STATE OBESITY RATES: AN ECONOMIC EXPLANATION

Kathryn McKeon, Siena College

ABSTRACT

Obesity is an issue of increasing importance within the United States, as it has serious health, social, and economic implications. This empirical paper attempts to explain the variations per state in obesity rates, with a focus on economic indicators. Not only does it look at different health issues associated with obesity but also attempts to determine the relationship between obesity and economic factors.. Using the method of Ordinary Least Squares, the relationship between state obesity rates and five quantitative variables is tested. These variables are: exercise, smoking, nutrition, income, and annual per person spend on fast food. The model also contains qualitative variables to explain state location. Dummy variables were set up for the 8 different regions in the United States: New England, Mid-East, Great Lake, Plains, South East, South West, Rocky Mountain, Far West. Analysis concluded that the variables exercise; spend per person on fast food, Great Lakes, Plains, South West, Rocky Mountain, and Far West. Smoking, nutrition, income, Mid-East, and South East were found to be significant in explaining state obesity rates. The overall model explains 71.9% of the variation in obesity amongst states.

INTRODUCTION

Obesity is an issue of increasing importance within the United States, as it has serious health, social, and economic implications. As stated in a report by the United States Department of Agriculture (2004), since the mid-1970s obesity has been a rapidly growing trend and today, nearly two out of three adults are either overweight or obese. While it is widely understood that the cause of obesity is a result of one's own choices, the purpose of this paper is to determine if economic factors impact a person's decision making process, leading them towards a life of obesity. As stated in the at the Economics of Obesity Workshop, "…economics, as a discipline that studies how individuals use limited resources to attain alternative ends, can provide unique insight into the actions and forces that cause individuals to gain excessive weight" (U.S. Department of Agriculture, 2004).

One of the leading forces in increased weight gain has been technology. Not only have there been innovations in the growth and production of food, technology has also created an increasingly sedentary lifestyle. People no longer have to exert the effort they once did to complete simple tasks. Eric Finkelstein, author of *The Fattening of America* in an interview with the New York Times (2008), gave an example from her own life, "I recently had an inexpensive printer installed in my office. So now I don't even have to walk the 100 steps to the community printer down the hall a few times a day." This loss of exercise coupled with cheap high caloric prepared foods are at the heart of the United States obesity problem. The model of this paper will test the significance between exercise and obesity, believing the relationship between the two to be inversely related. To test the theory of technological advances on the production of high fat, cheap, and quick food, the model will also look at the relationship linking obesity to per capita spend in fast food restaurants per state. It is the hypothesis of this paper that the relationship will be positively correlated, meaning: the higher the per capita fast food spend, the more obese people living in the state.

Smoking has often been considered an appetite suppressant, and is assumed to have an inverse effect on obesity, as smoking can be a substitute for eating. Nicotine, as tested in 2004 by Jessen, Bueman, Toubro, Skovgaard, and Astrup, is found to be negatively associated with hunger and potential food consumption and positively associated with fullness (p. 327). In a study by Cho, Grossman, and Saffer in 2004, it was discovered that as the real price of cigarettes goes up (therefore causing their quantity demanded go down) obesity increases significantly (p. 565). With increased taxes and new laws prohibiting the smoking of cigarettes indoors it is likely that because people are smoking less they are eating more. This model will also test for the relationship between a states number of smokers and obesity, believing the more smokers the less obesity.

Furthermore, the model will take into account how nutritious a person's diet is. This will be measure by the fruit and vegetable consumption per state. It is the hypothesis of this paper that obesity will be negatively related to the amount of healthy food a person is consuming. Also, since produce is more expensive than highly processed, fattening, foods, the model will look at state median per capita income. Someone who has limited funds is more likely to purchase cheaper foods that are more filling. Unfortunately, these cheaper mass-produced substitutes lack the vitamins and minerals contained in fruits and vegetables. In a study conducted by Gelbach, Klick, and Stratmann (2007) the effect of changes in relative food prices on individual behavior are analyzed. The results show that the change in price of healthy foods and unhealthy foods is directly related to the rise in obesity in the United States (p. 1). It is hypothesized that the relationship between income and obesity will be a negative one. People with higher incomes have the means to purchase food that is high in nutrients. It is under this assumption that because people have the incomes to buy healthy food they will be less obese.

Lastly, the regression model will look at the different U.S. regions. Using dummy variables, the United States will be divided into 8 areas to compare obesity rates by area. The reference region for this study will be New England. The reason for choosing New England is due to its recent success in a CalorieLab study, entitled *Mississippi is the Fattest State* (2008), showing it to be the healthiest place in which to live in the U.S.

REGRESSION MODEL AND HYPOTHESIS

$$\textbf{Obesity Rate}=\beta_1 + \beta_2\textbf{Excercise}+\beta_3\textbf{FastFood} + \beta_4\textbf{Smoker} + \beta_5\textbf{Nutrition} + \beta_6\textbf{Income} + D_1\textbf{MidEast} + D_2\textbf{GreatLakes} + D_3\textbf{Plains} + D_4\textbf{SouthEast} + D_5\textbf{SouthWest} + D_6\textbf{RockyMountains} + D_7\textbf{FarWest} + \mu_i$$

Variable	Definition	Hypothesized Slope	Source
Obesity Rate	% of adults per state with a Body Mass Index of 30kg/m^2 or higher 2007	-	Center for Disease Control and Prevention, Behavioral Risk Factor Surveillance System Survey Data
Exercise	% of Adults per state participating in moderate physical activity five or more days per week, or vigorous physical activity for 20+ minutes three or more days per week 2007	Negative	Center for Disease Control and Prevention, Behavioral Risk Factor Surveillance System Survey Data
Fast Food	Per capita adult spend on fast food per state. 2002 data adjust for 2007 prices.	Positive	Economic Census population data and limited-service eating place data found in Geographic Area Series
Smoker	% of adults per state that smoke cigarettes everyday 2007	Negative	Center for Disease Control and Prevention, Behavioral Risk Factor Surveillance System Survey Data
Nutrition	% adults per state eating fruits and vegetables 5 or more times a day 2007.	Negative	Center for Disease Control and Prevention, Behavioral Risk Factor Surveillance System Survey Data
Income	Per capita median income per state. 2 year average 2006-2007	Negative	U.S. Census Bureau- People and Households Income Status by State
Mid-East	Dummy Variable for U.S. region. Delaware, Maryland, New Jersey, New York, Pennsylvania	Unsure	-
Great Lakes	Dummy Variable for U.S. region. Illinois, Indiana, Michigan, Ohio,	Unsure	-

	Wisconsin		
Plains	Dummy Variable for U.S. region. Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota	Unsure	-
South East	Dummy Variable for U.S. region. Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia	Unsure	-
South West	Dummy Variable for U.S. region. Arizona, New Mexico, Oklahoma, Texas	Unsure	-
Rocky Mountain	Dummy Variable for U.S. region. Colorado, Idaho, Montana, Utah, Wyoming	Unsure	-
Far West	Dummy Variable for U.S. region. Alaska, California, Nevada, Oregon, Washington	Unsure	-

Regression

A.) This model observes all the assumptions and rules of the OLS Method.

B.) Obesity Rate per State is the dependent variable in the model and was regressed on the independent variables.

MODEL SUMMARY AND OVERALL FIT

Table 1: Descriptive Statistics

	Obesity	Exercise	Fast Food	Smokers	Nutrition	Income
Mean	26.31	49.884	494.954	14.982	23.714	37172.56
Standard Error	0.406	0.691	21.111	0.421	0.497	780.161
Median	26.3	49.2	532.864	14.55	24.3	35566.5
Mode	28.2	48.9	#N/A	15.1	28.5	#N/A
Standard Deviation	2.868	4.887	149.279	2.977	3.517	5516.573
Sample Variance	8.223	23.882	22284.339	8.860	12.369	30432577.11
Kurtosis	-0.274	0.154	4.330	0.867	-1.048	0.884
Skewness	-0.235	-0.206	-1.882	0.516	-0.139	0.974
Range	13.3	22.2	804.670	15.2	13.7	25272
Minimum	19.3	38.6	0	8.3	16.3	28845
Maximum	32.6	60.8	804.670	23.5	30	54117
Count	50	50	50	50	50	50

Model	Multiple R	R Squared	Adjusted R Squared	Standard Error of the Estimates	Significance F
1	0.887704	0.788019	0.719269	1.5194	4.24583E-09

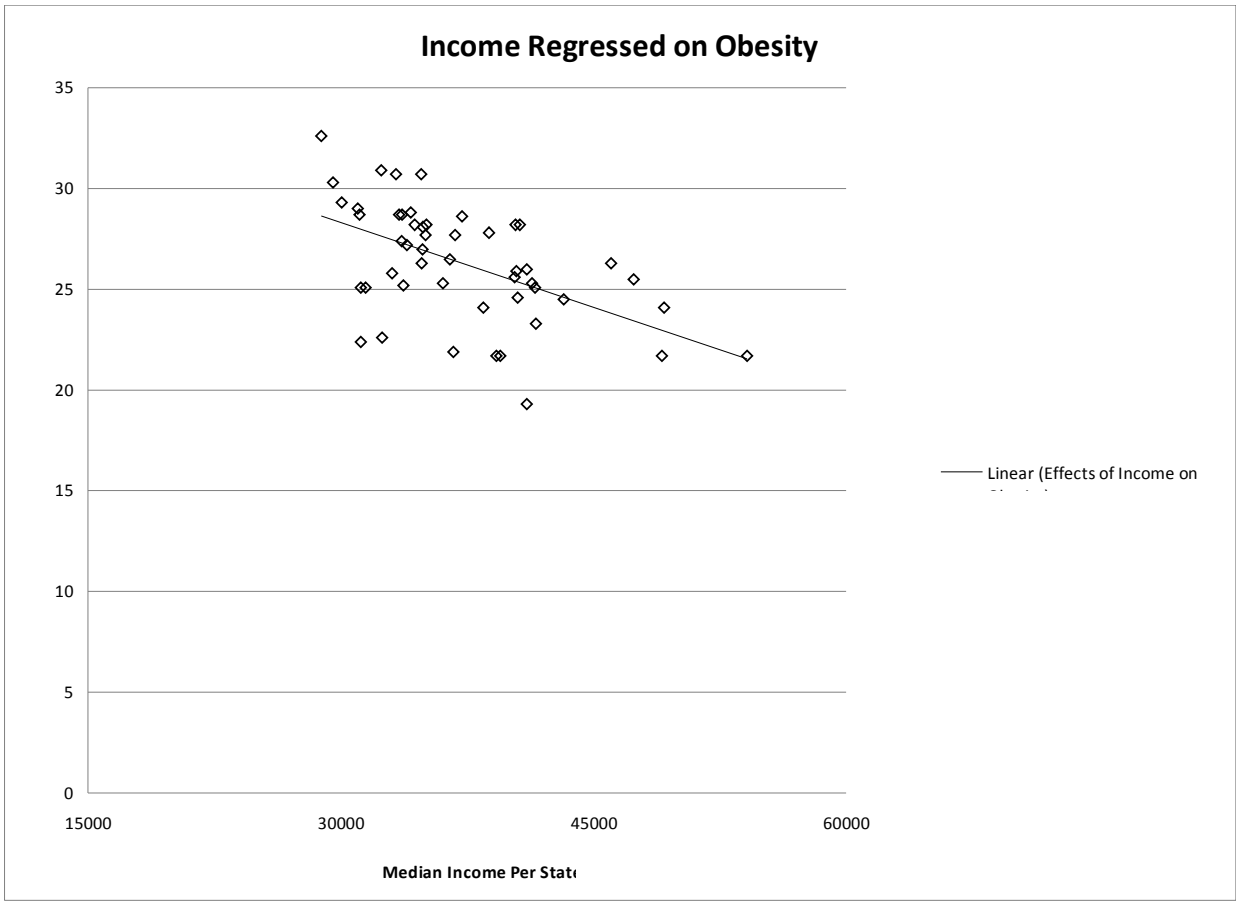
The overall goodness of fit for this model is 71.9%.

Table 2: Results

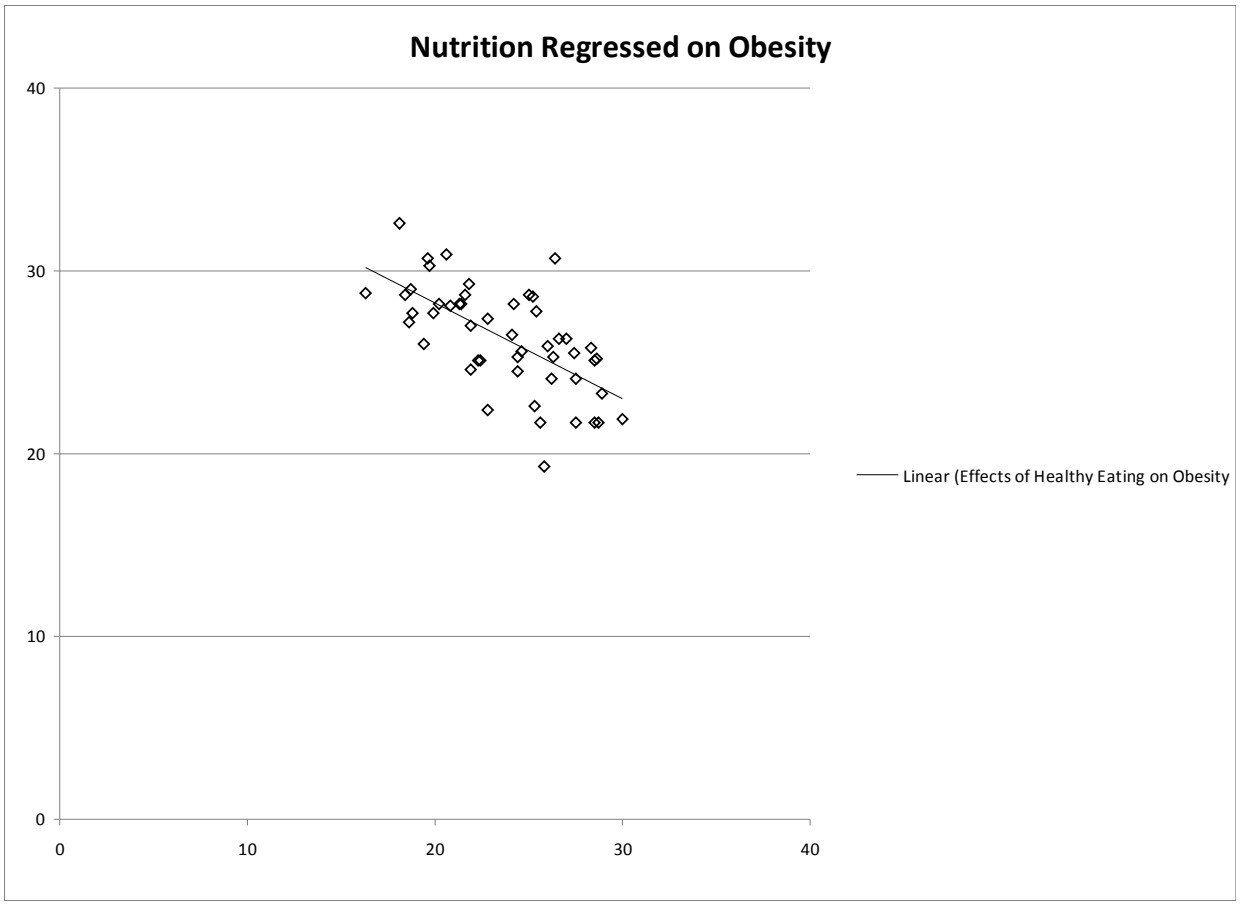
	Coefficients	Standard Error	t Stat	P-value
Intercept	33.67	6.05	5.56	2.48E-06
Exercise	-0.05	0.08	-0.71	0.48
FastFood	2.90-05	0.0017	0.017	0.99
Smokers	0.23	0.10	2.30	0.027
Nutrition	-0.18	0.11	-1.70	0.098
Income	-0.00013	6.06E-05	-2.22	0.033
Mid-East	3.05	1.019	2.99	0.0049
GreatLakes	1.40	1.09	1.29	0.21
Plains	1.46	1.13	1.29	0.20
SouthEast	2.26	1.21	1.89	0.068
SouthWest	1.52	1.19	1.29	0.21
RockyMnts	-1.62	1.12	-1.45	0.16
FarWest	1.41	0.92	1.52	0.14

The regression results show that the variables Smokers, Nutrition, Income, Mid-East, and SouthEast are significant based on their P-value and a 90% confidence interval. This means we can reject the null. These five variables have an impact on the obesity rate.

Partial Regression Graphs



This graph shows the relationship between Income per capita and Obesity. The two have an inverse relationship as hypothesized.



This graph shows the relationship between nutrition and obesity rates. The more people per state that consume the appropriate amount of fruits and vegetables a day, the lower the state obesity rate.

ANALYSIS

Overall, the model explains 71.8% of the variation in the obesity rate per state, which is significant. The variables: Smokers, Nutrition, Income, Mid-East, and SouthEast are significant in explaining the obesity variation by state, since they all have p-values smaller than 0.1. But in most cases the effect these variables have on the obesity rate is very small. The values predicted to have significance that proved insignificant after further analysis were: Exercise, Fast Food, Great Lakes, Plains, South West, Rocky Mountain, Far West. I feel their lack of significance can be explained by lack of available data.

It is no surprise that consuming fruits and vegetables has a significant effect on obesity rates. Someone who is making the conscious choice to eat healthy is most likely not going to be obese. Nutrition has a coefficient of -.181. This means for every one more person in a state eating 5 of ore fruits and vegetables a day, the obesity rate in that state will decrease by -.181. As hypothesized the relationship was negative. Along with this, median income per capita was also significant. As hypothesized it has a negative relationship with obesity rates per state. Its coefficient was -0.0001346.

Smoking on the other hand, while significant, did not affect obesity rates in the way originally hypothesized. The coefficient for smoking was .236. This shows a positive relationship between those who smoke and those who are obese. This may have something to do with a state's overall mindset, perhaps, people in a state that like to consume fattier foods are also not concerned about the risks associated with smoking.

It was surprising that the fast food variable proved to be insignificant. With a p-value of 0.987, the null is not rejected at a 90% confidence interval. The data for this estimator was not as accurate as hoped.[1]

The dummy variables were also interesting. The Mid-East and the South East seem to be much less healthy regions to live in than New England. Living in the Mid-East increases obesity rates 3.048 and in the South East by 2.26.

TESTS FOR VIOLATION OF ASSUMPTIONS

A.) Ramsey and BP test performed in STATA
B.) Residuals were graphed in Excel

Ramsey Reset Test
Null: model has no omitted variables
Prob > F = 0.7359

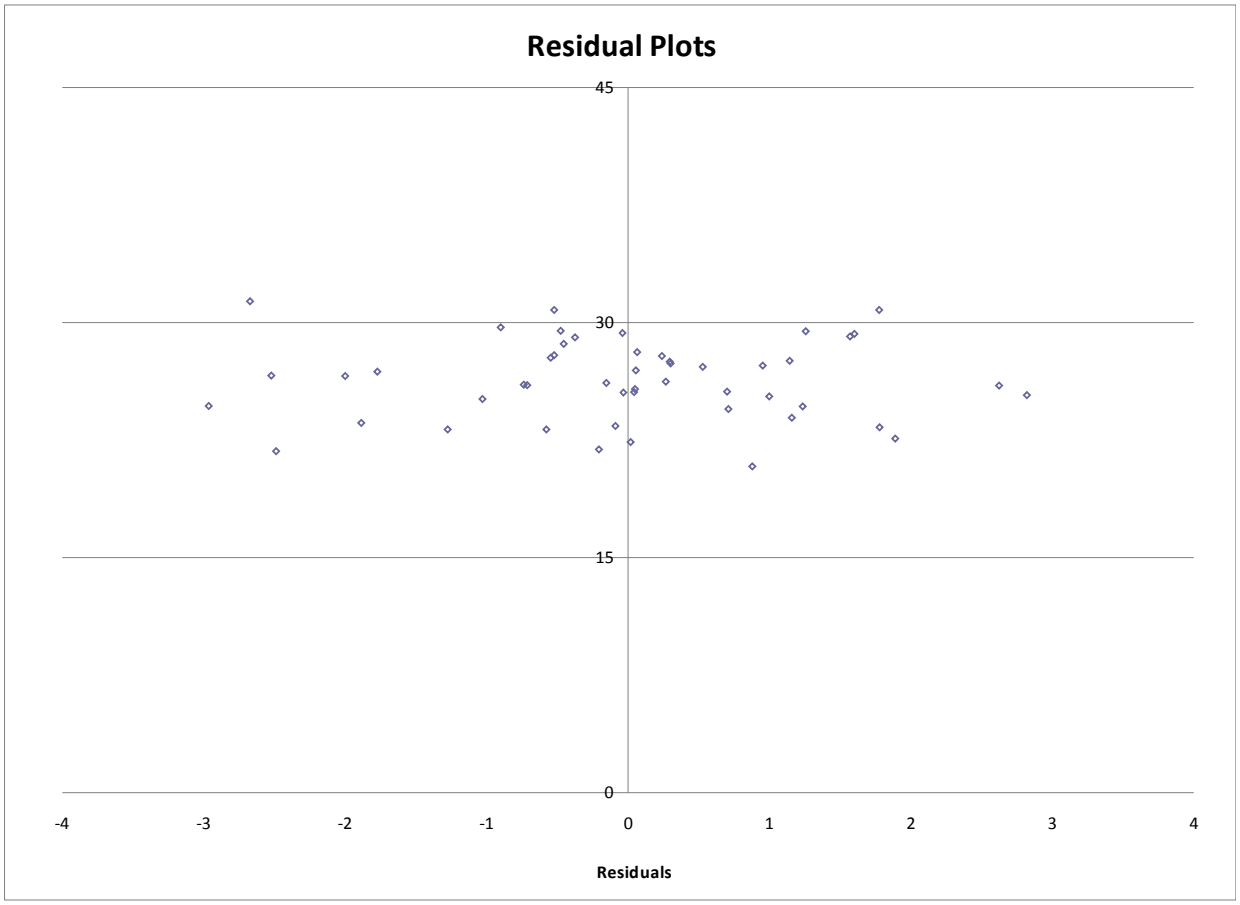
Since the null hypothesis is not rejected at a 90% confidence level, the model passes the Ramsey Test. This means no variables have been left out of the model, it is unbiased.

Breusch-Pagan / Cook-Weisberg Test for Heteroskedasticity
Null: Constant variance
Prob > chi2 = 0.6438

The null hypothesis is not rejected. This means the estimators are BLUE. The test proves homoscedasticity and that the estimators are efficient.

[1] 2007 Economic Census data will not be available until 2009. This forced the use of 2002 numbers.

Autocorrelation



There is no pattern to the plotted residuals proving there is no autocorrelation in the model. This was to be expected as the model was cross-sectional not time series.

CONCLUSION

In conclusion, of the variables that proved to be significant, not all had the relationship with the dependent variable originally hypothesized. The model explained 71.9% of the obesity rate per state but the impacts of these effects were not large. The significant variables in this model are: Smokers, Nutrition, Income, Dummy Mid-East, and Dummy South East. The insignificant variables are: Exercise, Fast Food, Dummy Great Lakes, Dummy Plains, Dummy South West, Dummy Rocky Mountains, Dummy Far West.

If given the opportunity to run the regression again the same variables would be used but different data explaining them. Exercise should have an impact on obesity rates. Given the small amount of information available on this variable the numbers used to estimate state exercise rates may not be the most accurate or best valuation of how many and how much people are exercising. Also, there are variables that are hard to quantify that would most likely have a large impact on obesity. Family member weights could indicate a lack concern on a person's part towards health and obesity. These behavioral estimators can be hard to measure but would probably show to be important in explaining obesity rates. As this was a cross-sectional model, it would be interesting to run another test using time series data.

Lastly, the United States' increasing obesity rates is not a problem to be ignored. With causes like low incomes and effects like increased health costs to the population, obesity and its spread is decreasing the overall well being of the country.

REFERENCES

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

CalorieLab. (2008). Mississippi is the Fattest State. Retrieved from http://calorielab.com/news/2008/07/02/fattest-states-2008/

Gelbach, J., Klick, J., and Stratmann, T., (2007). Cheap Doughnuts and Expensive Broccoli: The Effect of Relative Prices on Obesity. *FSU College of Law, Public Law Research Paper No. 261*, 1.

Finkelstein, E. interview by Melissa Lafsky, M. (2008). Eric Finkelstein Interview, February 8. Retrieved from http://nytimes.com

Jessen, A., Buemann, B., Toubro, S., Skovgaard, I. M., and Astrup, A. (2004). The appetite-suppressant effect of Nicotine is Enhanced by Caffeine. *Journal of Diabetes, Obesity and Metabolism* 7(4), 327.

United States Department of Agriculture, (2004). The Economics of Obesity: A Report on the Workshop Held at USDA's Economic Research Service. Retrieved from http://www.ers.usda.gov/Publications/EFAN04004/