study has, so far, established a cohesive relationship between exogenous demographic variables and current action to lose weight conditional on present weight. Obese and overweight individuals who try to lose weight are most likely more motivated, more aware, have access to more resources or a combination of all three. Hausman's specification tests which compare random effects to fixed effects conclude absence of unobserved heterogeneity in all BMI categories. However, this test assumes that all latent effects are time-invariant. Future research entails application of more sophisticated econometric techniques to study this problem and to control for self-selection bias.

Regression results show that those who reported trying to lose weight in the 2002 and 2004 NLSY79 surveys, were significantly more likely to read nutritional and ingredients information on the food labels than those who were trying to maintain the

current weight. Note that there is a distinction between trying to maintain the same weight and not trying anything at all. The former involves some level of effort, while the latter does not. Those who were not trying anything at all regarding their weight were less likely to read food labels when buying a food for the first time. Those who reported trying to lose weight in the 2002 round of survey and not trying to lose weight anymore in 2004, continuing to read food labels produced more fruitful results – in other words, individuals who continued reading food labels between the two waves were more likely to lose weight than those who either stopped reading the labels or never read the information. On the other hand, those who reported trying to lose weight in 2004 also, found starting to read the nutritional information and ingredients list to be a helpful tool towards weight loss.

References

Brown, D.J., and L. Schrader. "Information on Cholesterol and Falling Eggshell Consumption." *American Journal of Agricultural Economics.* 72(1990): 548-555.

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

Cawley, J. "The Impact of Obesity on Wages." *Journal of Human Resources*. 39(2004): 451-474.

Chang, H.S., and H. Kinnucan. "Advertising, Information and Product Quality: The Case of Butter." *American Journal of Agricultural Economics.* 73 (1991):1195-1203.

Guthrie, J., B. Lin, and E. Frazão. "Role of Food Prepared Away from Home in the American Diet, 1997-78 versus 1994-96: Changes and Consequences." *Journal of Nutrition Education and Behavior.* 34(2002): 11-16.

Ippolito, P.M., and A.D. Mathios. "Information, Advertising and Health Choices: A Study of the Cereal Market." *RAND Journal of Economics.* 21(1990):459-480.

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

Kim, S.Y., R.M. Nayga, Jr., and O. Capps, Jr. "The Effect of Food Label Use on Nutrient Intakes: An Endogenous Switching Regression Analysis." *Journal of Agricultural and Resource Economics.* 25(2000): 215-231.

Kristal, A.R., M.M. Hedderson, R.E. Patterson, and M.L. Neuhouser. "Predictors of Self-Initiated, Healthful Dietary Change." *Journal of the American Dietetic Association*. 101(2001): 762-766.

Kuchler, F., and E. Golan. "Is There a Role for Government in Reducing the Prevalence of Overweight and Obesity?" *Choices*, an online magazine of the American Agricultural Economics Association, 19 (3rd quarter), 2004.

Mathios, A.D., and P.M. Ippolito. "Health Claims in Food and Advertising and Labeling: Disseminating Nutrition Information to Consumers." In *America's Eating Habits: Changes & Consequences,* pp. 189-212. E. Frazao (ed.), Agriculture Information Bulletin No. 750, U.S. Department of Agriculture, Economic Research Service, April 1999.

McGinnis, J., and W. Foege. "Actual Causes of Death in the United States." *Journal of the American Medical Association*. 270(1993): 2207-2212.

Neuhouser, M.L., A.R. Kristal, and R.E. Patterson. "Use of Food Nutrition Labels is Associated with Lower Fat Intake." *Journal of the American Dietetic Association*. 99(1999): 45-53.

Putler, D.S. *The Effect of Health Information on Shell Egg Consumption.* Department of Agricultural and Resource Economics Working Paper #448, California Agricultural Experiment Station, 1987.

Taubes, G. "What if it's all been a Big Fat Lie?" *New York Times Magazine*. Issue July 7, 2002.

Teisl, M., N.E. Bockstael, and A. Levy. "Measuring the Welfare Effects of Nutrition Information." *American Journal of Agricultural Economics.* 83(2001): 133-149.

Variyam, J. N. *Nutrition Labeling in the Food-Away-From-Home Sector: An Economic Assessment.* Economic Research Report No. 4, Economic Research Service, USDA, Washington, DC, 2005. http://www.ers.usda.gov/publications/ERR4/

Variyam, J.N., and J. Cawley. "Nutrition Labels and Obesity." NBER Working Paper Series, Working Paper 11956, 2006.

Wolf, A., and G. Colditz. "Current Estimates of the Economic Costs of Obesity in the United States." *Obesity Research*. 6(1998): 97-106.

Zarkin, G.A., N. Dean, J.A. Mauskopf, and R. Williams. "Potential Health Benefits of Nutrition Label Changes." *American Journal of Public Health.* 83(1993): 717-724.

Table 1: Demographic Distributions – Sample Mean (Standard Deviation) or Sample %

Variables		2002	2004
Age		40.88 (2.24)	43.16 (2.25)
Male %		48.60	47.82
Female %		51.40	52.18
White non-Hispanic %		50.76	50.36
Black non-Hispanic %		30.57	30.50
Hispanic %		18.67	19.14
Education		13.18 (2.49)	13.22 (2.52)
Number of Respondents		7644	7576

Note: Number of common respondents between the two survey waves is 7158.

Table 2: Distributions of Covariates – Sample Mean (Standard Deviation) or Sample %

Variables		2002	2004
Income (in $100,000)		0.61 (0.62)	0.64 (0.65)
Urban residency %		75.48	75.04
Weeks employed (since last interview)		85.38 (50.45)	85.08 (57.85)
Health limitation %		12.78	13.86
Couple household %		58.39	57.91
Number of children		1.38 (1.30)	1.28 (1.25)

Table 3: Distribution of Nutritional Information Use by Gender-Race Cohorts (%)

Cohorts		2002		2004	
		High	Low	High	Low
Male, white non-Hispanic		53.14	46.86	57.82	42.18
Female, white non-Hispanic		74.95	25.05	78.27	21.73
Male, black non-Hispanic		53.73	46.27	54.67	45.33
Female, black non-Hispanic		65.03	34.97	68.83	31.77
Male, Hispanic		49.92	50.08	55.19	44.81
Female, Hispanic		67.95	32.05	68.87	31.13

Notes:

1. 'High' corresponds to 'always, 'often' or 'sometimes', and 'low' corresponds to 'rarely' or 'never' on the original scale of the survey instrument.
2. In 2002, of all respondents, 62.05% read the nutritional information when buying a food item for the first time. In 2004, 65.33% reported reading the information.
3. Rows sum to 100% for each survey year.

Table 4: Distribution of Ingredients Information Use by Gender-Race Cohorts (%)

Cohorts		2002		2004	
		High	Low	High	Low
Male, white non-Hispanic		48.18	51.82	51.39	48.61
Female, white non-Hispanic		63.19	36.81	67.63	32.37
Male, black non-Hispanic		56.00	44.00	58.54	41.46
Female, black non-Hispanic		65.11	34.89	67.49	32.51
Male, Hispanic		48.20	51.80	51.85	48.15
Female, Hispanic		60.55	39.45	64.91	35.09

Notes:

1. 'High' corresponds to 'always, 'often' or 'sometimes', and 'low' corresponds to 'rarely' or 'never' on the original scale of the survey instrument.
2. In 2002, of all respondents, 57.22% read the ingredients information when buying a food item for the first time. In 2004, 60.71% reported reading the information.
3. Rows sum to 100% for each survey year.

Table 5: Distribution of Body Mass Index by Gender-Race Cohorts (%)

Cohorts		2002				2004			
		Obese	Overwt	Healthy	Underwt	Obese	Overwt	Healthy	Underwt
Male, white non-Hispanic		28.21	46.11	25.36	0.32	29.04	46.63	23.98	0.35
Female, white non-Hispanic		26.66	26.82	44.33	2.19	26.80	28.80	42.56	1.84
Male, black non-Hispanic		32.19	38.70	28.48	0.63	33.66	38.54	26.71	1.09
Female, black non-Hispanic		48.39	30.02	20.52	1.07	52.50	27.26	18.95	1.29
Male, Hispanic		31.76	48.30	19.50	0.44	33.33	46.04	20.13	0.50
Female, Hispanic		40.38	33.09	25.22	1.31	40.87	32.63	25.90	0.60

Notes:

1. Individuals were categorized into obese, overweight, healthy and underweight using self-reported weight and height and Cawley's (2004) method to correct for the reporting bias. See text for more details.
2. Underweight individuals were later dropped from the analysis.
3. Rows sum to 100% for each survey year.

Table 6: Distribution of Weight Preferences by BMI Category (%)

Cohorts		2002				2004			
		Obese	Overwt	Healthy	Underwt	Obese	Overwt	Healthy	Underwt
Trying to lose weight	Rows	49.43	36.34	14.17	0.06	51.56	35.77	12.42	0.25
	Cols	64.56	42.75	20.75	2.56	68.65	45.01	20.06	11.26
Trying to stay the same	Rows	16.98	41.08	40.84	1.10	15.21	39.98	43.53	1.28
	Cols	14.77	32.17	39.83	29.49	13.07	32.45	45.36	36.62
Trying to gain weight	Rows	4.07	20.00	65.19	10.74	3.92	23.04	62.75	10.29
	Cols	0.46	2.02	8.19	37.18	0.34	1.88	6.57	29.58
Not trying anything	Rows	27.07	34.30	37.30	1.33	28.21	34.40	36.33	1.06
	Cols	20.21	23.06	31.23	30.77	17.94	20.66	28.01	22.54

Notes:

1. Individuals were categorized into obese, overweight, healthy and underweight using self-reported weight and height and Cawley's (2004) method to correct for the reporting bias. See text for more details.
2. In 2002, of all respondents, 43.44% were trying to lose weight, 28.47% were trying to stay the same, 3.65% were trying to gain weight, and 24.44% were not trying anything. In 2004, the corresponding frequencies were 45.98%, 29.01%, 3.02%, and 21.99%.
3. Underweight individuals and those trying to gain weight were later dropped from the analysis.
4. Row cells sum to 100% across weight preference categories. Column cells sum to 100% across BMI categories.

Table 7: Marginal Effects (t-statistics) of Willingness to Lose Weight

Variables		Obese		Overweight		Healthy	
		Excluding covariates	Including covariates	Excluding covariates	Including covariates	Excluding covariates	Including covariates
Age		0.016 (3.64)	0.018 (3.73)	0.006 (1.09)	0.006 (1.06)	0.003 (0.95)	0.001 (0.22)
Male*		– 0.214 (– 8.27)	– 0.213 (– 7.49)	– 0.512 (– 17.71)	– 0.524 (– 16.53)	– 0.237 (– 12.57)	– 0.235 (– 11.23)
Black non-Hispanic*		0.019 (0.69)	0.024 (0.79)	– 0.116 (– 3.17)	– 0.102 (– 2.45)	– 0.068 (– 4.37)	– 0.061 (– 3.35)
Hispanic*		0.095 (3.17)	0.082 (2.45)	0.162 (3.50)	0.148 (2.89)	– 0.008 (– 0.34)	0.004 (0.15)
Education[§]		0.044 (3.20)	0.039 (2.59)	0.072 (4.50)	0.057 (3.20)	0.007 (0.96)	0.009 (1.17)
Income (in $100,000)			0.045 (1.77)		0.056 (2.20)		– 0.001 (– 0.13)
Urban residency*			0.098 (2.86)		0.052 (1.39)		0.021 (1.18)
Health limitation*			0.056 (1.76)		0.021 (0.38)		– 0.010 (– 0.40)
Couple household*			0.043 (1.41)		0.038 (0.96)		0.035 (1.90)
Number of children			0.015 (1.46)		– 0.029 (– 2.05)		– 0.011 (– 1.58)
Specification test (p-value)		1.93 (0.16)	9.10 (0.25)	0.29 (0.59)	12.88 (0.08)	0.28 (0.59)	7.60 (0.37)

Notes:

1. t-statistic > 1.68 implies significance at 90% level of confidence at least, and t-statistic > 1.96 implies significance at 95% level of confidence at least.
2. * denotes categorical variable. Categories that were dropped are female, white non-Hispanic, non-urban residency, no health limitation, and not a couple household.
3. For categorical variables, the marginal effect is for discrete change of transitioning from 0 to 1. Negative marginal effect implies negative effect at the mean.
4. [§] denotes use of predicted value of education instead of observed education of the respondent from the survey.
5. 'Number of weeks employed since the last interview' was used as a covariate, but is not shown here in the table. Its marginal effect is about 0.0001 with corresponding t-statistic of 0.41 across all BMI categories.
6. Individuals were categorized into obese, overweight, healthy and underweight using self-reported weight and height and Cawley's (2004) method to correct for the reporting bias. See text for more details.

Table 8: Marginal Effects (t-statistics) of Nutritional Information Use

Variables		Obese		Overweight		Healthy	
		Observed data	Predicted value	Observed data	Predicted value	Observed data	Predicted value
Trying to lose weight*		0.142 (3.81)	0.823 (6.89)	0.132 (4.74)	0.455 (6.19)	0.086 (2.85)	0.869 (4.88)
Not trying anything*		– 0.216 (– 4.29)	– 0.793 (– 2.03)	– 0.191 (– 4.61)	– 1.476 (– 3.78)	– 0.241 (– 5.93)	– 1.082 (– 5.04)
Age		0.015 (3.43)		0.015 (3.10)		0.008 (1.47)	
Male*		– 0.174 (– 6.23)		– 0.236 (– 8.77)		– 0.326 (– 7.31)	
Black non-Hispanic*		– 0.066 (– 2.00)		– 0.019 (– 0.51)		– 0.059 (– 1.15)	
Hispanic*		– 0.054 (– 1.28)		– 0.010 (– 0.24)		– 0.019 (– 0.36)	
Education[§]		0.033 (2.34)		0.043 (2.98)		0.061 (4.00)	
Income (in $100,000)		0.059 (2.38)		0.067 (3.07)		0.087 (3.75)	
Urban residency*		0.028 (0.92)		0.096 (2.73)		0.056 (1.44)	
Health limitation*		– 0.083 (– 1.99)		0.043 (1.12)		– 0.042 (– 0.77)	
Couple household*		0.059 (1.99)		0.061 (1.85)		0.041 (1.08)	
Number of children		– 0.026 (– 2.65)		0.005 (0.50)		– 0.014 (– 1.03)	
Specification test (p-value)		31.12 (0.00)	0.72 (0.69)	31.60 (0.00)	3.49 (0.17)	26.88 (0.00)	1.71 (0.43)

Notes:

1. t-statistic > 1.68 implies significance at 90% level of confidence at least, and t-statistic > 1.96 implies significance at 95% level of confidence at least.
2. * denotes categorical variable. Categories that were dropped are female, white non-Hispanic, non-urban residency, no health limitation, not a couple household, and trying to stay about the same. However, predicted weight preference is continuous.
3. [§] denotes use of predicted value of education instead of observed education of the respondent from the survey.
4. 'Number of weeks employed since the last interview' was used as a covariate, but is not shown here in the table. Its marginal effect is about 0.0001 with corresponding t-statistic of 0.28 across all BMI categories.
5. Individuals were categorized into obese, overweight, healthy and underweight using self-reported weight and height and Cawley's (2004) method to correct for the reporting bias. See text for more details.

Table 9: Marginal Effects (t-statistics) of Ingredients Information Use

Variables		Obese		Overweight		Healthy	
		Observed data	Predicted value	Observed data	Predicted value	Observed data	Predicted value
Trying to lose weight*		0.186 (4.37)	0.509 (4.06)	0.061 (1.79)	0.137 (1.66)	0.067 (1.62)	0.179 (0.85)
Not trying anything*		– 0.146 (– 2.92)	– 1.284 (– 3.13)	– 0.174 (– 4.21)	– 2.976 (– 6.26)	– 0.141 (– 3.38)	– 1.506 (– 5.75)
Age		0.025 (4.45)		0.024 (4.17)		0.020 (2.79)	
Male*		– 0.085 (– 2.66)		– 0.212 (– 6.63)		– 0.242 (– 5.23)	
Black non-Hispanic*		0.089 (2.50)		0.140 (3.84)		0.098 (2.01)	
Hispanic*		0.024 (0.52)		0.026 (0.55)		– 0.015 (– 0.23)	
Education[§]		0.027 (1.52)		0.017 (1.04)		0.072 (3.83)	
Income (in $100,000)		0.009 (0.30)		0.077 (3.04)		0.062 (2.35)	
Urban residency*		0.015 (0.42)		0.055 (1.48)		0.017 (0.37)	
Health limitation*		– 0.037 (– 0.83)		0.027 (0.53)		0.035 (0.59)	
Couple household*		0.061 (1.72)		– 0.024 (– 0.64)		0.025 (0.55)	
Number of children		– 0.017 (– 1.38)		0.011 (0.85)		0.002 (0.10)	
Specification test (p-value)		11.84 (0.22)	0.30 (0.86)	16.73 (0.05)	1.17 (0.56)	17.62 (0.04)	2.43 (0.30)

Notes:

1. t-statistic > 1.68 implies significance at 90% level of confidence at least, and t-statistic > 1.96 implies significance at 95% level of confidence at least.
2. * denotes categorical variable. Categories that were dropped are female, white non-Hispanic, non-urban residency, no health limitation, not a couple household, and trying to stay about the same weight. However, predicted weight preference is continuous.
3. [§] denotes use of predicted value of education instead of observed education of the respondent from the survey.
4. 'Number of weeks employed since the last interview' was used as a covariate, but is not shown here in the table. Its marginal effect is about 0.0001 with corresponding t-statistic of 0.15 across all BMI categories.
5. Individuals were categorized into obese, overweight, healthy and underweight using self-reported weight and height and Cawley's (2004) method to correct for the reporting bias. See text for more details.

Table 10: Marginal Effects (t-statistics) from First Difference Estimation of Weight Loss Categories

Variables (change from 2002 to 2004)		Not trying to lose weight anymore		Trying to lose weight	
		Nutritional information	Ingredients information	Nutritional information	Ingredients information
Started reading information*		0.022 (0.31)	0.050 (0.82)	0.101 (2.30)	0.094 (2.56)
Continued to read information*		0.097 (2.25)	0.064 (1.49)	0.024 (0.87)	0.016 (0.63)
Change in income (in $100,000)		0.006 (0.13)	0.013 (0.28)	0.021 (0.84)	0.020 (0.79)
Change in percentage of weeks employed		– 0.021 (– 0.38)	– 0.018 (– 0.34)	0.015 (0.49)	0.013 (0.43)
Onset of health limitation*		0.083 (0.83)	0.088 (0.89)	– 0.038 (– 0.74)	– 0.034 (– 0.66)
Cessation of health limitation*		0.191 (1.71)	0.187 (1.66)	0.098 (1.59)	0.101 (1.63)

Notes:

1. t-statistic > 1.68 implies significance at 90% level of confidence at least, and t-statistic > 1.96 implies significance at 95% level of confidence at least.
2. See text for information regarding the dependent variable.
3. * denotes categorical variable. Categories that were dropped are stopped reading information on continued to not read information, and no change in health limitation status.