Fast Food Availability and Its Impact on Obesity

William B. Russell

ECMT 463

Professor Haeshin Hwang

30 April 2019

Abstract:

This paper questions whether obesity is impacted by fast food per capita. Data is taken from all fifty states and analyzed in comparison with several regressors derived from multiple government and business databases . With this research, I found significant variables that help explain the incidence of obesity and I have also found that differences in obesity rates have higher incidence among certain ethnic and racial groups.

I. Introduction

The question this paper seeks to answer is: Is Obesity impacted by the availability of fast food in an area (i.e. fast food restaurants per capita)? While obesity does not appear to be an economic issue on the surface, there are many economic issues that occur as a result of obesity. There are many costs associated with obesity, including both personal and societal costs. According to the U.S. Center for Disease Control's "Adult Obesity Causes and Consequences", the medical care costs of obesity in 2008 alone were estimated at $147 billion. An indirect cost to the economy that is not often considered is obesity-related absenteeism from the workplace, which the CDC estimates cost somewhere between $3.38 billion and $6.38 billion of possible nationwide economic production in 2008. Even in case where someone with obesity does continue to participate in the workforce, they are more likely to decrease in productivity compared to healthier employees. Along with these, there are many more direct and indirect costs associated with obesity that are negatively affecting the economy in a significant way. It is for this reason, that this paper seeks to focus on whether or not the availability of fast food restaurants in an area significantly impact obesity rates in given areas, as well as provide some applicable solutions to the continuously growing problem of obesity.

Economic theory indicates that there are several contributing reasons as to why obesity is such a hard issue to combat. The idea of opportunity costs plays majorly into an individual's decision to eat an unhealthy meal. Why? Because fast food is exactly that, fast. Not only is it fast, but it's also incredibly affordable. Our decisions naturally come from weighing the opportunity costs of situations, consider going to buy groceries: you drive to the store to buy groceries, you begin searching the store for those the specific items you're looking for, you then pay a sizeable price for each individual item, then you go home and the meal preparation takes 10-20 minutes. That is a lot of time and money spent that you could have been spending doing other, more enjoyable activities. In this case a 5 minute drive to the nearest fast food drive thru seems like a pretty tantalizing option. Along with opportunity costs, one of the biggest fundamentals of economics is the law of demand. The lower the cost of an item, the more likely you are to buy more of it.

The hypothesis of this paper is that if the fast food restaurants per capita increase, then the rates of obesity will also increase. To test this hypothesis, the null hypothesis will be that an increase in a state's fast food restaurants per capita will have no effect on obesity rates. We will be testing these hypotheses using Operation of Least Squares estimations as well as using hypothesis tests.

There has actually been quite a lot of research on obesity statistics, the contributing factors to obesity, and the costs that obesity has to ourselves and society alike. The article "Societal and Personal Costs of Obesity" discusses the underlying economic issues seen with the large growth of the incidence rate of overweight adults. For example, Dr. Jacob Seidell states that there are three levels of economic costs of obesity: direct costs, societal or indirect costs, and personal costs. His data determines that while obesity may seem to only affect the individual, it is also indirectly affecting the economy. Seidell concludes that, "overweight and obesity are likely to be major contributors to the total healthcare costs in affluent societies" (1998). And it seems that he was right.

The article "Social Dynamics of Obesity" discusses economic elements that could be congruent with the rapidly increrasing rate of overweight Americans. These include "falling food prices, the increased convenience of obtaining ready-to-eat food, and reductions in physically demanding labor at work (Burke & Heiland, 2006). The overall findings by the authors suggest that the levels of obesity should reach a stopping point and no longer increase. Burke and Heiland suggest the most impactful method for deterring Americans from extreme weight gain is, "better public education as well as better medical counseling about the relationship between body weight, body composition, and calorie-burning" (2006).

Finally, one piece of literature was actually quite similar to some aspects of this paper, "Weight Status and Restaurant Availability: A Multilevel Analysis" written by Neil K. Mehta and Virginia W. Chang. Their paper focused on the "restaurant environment" and its impact on factors contributing to obesity. They too hypothesized that the greater availability of fast food restaurants would increase the incidence of obesity (2009).

Now, even though this paper and the one from Mehta and Chang both seek to prove the same general hypothesis, they have many stark differences that make them quite unique from each other. Their data is derived from specific counties across America, whereas the data in this paper comes from all available reports from all 50 states. Their paper also attempted to focus on the distribution of full-service restaurants as well, finding them to be indicative of a healthy way to eat. Although full-service restaurants are more than likely healthier than fast food options, I do not believe this is justification for claiming full-service restaurants as a healthy way to eat. This paper will instead, solely focus on the distribution of fast-food restaurants, and this variable's impact on obesity rates by state. Along with this focus, this paper analyzes several explanatory variables that do not seem to have previously been associated with economic research.

The vast majority of the data used in this paper was accessed from official government databases or organizations compiling data from official government databases. There are many regressors being used in order to get the most accurate representation of how much impact the variable we are testing, fast food restaurants per capita, has on obesity rates measured over all fifty states. The data is measured across state populations, and specific subsets of those populations as well.

The methods of data analysis in this paper include operation of least squares, z-tests, and several regression models tested over various explanatory variables and categorical/dummy variables. The results of this study find that there is a positive relationship between the number of fast food restaurants per capita and obesity rate that is statistically significant, as well as statistically significant differences in obesity rates for different racial and ethnic groups. Through this paper, I will be discussing the general model and variables used in the data, the empirical results calculated from regression outputs, and finally what conclusions can be drawn from the academic addition from this paper.

II. Model Specification

In this paper I will be using three models for the purpose of determining whether or not there are any significant changes in the regression function when specific subsets of the population are being measured. Before discussing the models in detail, the variables abbreviations and interpretations are listed below:

Table 1: Variable Interpretations

Variable Name:	**Interpretation**
obesrate	The obesity rate. The percentage of a population that is classified as obese.
ffpcap	Fast food restaurants per capita. How many fast food restaurants there are in a particular area for every 10,000 people in a population.
inactiv	Rate of physical inactivity. This is a percentage of the total population that has not engaged in any physical exercise outside of their daily work schedules for at least the past 30 days.
livpov	Percentage of people living in poverty. This is the percentage of the total population that lives at or below the income level associated with poverty.
fem	Female population observation. Fem = 1 if the observed population is female and fem = 0 if the observed population is male.
black	Black population observation. Black = 1 if the observed population is categorized as black, and black = 0 otherwise (white or latino).
latino	Latino population observation. Latino = 1 if the observed population is categorized as latino, and latino = 0 otherwise (white or black).

Model 1: *The impact of fast food restaurants per capita on obesity rate:*

$$obesrate = \beta_0 + \beta_1(ffpcap) + \beta_2(inactiv) + \beta_3(livpov) + \varepsilon$$

Summary Statistics:

Variable	Obs	Mean	Std. Dev.	Min	Max
obesrate	50	30.752	3.733425	22.6	38.1
ffpcap	50	3.966	.9896629	1.9	6.3
inactiv	50	26.578	3.886202	19.2	34.4
livpov	50	12.232	2.903653	6.5	20.8

Model 2: *The impact of fast food restaurants per capita on obesity rate, including gender differences:*

$$obesrate = \beta_0 + \beta_1(ffpcap) + \beta_2(inactiv) + \beta_3(livpov) + \beta_4(fem) + \varepsilon$$

Summary Statistics:

Variable	Obs	Mean	Std. Dev.	Min	Max
obesrate	100	30.562	4.144164	18.4	39.4
ffpcap	100	3.966	.9846519	1.9	6.3
inactiv	100	26.578	3.866525	19.2	34.4
livpov	100	12.232	2.888951	6.5	20.8
fem	100	.5	.5025189	0	1

*Where fem = 0 if male, fem = 1 if female

<u>Model 3:</u> <u>*The impact of fast food restaurants per capita on obesity rate, including race/ethnicity differences:*</u>

$$obesrate = \beta_0 + \beta_1(ffpcap) + \beta_2(inactiv) + \beta_3(livpov) + \beta_4(black) + \beta_5(latino) + \varepsilon$$

<u>Summary Statistics:</u>

Variable	Obs	Mean	Std. Dev.	Min	Max
obesrate	138	32.10725	5.443467	17.5	46.4
ffpcap	138	3.963043	1.023168	1.9	6.3
inactiv	138	26.59783	3.988935	19.2	34.4
livpov	138	12.27826	3.002547	6.5	20.8
black	138	.3333333	.4731218	0	1
latino	138	.3333333	.4731218	0	1

*where black= 1 if black, 0 for anything else
*where latino =1 if latino, 0 for anything else

The specific hypothesis that I am going to test is that fast food restaurants per capita have an impact on the obesity rate of any given state. The assumed hypothesis is that there is that a change in fast food restaurants per capita has no effect on obesity rate. This is represented by the null hypothesis below. My proposed hypothesis is that the effect of fast food restaurants per capita on obesity rate is not zero. This is represented by the alternative hypothesis below.

Null hypothesis: $H_0 : \beta_1 = 0$
Alternative hypothesis: $H_1 : \beta_1 > 0$

The implications of the null hypothesis would be that the fast food restaurants per capita have no impact on changing the obesity rate, so the cause or at least a large contributing factor of obesity is still unknown to us, and we must look towards other factors in order to find a connection and eventually a solution. The implication of the alternative hypothesis is that fast food restaurants per capita do have a contributing effect on obesity, and depending on how significant this effect is, we may be able to use this knowledge to work towards a solution to the obesity epidemic.

As for the variables that make up these models, the variable being tested, the dependent variable in this equation, is *Obesity Rate* (obesrate). Obesity Rate is measured by the percentage of the total population of each state that was considered "obese" in the year 2017. This information was accessed from stateofobesity.org, an organization dedicated to educating the public about obesity, as well as advocating for and proposing better policies to answer the obesity problem the United States faces. All of the data information I accessed through them was originally compiled from major surveys including: the National Health and Nutrition Examination Survey and the Behavioral Risk Factor Surveillance System.

Other variable data I accessed through stateofobesity.org was information on obesity rates by gender for different states, obesity rates by race/ethnicity for different states, and the physical inactivity rate for different states. These variables were also measured during 2017. The physical inactivity rate is an explanatory variable which represents the percentage of inactivity among adults for any given population. Inactivity in this case was defined as adults who had not engaged in physical activity or exercise during the previous 30 days other than activity required for their regular job. The physical inactivity rate was also specifically taken from the Behavioral Risk Factor Surveillance System survey.

The variable being manipulated, the independent variable in this equation, was the number of *Fast Food Restaurants per capita*. This variable represents how many fast food restaurants there were per 10,000 people in a population of any measured state. This data was gathered from datainfiniti, a company dedicated to gathering business, consumer, product, and property information that they then compile into an accessible database. Datainfiniti sells access to their database, so I was able to get this information through an article published by "Thrillest.com", in which they list out the statistics originally gathered by datainfiniti.

The other explanatory variable in this equation were the percentage of people living in poverty. The percentage of people living in poverty variable data was accessed from the United States Census Bureau's database, and represents the percentage of the population living with an income level below the poverty line. This data was taken from the Census Bureau's data most recently published data regarding poverty, which was in 2017.

Lastly some categorical variables and dummy variables are taken into account. Gender and Race/Ethnicity are both variables used to differentiate between specific subsets of the overall population. These variables are being used to show whether or not certain groups of people are more affected by obesity than others. As for how the categories of each variable will be represented, when you have *k* possible options for a variable, you can only have *k-1* regressors in the equation. In the case for the gender variable, it is a dummy variable that is binary, with options Male or Female (yes or no). If the data reflects male population data, the X for the gender variable will be equal to 0, whereas if it reflects a female, the variable will be equal to 1. By setting the Male population to result in a 0, we are leaving the male category out of the equation/calculation in order to set a base for the female category to compare to. The race and ethnicity variable is made up of three categories: Black (black is 1, it is 0 otherwise), Latino

(latino is 1, it is 0 otherwise), and White. The white category is purposefully left out of the equation to set a base for the other two categories.

Other variables that were originally calculated as part of the regression were explanatory variables: percentage of population diagnosed with diabetes, high school graduation rate, and high school graduation rate for low-income students, and the categorical variable: Region. Region was divided into 5 categories: West, Midwest, Northeast, Southwest, and Southeast. The reason these variables are no longer being calculated as part of the regression because the estimated coefficients were consistently statistically insignificant in all cases, and many times their inclusion actually made the adjusted R-squared worse.

III. Empirical Results

Model 1: *The impact of fast food restaurants per capita on obesity rate:*

Regression Results:

Source	SS	df	MS		
				Number of obs	= 50
				F(3, 46)	= 18.23
Model	370.937561	3	123.645854	**Prob > F**	= 0.0000
Residual	312.047239	46	6.78363564	**R-squared**	= 0.5431
				Adj R-squared	= 0.5133
Total	682.9848	49	13.9384653	**Root MSE**	= 2.6045

obesrate	Coef.	Std. Err.	t	P>t	[95% Conf.Interval]	
ffpcap	1.002041	.4165097	2.41	0.020	.16365	1.840431
inactiv	.37888	.1238219	3.06	0.004	.1296393	.6281206
livpov	.3376915	.1564505	2.16	0.036	.022773	.65261
_cons	12.57739	2.612127	4.82	0.000	7.319455	17.83533

Predicted value of obesrate:

$$\widehat{obesrate} = 12.577 + 1.002(\widehat{ffpcap}) + .379(\widehat{inactiv}) + .338(\widehat{livpov}) + \hat{\varepsilon}$$

Model Interpretations:

Every coefficient in this case is statistically significant as indicated by the p-values. Both ffpcap and livpov are statistically significant at a 95% confidence level (due to their p-values being below 5%, but above 1%), while inactiv and the constant for this regression are both statistically significant at a 99% confidence level.

Model 2: *The impact of fast food restaurants per capita on obesity rate, including gender differences:*

<u>Regression Results:</u>

Source	**SS**	**df**	**MS**		
				Number of obs	= 100
				F(4, 95)	= 17.41
Model	719.213127	4	179.803282	**Prob > F**	= 0.0000
Residual	981.022473	95	10.3265523	**R-squared**	= 0.4230
				Adj R-squared	= 0.3987
Total	1700.2356	99	17.174097	**Root MSE**	= 3.2135

obesrate	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.Interval]**	
Ffpcap	1.061417	.3633763	2.92	0.004	.3400237	1.78281
Inactiv	.265295	.1080262	2.46	0.016	.0508359	.4797541
Livpov	.4631756	.1364924	3.39	0.001	.1922041	.7341472
fem	-.4	.6426991	-0.62	0.535	-1.675919	.875919
_cons	13.83585	2.301448	6.01	0.000	9.266895	18.4048

Predicted value of obesrate:

$$\widehat{obesrate} = 13.836+1.061(\widehat{ffpcap})+.265(\widehat{inactiv})+.463(\widehat{livpov})-0.4(\widehat{fem})+\hat{\varepsilon}$$

Model Interpretations:

With this model, every coefficient is statistically significant except for “fem”. The estimated coefficient for “fem” is -0.4, which means we could expect a drop in the percentage of obese people within a given population by 0.4 whenever “fem” was equal to 1, i.e. whenever the model were predicting obesity rate and the gender of the person being measure were female. However, the p-value for “fem” is over 50%, when we really want p-values of less than 5% in order to be confident about our estimates. “Fem”s p-value indicates that there is a 53.5% chance that the estimated coefficient for “fem” occurred due to chance, rather than due to relation to the model, and for this reason, gender cannot be proved to have any significant effect on obesity. We must omit “fem” from this regression equation. In a new prediction equation, every coefficient stays the same, except for the constant, which decreases to 13.63585.

Model 3: *The impact of fast food restaurants per capita on obesity rate, including race/ethnicity differences:*

Regression Results:

Source	**SS**	**df**	**MS**			
				Number of obs	=	138
				F(5, 132)	=	20.40
Model	1769.43181	5	353.886363	**Prob > F**	=	0.0000
Residual	2290.06094	132	17.3489465	**R-squared**	=	0.4359
				Adj R-squared	=	0.4145
Total	4059.49275	137	29.631334	**Root MSE**	=	4.1652

obesrate	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.Interval]**	
ffpcap	1.0096	.3755508	2.69	0.008	.2667238	1.752477
Inactiv	.2495086	.1144984	2.18	0.031	.0230196	.4759977
livpov	.2326746	.1447784	1.61	0.110	-.0537114	.5190605
black	6.993478	.8685056	8.05	0.000	5.275488	8.711468
latino	2.432609	.8685056	2.80	0.006	.7146187	4.150599
_cons	15.4709	2.508257	6.17	0.000	10.50932	20.43248

Predicted value of obesrate:

$$\widehat{obesrate} = 15.471+1.01(\widehat{ffpcap})+.25(\widehat{inactiv})+.233(\widehat{livpov})+6.993(\widehat{black})+2.433(\widehat{latino}) + \widehat{\varepsilon}$$

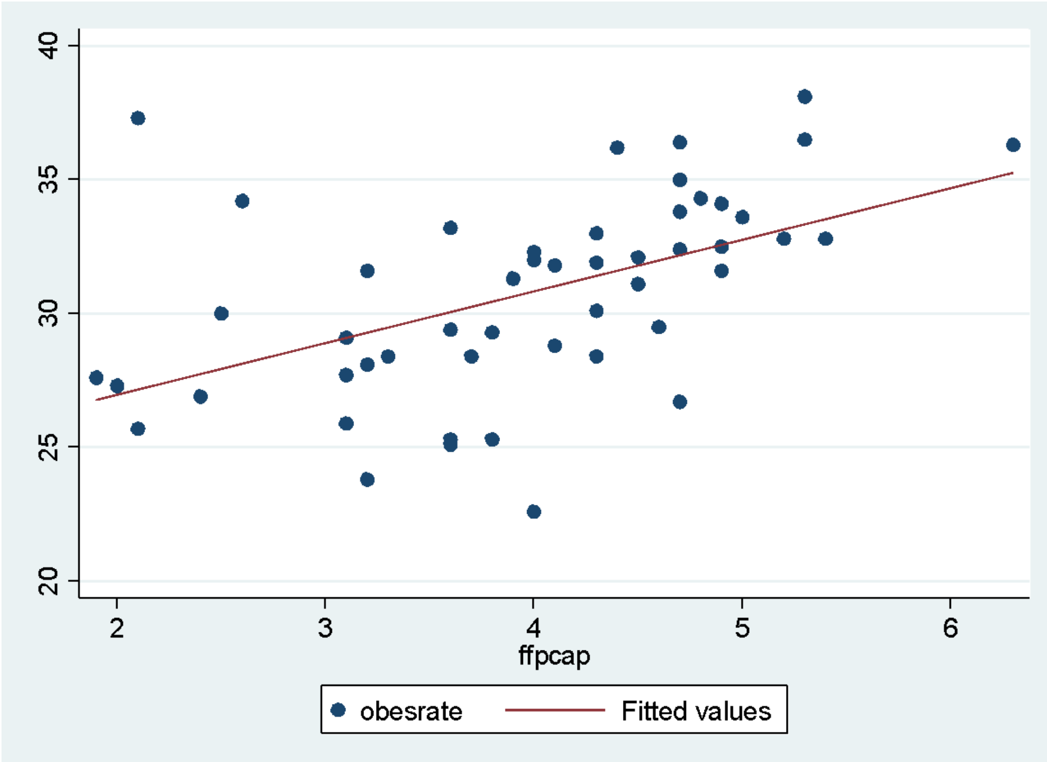
Model Interpretations:

All of the coefficients are statistically significant except for "livpov" now. The p-value for "livpov" is above 10%, which does not give us enough confidence in our estimate. Even more important, estimated confidence interval for "livpov" has a negative minimum and a positive maximum. Any confidence interval that contains 0 runs the risk of a coefficient being 0, when we believe there to be an effect. For these reasons, we must omit "livpov" from this regression. The new prediction equation is represented by:

$$\widehat{obesrate} = 15.63808+1.020462(\widehat{ffpcap})+.3490137(\widehat{inactiv})+6.993478(\widehat{black})+2.432609(\widehat{latino})$$

The estimated coefficients of this new regression equation are statistically significant to the 99% confidence level, whereas the original had all but two coefficients being significant to the 99% Confidence level. Overall, this regression shows that race and ethnicity may play important roles in determining risks for obesity. The coefficient for black = 6.993478, for example, indicates that a black person is 6.99% more likely to be obese than a white person.

<u>Figure 1: Fitted Regression Line Y=obesity rate X=fast food per capita</u>



This figure demonstrates the positive correlation between obesity rates and fast food restaurants per capita.

<u>Table 2: Regression Results</u>

Dep Var: Obesity Rate (in %)	**(1)**	**(2)**	**(3)**
Intercept	12.577	13.835	15.47
	(2.61)**	(2.301)**	(2.51)**
Ffpcap	1.002	1.061	1.01
	(.417)*	(.363)**	(.375)**
Inactiv (%)	.378	.265	.25
	(.123)**	(.108)*	(.114)*
Livpov (%)	.338	.463	.233

	(.156)*	(.136)**	(.145)
Female		-.4 (.643)	
Black			6.993 (.869)**
Latino			2.433 (.869)**
R-Squared	0.54	0.42	0.44
Adj. R-squared	0.51	0.4	0.41

Because the sample sizes were greater than 30 in all of the models cases, I will be using z-statistics distributed normally in order to test the hypotheses. We will be looking at a right facing z-table in order to take into account that the alternative hypothesis is that beta1 with be greater than 0 (not negative). The alpha level being used will be 0.05 (95% confidence), and since we're only looking on the right side of the distribution, we need to find the z-score associated with 50%(right side of z-distribution) minus 5%(alpha), 45%. This z-score is 1.645. The calculated z-statistic 28.04, which is much larger than the z-score 1.645. The fact that the calculated score is larger indicates that the null hypothesis: $H_0 : \beta_1 = 0$, falls into the rejection region, and the null hypothesis is therefore rejected.

IV. Conclusion

The central model of this paper is:

$$obesrate = \beta_0 + \beta_1(ffpcap) + \beta_2(inactiv) + \beta_3(livpov) + \varepsilon$$

And the central obesity rate estimation model of this paper was calculated as:

$$\widehat{obesrate} = 12.577 + 1.002(\widehat{ffpcap}) + .379(\widehat{inactiv}) + .338(\widehat{livpov}) + \widehat{\varepsilon}$$

This paper’s models all show that when fast food restaurants per capita increase by one, the obesity rate is also estimated to increase by about one (obesity rates are documented in percentage). Although that doesn’t seem like much, increasing/decreasing obesity rates by about one is the highest impact that any of the regressors have on obesity rate.

The motivation behind this paper is to address some of the contributing factors to obesity, demonstrating their actual impact, and theorizing possible solutions to what has become widely regarded as an epidemic. The percentage of the population classified as obese has been continually increasing over the years, and obesity results in billions of dollars of indirect costs to the economy. Whether it’s due to absenteeism, or All three of the models examined in this paper demonstrated that fast food restaurants per capita do indeed have an impact on obesity rates in a statistically significant way. As a result of this further found impact on obesity rates, future studies should focus on how to significantly change contributing factors of obesity, especially non-genetic risks like fast food restaurants per capita. The wide availability of of fast food restaurants that provide cheaper meals in a quicker time, is a very large contributing factor to obesity risk and local government policies should be considered and tested in order to lessen these impacts on populations.

V. References:

Adult Obesity Causes & Consequences | Overweight & Obesity | CDC. (n.d.). Retrieved April 30, 2019, from https://www.cdc.gov/obesity/adult/causes.html

Burke, M., & Heiland, F. (2007, February 9). Social Dynamics of Obesity. Retrieved April 30, 2019, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1465-7295.2007.00025.x

Mehta, N. K., & Chang, V. W. (2008, February). Weight status and restaurant availability a multilevel analysis. Retrieved April 30, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2440344/

Seidell, J. C. (1998). Societal and personal costs of obesity. Retrieved April 30, 2019, from https://www.ncbi.nlm.nih.gov/pubmed/9792474

State Briefs. (2017). Retrieved April 30, 2019, from https://www.stateofobesity.org/states/

US Census Bureau. (2018, August 16). Poverty Status: POV-46. Retrieved April 30, 2019, from https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pov/pov-46.html