An Econometric Analysis of Obesity

Debarshi Dutta
Fordham University
The Graduate School of Arts and Sciences
The Bronx, New York City
Email: ddutta@fordham.edu

***Abstract*—There are multiple factors that affect obesity. Amongst them are- poverty, density of fast food restaurants in the locality and low proximity to a conventional grocery store. Conversely, people who live in the vicinity of fitness centers are more likely to participate in recreational activities and have shown to be less obese. The paper analyzes the role these conditions which constitute the standard of living and living conditions play in the cause of obesity.**

1. Introduction

According to the Centers for Disease Control (CDC), a person is obese [1] when their calculated Body Mass Index (BMI) is 30.0 or higher. There have been substantial efforts made to curb the rise of obesity but there is a lack of evidence of clear success thus far. Data collected by the CDC in 2017-2018 indicate that a whopping 42.4% of American adults are considered obese, while obesity affects 13.7 million children and adolescents aged 2-19 years, which is about [2] 18.5% . Despite an increasing recognition of the problem, obesity continues to grow in the United States and the world at large. The problem is pervasive. The convoluted nature of obesity makes it one of the most difficult public health issues that confronts our society at this point.

1.1. An Overview of Obesity

Besides the CDC, the most comprehensive data regarding the proliferation of the epidemic has been provided by the National Health and Nutrition Examination Surveys (NHANES). An analysis of the NHANES data published during 2007-2008 was done by [3] Katherine M Flegal et al. The data indicates obesity rates for adults have been gradually on the rise over the past 30 years or so. In fact, the NHANES data indicates that approximately 68% were overweight or obese, and approximately 34% were obese. So there has been an approximate increase of 8% in obesity in adults in the United States over a span of twelve to thirteen years.

Obesity by itself is a recognized disease, but can be the source of a number of other diseases. Some of these [4]are-

1) Type 2 diabetes and prediabetes
2) Coronary artery disease (CAD)
3) Cancer
4) Sleep apnea
5) [5] With the rise of obesity amongst children and adolescents, they are beginning to develop risk factors for chronic diseases usually seen much later in life, such as dyslipidemia, hypertension, and hyperinsulinemia
6) An increased number of obese children and adolescents are now being diagnosed with type 2 [6] diabetes. This is especially troubling because this disease was virtually nonexistent in this particular population a few generations ago.

2. Literature Review

[7]David Hughes et al analyzes obesity from the economic point of view in the United Kingdom. Besides the conventional economic analysis prevalent in healthcare studies namely cost minimization, cost effectiveness, cost benefit and cost utility, the paper introduces burden of illness and cost of illness analysis in the disease area of obesity. These two metrics gauge the finances and resources required to treat an illness over a given period. They indicate the full cost (measured in terms of mortality and morbidity as well as treatment resource costs) imposed by a disease on society. Here the approach introduced by [8] Colditz in 1986 and then revised in 1990 by Wolf and Colditz [9] to find the cost of treating obesity as a disease is of particular interest. The total cost of treating obesity in the US in 1986 was $39.3 billion, while the number rose to $45.8 billion by 1990, which was 6% of the overall health care expenditure. The enormity of this number has been attributed to the fact that obesity invariably leads to other diseases and the cost of treating these co-morbidities gets added to the amount. For instance, 27% of Cardiovascular diseases is diagnosed in obese individuals and that, among the obese, 70% of CVD is attributable to obesity. 19% (0.27 x 0.70) then of the estimated aggregate cost of treating CVD in the US can be attributed to the disease obesity.

The paper however merely glosses over the indirect costs associated with obesity. For instance, because of the nature of the disease and the co-morbidities it leads individuals may need to take time off work. The lose of revenue arising from

lost productivity to $20 billion for the US in 1986. There is also the significant psychological and social toll on obese individuals, which is recognized but unfortunately cannot be quantified. The author completely fails to mention relevant factors like prescriptions, operations or support groups. These factors may be difficult to quantify but are extremely useful in treating obesity and therefore significant. Isolated cases like an individual born with hereditary diseases like diabetes and an overall healthy individual contracting CVD are not taken into account as well. The paper is limited because it assumes the co-morbidities are primarily weight related.

[10]Revels et al predicts obesity rate and obesity related healthcare costs using data analytic techniques specifically the autoregressive integrated moving average or the ARIMA model trained using data from the CDC. The ARIMA model was designed using the Box Jenkins method. Prediction was done with 95% confidence levels. Their results predict around 45% of the population will be obese by the year 2035. The number becomes nearly 75% if the percentage of the population that is overweight is taken into consideration. The predicted expenditure rises by $2.2 billion from 2016 to 2035, which is an annual increase of around $118 million of healthcare expenditure associated with obesity.

The study is affected adversely by the fact that time series analyses using the ARIMA model will include noise in the computation by definition, though some of the methods implemented by the Box Jenkins method negates some of the noise. The study did not include indirect costs and all the co-morbidities associated with obesity. Finally, the costs related to the successful treatment of obesity was not considered.

Allison et al [11] provides an analysis on whether the medical cost of people who had returned to a normal-weight and lived longer differed from the medical costs of people who simply lived longer. In this approach the population was divided in 5-year increments starting from 20-24 age range. For each range, the probability of obesity, a base mortality rate due to obesity, and the per capita healthcare costs were calculated. After that, two sensitivity analyses were conducted: one that included mortality rates associated with obesity, and one that which didn't feature obesity. Their results indicate that when including the mortality rates associated with obesity in the model, the healthcare costs were about 25% less than the projected healthcare costs for non-obese individuals.

There were a number of limitations in this study. The indirect costs of obesity do not feature at all. These costs however difficult to quantify are significant. The paper also assumes all co-morbidities are weight related and does not account for the isolated cases mentioned earlier. In addition the confidence intervals also appear to be too wide.

Thorpe et al [12] comments on the difference in healthcare spending between 1987 to 2001 on obesity, diabetes and hyperlipidemia for those considered obese and overweight. Healthcare spending on obesity increased by 51% relative to other categories over that period. In 2001, the spending on obese was 37.2% more than on individuals who fall within the normal weight range. This is an increase of 15.2% for the same category measured in 1987.

The paper does not take into account the issues associated with obesity can affect normal-weight individuals too. The co-morbidity list related to obesity included in the study does not consider gallbladder disease. The risk of gall bladder disease increases with a rise in BMI [13].

Finkelstein et al [14] used the Medical Expenditure Panel Surveys (MEPS) data from 1998 to 2006 and came to the conclusion that the approximate healthcare cost for an obese person is 42% higher than an individual considered within a the normal-weight range. One significant aspect of the paper is that it takes into account the indirect cost of prescription drugs- that amount is estimated to be around $7 billion for non-institutionalized obese patients in 1998.

There is an inherent flaw in the study because MEPS requires respondents to self-report their respective height and weights to calculate the BMI. There is also a bias in the reported numbers because MEPS only reports the non-institutionalized patients.

3. Need for Study

Obesity has been recognized as a disease by itself and it leads to other complications and illnesses as well. It is abundantly clear that the cost of treating obesity is a major draw on finances and resources. There are both direct and indirect costs involved, the latter of which are difficult to quantify but extremely significant. The expenditure to treat obesity and its related diseases have been rising for a while and the projections indicate the trend will continue. The situation is widespread across the world and the United States is not exempt from this.

It might be beneficial at this juncture to take a step back to survey and pinpoint the overarching factors that lead to obesity. This will lead to suitable policies being put into place and subsequent diversion of funds and resources to tackle the problem at a more grassroots level. The data published by the CDC [15] [16] indicates low income households- with no cars and low access to grocery stores, families with children eligible for free lunch with the SNAP program, WIC participants are likely to be overweight and obese. So it is interesting to identify the significant socioeconomic factors that leads to the prevalence of obesity.

4. Data

The US Department of Agriculture publishes the Food Environment Atlas. The data from the USDA is collected from multiple sources and cover a range of years and geographic levels. [1] Obesity rates in the data set used in the paper are from 2010 released in 2011 by the USDA. The

1. Source: The United States Department of Agriculture (USDA)- Food Environment Atlas.

- The link for the data download available at https://www.ers.usda.gov/data-products/food-environment-atlas/data-access-and-documentation-downloads/

data is published in excel format with variables aggregated on the respective worksheet.Each row in the data set contains food environment indicators for one of the U.S. counties. The data comprises of two hundred and eleven variables in three categories. They are-

1) Food Choices- the proximity of the population divided by counties to healthy, affordable food
2) Health and Well being- describes diabetes, and obesity rates
3) Community Characteristics- demography divided by age and race, income and poverty, metro and non metro status of a county, availability of recreational facilities.

4.1. Pre Processing

The data is available across multiple excel sheets. So, the data was combined into one comma separated value file. The general principle of operation for combining the data was, that after importing a spreadsheet, the feature correlation was collected. The features with low correlation were dropped. It is to be taken into consideration that the year of interest for this particular study is 2010. Data collected after 2010 is not a part of the study. This is because the dependent variable, that is adult obesity rate, is for the year 2010. For data missing for the year 2010, the available data closest to 2010 was used. The columns with inconsistent data, that is too many NULLs were generally dropped. More information regarding the unique cases for each spreadsheet and their features have been tabulated below

1) The health and physical activity data
    - The variable of choice- adult obesity for the year 2010 was extracted
    - The health variables were extracted
    - The diabetes data was dropped.
    - The data for the year 2009 was insufficient as it had too many NULL's and was consequently dropped
    - Any data beyond 2010 was not considered
2) Access and Proximity to Grocery Store data
    - the correlation of access features was calculated and the results indicate that all the columns are highly correlated with the exception of PCT_LACCESS_HHNV10
3) Store Availability data
    - The correlation between the features were calculated. Based on the results, the "State and county Federal information processing standards" column was dropped because of a low correlation value.

- The complete Food Environment Atlas Data Documentation available at- https://www.ers.usda.gov/webdocs/DataFiles/80526/archived_documentation_2011.pdf?v=3604.3

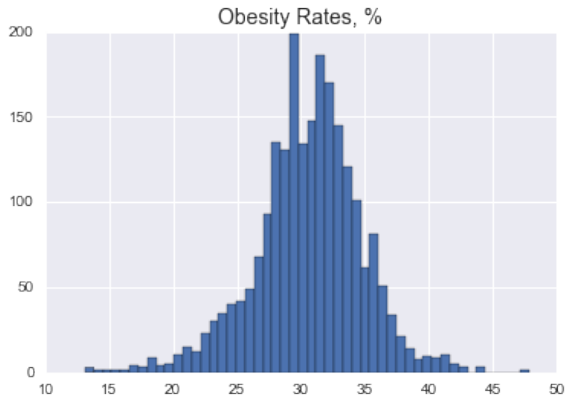

Figure 1. Histogram of response variable

4) The Restaurant Availability and Expenditures data
    - The columns that contain data from 2002 were dropped. This is because the same data exists for 2007, which was chosen as it is closer to 2010.
5) Food Assistance data
    - The most recent data in this table for some features were from 2008

A dataset was constructed out of sixty six relevant columns and the stored in comma separated values format.

4.2. Data Split

The dataset was split into the training and testing using the 80-20 rule.

5. Explanatory Data Analysis of Training Data

Before the EDA the following operations were performed

- Rows where the response variable (Adult Obesity Rate) = NULL were removed from the dataset. The number of NULLs were less than ten percent of the field.
- There were cases where numerical fields contained string values. That particular datatype was converted to float and NaN's in a case by case basis

5.1. Outliers

A histogram analysis (5.1) was performed for the outliers of the dependent variable (PCT_OBESE_ADULTS10) 1. Any value beyond three standard deviations of the mean of the dependent variable is defined as an outlier.

There are 23 rows with outliers in the outcome variable in the training data.

Figure 2. Scatterplot of the Correlation of the Adult Obesity to Low Income Households

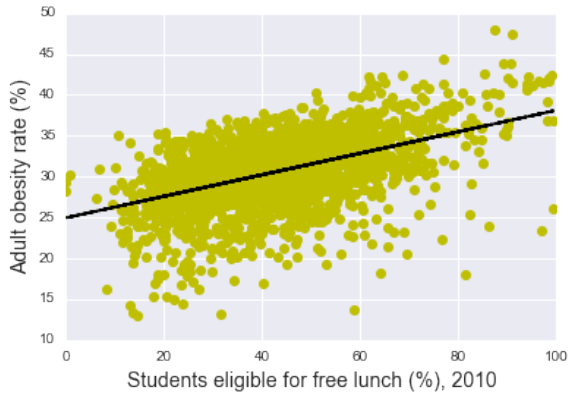

Figure 3. Scatterplot plot of the correlation of adult obesity rate to the students who qualify for free lunch

5.2. Correlation of the Adult Obesity Percent to Low income households

Low Income Households have been defined as

- No car and low access to grocery stores
- Households with students eligible for free lunch

The overall data with the exception of PCT_LACCESS_HHNV10 is highly correlated 5.2. As an addendum, the correlation values of all the variables is calculated with respect to the response variable. The highest correlation to the response variable is in fact with the students eligible for free lunch, that is, who have qualified for the SNAP program 5.2.

5.3. Check for Multicollinearity

The response variable is tested for multicollinearity.5.5 Eigen values and eigen vectors were used for this test. The collinear features were removed.

	Feature	PCT_WIC09
0	PCT_LACCESS_POP10	0.014154
14	PC_FSRSALES07	0.011043
15	PCT_SNAP09	0.011593
22	PCT_NSLP09	0.020444
25	PCT_SBP09	0.020140
29	FOODINSEC_07_09	0.018293
30	VLFOODSEC_07_09	0.016012
32	MILK_PRICE10	0.026192
33	SODA_PRICE10	0.013417
34	MILK_SODA_PRICE10	0.028365
48	PCT_NHWHITE10	0.685061
49	PCT_NHBLACK10	0.499488
50	PCT_HISP10	0.457496
51	PCT_NHASIAN10	0.076047
52	PCT_NHNA10	0.248387
57	POVRATE10	0.011948
59	CHILDPOVRATE10	0.011661

Figure 4. Some of the collinear variables

5.4. Mean Squared Error

The Mean Squared Error was calculated for the linear regression model. The performance threshold MSE is 17.877 for the model.

5.5. NULL values

All the NULL values were replaced by zeros

6. Modeling

Different modeling techniques have been used. The performance of these are then compared, using the MSE value.

6.1. Random Forest

The random forest technique is used to identify the relevant important features. The number of estimators (n) was set at hundred (100). The top 10 features from the Random Forest selection model were-

1) Students eligible for free lunch, 2010
2) SNAP participants, 2009
3) ERS natural amenity index, 1999
4) Minority population (African American), 2010
5) Expenditures per capita, restaurants, 2007
6) National School Lunch Program participants, 2009.
7) Minority population (Asian), 2010
8) Full-service restaurants per 1,000 population, 2007
9) SNAP-authorized stores per 1,000, 2008
10) Percentage of the white population, 2010

6.2. Linear Regression

The linear Regression was initially modeled across all the sixty three features and the performance was compared with the Random Forest model. (6.3.3, 6.3.3)

6.3. Improvements of the Regression Model

The following techniques were used to improve the linear regression model.

	Features	Importance Score	Running Sum
23	PCT_FREE_LUNCH10	0.1057	0.1057
47	NATAMEN	0.0962	0.2019
15	PCT_SNAP09	0.0948	0.2967
49	PCT_NHBLACK10	0.0788	0.3755
14	PC_FSRSALES07	0.0614	0.4369
51	PCT_NHASIAN10	0.047	0.4839
22	PCT_NSLP09	0.0418	0.5257
12	FSRPTH07	0.0309	0.5566
9	SNAPSPTH08	0.0305	0.5871
52	PCT_NHNA10	0.021	0.6081
48	PCT_NHWHITE10	0.0209	0.629
50	PCT_HISP10	0.0201	0.6491
26	PCT_SFSP09	0.0179	0.667
4	PCT_LACCESS_HHNV10	0.0161	0.6831
55	PCT_18YOUNGER10	0.0148	0.6979
7	CONVSPTH07	0.0144	0.7123
35	PCT_LOCLFARM07	0.0134	0.7257
57	POVRATE10	0.0131	0.7388
56	MEDHHINC10	0.0129	0.7517
54	PCT_65OLDER10	0.0127	0.7644
46	RECFACPTH07	0.0125	0.7769
13	PC_FFRSALES07	0.0124	0.7893

Figure 5. Top 80% features used for linear regression

6.3.1. Using top 80% features. The model was improved choosing the top twenty two features (6.3.3). The graph 6.3.3 suggests this is a better fit.

6.3.2. Outliers Removed. The model WAS redone by removing the outliers. Outliers defined as any data that lies beyond three standard deviations of the from the mean of the response variable. The fit in this case is very close to the previous technique (6.3.3).

6.3.3. MSE with outliers removed. This is the technique that produced the best result. The methodology was

- Remove response variables outliers from data
- Get coefficients for collinear features
- Check the correlation between the best features out of collinear feature sets
- Test Linear Regression with all non-collinear features
- Find the best feature out of the collinear feature set
- The MSE of the linear regression for the non-collinear feature set is 9.3571769776865317

6.3.4. Performance of the models. The figure 9 (6.4) shows the performance of the models ranked in terms of the MSE value. 'n' indicates the number of datapoints.

6.3.5. Testing. The trained model was tested with all non-collinear features and best feature out of collinear feature set. The MSE from the test data is calculated to be 9.3571769776865317

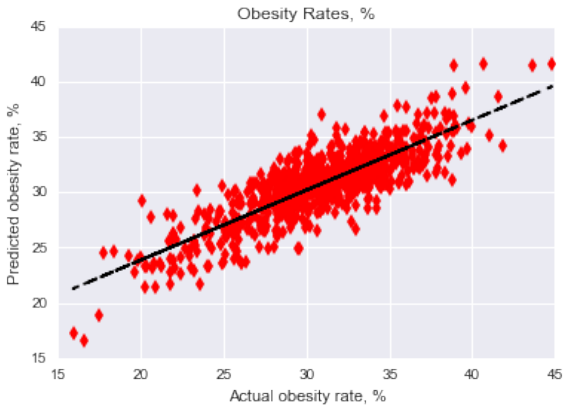

Figure 6. Random Forest

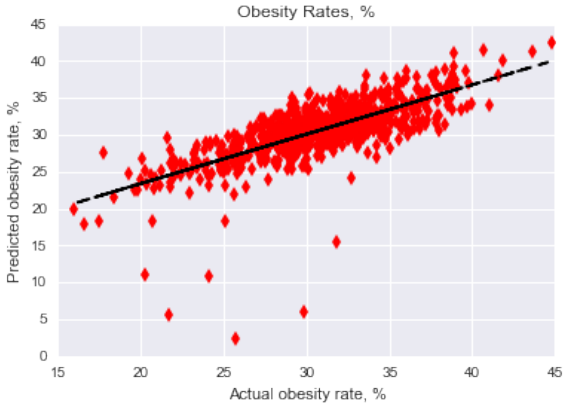

Figure 7. Linear Regression for all features

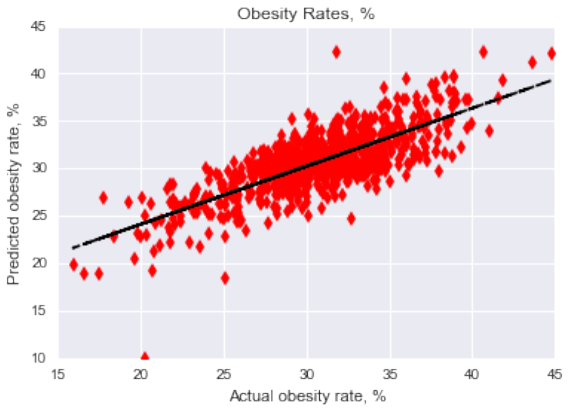

Figure 8. Linear Regression using top 80% features

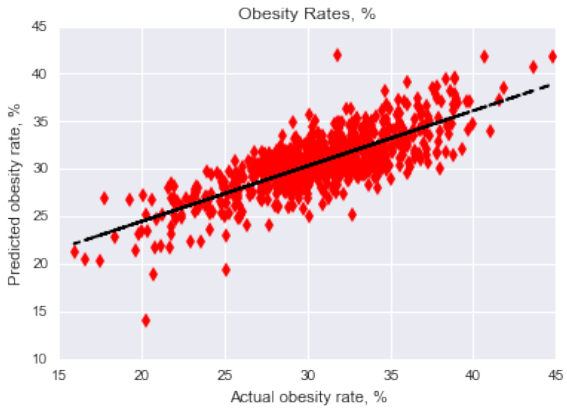

Figure 9. Linear Regression with outliers removed

6.4. Feature Selection Tests

The technique used to select features from the Linear Regression model is Recursive Feature Elimination included in the scikit learn library. The benefit of this function is that it produces the features by rank 6.4.

7. Conclusion

Some of the interesting conclusions of the study are

1) Counties with a high poverty rate have around 4-5% higher obesity rate than others. This corroborates the CDC findings and the working hypothesis of this study.
2) Percentage of students in a county who are eligible for free lunch program have the strongest correlation with a predictor variable, that is they are at a higher risk than others to be obese- which is the other hypothesis.
3) The feature selection from the Linear Regression model shows that child food insecurity percentage in households aggregated over the years 2001-07 is the highest ranked feature. This is significant because the response variable adult obesity rate is for the year 2010. So this verifies the initial hypothesis as well.
4) Inferences made from the other top features such as- WIC participants from the year 2009, state wide food insecurity aggregated over 2007-09 and SNAP participants from year 2009 also verify the initial hypothesis.

8. Limitations and Future Scope

8.1. Limitations

One of the limitations of the study is that the co-morbidities such as diabetes analyzed in detail earlier had to be ignored. The dataset unfortunately did not provide sufficient and consistent data for such diseases that warranted

	Feature	Rank
31	FOODINSEC_CHILD_01_07	1
27	PCT_WIC09	1
29	FOODINSEC_07_09	1
17	SNAP_OAPP10	1
15	PCT_SNAP09	1
35	PCT_LOCLFARM07	1
37	VEG_FARMS07	1
12	FSRPTH07	1
22	PCT_NSLP09	1
47	NATAMEN	1
38	FRESHVEG_FARMS07	1
49	PCT_NHBLACK10	1
7	CONVSPTH07	1
50	PCT_HISP10	1
51	PCT_NHASIAN10	1
52	PCT_NHNA10	1
55	PCT_18YOUNGER10	1
56	MEDHHINC10	1
48	PCT_NHWHITE10	1
23	PCT_FREE_LUNCH10	1
34	MILK_SODA_PRICE10	2
33	SODA_PRICE10	3
32	MILK_PRICE10	4

Figure 10. Feature Selection from Linear Regression Model

	Model	n_features	MSE
5	Random Forest	63	5.418264
3	Linear Regression _all_features	63	8.619312
2	Linear Regression	22	7.660215
1	Linear Regression_no outliers	22	6.348386

Figure 11. Comparison of the performances of the models

analysis and consequent inclusion. A similar study perhaps can be done from a different data source that included similar aspects.

8.2. Future Scope

This study can be improved on the following ways-

- Both LASSO and Ridge regressions can be implemented in the study and then their performance and feature selections can be compared with the linear model.
- A identical study can be done for the ATLAS data published in 2020. It is safe to assume that resources have been put in place and measures have been improved to tackle obesity since 2010. Ten years is a significant period. A comparison of this study and the one from 2020 data would indicate which measures have proved to be efficient and what needs to be addressed more moving forward.

List of Figures

References

[1] "Defining adult overweight and obesity — overweight & obesity — cdc," https://www.cdc.gov/obesity/adult/defining.html, (Accessed on 12/01/2020).

[2] "Childhood obesity facts — overweight & obesity — cdc," https://www.cdc.gov/obesity/data/childhood.html, (Accessed on 12/01/2020).

[3] "Prevalence and trends in obesity among us adults, 1999-2008 - pubmed," https://pubmed.ncbi.nlm.nih.gov/20071471/, (Accessed on 12/02/2020).

[4] "Obesity: Overview of an epidemic," https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3228640/, (Accessed on 12/01/2020).

[5] "Cardiovascular risk factors and excess adiposity among overweight children and adolescents: The bogalusa heart study - sciencedirect," https://www.sciencedirect.com/science/article/abs/pii/S0022347606008171, (Accessed on 11/29/2020).

[6] "Type 2 diabetes among north american children and adolescents: an epidemiologic review and a public health perspective - pubmed," https://pubmed.ncbi.nlm.nih.gov/10802501/, (Accessed on 11/29/2020).

[7] "A review of the economic analysis of obesity - pubmed," https://pubmed.ncbi.nlm.nih.gov/9246835/, (Accessed on 11/26/2020).

[8] "Economic costs of obesity and inactivity - pubmed," https://pubmed.ncbi.nlm.nih.gov/10593542/, (Accessed on 12/07/2020).

[9] "Current estimates of the economic cost of obesity in the united states - pubmed," https://pubmed.ncbi.nlm.nih.gov/9545015/, (Accessed on 12/07/2020).

[10] "Predicting obesity rate and obesity-related healthcare costs using data analytics," file:///Users/debarshi/Downloads/Predicting_Obesity_Rate_and_Obesity-Related_Health.pdf, (Accessed on 11/29/2020).

[11] "The direct health care costs of obesity in the united states. — ajph — vol. 89 issue 8," https://ajph.aphapublications.org/doi/10.2105/AJPH.89.8.1194, (Accessed on 11/29/2020).

[12] "The impact of obesity on rising medical spending — health affairs," https://www.healthaffairs.org/doi/full/10.1377/hlthaff.W4.480, (Accessed on 11/29/2020).

[13] "Relation of regional gray and white matter volumes to current bmi and future increases in bmi: a prospective mri study — international journal of obesity," https://www.nature.com/articles/ijo2011175, (Accessed on 12/08/2020).

[14] "Annual medical spending attributable to obesity: Payer-and service-specific estimates — health affairs," https://www.healthaffairs.org/doi/10.1377/hlthaff.28.5.w822, (Accessed on 12/02/2020).

[15] "Prevalence of obesity among adults, by household income and education — united states, 2011–2014 — mmwr," https://www.cdc.gov/mmwr/volumes/66/wr/mm6650a1.htm?s_cid=mm6650a1_w, (Accessed on 12/09/2020).

[16] "Obesity among wic-enrolled young children — overweight & obesity — cdc," https://www.cdc.gov/obesity/data/obesity-among-WIC-enrolled-young-children.html, (Accessed on 12/09/2020).

Appendix

TABLE 1: List of Variables

First column	Second column
LACCESS_POP10	Population, low access to store, 2010
PCT_LACCESS_POP10	Population, low access to store (%), 2010
LACCESS_LOWI10	Low income & low access to store, 2010
PCT_LACCESS_LOWI10	Low income & low access to store (%), 2010
LACCESS_CHILD10	Children, low access to store, 2010
PCT_LACCESS_CHILD10	Children, low access to store (%), 2010
LACCESS_SENIORS10	Seniors, low access to store, 2010
PCT_LACCESS_SENIORS10	Seniors, low access to store (%), 2010
LACCESS_HHNV10	Households, no car & low access to store, 2010
PC_FSRSALES07	Expenditures per capita, restaurants, 2007*
GROC07	Grocery stores, 2007
GROC12	Grocery stores, 2012
PCH_GROC_07_12	Grocery stores (% change), 2007-12
GROCPTH07	Grocery stores/1,000 pop, 2007
GROCPTH12	Grocery stores/1,000 pop, 2012
PCH_GROCPTH_07_12	Grocery stores/1,000 pop (% change), 2007-12
SUPERC07	Supercenters & club stores, 2007
SUPERC12	Supercenters & club stores, 2012
PCH_SUPERC_07_12	Supercenters & club stores (% change), 2007-12
SUPERCPTH07	Supercenters & club stores/1,000 pop, 2007
SUPERCPTH12	Supercenters & club stores/1,000 pop, 2012
PCH_SUPERCPTH_07_12	Supercenters & club stores/1,000 pop (% change), 2007-12
CONVS07	Convenience stores, 2007
CONVS12	Convenience stores, 2012
PCH_CONVS_07_12	Convenience stores (% change), 2007-12
FSRPTH12	Full-service restaurants/1,000 pop, 2012
CONVSPTH12	Convenience stores/1,000 pop, 2012
PCH_CONVSPTH_07_12	Convenience stores/1,000 pop (% change), 2007-12
SPECS07	Specialized food stores, 2007
SPECS12	Specialized food stores, 2012
PCH_SPECS_07_12	Specialized food stores (% change), 2007-12
SPECSPTH07	Specialized food stores/1,000 pop, 2007
SPECSPTH12	Specialized food stores/1,000 pop, 2012
PCH_SPECSPTH_07$_1$2	Specialized food stores/1,000 pop (% change), 2007-12
SNAPS08	SNAP-authorized stores, 2008
SNAPS12	SNAP-authorized stores, 2012
PCH_SNAPS_08_12	SNAP-authorized stores (% change), 2008-12
SNAPSPTH08	SNAP-authorized stores/1,000 pop, 2008
NATAMEN	ERS natural amenity index, 1999
PCH_SNAPSPTH_08_12	SNAP-authorized stores/1,000 pop (% change), 2008-12
WICS08	WIC-authorized stores, 2008
WICS12	WIC-authorized stores, 2012
PCH_WICS_08_12	WIC-authorized stores (% change), 2008-12
WICSPTH08	WIC-authorized stores/1,000 pop, 2008
WICSPTH12	WIC-authorized stores/1,000 pop, 2012
PCH_WICSPTH_08_12	WIC-authorized stores/1,000 pop (% change), 2008-12
FFR07	Fast-food restaurants, 2007
FFR12	Fast-food restaurants, 2012
PCH_FFR_07_12	Fast-food restaurants (% change), 2007-12
FFRPTH07	Fast-food restaurants/1,000 pop, 2007
FFRPTH12	Fast-food restaurants/1,000 pop, 2012

Continued on next page

TABLE 1 – continued from previous page

Variable Code	Variable Name
PCH_FFRPTH_07_12	Fast-food restaurants/1,000 pop (% change), 2007-12
FSR07	Full-service restaurants, 2007
FSR12	Full-service restaurants, 2012
PCH_FSR_07_12	Full-service restaurants (% change), 2007-12
FSRPTH07	Full-service restaurants/1,000 pop, 2007
CONVSPTH07	Convenience stores/1,000 pop, 2007
PCH_FSRPTH_07_12	Full-service restaurants/1,000 pop (% change), 2007-12
PC_FFRSALES02	Expenditures per capita, fast food, 2002*
PC_FFRSALES07	Expenditures per capita, fast food, 2007*
PC_FSRSALES02	Expenditures per capita, restaurants, 2002*
PCT_LOCLFARM07	Farms with direct sales (%), 2007
REDEMP_SNAPS08	SNAP redemptions/SNAP-authorized stores, 2008
REDEMP_SNAPS12	SNAP redemptions/SNAP-authorized stores, 2012
PCH_REDEMP_SNAPS_08_12	SNAP redemptions/SNAP-authorized stores (% change), 2008-12
PCT_SNAP09	SNAP participants (% pop), 2009*
PCT_SNAP14	SNAP participants (% pop), 2014*
PCH_SNAP_09_14	SNAP participants (change % pop), 2009-14*
RECFACPTH12	Recreation & fitness facilities/1,000 pop, 2012
PC_SNAPBEN10	SNAP benefits per capita, 2010
PCH_PC_SNAPBEN_08_10	SNAP benefits per capita (% change), 2008-10
SNAP_PART_RATE08	SNAP participants (% eligible pop), 2008*
SNAP_PART_RATE10	SNAP participants (% eligible pop), 2010*
SNAP_OAPP00	SNAP online application, 2000*
SNAP_OAPP05	SNAP online application, 2005*
SNAP_OAPP10	SNAP online application, 2010*
SNAP_FACEWAIVER00	SNAP face interview waiver, 2000*
SNAP_FACEWAIVER05	SNAP face interview waiver, 2005*
SNAP_FACEWAIVER10	SNAP face interview waiver, 2010*
SNAP_VEHEXCL00	SNAP vehicle exclusion, 2000*
SNAP_VEHEXCL05	SNAP vehicle exclusion, 2005*
SNAP_VEHEXCL10	SNAP vehicle exclusion, 2010*
SNAP_BBCE00	SNAP Broad-based Categorical Eligibility, 2000*
SNAP_BBCE05	SNAP Broad-based Categorical Eligibility, 2005*
SNAP_BBCE10	SNAP Broad-based Categorical Eligibility, 2010*
SNAP_REPORTSIMPLE00	SNAP simplified reporting, 2000*
SNAP_REPORTSIMPLE05	SNAP simplified reporting, 2005*
SNAP_REPORTSIMPLE10	SNAP simplified reporting, 2010*
PCT_NSLP09	National School Lunch Program participants (% pop), 2009*
PCT$_N SLP14$	National School Lunch Program participants (% pop), 2014*
PCH_NSLP_09_14	National School Lunch Program participants (change % pop), 2009-14*
PCT_FREE_LUNCH06	Students eligible for free lunch (%), 2006
SODA_PRICE10	Price of sodas/national average, 2010**
PCT_REDUCED_LUNCH06	Students eligible for reduced-price lunch (%), 2006
PCT_REDUCED_LUNCH10	Students eligible for reduced-price lunch (%), 2010
PCT_SBP09	School Breakfast Program participants (% pop), 2009*
PCT_SBP14	School Breakfast Program participants (% pop), 2014*
PCH_SBP_09_14	School Breakfast Program participants (change % pop), 2009-14*
PCT_SFSP09	Summer Food Service Program participants (% pop), 2009*
PCT_SFSP14	Summer Food Service Program participants (% pop), 2014*
PCH_SFSP_09_14	Summer Food Program participants (change % pop), 2009-14*
PC_WIC_REDEMP08	WIC redemptions per capita, 2008
PC_WIC_REDEMP12	WIC redemptions per capita, 2012
PCH_PC_WIC_REDEMP_08_12	WIC redemptions per capita (% change), 2008-12
REDEMP_WICS08	WIC redemptions/WIC-authorized stores, 2008
REDEMP_WICS12	WIC redemptions/WIC-authorized stores, 2012

Continued on next page

TABLE 1 – continued from previous page

Variable Code	Variable Name
PCH_REDEMP_WICS_08_12	WIC redemptions/WIC-authorized stores (% change), 2008-12
PCT_WIC09	WIC participants (% pop), 2009*
PCT_WIC14	WIC participants (% pop), 2014*
PCH_WIC_09_14	WIC participants (change % pop), 2009-14*
PCT_CACFP09	Child & Adult Care (% pop), 2009*
PCT_CACFP14	Child & Adult Care (% pop), 2014*
PCH_CACFP_09_14	Child & Adult Care (change % pop), 2009-14*
FDPIR12	FDPIR Sites, 2012
FOODINSEC_00_02	Household food insecurity (%, three-year average), 2000-02*
FOODINSEC_07_09	Household food insecurity (%, three-year average), 2007-09*
FOODINSEC_10_12	Household food insecurity (%, three-year average), 2010-12*
CH_FOODINSEC_02_12	Household food insecurity (change %),2000-02 to 2010-12*
CH_FOODINSEC_09_12	Household food insecurity (change %),2007-09 to 2010-12*
VLFOODSEC_00_02	Household very low food security (%, three-year average), 2000-02*
VLFOODSEC_07_09	Household very low food security (%, three-year average), 2007-09*
VLFOODSEC_10_12	Household very low food security (%, three-year average), 2010-12*
CH_VLFOODSEC_02_12	Household very low food security (change %),2000-02 to 2010-12*
CH_VLFOODSEC_09_12	Household very low food security (change %),2007-09 to 2010-12*
FOODINSEC_CHILD_01_07	Child food insecurity (% households, multiple-year average), 2001-07*
FOODINSEC_CHILD_03_11	Child food insecurity (% households, multiple-year average), 2003-11*
MILK_PRICE10	Price of low-fat milk/national average, 2010**
PCT_LACCESS_HHNV10	Households, no car & low access to store (%), 2010
MILK_SODA_PRICE10	Price of low-fat milk/price of sodas, 2010**
SODATAX_STORES11	Soda sales tax, retail stores, 2011*
SODATAX_VENDM11	Soda sales tax, vending, 2011*
CHIPSTAX_STORES11	Chip & pretzel sales tax, retail stores, 2011*
CHIPSTAX_VENDM11	Chip & pretzel sales tax, vending, 2011*
FOOD_TAX11	General food sales tax, retail stores, 2011*
DIRSALES_FARMS07	Farms with direct sales, 2007
SNAPSPTH12	SNAP-authorized stores/1,000 pop, 2012
PCT_LOCLSALE07	Direct farm sales (%), 2007
DIRSALES07	Direct farm sales, 2007
PC_DIRSALES07	Direct farm sales per capita, 2007
FMRKT09	Farmers' markets, 2009
FMRKT13	Farmers' markets, 2013
PCH_FMRKT_09_13	Farmers' markets (% change), 2009-13
FMRKTPTH09	Farmers' markets/1,000 pop, 2009
FMRKTPTH13	Farmers' markets/1,000 pop, 2013
PCH_FMRKTPTH_09_13	Farmers' markets/1,000 pop (% change), 2009-13
FMRKT_SNAP13	Farmers' markets that report accepting SNAP, 2013
PCT_FMRKT_SNAP13	Farmers' markets that report accepting SNAP (%), 2013
FMRKT_WIC13	Farmers' markets that report accepting WIC, 2013
PCT_FMRKT_WIC13	Farmers' markets that report accepting WIC (%), 2013
FMRKT_WICCASH13	Farmers' markets that report accepting WIC Cash, 2013
PCT_FMRKT_WICCASH13	Farmers' markets that report accepting WIC Cash (%), 2013
FMRKT_SFMNP13	Farmers' markets that report accepting SFMNP, 2013
PCT_FMRKT_SFMNP13	Farmers' markets that report accepting SFMNP (%), 2013
FMRKT_FRVEG13	Farmers' markets that report selling fruit & vegetables, 2013
PCT_FRMKT_FRVEG13	Farmers' markets that report selling fruit & vegetables (%), 2013
FMRKT_ANMLPROD13	Farmers' markets that report selling animal products, 2013
PCT_FRMKT_ANMLPROD13	Farmers' markets that report selling animal products (%), 2013
FMRKT_OTHER13	Farmers' markets that report selling other products, 2013
PCT_FMRKT_OTHER13	Farmers' markets that report selling other products (%), 2013
VEG_FARMS07	Vegetable farms, 2007
VEG_ACRES07	Vegetable acres harvested, 2007

Continued on next page

TABLE 1 – continued from previous page

Variable Code	Variable Name
VEG_ACRESPTH07	Vegetable acres harvested/1,000 pop, 2007
FRESHVEG_FARMS07	Farms with vegetables harvested for fresh market, 2007
FRESHVEG_ACRES07	Vegetable acres harvested for fresh market, 2007
FRESHVEG_ACRESPTH07	Vegetable acres harvested for fresh market/1,000 pop, 2007
ORCHARD_FARMS07	Orchard farms, 2007
ORCHARD_ACRES07	Orchard acres, 2007
ORCHARD_ACRESPTH07	Orchard acres/1,000 pop, 2007
BERRY_FARMS07	Berry farms, 2007
BERRY_ACRES07	Berry acres, 2007
BERRY_ACRESPTH07	Berry acres/1,000 pop, 2007
GHVEG_FARMS07	Greenhouse vegetable and fresh herb farms, 2007
GHVEG_SQFT07	Greenhouse veg and fresh herb sq feet, 2007
GHVEG_SQFTPTH07	Greenhouse veg and fresh herb sq feet/1,000 pop, 2007
SLHOUSE07	Small slaughterhouse facilities, 2007
FOODHUB12	Food hubs, 2012
CSA07	CSA farms, 2007
AGRITRSM_OPS07	Agritourism operations
AGRITRSM_RCT07	Agritourism receipts
FARM_TO_SCHOOL	Farm to school program, 2009
PCT_DIABETES_ADULTS09	Adult diabetes rate, 2009
PCT_DIABETES_ADULTS10	Adult diabetes rate, 2010
PCT_OBESE_ADULTS09	Adult obesity rate (county), 2009
PCT_OBESE_ADULTS10	Adult obesity rate (county), 2010
PCT_OBESE_ADULTS13	Adult obesity rate, 2013*
PCT_OBESE_CHILD08	Low-income preschool obesity rate, 2006-08
PCT_OBESE_CHILD11	Low-income preschool obesity rate, 2009-11
PCH_OBESE_CHILD_08_11	Low-income preschool obesity rate (% change), 2006-08 to 2009-11
PCT_HSPA09	High schoolers physically active (%), 2009*
RECFAC07	Recreation & fitness facilities, 2007
RECFAC12	Recreation & fitness facilities, 2012
PCH_RECFAC_07_12	Recreation & fitness facilities (% change), 2007-12
RECFACPTH07	Recreation & fitness facilities/1,000 pop, 2007
PCT_FREE_LUNCH10	Students eligible for free lunch (%), 2010
PCH_RECFACPTH_07_12	Recreation & fitness facilities/1,000 pop (% change), 2007-12
PC_SNAPBEN08	SNAP benefits per capita, 2008
PCT_NHWHITE10	% White, 2010
PCT_NHBLACK10	% Black, 2010
PCT_HISP10	% Hispanic, 2010
PCT_NHASIAN10	% Asian, 2010
PCT_NHNA10	% American Indian or Alaska Native, 2010
PCT_NHPI10	% Hawaiian or Pacific Islander, 2010
PCT_65OLDER10	% Population 65 years or older, 2010
PCT_18YOUNGER10	% Population under age 18, 2010
MEDHHINC10	Median household income, 2010
POVRATE10	Poverty rate, 2010
PERPOV10	Persistent-poverty counties, 2010
CHILDPOVRATE10	Child poverty rate, 2010
PERCHLDPOV10	Persistent-child-poverty counties, 2010
METRO13	Metro/nonmetro counties, 2010
POPLOSS00	Population-loss counties, 2000