RELATIONSHIP BETWEEN BODY WEIGHTS AND PROXIMITY TO RESTAURANTS IN THE UNITED STATES

by

JYOTSNA N. GHIMIRE

(Under the Direction of Cheolwoo Park)

ABSTRACT

The objective of this study is to analyze the relationship between availability of restaurants in the neighborhood and the prevalence of obesity in the United States (U.S.). The study sample consists of 72,691 individuals participating in the United States Department of Agriculture (USDA) National Household Food Acquisition and Purchase Survey (FoodAPS), a nationally representative cross-sectional survey of American households. Findings of this study indicate a positive and significant relationship between availability of both fast-food and non-fast-food restaurants in the neighborhood and the prevalence of body mass index (BMI). We also find that the probability of being obese is significantly higher for individuals who live within 0.25-mile radius of the restaurants. These findings can be used by local, state, and federal government to inform their health policy decisions.

INDEX WORDS: BMI, Obesity, Food Away from Home (FAFH), United States Department of Agriculture National Household Food Acquisition and Purchase Survey (USDA FoodAPS)

RELATIONSHIP BETWEEN BODY WEIGHTS AND PROXIMITY TO RESTAURANTS IN THE UNITED STATES

by

JYOTSNA N. GHIMIRE

M.S., University of Georgia, 2019

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2022

RELATIONSHIP BETWEEN BODY WEIGHTS AND PROXIMITY TO RESTAURANTS IN THE UNITED STATES

by

JYOTSNA N. GHIMIRE

Major Professor: Cheolwoo Park
Committee: Liang Liu
Cesar Escalante

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2022

DEDICATION

I would like to dedicate this thesis to my husband, and the happiness of my life "Ena and Aditi".

ACKNOWLEDGEMENTS

I want to express my sincere gratitude to my committee members Drs. Cheolwoo Park, Liang Liu, and Cesar Escalante for their continued support and invaluable guidance. I am particularly indebted to my committee chair, Dr. Cheolwoo Park, for sharing his extensive knowledge and insights during this research work. In addition, Dr. Park has always provided constructive comments, unrelenting support, and thorough research advice throughout the study.

I would like to thank the University of Georgia Department of Statistics for the opportunity to complete an MS degree in Statistics. Special thanks go to graduate coordinator Dr. Liang Liu for cheering and supporting me. I also would like to thank Dr. Cesar Escalante at the Department of Agriculture and Applied Economics. I am humbled by his consistent support and encouragement, which helped me grow academically.

I want to express my deepest gratitude to my parents and my family for their constant encouragement, love, and unyielding support over the years. I am very grateful to my husband, Ramesh Ghimire, for his continuous support and perseverance. None of this would have been possible without him being on my side. He is and will always be my inspiration. And lastly, I am thankful to my beautiful daughters, Ena and Aditi, whose smiles and positive attitudes constantly refresh and motivate me.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND AND STATEMENT OF THE PROBLEM

Consumption of restaurant food is an integral part of the American diet, and the number of people to rely on such convenience food is multiplying day by day. However, frequently eating restaurant foods could increase body mass index (BMI) and can make people overweight and obese because food prepared outside the home often contains excessive calories, fat, and sodium than the food prepared at home (Anderson & Matsa, 2011; Currie et al., 2010). This excessive calorie intake, coupled with a sedentary lifestyle, in general, increases the BMI of individuals or the risk of being overweight and obese. This situation may get worse if people have underlying or preexisting health conditions. Among various factors influencing behavioral aspects of calorie intake, the growth in the number of restaurant establishments corresponds with the increased consumption of such food (Oexle et al., 2015).

Easy access to restaurant foods, particularly fast foods near home, encourages people to take excessive calories, which can play a crucial role in dietary intake and obesity rates (Lee et al., 2015; Gustafson et al., 2016). It is argued that the availability of restaurants in close proximity to residential areas could encourage people to consume restaurant foods more often than preparing a meal at home. Accordingly, adults who live close to restaurants are expected to consume restaurant foods more frequently than their counterparts, increasing their body weights. This study tests this premise using USDA's Household Food Acquisition and Purchase Survey (FoodAps) data collected in 2016 (USDA, 2016). We examine whether or not individuals' body mass index (BMI) and obesity are associated with the prevalence of restaurants from their residences in the United States (U.S.). For this study, the prevalence of restaurants is defined as the number of restaurants within 0.25 miles radius of individuals' homes.

A study investigating the relationship between proximity to restaurants and BMI is policy-relevant for several reasons. The Centre for Disease Control and Prevention (CDC, 2019) reports that nearly 40 percent of the Americans who were age 20 or over were obese in 2016, and the prevalence of obesity is in increasing trend each year. Although overweight and obesity are a serious public health concern in the U.S., limited studies have investigated the socioeconomic and behavioral factors associated with overweight and obesity. Studies suggest that food prepared away from home (FAFH), particularly fast food, has an inferior nutritional quality compared to food prepared at home. Frequent consumption of such foods can be positively associated with higher fat and sodium consumption, leading to sustained positive energy balance and, thereby, a higher rate of obesity (Anderson & Matsa, 2011; USDA ERS, 2019; Lee et al., 2015).

Furthermore, location and distance, along with the restaurant density of both fast food and non-fast food, can influence people consuming such foods. All this occurs in a social, cultural, and economic environment that can aggravate the population's health unless active measures are taken to make the environment a health-promoting one (WHO, 2003). One such policy measure that could help improve diet quality and lower BMI is to discourage the opening of new restaurants, particularly fast-food restaurants, close to residential areas. Likewise, mandatory calorie posting by restaurants could be another remedy to offset the ever-growing obesity and BMI.

1.2 ORGANIZATION OF THE STUDY

Following this Introduction chapter, Chapter 2 reviews and summarizes existing literature related to this study, providing background information on the relationship between FAFH and body weights. This chapter also explores the relevant literature on the factors associated with changes in eating habits, the growth of fast food and non-fast-food restaurants in the U.S., and the relationship between eating habits and the growth of restaurants in the neighborhood. Chapter 3 introduces the conceptual framework, discusses the statistical/econometric model, explains the data source and the variables used in the study, and the data analysis process. Chapter 4 reports the study findings. Chapter 5 discusses the findings and summarizes the results, and Chapter 6 concludes this study.

CHAPTER 2

LITERATURE REVIEW

There has been a remarkable shift in eating culture in the United States over the past 40 years. The consumption of foods prepared outside the home has steadily increased from 16 percent to almost 33 percent of an individual's daily nutritional intake between 1970 and 2015 (Lin and Guthrie, 2012; Seguin et al., 2016). Studies suggest that there has been a notable decline in the amount of time devoted to food preparation at home and eating away from home has become a routine part of American eating culture (Warde et al., 2007). The decrease in time spent on meal preparation and cooking at home and the increase in the availability of high-calorie energy-dense food options in the neighborhood are argued to be historically associated with obesity prevalence (Lee et al., 2015; Seguin et al., 2016). According to the United States Department of Agriculture Economic Research Station (USDA ERS, 2019), the average American consumes one-third of their daily calorie intake from the food prepared outside the home.

Among various factors influencing the shift in consumer preferences toward eating away from home, the increased availability of FAFH options and the opportunity cost of time associated with food preparation at home are the most significant factors contributing to the shift in eating away from home (Smith et al., 2013; Saksena et al., 2018). Studies suggest that from 2000 to 2015, the number of quick-service restaurants operating in the United States increased by 20 percent, while the number of full-service restaurants remained unchanged (USDA ERS, 2020). In absolute numbers, this accounts for 340,000 new quick-service restaurant establishments, and restaurant spending accounts for more than half of total expenditures on food in the United States from 2000 to 2015 (USDA ERS, 2020). The increased spending, paired with increased geographic access to restaurant and supermarket locations, may nudge consumers to consume more calorie intense foods, contributing to a higher level of BMI and obesity, especially in individuals who primarily rely on such food (Inagami et al., 2009; Rahkovsky et al. 2018). Here, the underlying assumption is that the close proximity to restaurants lowers the consumption cost (e.g., the opportunity cost of time), resulting in the frequent use of such foods, leading to substantial body weight gain.

While eating away from home is increasingly popular, meals prepared away from home have different nutritional qualities than meals prepared at home. A research study by the U.S. Economic Research Service (ERS) found that food prepared away from home tends to have lower nutritional quality and is dense in calories than food prepared at home. An increase in the consumption of such food coupled with a sedentary lifestyle can significantly contribute to body weight gain and increased BMI. The situation worsens when greater fast-food access is positively associated with higher consumption of such foods at the community level. In an empirical study, Rosenheck (2008) found a link between the increased supply of fast-food restaurants and a substantial increase in obesity in the U.S. More recently, in an attempt to identify the effect of an increase in the supply of fast-food restaurants on obesity rates, Currie et al. (2010) reported a significant association between proximity to fast-food restaurants and the risk of being obese. Further, they reported that the prevalence of fast-food restaurants within a 0.1-mile radius of the school area results in a 5.2 percent increase in the obesity rate among ninth graders and the prevalence of such a restaurant within 0.5 miles of residence increases the probability of gaining 20 kilos weight by 1.6 percent among pregnant women. Moreover, they found that the effect of having a restaurant within a 0.1-mile distance is equivalent to an increase in daily calorie consumption by 100 calories due to the proximity of fast food.

Dunn (2010), using the 2004 to 2006 Behavioral Risk Factor Surveillance System (BRFSS) data and self-collected data on the number of fast-food restaurants, investigated the effect of fast-food availability on weight outcomes by geographic location in the U.S. The author found that the availability of such restaurants does not affect the weight outcomes in rural counties but does tend to increase the body mass index among females and non-whites in medium-density counties. Further, the author suggests that since the effect of access to fast food is concentrated only among females and minority groups in medium-density areas, intervention programs should be focused solely on these groups. This, in turn, could help reduce the cost of intervention while discouraging fast food consumption and reducing overall obesity levels.

Using Dietary Recall data between 1994 and 2004 from the National Health and Nutrition Examination Survey (NHANES) and using the first-difference estimator approach to analyze the causal relationship between FAFH and dietary intakes, Macino, Todd, & Lin (2010) found that FAFH does contribute to increase the caloric intake and reduce diet quality. Additionally, their study found that each meal away from home is estimated to add 130 more calories to the total daily calorie intake and lower the healthy eating index by two points. A similar study conducted by Todd, Masino, and Lin (2010) analyzed the association between FAFH and poor diet quality. They used 1994-96 and 2003-04 NHANES data and applied the fixed-effect estimation approach to account for unobservable influences. They found that FAFH increases daily caloric intake and reduces diet quality for an average adult. Further, their study suggests that the effect varies depending on meal type consumed away from home, with dinner having a higher impact.

A similar study by Liu et al. (2013) found that policy interventions that are designed to address rising obesity levels and encourage consumers to decrease their calorie intake have little to modest impact in achieving goals such as reducing calorie intake. Further, consumers' education level, their understanding of the information provided, and motivation and self-control are the major factors encouraging consumers to lower their calorie intake. Consumers' sociodemographic characteristics were also associated with paying attention to calorie labels posted or provided by restaurants. A Gallup study conducted in 2013 reported that 68 percent of people surveyed pay attention (great or fair attention) to calorie labels at restaurants (43 percent of people pay a "great deal" or a "fair amount" of attention to calorie labels at restaurants). The survey also found that sociodemographic characteristics matter – women, college graduates, and people making $75,000 or more are far more likely to pay attention to nutrition labels provided by restaurants compared to other sociodemographic groups (Brown, 2013).

Although previous studies have made significant contributions to the literature by investigating the potential association between the prevalence of restaurants and BMI, their outcomes have yielded different results. For example, Dunn (2010) and Anderson & Matsa (2011) found no causal link between restaurant food consumption and obesity at various levels of distance to the nearest restaurants. However, a study by Currie et al. (2010) and Rashad et al. (2005) found a strong association between the increased number of per capita restaurants and consumption of such restaurant food on increased BMI and obesity. Hence, this study uses the USDA's nationally representative FoodAPS data and investigates the potential association between the proximity of restaurant outlets and body mass index (BMI) in the United States.

CHAPTER 3

METHOD

3.1 DATA

This study utilizes the United States Department of Agriculture (USDA) National Household Food Acquisition and Purchase Survey (FoodAPS) data collected in 2016. The USDA FoodAPS is a nationally representative survey of American households that collects comprehensive data about household food purchases and acquisitions (USDA FoodAPS, 2018). Detailed information was collected about food purchased for consumption at home and away from home. The survey also contains a wide range of information about households and individuals, including their socioeconomic and demographic characteristics such as age, gender, education, race/ethnicity, employment status, and participation in federal food assistance programs.

The consolidated data used in this study include household-level public use file, individual-level public use file, Food at Home (FAH) events public use file, FAH item data, food Away from Home (FAFH) event data, FAFH item data, FAH nutrient information, FAFH nutrient information, meals and snacks information, and access data.

3.2 RESPONSE VARIABLES

Our response variable, body mass index (BMI), is defined as weight in pounds divided by height in inches squared. The FoodAPS individual-level survey collects BMI information for individuals who are two years or older. However, since the scope of our study is limited to adults 18 years and older, we exclude observations corresponding to individuals who are less than 18 years.

We conduct robustness checks by analyzing an alternative model where the response variable (obese – a bodyweight category) is a binary measure. The variable obese equals one if the individual's BMI is greater than or equal to 30, and zero otherwise. These measures are standard in the obesity literature (Solomon & Manson,1997; Horwich et al., 2001; Calle et al., 2005).

The visualization of the response variables (both BMI and bodyweight categories) is summarized in the Figure 1. The left panel of Figure 1 suggests that 75 percent of the respondents in our sample have BMI between 19.01 and 28.29. It also indicates that the mean (24.37) and the median (23.67) BMI are close, suggesting that our sample has an approximately normal distribution, and no transformation is necessary for statistical analysis. The figure on the right panel presents the categories of body weight distribution by the average number of restaurants in the neighborhood.

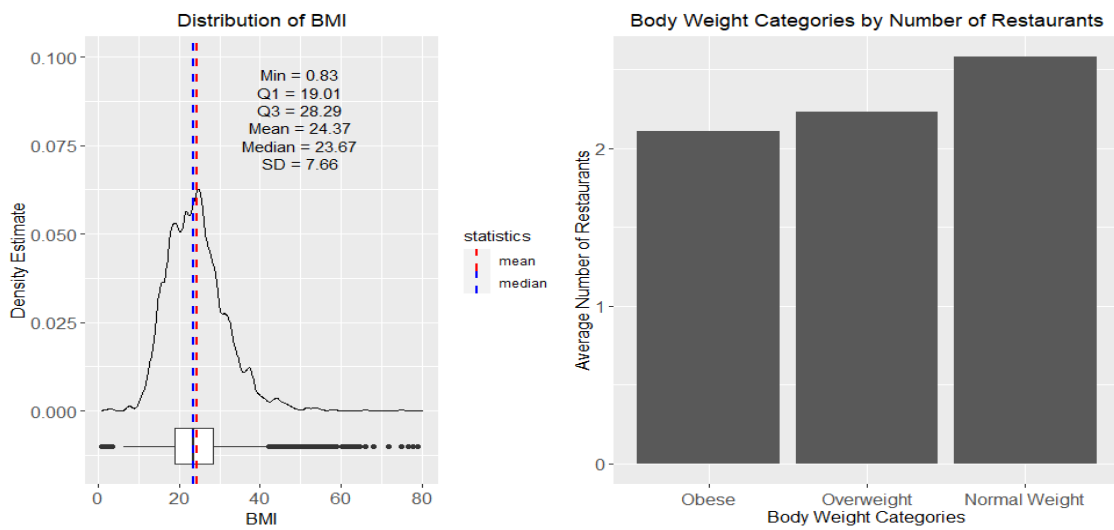

Figure 1: Distribution of response variables

3.3 PREDICTOR VARIABLES

This analysis used two types of predictor variables: numerical and categorical. The numerical variables are age, the total number of restaurants within 0.25 mi distance, frequency of eating food Away from Home (FAFH), and family income. The number of restaurants within a quarter-mile distance from the respondent's residence and FAFH are the variables of interest in this study. Table 1 summarizes the variables (both response and predictor) used in this study.

Table 1: Descriptions of the variables

VARIABLES	DESCRIPTION
BMI	A numeric variable measuring respondent's calculated Body Mass Index
No. of restaurants	A numeric variable that counts the number of fast-food and non-fast-food restaurants within 0.25 of residence
Sex	A categorical variable equals '1' if the respondent is male and '2' if the respondent is female.
Frequency of FAFH	A numeric variable that counts the number of times per week an individual eats dinner out.
Age group	A categorical variable equals '0' if the respondent's age is between 18-34 years, '1' if the respondent's age is between 35-64 years, and '2' if the respondent's age is above 65 years.
Race	A categorical variable equals '0' if the respondent is White, '1' if the respondent is Black, '2' if the respondent is Asian, and '3' if the respondent belongs to another race.
Education	A categorical variable equals '0' if the respondent had less than or equal to a high school degree, '1' if the respondent had some college degree, '2' if the respondent had a bachelor's degree, and '3' if the respondent had master's degree and above.
Marital status	A categorical variable equals '0' if the respondent was married, '1' if the respondent was never married, and '2' if the respondent was widowed, divorced, or separated.
Employment status	A binary variable equals '1' if the respondent had a job in the past week and '0' otherwise.
Hispanic	A categorical variable equals '1' if the respondent is of Hispanic ethnicity and '0' otherwise.
Reported Health status	A categorical variable equals '0' if the respondent reported health status is excellent or very good, '1' if the respondent reported health status is good, and '2' if the respondent reported health is fair or poor.
Family Income	A numeric variable that measures a family's average monthly income as a sum of income per member.

Obese	A binary variable equals '1' if the respondent's BMI is ≥ 30 and '0' otherwise.
Rural	A categorical variable equals '1' if the household is in the rural census tract and '0' otherwise.

FoodAPS, access data codebook, defines the number of restaurants within the quarter-mile distance as the count of FAFH outlets. Such outlet includes all full-service restaurants, limited-service restaurants, drinking places, hotels and motels, retail stores and vending machines, recreational places, schools and colleges, other FAFH sales, and food furnished and donated (USDA ERS, 2020). Likewise, FoodAPS individual public use file defines the frequency of FAFH as the usual times per week an individual eats dinner out.

We calculate the Pearson Correlation between all pairs of four numerical predictors and their relationship with BMI. The result is presented in Figure 3.

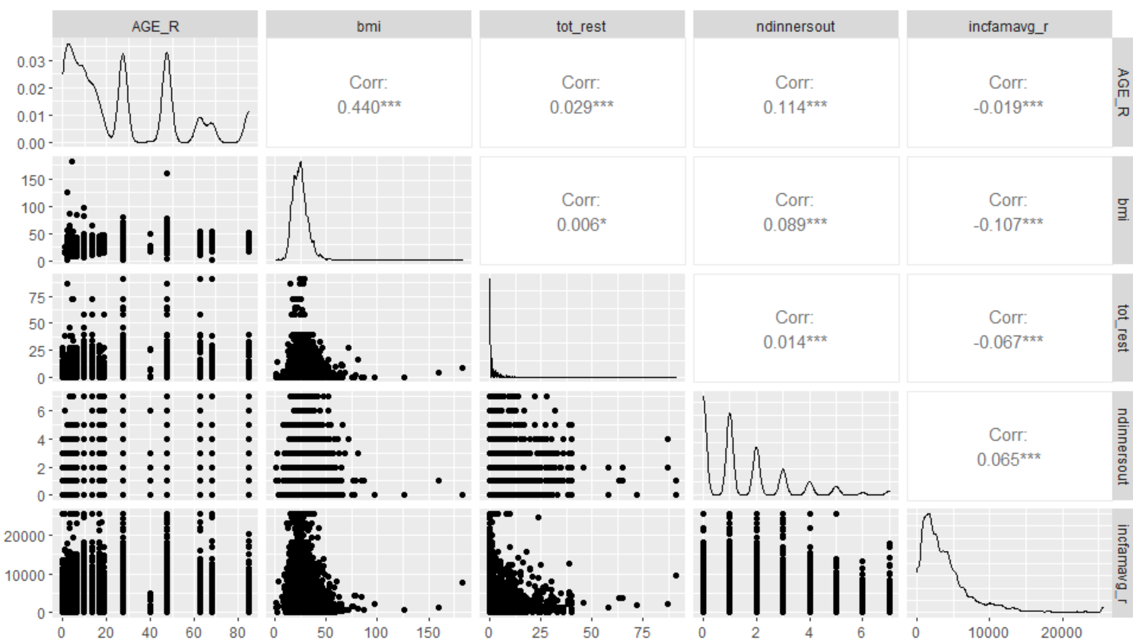

Figure 2. Pairwise correlations among numerical variables

We also include a number of sociodemographic characteristics of the respondents as control variables collected at the individual and household levels. Information on the sociodemographic characteristic comprises age, gender, marital status, education, homeownership, employment status, children in the household, household income, and race/ethnicity. The majority of these control variables are categorical. We calculated the frequencies of each level of these categorical variables in Figure 3.

Figure 3. Distribution of categorical predictors

The result summarized in the Figure 3 suggests that each factor level utilized in this study has enough observations to ensure the validity of the statistical model. The top panel in Figure 3 indicates nearly an equal number of males and females in the sample, but working-age people have a higher representation. Most of the respondents are non-Hispanic married white with high school or less education. There are nearly equal employed and unemployed survey respondents in the sample, with most reporting excellent health conditions and living in the urban census tract.

The box plots summarized in Figure 4 visualize the relationship between categorical predictor variables and the response variable, BMI. For instance, the average BMI of the respondents from both Urban and Rural census tract regions is approximately similar. The BMI of respondents whose perceived health status if fair or poor is significantly higher than those whose perceived health is in excellent condition. The millennial's BMI is slightly lower than adults in the working-age group, and Asians have comparatively lower BMI than other racial groups.

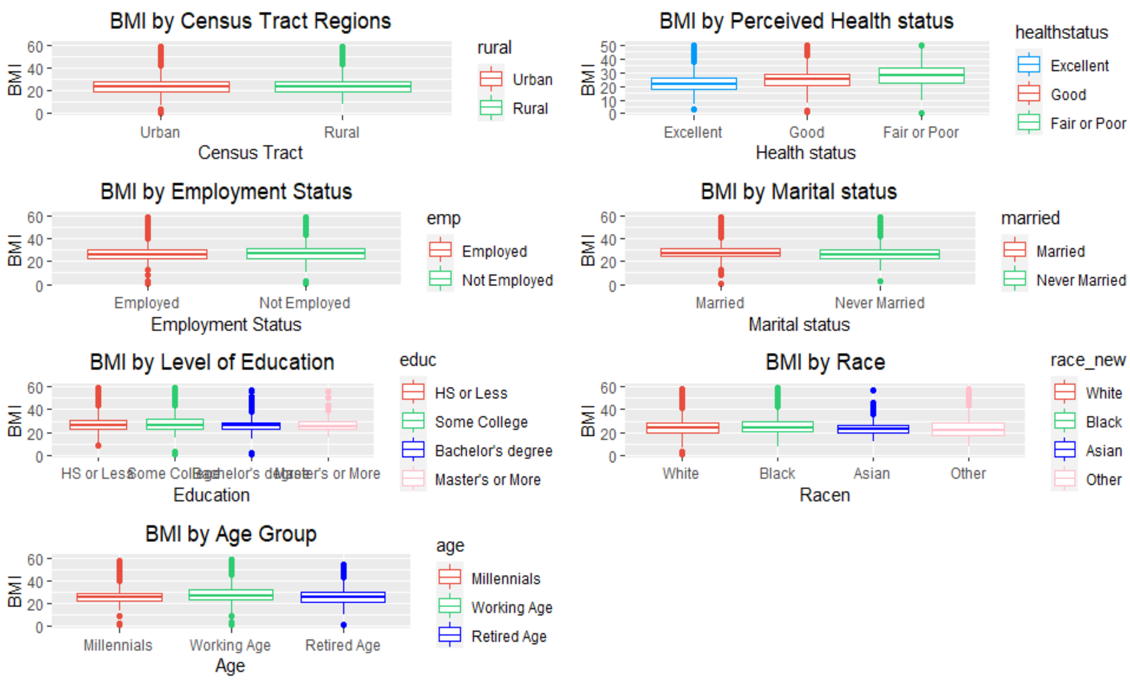

Figure 4. Relations between categorical predictors and BMI

3.4 STATISTICAL/ECONOMETRIC MODEL

We model that the BMI of an individual (*BMI*) depends on the number of restaurants within 0.25-mile proximity of his/her residence (*R*) while controlling for his/her socioeconomic (*S*) and demographic characteristics (*D*), and the frequency of eating out (*F*), as summarized in equation (1):

$$BMI = f(R, S, D, F). \qquad (1)$$

Since our response variable (*BMI*) is continuous, an ordinary least square (OLS) regression is used to estimate the association between BMI and the number of restaurants within

0.25 miles of proximity to individuals' residences. The OLS regression model estimates equation (1) by minimizing the sum of squares in the difference between the observed and predicted values of the response variable. The OLS formulation may be outlined in equation (2):

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \qquad (2)$$

where *Y* is the response variable, and $\boldsymbol{X}$ is a vector of the predictor variables (number of restaurants, frequency of FAFH, race, age, education, income, employment, marital status, and health status), $\boldsymbol{\beta}$ is the vector of the parameters to be estimated, and $\boldsymbol{\varepsilon}$ is an unobserved error term assumed to be approximately normally distributed. The OLS estimator assumes that all coefficients in the model are unknown and are to be estimated from the given data using the algebraic operation, summarized in equation (3).

$$\widehat{\boldsymbol{\beta}} = \quad (X'X)^{-1}X'Y \qquad (3)$$

A binary measure – obese, which is associated with BMI is also constructed. We also model whether the binary variable obese is associated with an increased number of restaurants within 0.25 miles of their residence. Since this response variable is binary, we use a logistic regression model to estimate equation (1). The logistic regression model may be summarized in equation (4):

$$ln\left(\frac{p(y=1)}{1-p(y=1)}\right) = ln\big(Odds(y=1)\big) = \alpha \ + \boldsymbol{\beta X} \qquad (4)$$

In this equation, *y* is the response variable indicating the obesity status of an individual in the sample. That is, *y* =1 if BMI ≥ 30 or 0 otherwise. X is the vector of explanatory variables, and β is the vector of parameters to be estimated.

CHAPTER 4

FINDINGS

4.1 SUMMARY STATISTICS

The FoodAPS dataset used in this study consist of 155,911 observations collected in 2016. However, after excluding irrelevant covariates, missing values and outliers, a total of 72,691 observations are used for this analysis. Table 2 presents the summary statistics of the variables used in the study.

Table 2: Summary statistics (n = 72,691)

Variable	Mean	Std. Dev.	Minimum	Maximum
Body Mass Index (BMI)	27.71	7.09	0.83	160.15
Obsess (BMI ≥ 30)	0.28	0.45	0	1
Number of restaurants (in 0.25 miles)	2.67	6.24	0	91
Female	0.44	0.49	0	1
Average HH income (Monthly)	4139	3950	1	25520
Frequency of FAFH per week	1.79	1.77	0	7
Millennials (17 – 34)	0.38	0.34	0	1
Working Age (35 – 65)	0.43	0.49	0	1
Retired Age (≥ 65)	0.19	0.39	0	1
White	0.75	0.33	0	1
Black	0.11	0.31	0	1
Asian	0.05	0.21	0	1
Other race	0.09	0.29	0	1
High school or less	0.47	0.35	0	1
Some college or associated degree	0.29	0.46	0	1
Bachelor’s degree	0.14	0.35	0	1
Master’s degree and above	0.09	0.29	0	1
Married	0.36	0.48	0	1
Employed	0.58	0.49	0	1
Hispanic	0.15	0.36	0	1
Perceived health (excellent or very good)	0.49	0.49	0	1

Perceived health (good)	0.38	0.49	0	1
Perceived health (poor)	0.21	0.41	0	1

The body mass index (BMI) for average respondents is nearly 28, twenty-eight percent of them are obese, and there are, on average, three restaurants (all restaurants -- fast food and non-fast food) within a quarter-mile radius of individuals' residence. The income variable reported in the table indicates that the monthly household income is $4,139, and participants eat dinner away from home an average of two times per week. In the sample, 38 percent of the participants are millennials (ages between 18 to 35), 43 percent are working age (ages between 35 to 65), and 19 percent are retired (age 65 and older). Regarding race/ethnicity, 75 percent of the survey respondents were white, 11 percent were black, 5 percent were Asian, 15 percent were Hispanic, and 9 percent belonged to other races. Forty-seven percent of the respondent have a high school or less education, 30 percent have some college or an associate degree, 14 percent have a bachelor's degree, and 9 percent have a master's degree or higher.

Regarding the demographic structure, 44 percent of the survey participants were female, and 36 percent were married. Of the participants, on average, 58 percent were employed. Regarding the general health condition of the survey participants, 41 percent perceived their health as in excellent or very good condition, 38 percent perceived in good condition, and 21 percent perceived in poor condition.

4.2 RELATIONSHIP BETWEEN PROXIMITY TO RESTAURANTS AND BMI: OLS REGRESSION RESULTS

Since the objective of this study is to investigate if proximity to restaurants and frequency of eating food away from home (FAFH) affect an individual's overall body mass index (BMI), we analyze the determinants of BMI using ordinary least square (OLS) model, and the findings are reported in Table 3. The regression equation is statistically significant, as indicated by the F statistic and corresponding p-values, to explain the variation in BMI. As indicated by R^2, 11 percent of the variation in BMI is explained by the regression equation. The decision criteria for hypothesis testing are based on a 5 percent significance level (Greene, 2012).

Table 3: OLS regression results on BMI and the number of restaurants in 0.25 mi

					95% CI	
Body Mass Index (BMI)	Coefficients	Std. Error	t-value	P>|t|	Lower	Upper
No. of restaurants (in 0.25 miles)	0.022	0.008	2.85	0.00	0.007	0.037
No. of restaurants (in 0.25 miles)^2	-0.001	0.000	-2.98	0.00	-0.001	0.000
Female	-0.059	0.052	-1.12	0.26	-0.161	0.044
Average HH income (monthly)	-0.0001	0.000	-18.75	0.00	-0.00015	-0.00012
Frequency of FAFH per week	-0.010	0.015	-0.66	0.51	-0.039	0.019
Working age (35 – 65)	1.891	0.062	30.64	0.00	1.770	2.012
Retired age (≥ 65)	1.116	0.085	13.09	0.00	0.949	1.283
Black	0.048	0.084	0.58	0.56	-0.116	0.213
Asian	-2.996	0.131	-22.83	0.00	-3.253	-2.739
Other Race	-0.217	0.102	-2.13	0.03	-0.417	-0.017
Some college or associated degree	0.328	0.060	5.46	0.00	0.210	0.446
Bachelor’s degree	-0.756	0.081	-9.32	0.00	-0.915	-0.597
Master’s degree and above	-0.384	0.101	-3.81	0.00	-0.581	-0.186
Married	0.658	0.059	11.17	0.00	0.543	0.774
Employed	-0.235	0.060	-3.93	0.00	-0.353	-0.118
Hispanic	-0.019	0.085	-0.23	0.82	-0.186	0.147
Perceived health (good)	1.805	0.059	30.65	0.00	1.689	1.920
Perceived health (poor)	4.116	0.072	57.22	0.00	3.975	4.257
Rural	0.041	0.061	0.66	0.51	-0.080	0.161
Constant	25.763	0.095	270.32	0.00	25.577	25.950

(R^2 = 0.1066, F-statistics = 443.94, Observations = 70,690)

The variable number of restaurants within 0.25 square miles, working-age dummy, retired-age dummy, some college or associated dummy, married dummy, perceived health (good) and perceived health (poor) were all positive and significant in explaining the variation in body mass index (BMI) at five percent or better significance level. But the average household income, race dummies such as Asian and others are negative and significant at a five percent or better significance level. Similarly, education dummy, which includes a bachelor's degree and master's degree and above, is also negative and significant. Likewise, the employment dummy variable is negative and significant in explaining the variation in the body mass index.

Our key predictor variable, our variable of interest - the number of restaurants (in 0.25 miles), is positive and significant in explaining the variation in BMI at a five percent or better significance level. However, its squared term is negative and significant at a one percent significance level, indicating that an increase in the number of restaurants within 0.25 miles is associated with an increase in BMI but at a decreasing rate. A $1,000 increase in monthly household income is associated with a 0.0001 unit decrease in BMI. Compared to millennials, working-age respondents who are defined to be in the age group 35 to 65 years have BMI two units higher. On the other hand, Asians are likely to have a 300 percent lower BMI, while other racial groups are likely to have a 21 percent lower BMI than the White respondents.

Regarding education, respondents with some college or associate degrees had a 33 percent higher BMI than respondents with high school or less education. But respondents with bachelor's degrees had 76 percent, and those with a master's degree or higher had 38 percent lower BMI than those who had high school or less education. Married respondents had 66 percent higher BMI than their unmarried counterparts, and employed respondents had a 24 percent lower BMI than unemployed respondents.

The FoodAPS survey also collects information about individuals' ratings of their general health status. The collected data were recorded into seven categories: excellent, very good, good, fair, poor, refused, and do not know. The observations in which respondents either did not know or refused to answer were excluded from the analysis, and the remaining five categories were merged into three. Excellent and very good health rating was merged into the excellent perceived health category. The rating good was kept as is, and fair and poor were combined into the poor perceived health category. Regarding their marginal effects, respondents with good perceived health conditions had 180 percent higher, and respondents with poor perceived health conditions had 412 percent higher BMI than their counterparts whose perceived health condition was excellent.

4.3 RESIDUAL ANALYSIS

This section assesses the fit of the linear regression model estimated in the previous section. Residuals are typically used to evaluate the fit and the validity of the statistical models and can also be used as a tool for model selection (Nobre & Singer, 2007). The basic assumption for the Ordinary Least Square (OLS) regression model is that the residuals follow a normal distribution, have a constant variance, independent, no outlier is present, and have a zero conditional mean (Greene, 2012). The expectation is that an adequate model should satisfy the above-stated assumptions and contain no obvious pattern in the residuals. If any assumptions are violated, it may lead to an unstable model. And the standard summary statistics such as the t or F test are not enough to ensure the models' accuracy. Further, mathematical literature indicates that violations of one or more linear regression assumptions might fail to represent the true nature of the situation and lead to serious problems with parameter estimations, mainly if the violations appear in combination (Wooldridge, 2015). Hence, to ensure the fit and validity of our model assumptions, we perform the residual analysis of our fitted model, and the result is presented in

Figure 5. The residuals will be randomly scattered around the center with no obvious pattern if the assumptions are met.

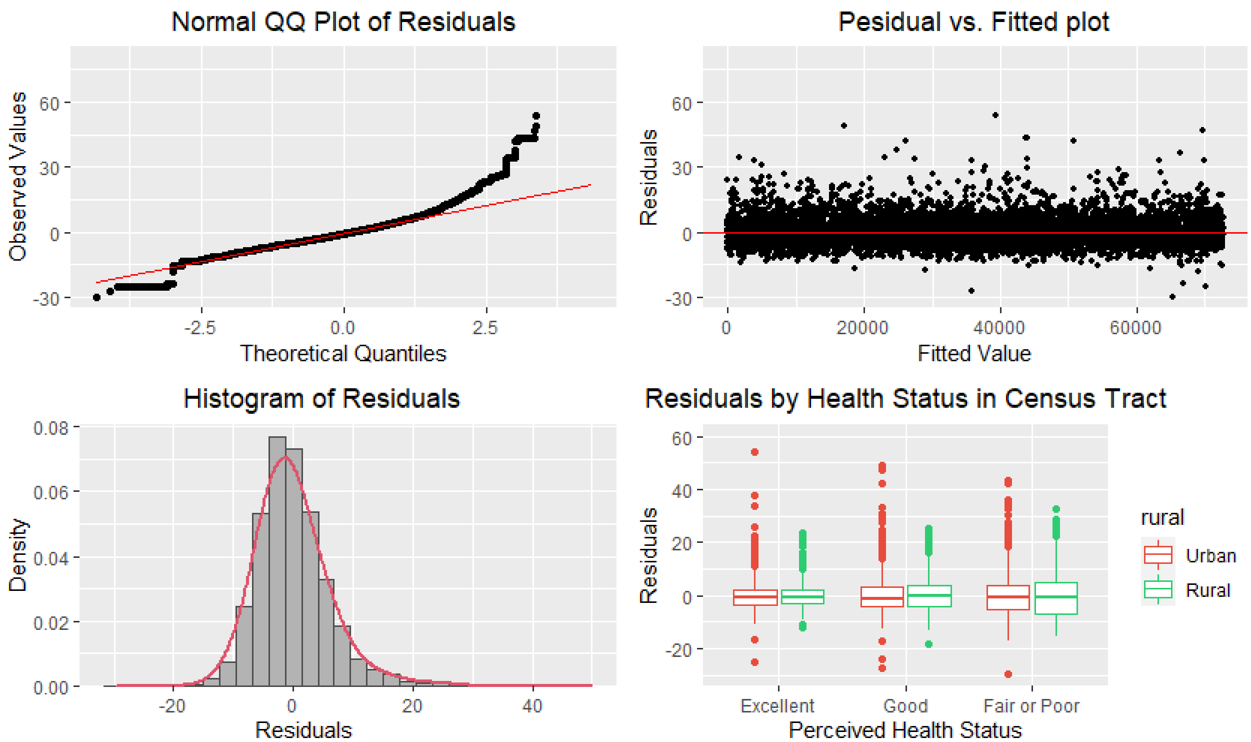

Figure 5. Residual plot of the fitted OLS model

The Q-Q plot, the residuals vs. fitted plot, and the histogram of the residual are the tools used to assess if the model satisfied the OLS assumptions. The diagnostic plots presented in Figure 1 show the residuals in four different ways. The Q-Q plot shows an approximately straight-line pattern in the middle part but heavy tails suggesting non-normality. However, the histogram of residuals is approximately symmetrical, we can consider the normality assumption not to be severely violated. The residual vs. fitted plot indicates that residuals are random and do not exhibit any pattern suggesting the residuals are independent of one another or mutually uncorrelated. Further, the residual vs. fitted plot shows a band of relatively constant width along the horizontal band, above and below the value 0, suggesting that the assumption of constant variance is satisfied. The boxplot of residuals by perceived health status means a conditional zero-mean assumption is held for different census tract respondents. Hence, we can conclude that the overall assumptions of linear regression have been satisfied.

4.4 RELATIONSHIP BETWEEN PROXIMITY TO RESTAURANTS AND THE PROBABILITY OF BEING OBESE: LOGISTIC REGRESSION RESULTS

Further, the FoodAPS survey collects individuals' weight category information based on their BMI and records it into three sub-categories. Adults aged 18 and older who have a BMI of 24.9 or less are not overweight, a BMI between 25.0 to 29.9 is defined as overweight, and a BMI higher than 30.0 is defined as obese. To make our analysis less complex, we regroup these three categories into two: obese (BMI ≥30) and not obese (BMI <30). Since BMI and obesity are closely related, we decided to estimate an alternative regression model where the response variable is obesity. All the predictor variables were the same as the baseline model. As the response variable is coded as a binary choice, logistic regression is used to model their response behavior (Wooldridge, 2010).

Table 4: Relations between proximity to restaurants and obesity: A logistic regression results

					95% CI	
Obsess	Odds Ratio	Std. Err.	z	P>|z|	Lower	Upper
No. of restaurants (in 0.25 miles)	1.007	0.003	2.04	0.041	1.0003	1.0130
No. of restaurants (in 0.25 miles)^2	1.000	0.000	-2.29	0.022	0.9996	1.0000
Female	1.190	0.021	9.67	0.000	1.1491	1.2333
Average HH income (Monthly)	1.000	0.000	-15.18	0.000	1.0000	1.0000
Frequency of FAFH per week	1.007	0.005	1.42	0.155	0.9972	1.0175
Working age (35 – 65)	1.784	0.039	26.58	0.000	1.7099	1.8623
Retired age (≥ 65)	1.329	0.040	9.55	0.000	1.2539	1.4093
Black	1.059	0.030	2.04	0.041	1.0023	1.1197
Asian	0.286	0.019	-19.28	0.000	0.2520	0.3250
Other race	1.240	0.042	6.32	0.000	1.1599	1.3256
Some college or associated degree	1.102	0.022	4.79	0.000	1.0590	1.1466
Bachelor’s degree	0.697	0.021	-11.78	0.000	0.6562	0.7400
Master’s degree and above	1.135	0.040	3.56	0.000	1.0586	1.2171

Married	1.246	0.025	10.92	0.000	1.1978	1.2962
Employed	0.896	0.018	-5.35	0.000	0.8610	0.9330
Hispanic	1.050	0.030	1.70	0.090	0.9925	1.1103
Perceived health (good)	2.295	0.049	38.80	0.000	2.2006	2.3933
Perceived health (poor)	4.057	0.099	57.56	0.000	3.8684	4.2556
Rural	1.066	0.022	3.06	0.002	1.0232	1.1100
Constant	0.151	0.005	-54.30	0.000	0.1410	0.1616

(n = 72,691, LR $\chi 2$ = 7,965.03, Prob > $\chi 2$ = 0.000, Pseudo R^2 = 0.092, Log Likelihood = -39,407.08)

The first column of Table 4 reports the odds ratio for the logistic regression, the second column reports the standard errors, and the last two columns present the confidence interval of the odds ratio. The model is significant, as indicated by the pseudo-likelihood and Wald chi-square. The decision criteria for hypothesis testing are based on a 5 percent level of significance. For an additional number of restaurants within a quarter miles distance, the odds of being obese increase by 1.0066 but at a decreasing rate as suggested by its square term. Likewise, the odds of being obese for a female respondent is 1.19 times higher than that of males. But the average household income and obesity are independent of each other.

Regarding the age variable, the odds of being obese for respondents in the working-age group are 1.785 times higher, and for the retired age group, the odds are 1.33 times higher than those who belong to the millennial age group. Compared to whites, Asians are 0.29 times, other races are 1.24 times, and Hispanics are 1.05 times more likely to be obese, but the significance level for the Hispanic race category is at 10 percent or better level. Similarly, compared to respondents who had high school or less education, people with some college or associated degree are 10 percent, people with a bachelor's degree are 0.70, and people with a master's degree or above education are 1.14 times more likely to be obese. Likewise, the odds of being obese are 1.245 times that of unmarried, and employed respondents have odds of being obese by 0.9 times that of unemployed. Regarding the health category, the odds of being obese are 2.29

times for people whose perceived health status is good, and the odds are 4.05 times for people whose health condition is poor compared to people whose perceived health condition is excellent. The logistic regression results are consistent with the OLS regression results to explain the variation in the individual's BMI and obesity.

CHAPTER 5

DISCUSSIONS

We used the USDA FoodAPS data collected in 2016 to analyze the association between body mass index and obesity and proximity to restaurants. We estimated two separate regression equations using OLS and logistic models. Both models shared common explanatory variables, but the response variables were different. Our estimates in the OLS model suggest that the availability of restaurants within a 0.25-mile radius of residential areas is associated with higher BMI. The existence of plentiful restaurants closer to the residential area not only encourages the residents to eat more but also makes them less active in preparing food at home or traveling a further distance to obtain healthy food (Rashad, Grossman, & Chou, 2005). However, the relationship between proximity to restaurants and BMI is non-linear, meaning as the supply of restaurants increases, BMI also increases but at a decreasing rate. Further, OLS results indicate that an increase in monthly household income is inversely associated with BMI implying that wealthy people are less vulnerable to the proximity of restaurants. However, the weekly frequency of food away from home is not significant to explain the variation in BMI.

Similarly, compared to white respondents, Asians and people of other races have lower BMI. However, there is not a significant association between proximity to restaurants and Black and Hispanic respondents. The regression results also indicate that people of working age (35 to 65 years) and retired age (65 years or more) have significantly higher BMI than their millennial counterparts. This could be because of the higher leisure time and stable income associated with these people compared to millennials could explain this result. Further, as expected higher level of educational attainment is inversely associated with body mass index. Consistent with the previous studies, our regression result indicates that people tend to have higher BMI at a lower level of education. Still, as the level of education increases toward advanced degrees such as bachelor's or master's degrees, people tend to have a lower level of BMI. For example, our result suggests that people with some college or associate degrees tend to have 32 percent higher BMI than high school dropouts, but people with bachelor's degrees have 76 percent lower, and people with master's degrees have 38 percent lower BMI than high school dropouts. Hence, this study supports previous findings that educated people are more health-conscious and tend to have lower BMI. In contrast, we found that married people had a significantly higher BMI than their unmarried counterparts.

Whether restaurants substantially contribute to increasing obesity is an empirical question and estimates of the linear OLS regression model are likely to give misleading results about the true relationship between eating out and obesity. In this connection, we decide to check the validity of the OLS regression result by fitting a logistic regression model. Since BMI is a continuous variable, and we cannot fit a logistic model with a continuous response variable, we choose obese as our new response variable. Obesity as a response variable is justified as it is strictly derived from BMI, and the comparison of OLS and logit regression results is less questionable.

Both OLS and logistic models focus on the number of restaurants available within 0.25 square miles of the respondents' residence as a relevant measure of BMI. This is because the neoclassical theory of consumer preferences suggests that consumers, as an optimizing agent, try to maximize their utility by consuming more when they are closer to restaurants (Anderson & Matsa, 2011). It is equally possible that the availability of a larger number of fast food and non-fast-food restaurants closer to the agents' residence could lower the monetary and nonmonetary costs of accessing unhealthy food and increase the calorie intake and thereby increase BMI (Currie et al., 2010).

CHAPTER 6

SUMMARY AND CONCLUSIONS

Using FoodAPS data from the USDA, we empirically investigated if BMI or the probability of being obese is associated with the number of restaurants within the 0.25-mile radius of individuals' residences. We find a positive and statistically significant relationship between BMI and the number of restaurants within 0.25-mile proximity. We also find that the probability of being obese is significantly higher for individuals who live within 0.25 miles radius of restaurants. Hence, educating residents about the excessive calory of restaurant meals may help them understand the consequences of eating restaurant meals, which can help to reduce their BMI. Policies encouraging residents involving on physical activities may also help reduce their BMI. Opening new public parks, improving physical safety of public parks, and other economic incentives (e.g., free or reduced-price pass for low-income residents) may help them involve more often in physical activities (Ghimire et al., 2014). Additionally, promoting the culture of cooking meals at home can also help reduce excessive calory intake and BMI.

REFERENCES

Anderson, M. L., & Matsa, D. A. (2011). Are Restaurants Really Supersizing America? *American Economic Journal: Applied Economics*, *3*(1), 152-88. DOI: 10.1257/app.3.1.152

Auchincloss, A. H., Mallya, G. G., Leonberg, B. L., Ricchezza, A., Glanz, K., & Schwarz, D. F. (2013). Customer Responses to Mandatory Menu Labeling at Full-Service Restaurants. *American Journal of Preventive Medicine*, *45*(6), 710-719. https://doi.org/10.1016/j.amepre.2013.07.014

Block, J. P., & Roberto, C. A. (2014). Potential Benefits of Calorie Labeling in Restaurants. *Jama*, *312*(9), 887-888. doi:10.1001/jama.2014.9239

Bollinger, B., Leslie, P., & Sorensen, A. (2011). Calorie Posting in Chain Restaurants. *American Economic Journal: Economic Policy*, *3*(1), 91-128. DOI: 10.1257/pol.3.1.91

Breck, A., Cantor, J., Martinez, O., & Elbel, B. (2014). Who reports noticing and using calorie information posted on fast food restaurant menus? *Appetite*, *81*, 30-36. https://doi.org/10.1016/j.appet.2014.05.027

Calle, E. E., Teras, L. R., & Thun, M. J. (2005). Obesity and Mortality. *New England Journal of Medicine*, *353*(20), 2197-2199. DOI: 10.1056/NEJM200511173532020

CDC (2016). 2016 Behavioral Risk Factor Surveillance System Survey. Available at: https://www.cdc.gov/brfss/annual_data/annual_2016.html

Chandon, P., & Wansink, B. (2007). Is Obesity Caused by Calorie Underestimation? A Psychophysical Model of Meal Size Estimation. *Journal of Marketing Research*, *44*(1), 84-99. https://doi.org/10.1509/jmkr.44.1.084

Consultation, W. H. O. (2003). Diet, Nutrition, and the Prevention of Chronic Diseases. *World Health Organization, Technical Report*. Geneva 2003. health.euroafrica.org

Currie, J., DellaVigna, S., Moretti, E., & Pathania, V. (2010). The Effect of Fast Food Restaurants on Obesity and Weight Gain. *American Economic Journal: Economic Policy*, *2*(3), 32-63. DOI: 10.1257/pol.2.3.32

Davis, A. (2013). Gallup Survey: In US, Less Than Half Look at Restaurant Nutrition Facts. https://news.gallup.com/poll/163904/less-half-look-restaurant-nutrition-facts.aspx

Din, N., Zahari, M. S. M., Othman, C. N., & Abas, R. (2012). Restaurant Operator's Receptiveness towards Providing Nutritional Information on Menu. *Procedia-Social and Behavioral Sciences*, *50*, 699-709. https://doi.org/10.1016/j.sbspro.2012.08.073

Dunn, R. A. (2010). The Effect of Fast-Food Availability on Obesity: An Analysis by Gender, Race, and Residential Location. *American Journal of Agricultural Economics*, *92*(4), 1149-1164. https://doi.org/10.1093/ajae/aaq041

Elbel, B., Kersh, R., Brescoll, V. L., & Dixon, L. B. (2009). Calorie Labeling And Food Choices: A First Look At The Effects On Low-Income People In New York City. *Health Affairs*, *28*(1), w1110-w1121. https://doi.org/10.1377/hlthaff.28.6.w1110

Fullerton, A. S. (2009). A Conceptual Framework for Ordered Logistic Regression Models. *Sociological methods & research*, *38*(2), 306-347. https://doi.org/10.1177/0049124109346162

Ghimire, R., Green, G. T., Poudyal, N. C., & Cordell, H. K. (2014). An analysis of perceived constraints to outdoor recreation. *Journal of Park and Recreation Administration*, *32*(4).

Gustafson, A., & Allen, J., & Schoenberg, N., & Swanson, M. (2016). The relationship between neighborhood food environment and food store choice on purchasing habits among SNAP and lower income households, USDA FoodAPS data. *University of Kentucky Center for Poverty Research Discussion Paper Series, DP2016-09.* http://www.ukcpr.org/research/discussion-papers

Horwich, T. B., Fonarow, G. C., Hamilton, M. A., MacLellan, W. R., Woo, M. A., & Tillisch, J. H. (2001). The Relationship Between Obesity and Mortality in Patients with Heart Failure. *Journal of the American College of Cardiology*, *38*(3), 789-795. https://doi.org/10.1016/S0735-1097(01)01448-6

Hruby, A., & Hu, F. B. (2015). The Epidemiology of Obesity: A Big Picture. *Pharmacoeconomics*, *33*(7), 673-689. https://doi.org/10.1007/s40273-014-0243-x

Inagami, S., Cohen, D. A., Brown, A. F., & Asch, S. M. (2009). Body Mass Index, Neighborhood Fast Food and Restaurant Concentration, and Car Ownership. *Journal of Urban Health*: Bulletin of the New York Academy of Medicine, *86*(5), 683-695. doi:10.1007/s11524-009-9379-y

Krieger, J. W., Chan, N. L., Saelens, B. E., Ta, M. L., Solet, D., & Fleming, D. W. (2013). Menu Labeling Regulations and Calories Purchased at Chain Restaurants. *American Journal of Preventive Medicine*, *44*(6), 595-604. https://doi.org/10.1016/j.amepre.2013.01.031

Lee A, Cardel M, Donahoo WT (2015). Social and Environmental Factors Influencing Obesity. *In: Endotext. MDText.com, Inc., South Dartmouth (M.A.); 2000. PMID: 25905211.*

Lee, K., & Lee, Y. (2018). Parents' meal choices for their children at fast food and family restaurants with different menu labeling presentations. *Nutrition Research and Practice, 12*(3), 243-250. DOI: https://doi.org/10.4162/nrp.2018.12.3.243

Lee-Kwan, S. H., Pan, L., Maynard, L., Kumar, G., & Park, S. (2014). Restaurant Menu Labeling Use Among Adults—17 states, 2012. *MMWR. Morbidity and mortality weekly report*, *63*(27), 581- 584.

Lin, B. H., & Guthrie, J. F. (2012). Nutritional Quality of Food Prepared at Home and Away From Home, 1977-2008, EIB-105. *U.S. Department of Agriculture, Economic Research Service, December 2012*. http://www.ers.usda.gov/topics/food-choiceshealth/diet-quality-nutrition.aspx

Lin, B. H., Guthrie, J., & Frazao, E. (1999). Nutrient contribution of food away from home. *America's America's Eating Habits: Changes and Consequences750*(5), 213-242.

Lin, C. A., Mou, Y., & Lagoe, C. (2011). Communicating nutrition information: Usability and usefulness of the interactive menus of national fast-food chain restaurants. *Journal of Communication in Healthcare*, *4*(3), 187-199. https://doi.org/10.1179/1753807611Y.0000000002

Liu, P. J., Wisdom, J., Roberto, C. A., Liu, L. J., & Ubel, P. A. (2014). Using Behavioral Economics to Design More Effective Food Policies to Address Obesity. *Applied Economic Perspectives and Policy*, *36*(1). https://doi.org/10.1093/aepp/ppt027

Mancino, L., Todd, J., & Lin, B. H. (2009). Separating what we eat from where: Measuring the effect of food food away from home on diet quality. *Food Policy*, *34*(6), 557-562. https://doi.org/10.1016/j.foodpol.2009.09.003

Mather, M., & Scommegna, P. (2015). Up to Half of U.S. premature Deaths are Preventable, Behavioral Factors Key. *Population Reference Bureau*. https://www.prb.org/us-premature-deaths/

Menu Labeling FAQ: All About the FDA Menu Labeling Law, available at: https://pos.toasttab.com/blog/fda-menu-labeling-law-faq (accessed date: 9/23/2019).

Morrison, R. M., Mancino, L., & Variyam, J. N. (2011). Will Calorie Labeling in Restaurants Make a Difference? *Economic Research Service/ USDA*. DOI 10.22004/ag.econ.121239

Nixon, J., & Ulmann, P. (2006). The relationship between health care expenditure and health outcomes. *The European Journal of Health Economics*, *7*(1), 7-18. https://doi.org/10.1007/s10198-005-0336-8

Oexle, N., Barnes, T. L., Blake, C. E., Bell, B. A., & Liese, A. D. (2015). Neighborhood fast food availability and fast-food consumption. *Appetite*, *92*, 227–232. https://doi.org/10.1016/j.appet.2015.05.030

Rashad, I., Grossman, M., & Chou, S. Y. (2005). The Supersize of America: An Economic Estimation of Body Mass Index and Obesity in Adults (No. w11584). *National Bureau of Economic Research.* DOI 10.3386/w11584

Restrepo, B., Minor, T., & Peckham, J. (2018). The Association Between Restaurant Menu Label Use and Caloric Intake. *United States Department of Agriculture, Economic Research Report.* https://www.ers.usda.gov/

Roberto, C. A., Larsen, P. D., Agnew, H., Baik, J., & Brownell, K. D. (2010). Evaluating the impact of menu labeling on food choices and intake. *American Journal of Public health*, *100*(2), 312-318. https://doi.org/10.2105/AJPH.2009.160226

Sacco, J., Lillico, H. G., Chen, E., & Hobin, E. (2016). The influence of menu labelling on food choices among children and adolescents: a systematic review of the literature.

Perspectives in Public Health, *137*(3), 173-181. https://doi.org/10.1177/1757913916658498

Saksena, M. J., Okrent, A. M., Anekwe, T. D., Cho, C., Dicken, C., Effland, A., Elitzak, H., Guthrie, J., Hamrick, K.S., Hyman, J., Jo, Y., Lin, B. H., Mancino, L., McLaughlin, P.W., Rahkovsky, I., Ralston, K., Smith, T.A., Stewart, H., Todd, J. & Tuttle, C. (2018). America's Eating Habits: Food Away from Home, Economic Information Bulletin 281119, *United States Department of Agriculture, Economic Research Service.* DOI: 10.22004/ag.econ.281119

Santos Nobre, J., & da Motta Singer, J. (2007). Residual Analysis for Linear Mixed Models. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, *49*(6), 863-875. https://doi.org/10.1002/bimj.200610341

Seguin, R. A., Aggarwal, A., Vermeylen, F., & Drewnowski, A. (2016). Consumption Frequency of Foods Aaway from Home Linked with Higher Body Mass Index and Lower Fruit and Vegetable Intake among Adults: A Cross-Sectional Study. *Journal of Environmental and Public Health*, *2016*. http://dx.doi.org/10.1155/2016/3074241

Smith, L. P., Ng, S. W., & Popkin, B. M. (2013). Trends in U.S. home food preparation and consumption: analysis of national nutrition surveys and time use studies from 1965–1966 to 2007–2008. *Nutrition Journal*, *12*(1), 1-10.

Taillie, L. S., Afeiche, M. C., Eldridge, A. L., & Popkin, B. M. (2017). The contribution of at-home and away-from-home food food to dietary intake among 2–13-year-old Mexican children. *Public Health Nutrition*, *20*(14), 2559-2568. https://doi.org/10.1017/S1368980016002196

Todd, J. E., Mancino, L., & Lin, B. H. (2010). The Impact of Food Away from Home on Adult Diet Quality. *USDA-ERS Economic Research Report Paper* (90). http://dx.doi.org/10.2139/ssrn.1557129

United States Department of Agriculture Economic Research Service, https://www.ers.usda.gov/publications/pub-details/?pubid=90227

U.S. Department of Agriculture Economic Research Service. (2016). National Household Food Acquisition and Purchase Survey (FoodAPS): User's Guide to Survey Design, Data Collection, and Overview of Datasets.

Wooldridge, J. M. (2015). Econometric Analysis of Cross Section and Panel Data. *MIT Press.*