The Review of Regional Studies 2009, Vol. 39, No. 2, pp. 149–169

Examining Diet Quality and Body Mass Index in Rural Areas using a Quantile Regression Framework*

Jebaraj Asirvatham+

ABSTRACT. Under- and overeating (based on calories) are different problems than being under- and overweight (based on BMI); they thus require different economic prescriptions. To study these differences, I use a quantile regression method that allows an examination of covariate effects at different points in the distribution. The examination focuses on differences in rural and urban areas, and also in characteristics such as education, age, and other economic, demographic and behavior variables. Results suggest that policy in rural areas and the South must emphasize recommendations based on dietary guidelines. I also find that the characteristics typical of rural areas have a large effect regardless of the region.

Keywords: Rural, region, health, quantile regression, diet, BMI, healthy eating index

JEL Classifications: R15, I10, C52

1. INTRODUCTION

One of the biggest public health concerns related to poor diet and physical inactivity is obesity, which reduces productivity, increases morbidity and is a risk factor for several other conditions such as hypertension, coronary heart disease and diabetes. In 2004, a press release from the Office of the Surgeon General reported that deaths caused by poor diet and physical inactivity had risen 33 percent in the past decade, which made those factors second only to tobacco as the leading cause of death in the United States (US).[1] Previous studies on diet and obesity have generally studied them in an Ordinary Least Squares (OLS) framework or have focused on specific groups. In this paper I model the factors affecting an important health outcome, the Body Mass Index[2] (BMI), in a quantile regression framework. I also examine diet quality, which in this study is measured by caloric intake and the Healthy Eating Index (HEI).

Special emphasis will be placed on urbanization and regional variables because of their unique characteristics. Even though rural areas are relatively more homogenous, their characteristics vary across regions. Therefore it is important to study not only the rural-urban differences in diet quality and BMI, but also regional differences in rural areas. Studying these differences in the conditional distribution of the three variables will be the first objective of this study. The second objective is to examine differences in diet quality and BMI across socio-economic characteristics typical of rural populations. This objective is designed to identify

* I thank Dr. Paul E. McNamara for discussions in the initial stages of this research. Excellent editorial comments from Maria Crawford and Joel Cuffey are gratefully acknowledged. Still, any errors in the writing, interpreting, and other parts of this paper are solely mine. The reviewers and organizers of this special section for this journal, including Matthew Fannin, have made valuable comments and suggestions that have improved this draft considerably for which I am very thankful. I am grateful to God for the opportunity to conduct this research.

+ Graduate Student, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, 326 Mumford Hall, 1301 West Gregory Drive, Urbana, IL 61801-3605. *Email:* ajebaraj@yahoo.com.

[1] March 9, 2004. http://www.hhs.gov/news/press/2004pres/20040309.html

[2] More information on BMI can be found at: http://www.cdc.gov/nccdphp/dnpa/healthyweight/assessing/bmi/index.htm

ISSN 1553-0892
SRSA, 1601 University Avenue, PO Box 6025, Morgantown, West Virginia 26506-6025, USA.

specific needs of given areas, whether they are resource allocation, consumer education, or other public health initiatives.

Research has shown that increase in caloric intake rather than lack of caloric expenditure played a larger role in enhancing body weight (Cutler, Glaeser, and Shapiro, 2003). Therefore, I focus on diet quality while accounting for differences in caloric intake and BMI due to physical activity. Moreover, this dataset does not have information on caloric expenditure other than the information on the frequency of physical activity. I use this variable to understand its effect on the three variables of interest.

The lowest quantile of caloric intake indicates hunger problems, while upper quantiles show excess intake problems. Similarly, in the case of BMI, individuals are underweight at the lower end of the distribution but are obese at the upper tail of the distribution. Each of the three variables contains information that shows problems of different natures at different points in the distribution. It is these differences in the distribution of energy intake—HEI and BMI—that make the use of quantile regression more compelling than OLS, which has the drawback of assuming the same effects across the entire distribution.[3] Quantile regression allows us to examine whether the covariate effects are different for each type of problem. This method has also proven useful in analyzing specific nutrients; for example, Variyam, Blaylock, and Smallwood (1998) used it to find different effects of education and income on the distribution of fat and cholesterol.

The next section discusses differences across levels of urbanization and regions. Section 3 briefly discusses the need for studying caloric intake, HEI and BMI. Section 4 introduces the empirical approach followed by a brief description of the econometric technique in Section 5. I then turn to a data description in Section 6 and the discussion of results in Section 7. Conclusions are drawn and limitations discussed in Section 8.

2. RURAL-URBAN AND REGIONAL DIFFERENCES

Regional and rural variables account for many differences including market access, population density and climate. These variables can also capture dietary habits or practices, dietary preferences, price of meals, and opportunity cost of time or other preferences that are specific to some areas. Thus "rural" is a mix of various socio-cultural, demographic, economic and locational characteristics that have been discussed in previous literature (for example, Edelman and Menz, 1996; Sobal, Troiano, and Frongillo, 1996) and hence will be only briefly discussed here.

Areas designated as rural, suburban or urban are based on population density and the degree of economic and social integration with urbanized areas[4]. The density criterion suggests that rural populations are more dispersed, implying lesser demand per unit area and thus making these areas less attractive for retailers (Kaufman, 1998). The integration criterion reflects access to markets for food products. While these two criteria together may indicate the availability of food items to some degree, they are likely a stronger indicator of the diversity of food items available in a rural area. Rural characteristics result in higher food prices, which influence or become a constraint on making good dietary choices (Nord and Leibtag, 2005; Olson et al.,

[3] Further discussion on quantile regression can be found in the econometrics section.

[4] The term "rural areas" used in this study is the "non-metropolitan statistical area" (non-MSA) as defined by the Office of Management and Budget.

2004). This forms the basis of our first hypothesis, that diet in rural areas is less varied and thus generally less healthy.

Rural areas generally have an older population, lower education levels, lower incomes, lower access to healthcare, and are less ethnically or racially diverse (cf., Arcury et al., 2005; Eberhardt and Pamuk, 2004). Most of the health disparities across regions are due to these differences in population characteristics. Among the well-established characteristics that explain these differences are the linkages between education and health (Cutler and Lleras-Muney, 2006) and income and health (Schnittker, 2004).

The rural population also has less access to healthcare because of the reduced supply of healthcare professionals as well as limited availability of specialized equipment and distant health care facilities (Bailey, 2009). The health disparities in rural areas show up in higher death rates of children and adults, higher rates of chronic disease and higher rates of unhealthy behavior such as alcohol consumption and smoking.

Further, there is much demographic and economic heterogeneity in rural areas across the country. For example, 91 percent of all rural African-American families live in the South, while three-fourths of rural Hispanic families live in the West or in the state of Texas (Fuguitt, 1995). In spite of these specific demographic characteristics differing in rural areas, some studies find that such patterns are more dynamic over longer periods of time (Fuguit, 1995; Johnson, 2006).

Technically, the categorization of areas into rural, suburban and urban is based on the definition of metropolitan statistical area (MSA) adopted by the Office of Management and Budget. The MSA central city is considered urban; the MSA non-central city is suburban; and the non-MSA area is considered rural. The central city and non-central city in an MSA are defined in Sections 5 and 6 of the *Revised Standards for Defining Metropolitan Areas in the 1990s* on the website of the U.S. Census Bureau.[5] The rest of the areas are termed rural.

3. DIET QUALITY AND HEALTH OUTCOMES

In this paper I suggest that the characteristics of rural areas could affect BMI and that socio-demographic factors accentuate overweight and obesity problems. The HEI was developed by the USDA to measure the healthfulness of diet by comparing an individual's intake to the dietary guidelines, evaluating an individual's diet composition based on Food Guide Pyramid (FGP) recommendations that emphasize variety, proportionality and moderation (Welsh, Davis and Shaw, 1993). For instance, it distinguishes between the calories obtained from fats and those from fruits or vegetables. Five of the ten components that make up HEI directly measure how well the intake of the five food groups—grains, vegetables, fruit, dairy products and meat—meets dietary guidelines. The other five components reflect the proportion of fat and saturated fat to total food energy, milligrams of cholesterol, sodium and dietary variety.[6] For example, if four servings of fruit are recommended, consuming four or more will result in an HEI score of 10, while two servings will receive a score of 5.

Several nutrition studies have reported that macronutrient distribution of the diet affects human body weight and composition (Samaha et al., 2003; Normand et al., 2001). In a clinical

[5] http://www.census.gov/population/www/metroareas/mastand.html

[6] More information on HEI can be found in documents available at: http://www.cnpp.usda.gov/publications/hei/hei94-96report.pdf, and http://www.cnpp.usda.gov/publications/hei/HEI89-90report.pdf.

trial, it was found that high-protein meals produce higher postprandial fat oxidation rates compared to a balanced protein diet in both obese and lean subjects (Labayen et al., 2004). Higher oxidation rates imply higher fat absorption and thus higher body mass. Caloric intake is specifically studied since HEI scores do not distinguish between under- and overconsumption.

Both excess calories and an imbalanced diet cause changes in body weight, although these changes may take different amounts of time depending on the amount of excess calories and the extent of deviation from dietary guidelines. Cutler, Glaeser, and Shapiro (2003) estimated that an increase of 10 to 12 pounds in median weight over two decades requires 100 to 120 excess calories per day. This also implies that a much higher caloric intake and a more unbalanced diet would increase the BMI in less time.[7] Studies on obesity have found variations or differences in diet quality across population based on individual characteristics such as education level, income, health knowledge, physical activity, genetics and other demographic and economic variables (Finke and Houston, 2003; Kim, Nayga, and Capps, 2001; and Variyam, Blaylock, and Smallwood, 1998). However, a more detailed study on diet and health focusing on regional patterns in rural areas is lacking in the literature.

4. EMPIRICAL APPROACH

Cawley (2004) presents an economic framework for understanding physical activity and eating behaviors. I use a similar framework, assuming that individuals seek to maximize utility from eating food and producing a health outcome[8]. The health outcome, BMI, is influenced by diet quality, physical activity, environment and genetic factors, among others. Diet quality, as discussed earlier, is evaluated by caloric intake and HEI. Low HEI scores, however, do not distinguish between under- and overconsumption. Therefore it is important to consider caloric intake especially when deciding on a policy to provide more nutrition or better nutrition education on meeting dietary guidelines. To understand the factors affecting energy intake, healthiness of diet, and BMI, the following analytical framework is adapted:

$$BMI = f\left(E(X,R), HEI(X,R), \varepsilon\right);$$
$$Energy = f\left(X,R,u\right);$$
$$HEI = f\left(X,R,v\right),$$

where E represents energy, the vector R includes urbanization and regional variables, and the vector X includes individual characteristics such as age, gender, race, ethnicity, education, income, employment, nutrition knowledge and label use. The random error terms in the models are ε, u and v respectively.

5. ECONOMETRIC FRAMEWORK

In addition to estimating the model using quantile regression (QR) to understand the different covariate effects in the distribution of energy intake, HEI and BMI, I compare these results to OLS results. The main advantage of quantile regression over linear regression methods is that quantile regression provides for a more complete statistical analysis of the stochastic relationship among random variables by estimating the entire conditional output distribution (Koenker and Hallock, 2001). Ordinary least squares regression (OLS) limits comparison to

[7] One unit BMI would be an equivalent of about 7 lbs. for an individual who is 70 inches tall.

[8] This dataset did not include a time diary of activities.

mean levels, while quantile regression characterizes the heterogeneous effects of the covariates at different quantiles of the dependent variable. Although the point of reference changes based on the quantile estimated, this method does not lose any observations. Instead, each observation is given different weights based on its distance from the quantile being estimated. For instance, while estimating quantiles at the lower end of the distribution, the observations at the upper end of the distribution receive lower weights based on the distance from the point of reference. This is achieved by assigning different weights to the residuals in the objective function, as explained below.

The estimates in the linear QR model have the same interpretation as those of an OLS model except for the respective quantiles where each is estimated. Quantile regression characterizes the conditional distribution in the presence of heteroskedasticity (Deaton, 1997), and its estimators are also more efficient than OLS when the error terms are non-normal (Buchinsky, 1998). Further, quantile regression solves a weighted sum of absolute deviations, which makes the estimators insensitive to outliers on the dependent variable and thus makes it a robust measure of location.

The econometric framework, based on Buchinsky (1998), is specified below. Let the true model be $y_i = x_i^{'}\beta_\theta + u_{\theta i}$, and $quant_\theta\left(y_i \mid x_i\right) = x_i^{'}\beta_\theta$ denote the conditional quantile of y_i, conditional on the regressor vector x_i. The distribution of the error term is not specified in quantile regression, but the error term is assumed to satisfy the exogeneity restriction in each quantile. Statistically, this could be written as $quant_\theta\left(u_{\theta i} \middle| x_i\right) = 0$. The estimator for β_θ of the θ^{th} quantile is obtained by minimizing a sum of asymmetrically weighted absolute residuals:

$$\min_\beta \frac{1}{n}\left\{ \sum_{i:y_i \geq x_i^{'}\beta} \theta \mid y_i - x_i^{'}\beta \mid + \sum_{i:y_i < x_i^{'}\beta} \left(1-\theta\right) \mid y_i - x_i^{'}\beta \mid \right\}$$

This framework allows the marginal effects or response of the covariates, given by β_θ, to vary across quantiles. As is standard in the literature, I obtain quantile regression estimates at the 10$^{\text{th}}$, 25$^{\text{th}}$, 50$^{\text{th}}$, 75$^{\text{th}}$ and 90$^{\text{th}}$ percentile which is referred to as 0.10, 0.25, 0.50, 0.75 and 0.90 quantiles, respectively. This gives a comparison of estimates at symmetrical ends of the distribution and the median can be compared to the mean estimates.

The criteria for the selection of the Continuing Survey of Food Intakes by Individuals (CSFII) sample were sex, age and income, which are all used as control variables in the econometric model. However, the complex survey design affects the estimates and their variances (Lee, Forthofer, and Lorimer, 1989). A bootstrapping method with 200 replications was employed to obtain unbiased estimates. The following section describes the data used in this study.

6. DATA DESCRIPTION

This article uses the Continuing Survey of Food Intakes by Individuals (CSFII) and the Diet and Health Knowledge Survey (DHKS), two nationwide surveys conducted by the Agricultural Research Service (ARS), USDA, from 1994-1996.[9] The target population of these

[9] There are more recent survey datasets on food consumption that do not have the full set of variables included in the CSFII. Leaving such variables out of the analysis would undoubtedly lead to omitted variable bias, which would likely lead to overestimates of coefficients for the variables of interest. Hence, these datasets were not considered in the present analysis.

surveys was non-institutionalized individuals in all 50 states and Washington, DC. The sample of individuals provided food intakes for two non-consecutive days. The DHKS was administered to one person per household age 20 or above within the CSFII sample to measure attitude and knowledge about diet and health. The overall DHKS response rate was 73.5 percent.

Studies on surveys have found that underreporting is a problem, especially in dietary intake and other self-reported health measures such as the BMI. To reduce underreporting, ARS adopted a multiple-pass approach, i.e., collecting information on dietary intake 3 to 10 days apart. To reduce respondent burden, an exclusive 24-hour recall method was administered by two interviewers.

A study in Spain found that survey respondents underestimate weight and overestimate height (Basterra-Gortari et al., 2007). However, comparing self-reported and measured BMI among the 10,639 participants of the National Health and Nutrition Education Study III (US study), McAdams, Van Dam, and Hu (2007) showed that the correlation between the two measures was 90 percent or higher. They further found that biomarkers were equally correlated with both measures. In a short review, Chiolero, Peytremann-Bridevaux, and Paccaud (2007) suggest that the use of BMI categories could yield biased estimates but using it as a continuous variable does not. In this study, I use BMI as a continuous variable and therefore do not expect considerable bias.

Table 1 shows that mean energy intake was slightly lower in rural areas than in urban areas, but the standard deviation was slightly higher. Rural residents scored lower on average in healthiness of their diet. This could also reflect less variety in the foods available in rural areas compared to urban areas that have more diverse population demographics. In the case of BMI, the index for rural areas was slightly higher than for urban or suburban areas.

TABLE 1. Weighted Summary Statistics of the Dependent Variables

Variable	Urbanization	mean	*SD*	Percentile				
				p10	p25	p50	p75	p90
Energy (kcal)								
	Urban	1977	*870*	1081	1435	1828	2347	3015
	Suburban	1991	*842*	1085	1408	1848	2421	3076
	Rural	1960	*895*	1027	1351	1767	2385	3170
	Overall	1979	*863*	1063	1399	1823	2386	3073
Healthy Eating Index (HEI)								
	Urban	63.32	*11.69*	48.44	55.03	63.44	72.19	78.61
	Suburban	63.86	*11.38*	49.02	56.10	63.52	72.20	79.11
	Rural	60.90	*11.06*	47.02	53.30	60.66	68.22	75.94
	Overall	63.01	*11.46*	48.43	54.90	62.96	71.36	78.38
Body Mass Index (BMI)								
	Urban	26.12	*5.36*	20.25	22.31	25.25	28.82	33.09
	Suburban	26.20	*5.26*	20.67	22.71	25.34	28.48	32.55
	Rural	26.62	*5.70*	20.41	22.71	25.69	29.68	33.47
	Overall	26.28	*5.40*	20.53	22.60	25.39	28.89	33.00

TABLE 2. Variable Definitions

Variable	Subcategories	Definition
Age		Number of years
Education	*Less than HS*	= 1 if less than high school else = 0
	High School graduate	= 1 if high school graduate only else = 0
	Some college	= 1 if one to three years in college else = 0
	College graduate	= 1 if college graduate else = 0
Percent above poverty		CPI-adjusted for income and household size
Employed		= 1 if full time or part time else = 0
Gender		= 1 if female else = 0
Race	*White*	= 1 if White else = 0
	Black	= 1 if Black else = 0
	Other race	= 1 if Other race else = 0
Ethnicity	*Hispanic*	= 1 if Hispanic else = 0
Smoke		= 1 if smoker else = 0
Label Use		= 1 if using nutrition labels else = 0
*Physical Inactivity**		Indicates frequency of activities
Diet and health knowledge score		Based on correct answers to 14 diet and health related questions
Urbanization	*Urban*	Metropolitan statistical area, central city
	Suburban	Metropolitan statistical area, non-central city
	Rural	Non-metropolitan statistical area
Region	*Northeast*	= 1 if residing in Northeast else = 0
	Midwest	= 1 if residing in Midwest else = 0
	South	= 1 if residing in South else = 0
	West	= 1 if residing in West else = 0
Year	*1994*	= 1 if surveyed in 1994 else = 0
	1995	= 1 if surveyed in 1995 else = 0
	1996	= 1 if surveyed in 1996 else = 0

* Physical inactivity scores were as follows: 1 = "Daily", 2 = "5 - 6 times per week", 3 = "2 - 4 times per week", 4 = "Once a week", 5 = "1 - 3 times per month", 6 = "Rarely or never"

Table 2 describes the variables used in this study, and Table 3 presents the descriptive statistics. The rural population is on average four years older and less educated than their urban counterparts. Only 20 percent of rural residents surveyed had a college degree, compared to 32 percent for urban and 27 percent for suburban residents. Rural areas have a smaller proportion of highly educated people and a larger proportion of people who have not finished high school (LHS). Income is reported in two ways: absolute income category and income relative to the poverty level. The percent above poverty level is preferred because it is based on CPI-adjusted

adjusted income and household size. Percent above poverty level, employment and gender distribution did not differ much between rural and urban areas. In terms of distribution of race groups, the rural and suburban populations were less diverse than the urban population. About 88 percent of the rural population was white, while only 5 percent was black. In contrast, the urban population was 68 percent white and 22 percent black.

I included regional dummy variables for Northeast, Midwest, South and West to examine regional differences. This regional classification is based on the nine Census divisions defined by the U.S. Census Bureau. Only 23 percent of the survey respondents lived in a rural area, while 45 percent lived in suburban areas. A majority of the sample was from the South (35 percent), while the fewest came from the Northeast region (19 percent). The southern region had the highest rural population (40 percent), followed by the Midwest (30 percent).

The behavioral variables also varied across urbanization levels. The frequency of exercise was lower in rural areas, indicating lower caloric needs relative to the urban areas. The proportion of rural residents who read labels for nutritional information was about nine percentage points lower than that of urban or suburban residents. Diet and health knowledge is calculated based on responses to 14 questions in the DHKS which are listed in Appendix A. Each correct answer received one point, which adds to a maximum attainable 14 points. This knowledge was higher in rural than in urban areas. About the same percentage of the population smoked in rural and urban areas.

7. RESULTS

Factors affecting caloric intake (energy model) and HEI were estimated to understand their association with BMI. To interpret the results, it is important to understand that the coefficients of the quantile regression in the three models depend on the quantile estimated. In the energy model, relatively low caloric intake at the lower end of the distribution indicates relatively more hunger problems, while higher intake at the upper levels of the distribution would indicate overconsumption. The BMI model has a similar interpretation. Relatively low BMI in the lower end of the distribution would thus indicate problems associated with being underweight, while higher BMI at the upper levels of the distribution would indicate more problems with being overweight. In the HEI model, however, the interpretation is different since the score increases with the healthiness of the diet. Therefore, relatively low scores at the lower levels of the distribution indicate a poor diet, but higher scores at the upper levels of the distribution indicate a better diet.

The coefficient estimates of the indicator or dummy variables reported in the tables are relative to the omitted category. Among the regional indicator variables, Northeast was excluded since it was the least heterogeneous across urbanization levels. The primary interest of this study is rural areas, so the rural variable was retained and the urban variable omitted in the regression models. In spite of controlling for most of the individual characteristics, we find significant rural-urban and regional differences in some points on the distribution.

7.1. Regional and urbanization variables

There were very few significant differences observed across regions and among rural areas in the different regions in the energy model. In Table 4, rural and suburban residents showed no difference in consumption relative to urban residents. Among the regions, Midwest showed the highest caloric intake and South the lowest. Based on the quantile regression

TABLE 3. Weighted Means of Explanatory Variables by Urbanization Levels

Variable	Subcategories	Urban		Suburban		Rural	
		Proportion or Mean	*Std error*	Proportion or Mean	*Std error*	Proportion or Mean	*Std error*
Age (years)		44	*0.71*	45	*0.51*	48	*0.49*
Education:	*Less than HS*	0.15	*0.01*	0.12	*0.01*	0.19	*0.03*
	High School graduate	0.30	*0.02*	0.36	*0.02*	0.38	*0.02*
	Some college	0.23	*0.01*	0.25	*0.01*	0.23	*0.01*
	College graduate	0.32	*0.02*	0.27	*0.02*	0.20	*0.04*
Income (percent above poverty)		227	*3.65*	253	*2.73*	222	*5.03*
Employed (full time or part time)		0.62	*0.01*	0.66	*0.01*	0.61	*0.02*
Gender		0.57	*0.02*	0.54	*0.01*	0.55	*0.02*
Race:	*White*	0.68	*0.03*	0.89	*0.01*	0.88	*0.05*
	Black	0.22	*0.03*	0.06	*0.01*	0.05	*0.02*
	Other race	0.05	*0.01*	0.02	*0.00*	0.00	*0.00*
Ethnicity:	*Hispanic*	0.09	*0.01*	0.08	*0.01*	0.08	*0.06*
Smoke		0.24	*0.01*	0.24	*0.01*	0.27	*0.02*
Label Use		0.59	*0.02*	0.61	*0.02*	0.51	*0.03*
*Physical Inactivity**		3.92	*0.08*	3.83	*0.06*	3.62	*0.06*
Diet and Health Knowledge Score		8.30	*0.15*	8.94	*0.09*	8.79	*0.10*
Region:	*Northeast*	0.18	*0.02*	0.25	*0.02*	0.09	*0.09*
	Midwest	0.18	*0.02*	0.25	*0.02*	0.30	*0.01*
	South	0.37	*0.04*	0.31	*0.03*	0.40	*0.08*
	West	0.26	*0.04*	0.19	*0.03*	0.21	*0.02*
Year:	*1994*	0.36	*0.03*	0.31	*0.01*	0.32	*0.01*
	1995	0.32	*0.02*	0.34	*0.01*	0.34	*0.01*
	1996	0.32	*0.02*	0.34	*0.01*	0.34	*0.01*

Notes: *Physical inactivity scores were as follows: 1 = "Daily", 2 = "5-6 times per week", 3 = "2-4 times per week", 4 = "Once a week", 5 = "1-3 times per month", 6 = "Rarely or never"

estimates, residents of the southern region consumed less only at the 0.25 quantile (=0.25*q*), by 60 calories, and those in the Midwest consumed more at the median quantile, by 113 calories.

Even though the OLS coefficient for West was insignificant relative to Northeast, the quantile estimates were significant only at the 0.90 quantile (= 0.90*q*) by 134 calories. Among the rural areas across regions, only a few differences were observed. The Midwest showed higher consumption by 159 calories in the OLS, but no difference in quantile regression estimates.

TABLE 4. Quantile Regression Estimates of the Energy Model (*n*=5,413)

Variables	Quantile					OLS
	0.1	0.25	0.5	0.75	0.9	
Age	-10.55	-11.24	-14.14	-16.19	-31.12	-16.87
	(0.01)*	(0.01)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Age squared	0.06	0.04	0.04	0.03	0.14	0.05
	(0.08)***	(0.26)	(0.22)	(0.54)	(0.05)***	(0.14)
Black	-4.31	-63.83	-68.53	3.82	-55.92	-38.54
	(0.92)	(0.03)**	(0.05)**	(0.94)	(0.39)	(0.23)
Other race	-59.50	-106.92	14.77	16.52	176.01	-21.67
	(0.35)	(0.01)**	(0.80)	(0.83)	(0.08)***	(0.65)
Hispanic	30.49	-64.87	-81.53	14.94	-22.08	-56.59
	(0.40)	(0.17)	(0.09)***	(0.86)	(0.79)	(0.17)
Female	-382.40	-474.45	-615.64	-805.19	-980.22	-693.66
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Employed	49.01	14.35	32.63	59.26	31.25	50.89
	(0.10)***	(0.57)	(0.14)	(0.10)	(0.48)	(0.02)**
% Poverty	1.41	0.55	0.52	1.10	0.21	0.71
	(0.06)***	(0.44)	(0.45)	(0.24)	(0.86)	(0.27)
% Poverty^2	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.14)	(0.88)	(0.75)	(0.35)	(0.93)	(0.40)
HS Grad	103.23	72.18	61.95	51.21	103.43	57.80
	(0.01)*	(0.02)**	(0.03)**	(0.26)	(0.08)***	(0.08)***
Some College	175.61	124.83	110.89	81.02	93.62	105.48
	(0.00)*	(0.00)*	(0.00)*	(0.08)***	(0.13)	(0.00)*
College Grad	181.87	116.19	96.26	39.01	43.92	60.04
	(0.00)*	(0.00)*	(0.01)*	(0.48)	(0.49)	(0.09)***
Smoke	-43.83	-11.11	23.04	66.25	237.39	66.11
	(0.17)	(0.68)	(0.40)	(0.15)	(0.00)*	(0.01)*
*Physical Inactivity**	-10.08	-6.70	6.12	-3.11	-14.81	-9.13
	(0.06)***	(0.22)	(0.26)	(0.70)	(0.12)	(0.06)***
Label Use	-30.33	-59.91	-102.52	-94.94	-109.37	-85.96
	(0.22)	(0.01)**	(0.00)*	(0.00)*	(0.01)*	(0.00)*
Knowledge	20.91	10.07	15.22	19.09	15.56	17.40
	(0.00)*	(0.02)**	(0.00)*	(0.00)*	(0.07)***	(0.00)*

Table 4 Cont'd

Table 4 Cont'd

Variables	0.1	0.25	0.5	0.75	0.9	OLS
Midwest	28.81	49.40	113.15	66.90	25.09	61.04
	(0.48)	(0.17)	(0.00)*	(0.17)	(0.69)	(0.06)***
West	-44.04	48.14	54.94	70.56	133.61	47.87
	(0.22)	(0.25)	(0.13)	(0.20)	(0.06)***	(0.17)
South	-37.48	-60.31	-40.67	-48.73	-72.09	-71.82
	(0.25)	(0.05)***	(0.14)	(0.25)	(0.21)	(0.02)**
Suburb	6.07	-30.61	-12.80	-0.75	-37.35	-7.88
	(0.84)	(0.24)	(0.60)	(0.98)	(0.44)	(0.76)
Rural	-26.72	-55.02	-31.75	81.12	57.88	-46.60
	(0.79)	(0.48)	(0.56)	(0.53)	(0.61)	(0.50)
Midwest-Rural	38.27	50.59	78.03	106.05	228.23	158.76
	(0.72)	(0.57)	(0.23)	(0.48)	(0.15)	(0.05)**
South-Rural	-62.76	-40.92	-96.23	-230.93	-176.07	-74.44
	(0.53)	(0.64)	(0.10)	(0.09)***	(0.19)	(0.31)
West-Rural	102.90	72.07	72.38	-75.43	-100.65	84.22
	(0.38)	(0.51)	(0.33)	(0.62)	(0.49)	(0.32)
Year 94	-23.47	5.75	29.03	56.14	33.03	19.86
	(0.40)	(0.83)	(0.19)	(0.12)	(0.49)	(0.41)
Year 95	-21.36	2.11	23.38	7.17	12.84	9.87
	(0.41)	(0.94)	(0.36)	(0.84)	(0.79)	(0.67)
Constant	1,327.36	1,968.02	2,470.74	3,121.46	4,277.37	2,747.85
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Pseudo-R^2	0.104	0.126	0.159	0.189	0.223	0.281

Note: *p* values in parentheses. *** significant at 10%; ** significant at 5%; * significant at 1%

Southern residents showed no significant differences in OLS but lower consumption at the 0.75*q* by 231 calories.

Concerning the healthiness of their diet, rural residents had very low HEI scores compared to urban residents (Table 5). Furthermore, the quantile regression estimates for rural residents were significantly lower in the lower quantiles up to the median and showed no significant differences at the upper quantiles of the distribution. Among the regions, the South, on average, had lower scores than Northeast in the OLS and all the quantiles except the lower tail, while Midwest had lower scores only at the median and 0.75q. No significant differences were observed in rural

TABLE 5. Quantile Regression Estimates of the Health Eating Index Model (*n* = 5524)

Variables	Quantile					OLS
	0.1	0.25	0.5	0.75	0.9	
Age	-0.19	-0.10	-0.13	-0.08	0.07	-0.07
	(0.03)**	(0.16)	(0.08)***	(0.25)	(0.43)	(0.18)
Age^2	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.31)	(0.00)*
Black	-3.52	-2.54	-3.13	-2.36	-3.19	-2.95
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Other Race	3.40	2.63	2.59	2.79	1.81	2.53
	(0.01)**	(0.01)**	(0.00)*	(0.00)*	(0.09)***	(0.00)*
Hispanic	2.08	1.33	1.48	1.86	2.15	1.96
	(0.09)***	(0.14)	(0.10)	(0.03)**	(0.07)***	(0.00)*
Female	0.99	1.36	1.88	2.02	2.09	1.71
	(0.03)**	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Employed	-0.50	-0.67	-1.08	-1.23	-1.99	-1.01
	(0.39)	(0.21)	(0.04)**	(0.02)**	(0.00)*	(0.00)*
% Poverty	0.02	0.03	0.01	0.00	-0.01	0.01
	(0.24)	(0.02)**	(0.24)	(0.83)	(0.30)	(0.09)***
% Poverty^2	-0.00	-0.00	-0.00	0.00	0.00	-0.00
	(0.94)	(0.19)	(0.92)	(0.48)	(0.07)***	(0.91)
HS Grad	1.23	1.23	1.65	1.49	1.82	1.61
	(0.05)***	(0.04)**	(0.01)*	(0.02)**	(0.01)*	(0.00)*
Some College	1.92	2.36	2.18	1.57	2.45	2.15
	(0.01)**	(0.00)*	(0.00)*	(0.03)**	(0.00)*	(0.00)*
College Graduate	4.73	4.78	4.88	4.81	4.13	4.81
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Smoke	-3.46	-3.68	-4.19	-4.60	-4.34	-4.15
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Physical	-0.20	-0.22	-0.31	-0.32	-0.33	-0.29
*Inactivity**	(0.16)	(0.07)***	(0.00)*	(0.01)*	(0.00)*	(0.00)*
Label Use	3.03	2.84	3.40	3.74	3.22	3.13
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Knowledge	0.54	0.58	0.54	0.48	0.42	0.50
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*

Table 5 cont'd.

Table 5 Cont'd.

Variables	0.1	0.25	0.5	0.75	0.9	OLS
Midwest	0.03	-1.15	-1.13	-1.55	-1.20	-1.17
	(0.97)	(0.13)	(0.10)***	(0.03)**	(0.13)	(0.02)**
West	-0.15	-0.54	-0.43	-0.27	-0.21	-0.26
	(0.87)	(0.51)	(0.60)	(0.72)	(0.73)	(0.63)
South	-0.26	-1.75	-1.97	-2.59	-1.94	-1.89
	(0.75)	(0.02)**	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Suburb	-0.29	-1.12	-0.78	-0.63	-0.65	-0.67
	(0.66)	(0.02)**	(0.17)	(0.22)	(0.21)	(0.05)***
Rural	-2.63	-3.82	-3.11	-1.28	-1.55	-2.41
	(0.07)***	(0.06)***	(0.07)***	(0.39)	(0.50)	(0.05)***
Midwest-Rural	3.08	3.01	1.03	0.35	-0.73	1.11
	(0.11)	(0.16)	(0.61)	(0.84)	(0.76)	(0.43)
South-Rural	0.64	1.19	-0.65	-1.67	-2.72	-0.62
	(0.66)	(0.57)	(0.71)	(0.31)	(0.26)	(0.64)
West-Rural	2.28	1.68	-0.27	-2.08	-1.14	-0.26
	(0.25)	(0.47)	(0.90)	(0.25)	(0.67)	(0.85)
Year 94	0.03	0.29	0.21	0.30	0.65	0.23
	(0.96)	(0.58)	(0.70)	(0.57)	(0.20)	(0.45)
Year 95	-0.28	0.32	0.69	0.46	0.42	0.42
	(0.66)	(0.56)	(0.17)	(0.37)	(0.43)	(0.19)
Constant	41.36	45.19	54.60	61.93	66.95	53.60
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Pseudo-R^2	0.098	0.105	0.122	0.141	0.143	0.219

Notes: p values in parentheses; *** significant at 10%;;** significant at 5%; * significant at 1%

areas across the regions (region-rural interaction dummies). Lower HEI quantiles suggest diets that have less variety and do not meet dietary guidelines regardless of under or overconsumption. Hence, the significantly lower HEI scores in lower quartiles for rural areas and the South suggest unhealthy dietary habits for these areas, habits that are unhealthy even compared to Northeast Urban areas.

In regard to average BMI, there were no differences across urbanization levels and regions except for Midwest (Table 6) in the OLS model. However, in the quantile regression the rural variable is significant at 0.75q and Midwest only at the median quantile. These results suggest that in the case of BMI, all regions and rural areas across all regions face similar problems of an overweight and obese population.

TABLE 6. Quantile Regression Estimates of the Body Mass Index Model (*n* = 5417).

Variables	0.1	0.25	0.5	0.75	0.9	OLS
Age	0.21	0.24	0.31	0.41	0.37	0.33
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Age^2	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Black	0.52	1.42	1.92	2.87	3.46	2.05
	(0.04)**	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Other race	-0.18	0.02	-0.60	-0.50	-0.95	-0.33
	(0.54)	(0.94)	(0.06)***	(0.35)	(0.32)	(0.37)
Hispanic	0.82	0.98	0.97	0.67	0.56	0.69
	(0.01)**	(0.00)*	(0.00)*	(0.18)	(0.46)	(0.03)**
Female	-1.86	-1.64	-1.15	-0.35	0.97	-0.66
	(0.00)*	(0.00)*	(0.00)*	(0.14)	(0.01)*	(0.00)*
Employed	0.23	0.24	0.08	-0.47	-0.46	-0.18
	(0.19)	(0.16)	(0.68)	(0.06)***	(0.30)	(0.36)
% Poverty	0.01	-0.00	-0.00	-0.01	-0.02	-0.00
	(0.19)	(0.73)	(0.88)	(0.27)	(0.08)***	(0.36)
% Poverty^2	-0.00	0.00	-0.00	0.00	0.00	0.00
	(0.42)	(0.68)	(0.72)	(0.75)	(0.28)	(0.98)
HS Grad	-0.06	-0.15	-0.63	-0.40	-0.78	-0.53
	(0.79)	(0.50)	(0.01)*	(0.23)	(0.11)	(0.02)**
Some College	-0.80	-0.70	-1.07	-0.87	-1.07	-1.09
	(0.00)*	(0.01)*	(0.00)*	(0.02)**	(0.05)***	(0.00)*
College Graduate	-0.97	-1.19	-1.89	-1.94	-2.25	-1.89
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Smoke	-1.56	-1.23	-1.27	-1.33	-1.08	-1.41
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.01)**	(0.00)*
Physical	-0.03	0.06	0.18	0.30	0.43	0.18
*Inactivity**	(0.41)	(0.12)	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Label Use	0.13	0.17	0.35	0.21	0.26	0.34
	(0.43)	(0.31)	(0.02)**	(0.34)	(0.48)	(0.02)**
Knowledge	0.04	0.07	0.05	0.07	0.11	0.08
	(0.21)	(0.05)**	(0.07)***	(0.21)	(0.13)	(0.01)*

Table 6 cont'd.

Table 6 cont'd.

Variables	0.1	0.25	0.5	0.75	0.9	OLS
Midwest	0.24	0.37	0.60	0.42	0.56	0.52
	(0.38)	(0.14)	(0.01)**	(0.15)	(0.33)	(0.02)**
West	-0.08	-0.34	-0.21	-0.15	0.06	-0.11
	(0.77)	(0.13)	(0.37)	(0.70)	(0.92)	(0.66)
South	0.16	0.06	0.20	0.09	-0.13	0.14
	(0.56)	(0.77)	(0.36)	(0.78)	(0.82)	(0.48)
Suburb	0.11	0.06	0.17	0.29	-0.02	0.08
	(0.58)	(0.76)	(0.31)	(0.25)	(0.96)	(0.64)
Rural	-0.26	0.17	0.51	1.58	0.67	0.61
	(0.71)	(0.76)	(0.30)	(0.03)**	(0.47)	(0.21)
Midwest-Rural	0.55	0.32	0.24	-0.19	0.83	0.18
	(0.49)	(0.61)	(0.67)	(0.81)	(0.46)	(0.73)
South-Rural	0.17	-0.02	-0.13	-0.69	0.02	-0.30
	(0.82)	(0.97)	(0.81)	(0.41)	(0.98)	(0.57)
West-Rural	-0.17	0.00	-0.47	-1.36	-0.43	-0.61
	(0.83)	(1.00)	(0.40)	(0.13)	(0.72)	(0.33)
Year 94	0.16	0.22	0.18	0.39	0.77	0.37
	(0.41)	(0.17)	(0.24)	(0.10)	(0.05)***	(0.04)**
Year 95	-0.20	-0.10	0.02	0.60	0.87	0.29
	(0.27)	(0.60)	(0.90)	(0.02)**	(0.03)**	(0.10)
Constant	16.28	17.41	18.50	19.98	25.86	19.28
	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*	(0.00)*
Pseudo R^2	0.070	0.062	0.057	0.061	0.077	0.086

Notes: *p* values in parentheses. *** significant at 10%; ** significant at 5%; * significant at 1%

7.2. Socioeconomic and demographic variables

In regard to the variables other than urbanization and region, the results were generally consistent with previous literature. Positive effects of education were observed in calorie intake, HEI and BMI. On average, high school graduates and college graduates consumed about 60 calories more than those with less than a high school diploma, and those with some college education consumed about 105 calories more (Table 4). However, the quantile regression estimates show that while consumption was higher at higher levels of education, it was about the same for high school and college graduates at the lower quantiles of caloric intake. Furthermore, the amount of calories decreased with the quantile up to the median, and thereafter no specific patterns are observed except that those with some college education consumed 81 calories in 0.75q, while the high school graduates consumed over 100 calories more than the less-than-high-school group in the upper tail of the distribution.

The income variable did not show any effect except at 0.1q. Even though being employed increased caloric intake by about 50 calories, this effect was seen only in the lower tail (0.1q). Women, compared to men, consumed considerably fewer calories; the differential increased from 400 calories in the lower tail to almost 1,000 calories in the upper tail of the distribution. Older individuals consumed on average 17 calories fewer than younger individuals, but the quantile regression estimates suggest that the differential increased across the entire distribution. There was almost no difference across racial groups except in select quantiles.

In terms of HEI, education levels showed bigger differences than all of the other variables except for the diet knowledge variable (Table 5). The more educated individuals consistently showed a healthier diet relative to their less-educated peers in each of the quantiles. College graduates in all quantiles were close to 5 index points higher than those who had not finished high school. Those with some college had about 1.5 to 2.5 more points; high school graduates had only 1 to 2 more points than those without a high school diploma. Even though the income variable showed an increase in HEI, the magnitude was very small. Those who were employed had about 1 to 2 fewer points than the unemployed, but only in the median and upper quantiles. Among the demographics, the race variable showed larger differences than other variables. African Americans consumed a less healthy diet and the Other Race group a more healthy diet compared to White Americans. Hispanics also ate a healthier diet than did non-Hispanics.

In terms of BMI, again a positive effect of education but not of income can be seen. At each quantile, the slope of the education gradient is negative, indicating a lower BMI with higher education. The average BMI of a college graduate was close to two points less than that of someone with less than a high school diploma, but the quantile estimates show that the BMI range was about 1 to 2.2 points lower. Among the demographics, the differential across race was the biggest, with African Americans on average two points higher than Whites. However, the quantile estimates show that the difference increased with levels in the BMI distribution.

7.3 Behavior and information variables

In regard to caloric intake, smokers consumed on average 66 calories more than did nonsmokers (Table 4). Nonetheless, the quantile regression estimates show that smokers consumed as much as 237 calories more than nonsmokers, but only in the upper tail of the calorie intake distribution. Those who read nutrition labels consumed about 86 fewer calories than those who do not, but this difference ranged from 60 in 0.25q to about 110 in 0.90q. Diet knowledge showed a positive association with caloric intake, showing an average effect of 17 calories but a range of 10-21 calories in the quantile regression estimates.

All the behavior variables showed some effect on HEI. Those who used nutrition labels had a healthier diet by about 3 points, and this varied very little across the entire distribution (Table 5). Diet knowledge caused a 0.5 index point increase on average in healthfulness of the diet. One must keep in mind that the knowledge variable ranged from 0 to 14, which implies that the effect is larger than for other binary behavior variables. However, the differential decreased slightly as the diet quality improved, i.e., at the higher levels of HEI. The estimates of the physical inactivity variable suggest that there was a negative association: those exercising less frequently had relatively lower HEI. Those who smoked had a considerably less healthy diet than did nonsmokers.

BMI was positively associated with nutrition label use and diet knowledge. Even though the OLS estimate is positive, implying that label use and diet-health knowledge are positively correlated with BMI, the quantile estimates show that the relationship is positive only at the median quantile for both; and at 0.25q for diet knowledge. Smokers have a BMI that is about 1.5 points lower than that of nonsmokers. The physical inactivity variable showed a positive effect on BMI: those who exercised less frequently had a higher BMI.

8. CONCLUSIONS

This study helps in understanding critical differences across urbanization levels and regions as well as the socioeconomic and demographic characteristics in the conditional distribution of diet choices—total caloric intake and the Healthy Eating Index (HEI)—and an important health outcome, Body Mass Index (BMI). Three regression models were analyzed using quantile regression techniques to understand differences across the conditional distribution.

Compared to the OLS, the use of the quantile regression method helped in more fully understanding the nature of the relationships across the conditional distributions of energy intake, HEI and BMI. For example, the OLS estimate was insignificant for West but significant for Midwest in the energy model. However, the quantile estimates were significant in the upper tail in the case of West and at the median for Midwest. This has important implications since higher consumption at lower levels of caloric intake is not a problem (example, Midwest), but higher consumption at upper levels is a problem that contributes to higher BMI (example, West). Another example from among the behavioral variables is that label use turns out to be positively correlated with the BMI, implying a negative health impact. However, the quantile regression estimates clearly show that this effect is statistically significant only at certain points in the distribution where it is of little or no concern. Further, the results were consistent with existing theory and empirical research on education, age and physical activity.

In general, the differences across urbanization levels, regions and rural areas across regions were the greatest in the case of HEI, but less for calories and least in the case of BMI. There are no rural-urban or regional differences in the upper tail of the distribution of calories, HEI or BMI after controlling for the various socioeconomic and demographic factors. The same pattern is observed in all the regions, except for Midwest in calories and South in HEI. In the case of calories, this suggests that there are fewer differences across regions or urbanization levels. Similarly, there are obesity issues in the US irrespective of region. However, in the case of HEI, people are making healthy choices regardless of urbanization levels and regions. From all three models, we can note that the characteristic features of the rural population such as age, income and education only accentuate the problem of low diet quality in rural areas.

Although the above discussion on rural areas shows some nutrition and health disadvantages, a positive note is that several modifiable behavior variables such as label use and physical activity have positive health effects. For nutrition labeling to have an impact, people need to have knowledge of the elements of a balanced diet Since label usage is lowest among rural respondents, more effort should be made to educate them to use it as a tool in making healthier dietary choices. The emphasis should be on helping people make informed choices and not just providing information. A long-term solution, however, would be to invest in education, which has significant positive effects on the healthiness of diet and on BMI. The positive effects were higher in magnitude at higher levels of education in the quantile regression models discussed above.

Although diet quality and physical activity are among the most important factors affecting BMI, environment and genetics should not be understated. Not considering genetics and some of the environmental factors such as family situation or parental characteristics represents some of the limitations of this study. However, regional and urbanization variables do capture general environmental differences such as road networks and nature of work. The dataset used in this study does not have a time diary, which limits us from controlling for caloric expenditure due to time spent on physical activity and other leisure activities as well as the intensity of these activities. However, we do control for the frequency of exercise in each of the models, which shows some effects.

In summary, the energy and HEI models in a quantile regression framework clearly show that rural residents have lower diet quality, indicating an unhealthy diet, but no differences in excess caloric intake compared to their urban counterparts. Among the regions, there were fewer differences in calorie intake and BMI but more differences in HEI. Therefore nutrition policies to help individuals make healthful choices that meet the dietary guidelines should be emphasized in rural areas and especially in the South. Any policies that are geared towards reducing excess caloric intake or maintaining a healthier BMI should be aimed at a national level. Still, it would be more useful to target policies that focus on specific groups, such as less educated people, within these areas.

REFERENCES

Arcury, Thomas A., Wilbert M. Gesler, John S. Preisser, Jill Sherman, John Spencer, and Jamie Perin. (2005) "The Effects of Geography and Spatial Behavior on Health Care Utilization among the Residents of a Rural Region," *Health Services Research*, 40, 135–156.

Bailey, Jon M. (2009) "The Top 10 Rural Issues for Health Care Reform," *A Series Examining Healthcare Issues in Rural America*, 2, March last accessed September 2010 at http://files.cfra.org/pdf/Ten-Rural-Issues-for-Health-Care-Reform.pdf.

Basterra-Gortari, F. Javier, Maira Bes-Rastrollo, Lluis Forga, Alfredo J. Martinez, and Miguel A. Martinez-Gonzalez. (2007) "Validación Del Índice De Masa Corporal Auto-Referido En La Encuesta Nacional De Salud," *Anales del Sistema Sanitario de Navarra*, 30, 373–381.

Buchinsky, Mosche. (1998) "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *Journal of Human Resources*, 33, 88–126.

Cawley, John. (2004) "An Economic Framework for Understanding Physical Activity and Eating Behaviors," *American Journal of Preventive Medicine*, 27(3, Supplement 1), 117–125.

Chiolero, Arnaud, Isabelle Peytremann-Bridevaux, and Fred Paccaud. (2007) "Associations between Obesity and Health Conditions may be Overestimated if Self-Reported Body Mass Index is used," *Obesity Reviews*, 8, 373–374.

Cutler, David M. and Adriana Lleras-Muney. (2006) "Education and Health: Evaluating Theories and Evidence." *National Bureau of Economic Research (NBER) Working Paper Series* 12352: Cambridge, MA.

Cutler, David M., Edward L. Glaeser, and Jesse M. Shapiro. (2003) "Why have Americans Become More Obese?" *Journal of Economic Perspectives*, 17, 93–118.

Deaton, Angus. (1997) *The Analysis of Household Surveys : A Microeconometric Approach to Development Policy*. Baltimore, MD: Published for the World Bank by Johns Hopkins University Press.

Eberhardt, Mark S. and Elsie R. Pamuk. (2004) "The Importance of Place of Residence: Examining Health in Rural and Nonrural Areas," *American Journal of Public Health*, 94, 1682–1686.

Edelman, Mark A. and Brian L. Menz. (1996) "Selected Comparisons and Implications of a National Rural and Urban Survey on Health Care Access, Demographics, and Policy Issues," *Journal of Rural Health*, 12, 197–205.

Finke, Michael S. and Sandra J. Huston. (2003) "Factors Affecting Probability of Choosing a Risky Diet," *Journal of Family and Economic Issues*, 24, 291–303.

Fuguitt, Glenn V. (1995) "The Impact of Migration on the Nonmetropolitan Population Age Structure, 1960-1990," *Population Research and Policy Review*, 14, 215–232.

Johnson, Kenneth M. (2006) "Demographic Trends in Rural and Small Town America," *Reports on America*, 1, 1–35.

Kaufman, Phil R. (1998) "Rural have Poor Access to Supermarkets, Large Grocery Stores" *Rural Development Perspectives*, 13(3), 19–26.

Kim, Sung-Yong, Rodolfo M. Nayga, Jr., and Oral Capps Jr. (2001) "Health Knowledge and Consumer use of Nutritional Labels: The Issue Revisited," *Agricultural and Resource Economics Review*, 30, 10–19.

Koenker, Roger and Kevin F. Hallock. (2001) "Quantile Regression," *Journal of Economic Perspectives*, 15, 143–156.

Labayen, Idoia, Nieves Díez, Dolores Parra, Alvaro González, and Alfredo J. Martínez. (2004) "Basal and Postprandial Substrate Oxidation Rates in Obese Women Receiving Two Test Meals with Different Protein Content," *Clinical Nutrition*, 23, 571–578.

Lee, Eun S., Ronald N. Forthofer, and Ronald J. Lorimer. (1989) *Analyzing Complex Survey Data*. Newbury Park, CA: Sage.

McAdams, Mara A., Rob M. Van Dam, and Frank B. Hu. (2007) "Comparison of Self-Reported and Measured BMI as Correlates of Disease Markers in U.S. Adults," 15, 188–196.

Nord, Mark and Ephraim Leibtag. (2005) "Is the "Cost of enough Food" Lower in Rural Areas?" *Review of Regional Studies*, 35, 291–310.

Normand, Sylvie, Yadh Khalfallah, Corinne Louche-Pelissier, Christiane Pachiaudi, Jean-Michel Antoine, Stéphane Blanc, Michel Desage, Jean P. Riou, and Martine Laville. (2001) "Influence of Dietary Fat on Postprandial Glucose Metabolism (Exogenous and Endogenous) using Intrinsically (13) C-Enriched Durum Wheat," *British Journal of Nutrition*, 86, 3–11.

Olson, Christine M., Kendra Anderson, Elizabeth Kiss, Frances C. Lawrence, and Sharon B. Selling. (2004) "Factors Protecting Against and Contributing to Food Insecurity among Rural Families," *Family Economics & Nutrition Review*, 16, 12–20.

Samaha, Frederick F. Nayyar Iqbal, Prakash Seshadri, Kathryn L. Chicano, Denise A. Daily, Joyce McGrory, Terrence Williams, Monica Williams, Edward J. Gracely, and Linda Stern. (2003) "A Low-Carbohydrate as Compared with a Low-Fat Diet in Severe Obesity," *New England Journal of Medicine*, 348, 2074–2081.

Schnittker, Jason. (2004) "Education and the Changing Shape of the Income Gradient in Health," *Journal of Health and Social Behavior,* 45, 286–305.

Sobal, Jeffery, Richard P. Troiano, and Edward A. Frongillo Jr. (1996) "Rural-Urban Differences in Obesity," *Rural Sociology*, 61, 289–305.

US Dept of Agriculture, Agricultural Research Service, ARS Food Surveys Research Group. Data and Documentation for the 1994-96 Continuing Survey of Food Intakes by Individuals (CSFII)-Diet and Health Knowledge Survey (DHKS) [data on CD-ROM]. US Department of Agriculture: Riverdale, MD:; 1998. National Technical Information Service Accession No. PB98-500457.

Variyam, Jayachandran N., James Blaylock, and David Smallwood. (1998) "Information Effects of Nutrient Intake Determinants on Cholesterol Consumption," *Journal of Agricultural and Resource Economics*, 23, 110–125.

Welsh, Susan O., Carole Davis, and Anne Shaw. (1993) *USDA's Food Guide, Background and Development, Miscellaneous Publication No. 1514*. Nutrition Education Division, Human Nutrition Information Service: Hyattsville, MD. September.

APPENDIX A

Q 1: Based on your knowledge, which has more saturated fat?

a. liver/t-bone
b. butter/margarine
c. egg white yolk
d. skim/whole milk

Q 2: Which has more fat?

a. Hamburger/ground round
b. Pork chops/spare ribs
c. Hot dogs/ham
d. Peanuts/popcorn
e. Yogurt/sour cream
f. Porterhouse/round

Q 3: Which kind of fat is more likely to be a liquid rather than a solid…

- saturated fats
- polyunsaturated fats
- are they equally likely to be liquids?
- Don't know

Q 4: If a food has no cholesterol is it also…

- low in saturated fat.
- high in saturated fat. or
- could it be either high or low in saturated fat?
- Don't know

Q 5: Is cholesterol found in...

- vegetables and vegetable oils.
- animal products like meat and dairy products.
- all foods containing fat or oil?
- Don't know

Q 6: If a product is labeled as containing only vegetable oil is it …

- low in saturated fat.
- high in saturated fat. or
- could it be either high or low in saturated fat?
- Don't know