Econometric Analysis of Rising Body Mass Index in the U.S.: 1996 versus 2002

Bidisha Mandal

Wen S. Chern

The Ohio State University
Department of Agricultural, Environmental & Development Economics
2120 Fyffe Road, Columbus, OH 43210

Correspondence:

Bidisha Mandal
mandal.7@osu.edu
(614-292-8339)

Wen Chern
chern.1@osu.edu
(614-292-6414)

Selected paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California, July 23-26, 2006.

Econometric Analysis of Rising Body Mass Index in the U.S.: 1996 versus 2002

Abstract

Currently over 30% of American adults are obese, more than twice the percentage prevalent in 1980 (American Obesity Association). At the same time, almost 65% adult Americans are said to be overweight. Such high prevalence levels are a major public health concern. Both overweight and obesity are associated with increased health risk for chronic diseases such as heart disease, type 2 diabetes, high blood pressure, stroke, fatty liver disease and some forms of cancer. In this paper we explore the factors that contribute to increasing rates of obesity and overweight, and study the differences in years 1996 and 2002. We use a multilevel econometric approach to model the four classifications of body mass index (BMI) – obese, overweight, healthy and underweight - as a function of individual characteristics, lifestyle indicators and external environment. The results are reasonably consistent within the two years and with findings from previous studies. However, three significant differences are found between the two years at the state-level. Two of them are completely new findings. Higher urban residency is associated with lower rates of overweight and obesity. On the other hand, higher participation in food-stamp programs in the more recent year is associated with increasing prevalence of obesity. Excise tax on cigarettes also has a positive association with obesity only. Previous studies have uses either per-capita sales of restaurants, or price of meals available in fast-food and full-service restaurants. We explored a new variable – density of fast-food and full-service restaurants serving meals over a wide price range. Such a variable is used to not only capture the importance of difference between fast-food restaurants and full-service restaurants, but to also distinguish between the effects of lower-priced and higher-priced meals. We find that lower-priced food from fast-food restaurants has positive effect, and higher-priced food from full-service restaurants has negative effect. Three new individual-level lifestyle predictors have been added, and they all seem to be significant in explaining the weight outcomes. Inadequate consumption of fruits and vegetables, irregular or no exercise, and poor self-reported health status are all significantly associated with increasing rates of overweight and obesity.

I. Introduction

Since the 1980s, Americans are talking about muscle tone, exercise routines, and being in shape (Cassell). Innumerable fitness centers promote the importance of taking control of one's health. People who are overweight are considered unfit. Media reflects this view as overweight people are usually seen in character or supporting roles, and fashion models splashed across a plethora of magazines are always thin. However, both overweight, and the greater evil obesity, are more than just cosmetic problems; they are the second leading cause of preventable death in the U.S., behind tobacco usage (McGinnis and Foege). Obesity is a complex chronic disease involving environmental, genetic, metabolic, and behavioral components. Overweight and obese individuals are at increased risk for many serious health conditions such as type 2 diabetes, coronary heart disease, stroke, gallbladder disease, and some types of cancer (Centers for Disease Control and Prevention).

Genes, understandably, are hard to control, but how easy is it to eat right and exercise regularly? In an economic model, our weight is a personal choice along several dimensions – occupation, leisure time activity, residence, and of course, food intake (Philipson). However, the persistent and steep increase in rates of overweight and obesity in the U.S. has become a major public health concern. Annual costs of direct health care and lost productivity resulting from obesity and its consequences have been estimated at more than $50 billion in 1995 dollars (Wolf and Colditz). In an agricultural society, physical activity is part of the occupation. But in a post-industrial developed society, like U.S., where most work is sedentary, one has to pay to stay in shape. Thus, the decline in work related physical activity seems to be one of the prime causes of obesity.

Another explanation may be provided by our increasing dependence on market-produced food as a substitute for household-produced food. Fast-food is often blamed for the rise in obesity. But, as the value of time increases, it is only natural to turn to food that is delivered faster. No wonder there is a higher demand for burgers than for the healthier sushi. Not only is this trend seen in the market, but also at home. Relatively inexpensive pre-cooked meals have flooded the grocery stores. So why spend half an hour over one meal, when it may be prepared in five minutes?

I.A. Previous Literature

Kuchler and Golan investigated 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 these days 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 did 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 found that among individuals whom professionals would classify as obese, 13% said that their weight is about right or even too low. Although the 1994 National Labeling and Education Act 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, Lin and Frazão). A 2003 Gallup Poll survey found that two thirds of consumers suspect that most food sold at fast-food restaurants was not good for them (Saad). However, consumers may not be able to gauge precisely the nutritional content of the food available in restaurants.

Science journalist Gary Taubes reports that the obesity epidemic started during late 70's when the obesity rates shot up from 12-14 % to about 22-25 %. He also adds that starting 1977, government started telling Americans to eat less fat. Since then a variety of diets such as low-fat and low-carb have hit the market.

Smoking habits are also, often, associated with body weight. Chou et al (2004) have shown that smoking cessation due to higher cigarette prices have resulted in increase in body mass index (BMI). The same paper does extensive analyses using other factors too, such as per capita number of fast-food and full-service restaurants, prices of a meal in each type of restaurant, food consumed at home, and alcohol, and clean indoor air laws. The data is mainly from the Behavioral Risk Factor Surveillance System (BRFSS) for sixteen continuous years from 1984 to 1999. BRFSS data is individual-specific. The researchers modify it into state-specific by converting the absolute numbers to percentage of occurrence in each state, and including sampling weights in the regression analyses. That is, instead of directly analyzing individual-specific characteristics such as the role of individuals race, gender, education, marital status, income, and age in explaining their BMI, the researchers analyze how percentage of Black non-Hispanic, Hispanic, or other races, males, less educated individuals and so on, is related to the prevalence of obesity in each state. Additional state-specific characteristics are used from other sources, such as Census of Retail Trade (CRT), American Chamber of Commerce Researchers Associations (ACCRA) Cost of Living Index, and Bureau of Labor Statistics (BLS) Consumer Price Index (CPI). Publications of CRT are available every five years only. Thus, the authors used interpolations and extrapolations of state-specific logarithmic time trends for other years. ACCRA publishes fast-food and food-at-home prices quarterly for between 250 and 300 cities. Annual state-specific prices were obtained as population-weighted averages of the city prices and then averaged over the four quarters in a given year. Fixed-effects models are created to test how much of the trend in the prevalence of the percentage of the population that is obese and in BMI can be accounted for by the state-specific variables. However, trend measures are omitted due to multicollinearity with the state-specific variables. This limits causal interpretation. The main results of this study are that the included state-specific variables have the expected effects on obesity. First, per capita number of restaurants (sum of fast-food and full-service) has a significant and positive effect in explaining the increase in obesity since 1978. Second, downward trend in food prices accounts for the upward trend in weight outcomes. And finally, as mentioned before, cigarette price has positive effect.

In general, awareness on obesity is growing. From October to December 1999 there were fewer than 50 articles in the American press about obesity and overweight, whereas during October to December 2002, there were more than 1,200 such articles (Wellness Junction). Thus we find it only natural to study the obesity scenario and compare two relatively recent years.

I.B. Objective

The objective of our paper is quite similar to that of Chou et al (2004) – what is causing the increasing rates of obesity and overweight. However, for practical purposes and relevance, we analyze cross-sectional data from two recent years; we also include more information from the individual-level, and do not convert them into state-specific percentages; finally we use different state-level variables that are easily and consistently available from all states for both years, and are modified as less as possible to avoid possible errors due to extrapolations. Our data is discussed in detail in the next section.

In this paper we investigate the factors that affect overweight and obesity, and conduct cross-sectional analysis of data from years 1996 and 2002. We believe that body weight is a function of both individual characteristics, lifestyle indicators such as smoking, exercising, health insurance coverage, and external factors, such as average cost per meal in fast-food and full-service restaurants, unemployment rate, unequal income distribution, and residence location. To this effect we employ a multilevel approach where individuals are nested within states. Thus, the methodology is one of the main differences from previous studies. Some of our variables are different based on their availability and necessity of use. However, we hope that our findings will provide continuity in establishing the factors that may account for the rising overweight and obesity.

In the next section we explain the data that has been used to conduct the analyses. In the subsequent sections we discuss the methodology and present the empirical results. Finally we present the main conclusions, and discuss scope for future research.

II. Data

Our goal is to examine several individual-level and state-level socio-economic factors that might explain the rising phenomenon of overweight and obesity in the United States. To address this we use individual-level data from BRFSS, augmented with state-level measures from the Economic Censuses (EC), Current Population Survey (CPS), Economic Research Service (ERS) and Bureau of Labor Statistics (BLS). We study two relatively recent time periods with sufficient gap in between to investigate the changes in the factors contributing to rising overweight and obesity, if at all. EC are conducted every five years, ending with 2s and 7s. Thus, years 1997 and 2002 are natural choices. However, due to inconsistent results from 1997's individual-level, we decided to use 1996 BRFSS data.

The BRFSS was established in 1984 by the Centers for Disease Control and Prevention (CDC). It conducts telephone surveys annually to monitor state-level prevalence of major behavioral risks among adults associated with premature morbidity and mortality, which are useful for planning, initiating, supporting, and evaluating health promotion and disease prevention programs. From BRFSS, we obtain the individual-level data; specifically, the dependent variable – individual's BMI; demographic variables - age, education, gender, have kids or not, marital status, race; lifestyle indicators – income, working status, smoking status, self-reported health status, health insurance coverage, participation in physical activity, and consumption of fruits and vegetables. In 2002 an additional race category is used by BRFSS – multiracial non-Hispanic. These surveys interview individuals who are 18 years of age or older only. Interviewers ask the height and weight of respondents, and then calculate the BMI themselves.

From preliminary analysis we found that the 1997 BRFSS data produces results that are contrary to the findings from previous studies (Chou et al). Those results were also extremely and impossibly different from our findings from 2002's data. Thus, instead of using 1997's individual-level data from BRFSS, we use 1996's data that agrees with previous research.

BMI is calculated as the ratio of weight in kilograms to the square of height in meters. At the individual-level, we retain only those respondents who provide complete information on the demographics and other variables of interest. Also, we discard information on respondents who are 95 years and older, since their BMI prove to be outliers more often. Specifically, 72 such observations are deleted from year 1996 and 44 observations from year 2002. After these considerations, we have 117,139 observations in year 1996, and 190,982 observations in year 2002. We then categorize BMI into underweight, healthy, overweight and obese. An individual whose BMI is less than or equal to 18.5 kg/m^2 is underweight. Those with BMI between 25 to 29.9 kg/m^2 are overweight, and those with BMI greater than or equal to 30 kg/m^2 are considered obese. BMI is the easiest and cheapest method of assessing overweight and obesity, and this classification is standard.

State-level annual average state unemployment rates and percentage of the state population living in metropolitan statistical areas are obtained from the BLS. EC gather information on industrial and business activities, and include Census of Retail Trade. From EC we obtained a wide range of average cost per meal in both full-service and fast-food restaurants. We convert these into per capita values by calculating the ratio of the numbers to the population estimates from each state. We analyze all 50 states and District of Columbia. These state-level predictors serve as proxies for availability of various types of food. EC also provides total sales and numbers of fast-food and full-service restaurants from each state, however we believe that our measure is quite appropriate because not only is it able to capture the essence of growth in numbers of both fast-food and full-service restaurants, but also the trend of food being served –

lower priced or higher priced. It is logical that decreasing price of ready-made food outside home is positively associated with increasing overweight and obesity rates.

The full-service industry comprises of establishments primarily engaged in (1) providing food services where patrons generally order or select items and pay after eating, or (2) selling a specialty snack or nonalcoholic beverage for consumption on or near the premises. Food and drink may be consumed on the premises, taken out, or delivered to the customer's location. Some establishments (except snack and nonalcoholic beverage bars) in this industry may provide these food services in combination with selling alcoholic beverages. The fast-food industry comprises of establishments primarily engaged in providing food services (except snack and nonalcoholic beverage bars) where patrons generally order or select items and pay before eating. Food and drink may be consumed on premises, taken out, or delivered to customers' location. Some establishments in this industry may provide these food services in combination with selling alcoholic beverages.

Gini coefficient is commonly used as a measure of inequality in income distribution. It ranges from 0 to 1, where 0 corresponds to perfect equality, that is everyone has the same income, and 1 means complete inequality, that is one person gets all the income and the rest get nothing. Many believe that this measure is inadequate in economies with some benefit system such as cash incentives, or food-stamp programs. Thus, additionally, we include percentage of households that participate in the food-stamp program from each state as a covariate. This information is obtained from ERS.

Table 1 gives the sample means and standard deviations for the variables of interest from both individual and state-levels for the two years. For either year, there are no restaurants that sell fast-food worth $20 or more on an average. Additionally, there were no fast-food restaurants in 1997 whose average cost per meal is $15 or higher. Similarly, there were no full-service restaurants in 2002 whose average cost per meal is less than $2.

Table 1
Descriptive Statistics

		1996-97 N = 117,139		2002 N = 190,982	
Variables		Mean	SD	Mean	SD
State-level:					
Density of Fast-food restaurants	< $2	0.000009	0.00001	0.000003	0.00001
serving meals worth	$2 - $4.99	0.000383	0.00001	0.000258	0.00007
	$5 - $6.99	0.000174	0.00006	0.000241	0.00007
	$7 - $9.99	0.000056	0.00004	0.000089	0.00004

	$10 – $14.99	0.000043	0.00002	0.000057	0.00003
	$15 – 19.99	-	-	0.000019	0.00002
Density of Full-service restaurants serving meals worth	< $2	0.000005	0.00001	-	-
	$2 - $4.99	0.000143	0.00008	0.000054	0.00005
	$5 - $6.99	0.000248	0.00010	0.000185	0.00009
	$7 - $9.99	0.000173	0.00006	0.000205	0.00006
	$10 – $14.99	0.000125	0.00007	0.000125	0.00007
	$15 – 19.99	0.000049	0.00006	0.000078	0.00005
	$20 - $29.99	0.000021	0.00002	0.000038	0.00003
	$30 and above	0.000013	0.00002	0.000021	0.00003
Gini coefficient		0.39	0.02	0.40	0.02
Food-stamp %		1.96	2.31	1.95	1.97
Unemployment rate		4.77	1.22	5.35	1.02
MSA residence %		69.02	21.76	70.35	21.82
Excise tax on cigarettes $		0.38	0.24	0.61	0.46
<u>Individual-level:</u>					
(continuous variables)					
Age		46.39	17.53	47.74	16.65
(categorical variables)		Percentage		Percentage	
BMI classifications					
	Underweight	2.6		1.8	
	Overweight	34.8		36.5	
	Obese	16.3		22.5	
	Healthy	46.2		39.2	
Education					
	College or higher	26.6		32.2	
	Some college	27.7		27.4	
	HS or lower	45.6		40.4	
Gender					
	Male	42.3		42.5	
	Female	57.7		57.5	
Children					
	No children	62.1		62.0	
	At least one child	37.9		38.0	
Race					
	White non-Hispanic	81.6		81.4	
	Black non-Hispanic	8.2		7.3	
	Other non-Hispanic	3.7		4.3	
	Hispanic	6.6		5.5	
	Multiracial non-Hispanic	-		1.5	
Marital					
	Never been married	17.0		15.8	
	Divorced/widowed/separated	26.1		26.5	
	Married/unmarried couple	57.0		57.7	
Work					
	Employed for wages	54.1		54.7	
	Self-employed	9.0		9.4	
	Unemployed	3.5		4.2	
	Unable to work	3.3		4.5	
	Retired/homemaker/student	30.1		27.2	
Income					
	$50,000 and above	24.9		43.9	

$20,000 - $49,999	49.9	36.6
$19,999 and less	25.2	19.5
Smoke		
Currently smokes	23.6	22.8
Former smoker	24.2	26.4
Never smoked	52.2	50.8
Self-reported health status		
Excellent/very good/good	85.4	84.5
Fair/poor	14.6	15.5
Health-plan		
Has health insurance	87.5	87.6
No health insurance	12.5	12.4
Fruits and vegetable consumption		
Less than 3 servings per day	34.9	38.4
3-4 servings per day	41.0	37.4
5 or more servings per day	24.1	24.2
Exercise (outside work)		
Exercise regular	70.3	76.5
Exercise irregularly or never	29.7	23.5

III. Methodology

Primarily, we are interested in investigating the factors that explain rising overweight and obesity in the United States, and if these factors contribute differently during the two years – 1996 and 2002, under study. In addition, we believe that body weight is not only a function of individual behavior, but is also a result of external environment. Thus we want to model the following:

$$\text{BMI classifications} = f\,(\underbrace{\text{individual behavior}}_{\uparrow\ \text{individual-level}} + \underbrace{\text{environmental factors}}_{\uparrow\ \text{state-level}})$$

To this extent, we use hierarchical or multilevel modeling (Raudenbush and Bryk) with two levels for our analyses since our datasets consist of individuals nested within states. As mentioned earlier, BMI may be categorized into underweight, healthy, overweight and obese. Using this classification instead of the continuous measurement as our dependent variable, we conduct a multilevel multinomial logit regression. We incorporate the state-level information into the individual-level by constructing random intercept hierarchical model. Using the notations from Raudenbush and Bryk (chapter 10), we present the details of our model below.

We are mainly concerned with a good model fit for explaining overweight and obesity. There are M = 4 possible outcomes – underweight, healthy, overweight and obese. Although this classification is itself ordered, we use 'healthy' as the comparison group. The response, R, takes on the value of m with probability $P(R = m) = \varphi_m$, for m = 1,…, M. With 4 outcomes,

$$\left.\begin{aligned} P(R_{ij} = 1) &= \varphi_{1ij} \\ P(R_{ij} = 2) &= \varphi_{2ij} \\ P(R_{ij} = 3) &= \varphi_{3ij} \\ P(R_{ij} = 4) &= \varphi_{4ij} = 1 - \varphi_{1ij} - \varphi_{2ij} - \varphi_{3ij} \end{aligned}\right] \tag{1}$$

where, $R_{ij} = 1$ implies i^{th} individual from j^{th} state is underweight, $R_{ij} = 2$ implies i^{th} individual from j^{th} state is overweight, $R_{ij} = 3$ implies i^{th} individual from j^{th} state is healthy, and $R_{ij} = 4$ implies i^{th} individual from j^{th} state is obese.

In other words, at level 1, we construct the dummy variables

$$\left.\begin{aligned} Y_{mij} &= 1, \;\; R_{ij} = m \\ Y_{mij} &= 0, \;\; otherwise \end{aligned}\right] \tag{2}$$

where, Y_{mij} is an indicator variable for category m for i^{th} individual in j^{th} state.

According to the multinomial distribution, the expected value and variance of Y_{mij}, given φ_{mij}, are then

$$\mathrm{E}(Y_{mij} \mid \varphi_{mij}) = \varphi_{mij}, \text{and } \mathrm{Var}(Y_{mij} \mid \varphi_{mij}) = \varphi_{mij}(1 - \varphi_{mij}) \tag{3}$$

The covariance between outcomes Y_{mij} and $Y_{m'ij}$ is

$$\mathrm{Cov}(Y_{mij}, Y_{m'ij} \mid \varphi_{mij}, \varphi_{m'ij}) = -\varphi_{mij}\varphi_{m'ij} \tag{4}$$

Using the notion of multinomial regression, outcome at level 1 is the log-odds of falling into category m relative to category M. We shall refer to category M, 'healthy', as the reference category. Thus, for each category m = 1,…., M-1, we have

$$\eta_{mij} = \log(\frac{\varphi_{mij}}{\varphi_{Mij}}) = \log(\frac{P(R_{ij} = m)}{P(R_{ij} = M)}), \quad for \;\; m = 1,2,3 \tag{5}$$

Thus, the multilevel multinomial regression may be written as follows:

Level 1 model (individual): $$\eta_{mij} = \beta_{0j(m)} + \beta X_{ij} \tag{6}$$

Level 2 model (state): $$\beta_{0j(m)} = \alpha_{00(m)} + \alpha Z_{0j(m)} + \mu_{0j(m)} \tag{7}$$

where $\mu_{0j(m)}$ has multivariate normal distribution with component means of 0 and variance-covariance matrix τ.

Note that in level 1 we do not include an error component because η_{mij} is already expressed as the expected values of the indicator variables for the BMI classifications. Needless to say, the underlying distribution at this level is multinomial logit as is evident from equation (5).

With $M = 4$, there are three level 1 equations, and three corresponding level 2 equations.

$\beta_{0j(m)}$ is the intercept in level 1 model for category *m*;

β is the vector of parameters corresponding to level 1 predictors given by the vector X_{ij};

$\alpha_{00(m)}$ is the intercept in level 2 model; and

α is the vector of parameters corresponding to level 2 predictors given by the vector $Z_{0j(m)}$;

Thus, the combined model, which is finally analyzed, may be written as:

$$\eta_{mij} = [\alpha_{00(m)} + \beta X_{ij} + \alpha Z_{0j(m)}] + [\mu_{0j(m)}] \qquad (8)$$

We would like to emphasize that this combined model is a sum of two parts – fixed and random – separated by brackets. The three terms in the first bracket, two alpha terms, and one beta term represent the fixed part. The term in the second bracket represent the random part, representing the variation in intercepts among states.

In level 1, we model individual's log-odds of belonging into one of the BMI categories, holding 'healthy' as the baseline group. The intercept at this level, $\beta_{0j(m)}$, is considered to be a random variable that is influenced by state-level predictors, and is expressed as a function of state specific variables $Z_{0j(m)}$ in level 2 equation. For simplicity, the slopes at all levels are assumed to be constant, and we use the same set of predictors for each category. Thus, this model provides a convenient framework for studying multilevel data and systematically analyzes how covariates measured at various levels of a hierarchical structure affect the outcome variable.

If we could get individual-level data with all relevant information then standard regression analyses would have sufficed. Historically, there are three approaches with OLS regression while dealing with hierarchical structure for a linear model. The first approach is to ignore this structure and give each individual the group or cluster values. Efficient estimation and accurate hypothesis testing based on the ordinary least squares (OLS) regression require that the random errors are independent, normally distributed, and have constant variance. Thus, if we use the continuous measure BMI, we can fit the model as: $y_{ij} = \mu + \gamma W_{ij} + r_{ij}$

where,

y_{ij} is the BMI of i^{th} individual in j^{th} state;

γ is the vector of parameters for corresponding explanatory variables given by the vector W_{ij};

W_{ij} consists of both state and individual-level characteristics;

r_{ij} is the random component;

Structurally, however, the data is hierarchical because individuals are nested within states. There are variables measured on individuals and each state. Because individuals tend to share certain state characteristics, the primary assumption of independence among observations no longer applies, i.e. individuals from a state are more homogenous than if randomly sampled from a larger population. Under the violation of this assumption, OLS regression produces standard errors that are too small. This, in turn, leads to a higher probability of rejection of a null hypothesis (Cohen et al.; Mandal and Chern).

The second approach is to obtain a mean on each predictor variable and the dependent variable for each cluster rather than individual-level values. This analysis, also called the aggregate analysis, fails to capture the within group information, leading to inaccurate conclusions (Raudenbush and Bryk). This is because the relations between aggregated variables are much stronger, and can be very different from the relations between the individual-level variables.

The third OLS approach is to analyze the regression of the dependent variable on predictors at the individual-level, but also include a set of dummy variables to represent the clusters. This method focuses on the relationship of the individual-level predictors to the dependent variable when differences among group means are taken care of (Cohen et al.). This is often called the fixed effect approach to clustering, and if the number of clusters is small, then this method is recommended for the analysis of nested data (Snijders and Bosker). This approach is the analysis of covariance (ANCOVA) model. We do not follow this procedure because not only we have 51 clusters, but also because we have several higher level covariates.

The multilevel or hierarchical model is a more precise solution to the issues discussed above, since it takes care of the violation of homoscedasticity. In such models, each cluster or group essentially has a different regression model, with its own intercept and slope. They express relationships among variables within a level, and specify how variables in different levels are associated, as they allow for the partitioning of variance into within-group and between-group components. Note that our model is more complex because the level 1 model is discrete choice whereas level 2 model is linear involving a continuous dependent variable. Parameter estimation in hierarchical generalized linear models is more complicated, involving approximations to a maximum likelihood. The most frequently used methods are based on a first- or second- Taylor series expansion around an estimate of the fixed and random portions of the model (Raudenbush and Bryk). This is referred to as penalized quasi-likelihood estimation. More precise methods are based on Gauss-Hermite quadrature or a Laplace approximation. We used 'Proc NLMixed' to carry out the computations, and the default estimation method is Adaptive Gaussian quadrature with quasi-Newton optimization technique. For detailed information on this procedure, we refer to SAS documentations.

IV. Analysis and Results

IV.A. Preliminary Analysis

We begin by comparing the data between 1996-97 and 2002, and compute the paired t-statistics for all state-level predictors, and the dependent variables at individual-level. Results are given in Table 2 below.

Table 2
Paired t-tests: 2002 values compared to 1996-97 values

Variables		t-statistic
Healthy %		– 16.12 *
Obese %		18.64 *
Overweight %		4.10 *
Underweight %		– 5.95 *
Density of Fast-food restaurants	< $2	– 3.17 *
serving meals worth	$2 - $4.99	– 12.69 *
	$5 - $6.99	5.97 *
	$7 - $9.99	4.79 *
	$10 – $14.99	3.06 *
	$15 – 19.99	9.14 *
Density of Full-service restaurants	< $2	– 2.76 *
serving meals worth	$2 - $4.99	– 9.22 *
	$5 - $6.99	– 6.04 *
	$7 - $9.99	3.65 *
	$10 – $14.99	3.05 *

	$15 – 19.99	3.32 *
	$20 - $29.99	4.68 *
	$30 and above	3.31 *
Gini coefficient		3.57 *
Food-stamp %		– 0.03
Unemployment rate		3.70 *
MSA residence %		5.27 *
Excise tax on cigarettes $		5.62 *

* $p < 0.01$

The differences between the two years are significant and in the expected direction. Percentages of overweight and obese have significantly increased in a span of six years, while the percentage of healthy has decreased. At the state level, there have also been significant increases in excise tax on cigarettes and urban residency. Unemployment rate is higher in 2002. Also, the disparity in income distribution is greater. For both types, density of restaurants selling higher-priced food has increased and lower-priced food has decreased. This is not surprising given inflation. However, because we conduct cross-sectional analyses for the two years separately, we do not use CPI deflators.

Next, we assess whether multilevel multinomial regression is indeed suitable for our data, that is, if hierarchical modeling is necessary. We test this by looking at the significance of the variance components at both levels. We note that both are significantly different from 0, thus suggesting that states do differ with respect to the various categories of BMI, and there is even greater variation among individuals within states. In Table 3, we present the estimates of the variance components. Variance component at level-1 corresponds to the error at the individual-level with multinomial logit distribution, and variance component at level-2 corresponds to the error at state-level with multivariate normal distribution. Prevalence of overweight and obesity, our primary interest, is better explained than the prevalence of underweight with our specification and model.

Table 3: Covariance parameter estimates (SE)

Variance component	Underweight 1996	Underweight 2002	Overweight 1996	Overweight 2002	Obese 1996	Obese 2002
Level-1	0.962*	0.979*	1.002*	1.002*	1.008*	1.004*
	(0.006)	(0.005)	(0.005)	(0.004)	(0.006)	(0.004)
Level-2	0.014**	0.010**	0.004*	0.003*	0.031*	0.011*
	(0.010)	(0.007)	(0.002)	(0.001)	(0.009)	(0.003)

** $p < 0.05$; * $p < 0.01$

IV.B. Main Results

In Table 4 below we present the estimates and corresponding t-ratios from the multilevel multinomial regression model. We had calculated the variance inflation factors apriori and did not find significant multicollinearity among the various predictors at the two levels.

Table 4
Estimates (t-ratios) from Multilevel Multinomial Regression
Dependent variable: BMI categories – Underweight, Overweight, Obese versus Healthy

Variables	Underweight		Overweight		Obese	
	1996	2002	1996	2002	1996	2002
Intercept	–0.70	0.75	–2.98	–2.23	–4.10	–2.32
	(–1.05)	(1.09)	(–9.93)	(–8.14)	(–6.44)	(–4.90)
<u>State level:</u>						
Density of Fast-food restaurants selling meals worth						
< $2	1676.13	–450.90	3729.20	25.42	5613.26	–3603.36
	(0.62)	(–0.09)	(3.07)	(0.01)	(2.13)	(–1.04)
$2 - $4.99	727.90	59.58	409.08	229.07	646.49	738.2
	(1.46)	(0.15)	(1.86)	(1.37)	(1.35)	(2.49)
$5 - $6.99	–292.10	739.56	51.18	–8.44	352.39	–114.91
	(–.38)	(1.63)	(0.15)	(–0.05)	(0.48)	(–0.36)
$7 - $9.99	1108.57	889.66	875.06	180.30	1908.63	184.04
	(1.12)	(0.92)	(1.95)	(0.46)	(1.97)	(0.27)
$10 – $14.99	–52.01	–1205.36	858.59	464.88	2162.22	414.51
	(–0.03)	(–1.32)	(0.94)	(1.25)	(1.09)	(0.63)
$15 – 19.99	-	-	-	610.28	-	1474.43
	-	-	-	(0.67)	-	(0.91)
Density of Full-service restaurants selling meals worth						
< $2	–3813.24	-	953.91	-	1653.93	-
	(–1.46)	-	(0.82)	-	(0.64)	-
$2 - $4.99	–172.02	439.70	–47.13	277.42	–597.71	286.06
	(–0.32)	(0.57)	(–0.20)	(0.89)	(–1.14)	(0.53)
$5 - $6.99	551.29	227.87	–431.40	–313.10	–311.46	–471.52
	(1.34)	(0.48)	(–2.40)	(–1.66)	(–0.80)	(–1.41)
$7 - $9.99	–2255.60	–1392.80	351.96	–79.98	–183.79	–226.80
	(–3.42)	(–2.56)	(1.24)	(–0.37)	(–0.30)	(–0.60)
$10 – $14.99	824.75	–179.53	–297.87	83.86	–297.95	69.71
	(1.11)	(–0.31)	(–0.91)	(0.35)	(–0.43)	(0.16)
$15 – 19.99	–287.29	–608.64	407.00	124.95	1779.52	861.41
	(–0.25)	(–0.65)	(0.79)	(0.33)	(1.57)	(1.30)
$20 - $29.99	1021.41	2593.59	–3189.84	–1101.73	–7062.95	–3063.74
	(0.40)	(1.65)	(–2.76)	(–1.70)	(–2.80)	(–2.65)
$30 and above	–2337.59	595.23	–544.39	–1913.70	–4021.69	–3097.01
	(–0.80)	(0.36)	(–0.41)	(–2.79)	(–1.34)	(–2.55)
Food-stamp %	0.002	0.001	0.003	0.02	0.01	0.03
	(0.13)	(0.03)	(0.42)	(2.25)	(0.76)	(2.65)
Gini Coefficient	0.58	–0.88	0.66	0.10	1.23	–1.26
	(0.41)	(–0.51)	(1.02)	(0.15)	(0.89)	(–1.04)

Unemployment rate	0.08	0.003	–0.003	–0.02	–0.02	–0.03
	(2.02)	(0.07)	(–0.16)	(–1.10)	(–0.46)	(–1.21)
MSA residence %	0.004	–0.001	–0.001	–0.003	–0.002	–0.004
	(1.97)	(–0.56)	(–0.87)	(–3.08)	(–0.91)	(–2.91)
Excise tax on cigarettes $	0.05	–0.07	0.09	0.06	0.21	0.17
	(0.30)	(–0.77)	(1.14)	(1.71)	(1.20)	(2.71)
Individual level:						
Age	–0.09	–0.10	0.09	0.08	0.16	0.15
	(–12.91)	(–15.53)	(27.33)	(35.95)	(34.87)	(50.48)
Age-squared	0.001	0.001	–0.001	–0.001	–0.002	–0.001
	(11.20)	(13.43)	(–23.22)	(–30.78)	(–33.27)	(–48.73)
Education:						
College or higher	0.12	0.04	–0.31	–0.28	–0.48	–0.50
	(2.07)	(0.76)	(–15.41)	(–18.73)	(–17.18)	(–27.92)
Some college	0.04	–0.09	–0.07	–0.06	–0.04	–0.04
	(0.84)	(–2.00)	(–3.44)	(–4.33)	(–1.81)	(–2.59)
HS or lower						
Gender:						
Male	–1.25	–1.01	0.92	0.90	0.42	0.53
	(–20.74)	(–20.82)	(56.90)	(76.71)	(19.66)	(38.12)
Female						
Children:						
No children	–0.08	0.03	–0.11	–0.06	–0.15	–0.05
	(–1.52)	(0.66)	(–6.01)	(–4.23)	(–6.02)	(–3.38)
At least one child						
Race:						
Hispanic	–0.32	–0.13	0.26	0.24	2.88	0.17
	(–3.03)	(–1.57)	(7.48)	(9.44)	(6.14)	(5.61)
Black non-Hispanic	–0.33	–0.26	0.52	0.53	0.78	0.84
	(–3.46)	(–3.07)	(16.65)	(21.30)	(21.21)	(32.01)
Other non-Hispanic	0.16	0.27	–0.17	–0.13	–0.13	–0.16
	(1.57)	(3.61)	(–3.83)	(–4.69)	(–2.20)	(–4.51)
Multiracial non-Hispanic	-	–0.03	-	0.19	-	0.43
	-	(–0.20)	-	(3.92)	-	(8.24)
White non-Hispanic						
Marital:						
Never been married	0.09	–0.32	–0.23	–0.21	–0.23	–0.28
	(0.62)	(–2.65)	(–3.51)	(–4.27)	(–2.57)	(–4.67)
Divorced/widowed/separated	–0.25	0.07	–0.23	–0.37	–0.53	–0.54
	(–1.62)	(0.56)	(–3.65)	(–7.61)	(–6.18)	(–8.98)
Married/unmarried couple						
*Age*Marital:*						
Age*Never	0.01	0.01	0.002	–0.001	0.004	0.01
	(1.91)	(3.30)	(1.07)	(–0.12)	(1.98)	(4.20)
Age*Div./wid./sep.	0.01	0.001	0.003	0.01	0.01	0.01
	(2.90)	(0.02)	(2.45)	(5.85)	(4.97)	(6.68)
Age*Married						

Work:						
Employed for wages	–0.30	–0.16	0.07	0.12	0.09	0.16
	(–5.24)	(–3.24)	(2.90)	(7.48)	(2.92)	(8.22)
Self-employed	–0.14	0.05	–0.02	0.004	–0.10	–0.09
	(–1.46)	(0.59)	(–0.72)	(0.19)	(–2.29)	(–3.09)
Unemployed	–0.11	–0.02	0.16	0.10	0.30	0.23
	(–0.98)	(–0.21)	(3.28)	(2.98)	(5.17)	(6.45)
Unable to work	0.48	0.16	0.06	0.02	0.32	0.23
	(4.25)	(1.81)	(1.04)	(0.46)	(5.68)	(6.53)
Retired/homemaker/student						
Income:						
\$50,000 and above	–0.18	–0.20	–0.10	0.01	–0.49	–0.13
	(–2.33)	(–4.32)	(–3.52)	(0.83)	(–13.62)	(–6.60)
\$20,000 - \$49,999	–0.10	–0.39	–0.02	–0.07	–0.23	–0.35
	(–1.80)	(–6.46)	(–1.09)	(–3.19)	(–8.48)	(–14.86)
\$19,999 and less						
Smoke:						
Currently smokes	0.53	0.46	–0.34	–0.33	–0.71	–0.68
	(10.59)	(10.71)	(–17.38)	(–22.59)	(–26.54)	(–38.11)
Former smoker	0.06	–0.16	–0.01	0.03	0.04	0.10
	(0.89)	(–3.05)	(–0.32)	(2.08)	(1.45)	(6.47)
Never smoked						
Self-reported health status:						
Excellent/very good/good	–0.45	–0.53	–0.11	–0.14	0.01	–0.65
	(–7.23)	(–10.58)	(–4.09)	(–7.36)	(0.48)	(–33.40)
Fair/poor						
Health-plan:						
Has health insurance	–0.27	0.02	0.03	0.06	–0.01	0.11
	(–4.52)	(0.43)	(1.35)	(3.21)	(–0.43)	(5.00)
No health insurance						
Fruits and veg. consump.:						
Less than 3 servings/day	0.04	0.02	0.09	0.14	0.19	0.24
	(0.61)	(0.39)	(4.42)	(9.21)	(6.81)	(13.33)
3-4 servings/day	–0.03	–0.07	0.07	0.09	0.07	0.13
	(–0.51)	(–1.52)	(3.70)	(6.41)	(2.69)	(7.73)
5 or more servings/day						
Exercise (outside work):						
Exercise regular	–0.28	–0.39	–0.04	–0.06	–0.38	–0.45
	(–5.75)	(–9.36)	(–2.39)	(–4.46)	(–17.18)	(–28.53)
Exercise irregularly/never						

$t_{0.01}$=2.58; $t_{0.05}$=1.96; $t_{0.10}$=1.64;

The results are consistent between the two years under study, and with Chou et al's study for the common variables. Overall, density of fast-food restaurants serving lower-priced meals is positively associated with increasing rates of overweight and obesity, whereas density of full-service restaurants serving higher-priced meals has negative association. No other state-level predictors turn out to be significant for 1996. In 2002, increasing excise tax on cigarettes is associated with increasing rates of overweight and obesity, consistent with previous findings. Additionally, higher food-stamp participation has positive associated, whereas increasing urban residency is associated with lower rates of overweight and obesity in 2002.

Ploeg et al (2006) find that food-stamp participation had positive association with increasing obesity among women till 1994, and by 1999-2002 this association has almost disappeared. For men, however, food-stamp participation had negative association in earlier years, though in recent years no such association is visible. Our finding is thus quite important. In 1996, we find no significance of food-stamp participation in explaining rising overweight and obesity, but in 2002, it is an important predictor for obesity though not for overweight.

At the individual-level the similarities between the two years are that men, less educated individual, individuals with children, Hispanic, Black non-Hispanic, multiracial non-Hispanic and individuals who do not eat adequate amounts of fruits and vegetables everyday are more likely to be overweight and obese. Other non-Hispanics (mainly Asians), individuals with higher income, those who report good health and those who exercise regularly are less likely to be overweight and obese. Higher prevalence of obesity is noted among individuals with health insurance. This could mean a greater burden on state and federal budgets as BMI continues to rise. Finkelstein, Fiebelkorn, and Wang found that Medicare and Medicaid pay for at least half of obesity-attributable medical expenses. This means that what would otherwise be a matter of personal choice has become a matter of concern for all taxpayers.

We included two quadratic terms - square of age, and interaction between age and marital status. Age is the only continuous predictor at the lowest level. From preliminary analysis we noticed a U-shaped relationship between age and BMI classifications. Thus, the age-squared term is included to capture the curvature. For both years, though overweight and obesity rates increase with age, it starts to decrease after a certain point. Also from preliminary analyses we found that not including the interaction between age and marital status resulted in opposite direction of association between obesity and ‘never been married’. Thus, we include the interaction to show that those who are single are less likely to be obese, but with age this association is reversed. Interestingly, the interaction is insignificant in explaining overweight. This was the only interaction causing any change at the lower level. Hence no other interaction term was included in the analysis. Divorced or widowed or separated individuals also tend to be less overweight and obese, but with age the prevalence increases in this category of marital status.

One interesting finding is that, for both years, current smokers are less likely to be both overweight and obese, whereas former smokers are more likely to be obese. Chen, Yen and Eastwood showed that such a result should be interpreted very carefully due to the endogeneity of smoking. In our case, the assumedly exogenous factor at the state-level, the increasing excise taxes on cigarettes can partially explain this. It is possible that many who smoked earlier ceased to do so as a result of increasing price of cigarettes, and the smoking cessation caused weight gain.

Individual work status is and important explanatory variable. We club together retired, homemakers and students, because from preliminary analyses these three groups show similar trends with respect to their weight outcomes. This is the base category in work status. At the state-level, unemployment turns out to be completely insignificant. However, from the lower-level we find that self-employed individuals are the healthiest. Both employed and unemployed individuals, the latter slightly more so, are more likely to be overweight and obese. This also explains the insignificance of unemployment at the state-level – two opposite characteristics showing the same effect.

V. Conclusions

We followed a multilevel approach to locate significant explanatory variables for the increasing trend in overweight and obesity. There are two levels under scrutiny – individual, and state. That is, our weight outcomes are not only decided by our demographics and lifestyle, but also influenced by external factors. We study two recent time periods, six years apart – 1996 and 2002.

At state-level we find that type of restaurants and average cost per meal somewhat account for the weight outcomes for both years. Lower-priced foods from fast-food restaurants have positive effect, and higher-priced foods from full-service restaurants have negative effect. This result could also be interpreted as increasing substitution of lower-priced ready made fast-foods in our day to day lives is leading to unhealthy weight outcomes. The effect of higher-priced food from full-service restaurants is less intuitive; it could be that higher-priced foods from these restaurants are of inherently better quality, thus not negatively affecting our body weights. Increasing excise tax on cigarettes and increasing number of households participating in food-stamp programs are more likely to be obese, though not overweight. The former result has been much discussed in previous literature; our finding is based on more recent years and supports previous studies. The latter is a significant outcome of this research because previous studies have shown that in earlier years, 1976-80 and 1988-94, this was positively associated with weight outcomes, and since then the significance of food-stamp participation has been diminishing in explaining the increasing rates of overweight and obesity. We show that though this is true for 1996, for the more recent year 2002, food-stamp participation is again positively associated with increasing prevalence of obesity. Thus, once again, this feature has come to the forefront and calls for attention.

Another important finding of our study is that increased urban residency is associated with lower rates of both overweight and obesity. In the absence of other supporting state characteristics, it is difficult to interpret this result, but we believe that it can be explained by presence of certain urban characteristics such more number of gyms, more health consciousness among people in the cities, and higher likelihood of walking to work in densely populated cities. Unemployment rate, however, does not seem to be a significant predictor for either year. But this can be explained the labor market status at the individual-level. We find that both employed and unemployed individuals are more likely to be overweight and obese. Self-employed individuals, on the other hand, are healthier. This is not surprising; most jobs are sedentary in developed countries. We also find evidence that higher income is associated with lower weight outcomes, supporting the belief that one has to pay to stay in shape. This could be either through gym membership, or consuming higher-priced yet good quality food.

Consumption of adequate amounts of fruits and vegetables and participation in regular physical activities imply lower weight outcomes. Those with health-insurance tend to be obese; this should worry the state and federal health budgeters. Finally, those who reported excellent, very good or good health, have significantly lower weight outcomes than those who reported fair or poor health, implying increasing awareness.

In general, increase in age signifies increase in overweight and obesity rates. However, as is shown by the quadratic term, after a certain age these rates start to decrease again. Men, Black non-Hispanics, multiracial non-Hispanics, Hispanics and less educated individuals are at significantly higher risk. Unfortunately, current smokers have significantly lower body weight than non-smokers, whereas former smokers are more likely to be obese. Comparing with the outcome from the state-level, it seems that higher price of cigarettes leads to more smoking cessation which might lead to weight gain as is argued in many medical studies.

Additional variables such as kind of job (blue collar or white collar), proximity of fast-food restaurants from work place, type of food consumed at home, whether or not parents and/or close relatives are obese, and frequency of consumption of fast-foods would have contributed greatly to this study. For simplicity we used only one type of multilevel structure – random intercept hierarchical model. However, one could also try random slopes and random intercept plus slopes hierarchical models. Such models would test if the characteristics at one particular level are affected by the characteristics from other levels. In spite of these issues we have many interesting findings.

First, as far as we know, this is a first multilevel approach to model individual's weight outcomes. Most of our findings for the individual-level are consistent with those of previous studies (Philipson; Chou et al.). Second, we have shown that densities of different types of restaurants serving lower-priced and higher-priced food are important predictors. Third, at the state-level we have shown the importance of urban residency, and food-stamp participation in recent years. Finally, we have included various lifestyle indicators not previously analyzed. Inadequate consumption of fruits and vegetables, irregular or no exercise, and worse self-reported health status are all significantly associated with increasing rates of overweight and obesity.

Thus, we can confidently state that more individuals today recognize overweight and obesity as health hazards. Those who exercise, and consume healthy amounts of fruits and vegetables, are suitably fit and healthy. However, we know that certain groups of people are inherently more susceptible to overweight and obesity, such as Black non-Hispanics, multiracial non-Hispanics, Hispanics, individuals who are unable to work, individuals from lower income categories, and less educated people. They need immediate attention given that this epidemic has been around for a while now.

References

American Obesity Association, available on the World Wide Web: http://www.obesity.org/

Cassell, J.A., (1995). Social Anthropology and Nutrition: A Different Look at Obesity in America. *Journal of the American Dietetic Association*, 424-427.

Centers for Disease Control and Prevention (CDC), available on the World Wide Web: http://www.cdc.gov/

Chen, Z., Yen, S., and Eastwood, D., (2005). Does Cigarette Consumption Have a Causal effect on Weight Reduction? Submitted to *Journal of Family Economics and Issues* in April 2005

Chou, S., Grossman, M., and Saffer, H., (2004). An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics*, 23, 565-587.

Cohen, J., Cohen, P., West, S.G., and Aiken, L.S., (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Third Edition). Mahwah, NJ: Earlbaum.

Guthrie, J.F., Lin, B.H., and Frazão, E., (2002). 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, 11-16.

Kuchler, F., and Golan, E., (2004). 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).

Mandal, B., and Chern, W., (2006). Changes in Factors Contributing to Rising Body Mass Index: 1997 versus 2002. Forthcoming in *Consumer Interests Annual*, 52.

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

Philipson, T., (2001). The World-wide Growth in Obesity: An Economic Research Agenda. *Health Economics*, 10, 1-7.

Ploeg, M.V., Mancino, L., and Lin, B., (2006). Food Stamps and Obesity: Ironic Twist or Complex Puzzle? *Amber Waves*, February 2006 Features Article, United State Department of Agriculture.

Raudenbush, S.W., and Bryk, A.S., (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (Second Edition). Newbury Park, CA: Sage.

Saad, L., (2003). *Public Balks at Obesity Lawsuits*. Princeton, NJ: The Gallup Organization.

Snijders, T.A.B., and Bosker, R.J., (1999). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Thousand Oaks, CA: Sage.

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

Wellness Junction, available on the World Wide Web: http://www.wellnessjunction.com/athome/selfcare/epidemic.htm

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