A Spoonful of Sugar Makes Public Health Go Down: An Analysis on the Impact of Soda Taxes on Diabetes and Obesity

Alexandra Baker

November 28th, 2020

Executive summary

Given the national public health crisis in the US with soaring public health obesity cases, interventions aimed at influencing consumers to make healthy food choices have arisen. This can be done via price manipulation through taxes on unhealthy food and subsidies on healthy food. The implementation tests their efficacy in lessening preventable lifestyle diseases such as obesity and diabetes. This study used 2008 and 2018 data from the BRFSS survey by the CDC and ERS survey by the USDA, aggregated to the state level, to examine the impact of soda taxes on diabetes and obesity outcomes. OLS and panel data regressions were conducted. In this study, the soda taxes did not have any impact on the diabetes or obesity outcomes with coefficients close to zero and no statistical significance. The results of this study imply that a soda tax might not be a sufficient intervention to influence consumer choice in the US to limit unhealthy food choices and effect public health.

Background

According to the Center for Disease Control, 40% of deaths, equating to 900,000 people each year in the United States, stem from the top five lifestyle diseases that are largely preventable. (CDC, 2020) It estimates that between 21 to 39% of the deaths caused by the respective diseases could be prevented with lifestyle changes in diet and exercise. (CDC, 2020) Recognizing the impact of such changes, there have been many studies on interventions to try to influence consumer choice in the US. One such line of interventions changes the prices to consumers. The general idea follows the line of reasoning of economics that if prices are raised on food that are detrimental to human health then the demand would lower for those foods. Conversely, if the prices were lowered for healthy food, demand would increase for these foods. The resulting impact would benefit public health, lessening the outcome of diseases such as obesity and diabetes.

Prices of food and public health

With that reasoning in mind, there are many studies that show that food prices have an impact on public health. In the case of obesity, the literature shows that while there are usually small impacts of food prices on weight, there can be larger impacts on people of lower socioeconomic status and people who are on the borderline of being overweight. (Powell & Chaloupka, 2009) That makes sense because people who have more money would be more insulated from price changes, enough to not change their behavior too drastically whereas people less insulated, would be more impacted. Because of the disproportionate health risks for obesity and diabetes for people of lower socioeconomic status, this could be a good change. It could potentially benefit the people who are most at risk. On the other hand, the raising of prices could negatively impact people of lower socioeconomic status and their purchasing power. It could push people into situations of greater food insecurity. Whether by psychological distress or other reasons, low socioeconomic status is associated with obesity.(Christiansen, 2019) The result of which means that interventions would also have to disproportionately impact people of lower socioeconomic status to be the most effective. (Christiansen, 2019)

Already, many places of lower socioeconomic status deal with environmental factors of food insecurity such as lower accessibility to supermarkets and fresh food. That said, the impact of prices remains consistent and is statistically significant even when comparing lower-end versus higher-end grocery stores. (Ghosh-Dastidar et al., 2014) Food desserts do not have as much of an impact as the in-store prices that consumers pay. In other words, the price of healthy and junk food matter more than the type of store. That said, the distance of healthy food is positively correlated with obesity. (Ghosh-Dastidar et al., 2014)

Interestingly, the impact of food prices on obesity can go the other way. The lowering in food prices for calorie-dense meals can increase body fat percentage even if body weight itself does not change much. (Lu & Goldman, 2010) Similarly, a study focused on diabetes found that high prices of healthy food led to higher blood sugar, especially for low-income people. (Anekwe & Rahkovsky, 2018) Dairy was particularly associated with high blood sugar. (Anekwe & Rahkovsky, 2018) These results indicate that the impact of food prices goes both ways. Low prices for unhealthy food have a role in worsening health and raising those unhealthy food prices can have a positive effect.

In the case of diabetes, taxes on unhealthy food and subsidies on healthy food can be effective in preventing deaths from diabetes, strokes, and coronary heart disease. (Penevalo et al., 2017) The impacts were more pronounced depending on the magnitude of the price change. Unhealthy food taxes are one of the many interventions that can be used to impact food prices. It is the primary intervention that this study will use as well.

Corrective Taxation

There are varying views on the impact of taxes and subsidies. One perspective is that tax and subsidies might not work well enough to be effective on their own. (Powell & Chaloupka, 2009) This could especially be the case if the interventions do not have enough of an impact on purchasing power. They might work better in combination with other similar interventions for most people. (Powell & Chaloupka, 2009; Finklestein, 2014) The impact of the tax on weight loss were underwhelming. There was an overestimation of its impact. (Lin, 2011)

Significant price changes could work on more susceptible groups, especially for teens and children, who usually have an inflexibility of income. (Powell & Chaloupka, 2009) Like minors, people of low socioeconomic status are more impacted by junk food taxes which tend to be regressive. Regressive taxes disproportionately impact people of lower income because although they are charged the same to everyone, it accounts for a larger part of expenditure for a person with less money. Similarly, other forms of corrective taxation have emerged such as a tax on fatty foods. It is also regressive even though fat consumption is the same across socioeconomic classes. (Leicester, 2004) Junk food taxes were shown to be just as effective as another consumer intervention, traffic light labeling. Both were effective and saved money. (Sacks et al., 2011)

As for diabetes, a sugary beverage tax was found to reduce consumption of sugary drinks by 15%. It estimated the prevention of 2.4 million cases of diabetes, thousands of early deaths, and billions of dollars-worth of medical expenses in the US within 10 years. (Wang, 2012) For the purposes of this study, a form of a sugary beverage tax will be used to determine impact on diabetes and obesity within the sample. This study hopes to provide an update on prior knowledge, with more recent 2008 and 2018 data, to examine if the trends hold true for diabetes and obesity with a soda tax. This study aims to examine the efficacy of improved health outcomes due to a form of corrective taxation as an intervention.

Data

This study uses self-report data from the long running BRFSS study administered by the Center for Disease Control (CDC). The independent variable, soda tax in percentages, comes from the United States Department of Agriculture's (USDA) long running ERS study. Both datasets were aggregated at the state level and included all fifty states plus the District of Colombia for the years 2008 and 2018. Per year there were 51 observations. See figure 4 in the appendix to see the racial and ethnic break up in the sample. This sample contained a large number of people who identified as white compared to other races (see figures 4, 9, 10). Diabetes was most prevalent in the Midwest with Wyoming having the highest cases (see figure 5). Obesity was also highest in the Midwest with Arizona having the most cases (see figure 6). In the year 2008, there were more women than men, more people under the age of 65, more non-Hispanics, more physically active, more college-educated, and more employed than not (see figure 9). In the year 2018, there main difference is that there were more inactive people than active (see figure 10).

The BRFSS data included the control variables for sex, income, education, employment, physical activity, race, and ethnicity. All of the values for the answers “refuse to respond” and “prefer not to say” were made blank so as to not interfere with the analysis. The sex variable was turned into a binary variable leaving only male (0) and female (1). Binary means that dummy variables were created to make categorical variables numeric and given the values of 1 and 0 to represent the two categories. The income variable was initially broken up into ranges of $15,000. It was converted into a binary variable with the categories of $35,000 and above (1) and below that amount (0). Education was converted to college and above (1) and less than college (0). The employment variable had a lot of separate categories including employed, self-employed, unemployed, student, etc. Self-employed and employed were combined to make the (1) value for the new variable, job. The (0) value consisted of the other categories besides employed and self-employed. Age65, meaning 65 years or older was one category with the other category being 64 and younger. It was already defined in the original data. The question regarding the participants getting exercise within the past thirty days was converted into a new variable, active, in which participants getting regular or somewhat regular exercise within the last 30 days

(1) was compared to (0) irregular physical activity and no physical activity. Smoke has the value of (1) for yes and (0) for not smoking. _Hispanc is also a yes (1) or no (0) response for the question, are you Hispanic? The race variable was broken up into main race categories including Black, White, Asian or Pacific Islander, and Native American or Indigenous. Obese was a variable defined as having a BMI of 30 and over (1) versus under that amount (0). Similarly, _diabete3 was the variable created for blood sugar in the diabetic range (1) versus under that amount (0), excluding gestational diabetes. The variables _ststr, _finlwght, and _psu were used to weight the dependent and control variables.

From the USDA ERS, only the soda tax variable was used. It reported soda taxes in percentages at the county level for all of the fifty states and the District of Colombia. The final variables used were the following: Sodatax as the main independent variable, _ststr, _fnlwght, & _psu for weighting, obese & _diabete3 as dependent variables, and finally as control variables _35kplus, _age65yr, _hispanc, active, Asian, Black, college, female, job, native, smoke, and White.

Methodology

After all of the BRFSS variables were made categorical and binary, I used *svyset* to set the appropriate weights and let STATA know that it was survey data. After that, I used *svy: proportion* to make proportions of all of my variables from the BRFSS dataset to aggregate them at the state level. The soda tax variable also had to be aggregated at the state level by averaging the county level soda taxes by state. I then exported outputs to Microsoft Excel to combine the USDA and CDC datasets.

To analyze the data, I decided to run separate Ordinary Least Squares regressions separating by year and dependent variable. The variables for obesity and diabetes were interchanged as control variables depending on which one was used as the dependent variable. I then ran the 2-year data as panel data using xtreg for both diabetes and obesity. I included fixed effect and random effect regressions.

For the descriptive statistics, I used the summarize function and then used graphs including scatterplots and histograms.

Results

Looking at the output for the simple linear regressions in figures 1 and 2, soda tax was not found to be a statistically significant predictor of diabetes or obesity. The same was true for the panel data regressions in figure 3. According to the correlations in figure 7, obesity and diabetes in the sample were strongly positively correlated at 0.829 on a scale of -1 to 1. Despite the strong positive relationship, the regressions have slightly opposing signs on their coefficients although close to 0.

Interestingly, having an income of $35,000 or more had a slight positive relationship with the incidence of diabetes and obesity. That could be because the income cut off was too low to effectively examine people of low socioeconomic status versus those not. Its relationship is not strong enough to be statistically significant for 2008 or in the panel data regressions (see figures

1, 3,). The income variable was very statistically significant for the 2018 OLS regression, meaning it was significant at the 1% significance level (see figure 2).

Being above the age of 65 years old was statistically significant at the 1% significance level for diabetes in the year 2008 but not at all significant for the year 2018 or for either year for obesity for the OLS regressions. Its results were the inverse of the prior mentioned income variable. This is one of the most statistically significant predictor variables in the panel regression (see figure 3). It was statistically significant at the 1% level for the regular panel regression for diabetes and obesity, and the fixed effects regression for obesity. It was significant at the 5% level for the random effect regression for obesity. Considering that the correlation with both obesity and diabetes for age over 65 is over 0.54 for both, it makes sense that age would be a good predictor for these lifestyle diseases (see figure 7).

What is interesting for the Hispanic variable is that it has an inverse relationship with both diabetes and obesity. Asian also had that at the 1% significant level for the obesity OLS in 2008 and all of the obesity panel regressions. Having a high proportion of people who identify as Asian seem to be a strong predictor of lower obesity rates. That was not the case for diabetes. Conversely, people who identify as Black are statistically significant at the 5% level for obesity for the 2018 OLS and the panel data regression for fixed effects with a positive relationship. That mean a 1 unit increase in Black people is related to a 0.122 unit increase in obesity for the fixed effects regression.

Much like the other variables, the coefficient for Native American or Indigenous had opposing signs for diabetes and obesity. It had a negative relationship with diabetes and a positive one with obesity. That means that higher proportions of Native Americans are associated with lower incidence of diabetes and higher incidence of obesity. These results are significant at the 5% level for the 2008 OLS for both obesity and diabetes and for obesity regular panel and random effect regressions. It is significant at the 1% level for the diabetes panel data except for the fixed effect.

As expected, having a college education or higher was associated with lower incidence of diabetes and obesity. This result was consistent with the literature. The relationship was statistically significant at the 5% significance level for the 2008 Diabetes OLS, meaning that for a 1 unit increase in college educated people, there was a 0.053 unit decrease in diabetes. Similarly, there was the same significance for this variable for the panel data random effects regression. Although physical activity had a consistently negative relationship with both diabetes and obesity (see figures 1, 2, & 3), this variable had no statistical significance and the coefficient magnitudes were fairly small. Similarly, smoking did not have any visible trends except for its coefficients being negative for all of the panel data regressions. It did not have any statistical significance for any of the regressions.

One of the more surprising results was the strong relationship between female and obesity. Compared to diabetes that had a positive relationship, obesity had a negative one. At the 1% level were all of the obesity regressions except for the 2018 OLS, which was only significant at the 5% level. More than that, the coefficients are some of the largest in the sample. For a 1 unit increase in the proportion of females, there was a 1.3 to 1.8 unit decrease in the incidence of obesity. This result is consistent with current national data in which 43% of men were obese while only 41% women were. (NIDDK 2020)

Having a job was statistically significant at the 1% level with a negative relationship for almost all of the regressions for diabetes (see figures 1, 2, & 3). The coefficients for this variable were also somewhat bigger in magnitude than the other variables but not to the point of female variable. The Hausman Test (see figures 11 & 12) failed to reject the null hypothesis when comparing the results from the fixed effects regressions and the random effects regressions for both diabetes and obesity. Based on those results, it is better to look at the results from the random effects regression rather than the fixed effects regressions because it would be more reliable.

Figure 1. 2008 Ordinary Least Squares Regression

Regressor	OLS Diabetes	OLS Diabetes	OLS Obesity	OLS Obesity
Sodatax	-0.001	-0.000	0.000	0.001
	(0.001)	(0.000)	(0.002)	(0.001)
_35kplus		0.009		0.050
		(0.014)		(0.034)
_age65yr		0.094		0.174
		(0.032)**		(0.103)
_hispanc		-0.022		-0.044
		(0.012)		(0.035)
active		-0.014		-0.051
		(0.031)		(0.085)
asian		0.022		-0.205
		(0.020)		(0.047)**
black		-0.011		0.073
		(0.015)		(0.038)
college		-0.053		-0.069
		(0.026)*		(0.075)
female		0.131		-1.313
		(0.127)		(0.350)**
job		-0.153		0.149
		(0.029)**		(0.092)
native		-0.091		0.255
		(0.037)*		(0.110)*
obese		0.245		
		(0.030)**		
smoke		0.000		-0.089
		(0.030)		(0.091)
white		-0.045		0.022
		(0.011)**		(0.027)
_diabete3				1.786
				(0.240)**
_cons	0.103	0.110	0.283	0.701
	(0.004)**	(0.080)	(0.007)**	(0.221)**
SER	0.02	0.01	0.04	0.02
R^2	0.006	0.934	0.001	0.867
Adjusted R^2	-0.004	0.924	-0.009	0.846
N	102	102	102	102

* p<0.05; ** p<0.01

Figure 2. 2018 Ordinary Least Squares Regression

Regressor	OLS Diabetes	OLS Diabetes	OLS Obesity	OLS Obesity
Sodatax	-0.001	-0.000	0.001	0.001
	(0.001)	(0.000)	(0.002)	(0.001)
_35kplus		-0.031		0.120
		(0.017)		(0.043)**
_age65yr		-0.013		0.059
		(0.046)		(0.147)
_hispanc		-0.018		0.105
		(0.027)		(0.069)
active		-0.006		-0.073
		(0.040)		(0.128)
asian		0.066		-0.084
		(0.045)		(0.145)
black		-0.018		0.157
		(0.029)		(0.065)*
college		-0.064		-0.007
		(0.035)		(0.103)
female		0.462		-1.412
		(0.252)		(0.570)*
job		-0.183		0.311
		(0.039)**		(0.153)*
native		0.010		0.028
		(0.053)		(0.166)
obese		0.260		
		(0.046)**		
smoke		0.027		0.156
		(0.052)		(0.155)
white		-0.027		0.122
		(0.026)		(0.068)
_diabete3				1.737
				(0.360)**
_cons	0.119	-0.016	0.310	0.488
	(0.005)**	(0.147)	(0.010)**	(0.384)
SER	0.02	0.01	0.04	0.02
R^2	0.013	0.926	0.006	0.861
Adjusted R^2	-0.007	0.897	-0.015	0.807
N	51	51	51	51

* p<0.05; ** p<0.01

Figure 3. Panel Regressions

Regressor	XT Diabetes	XT FE Diabetes	XT RE Diabetes	XT Obese	XT FE Obese	XT RE Obese
Sodatax	-0.000		-0.000	0.001		0.001
	(0.000)		(0.000)	(0.001)		(0.001)
_35kplus	0.009	0.004	0.009	0.051	0.116	0.051
	(0.013)	(0.041)	(0.013)	(0.032)	(0.092)	(0.032)
_age65yr	0.090	0.023	0.090	0.200	0.345	0.200
	(0.030)**	(0.054)	(0.030)**	(0.094)*	(0.094)**	(0.094)*
_hispanc	-0.022	-0.023	-0.022	-0.041	-0.015	-0.041
	(0.013)	(0.017)	(0.013)	(0.031)	(0.031)	(0.031)
active	-0.013	-0.002	-0.013	-0.067	-0.167	-0.067
	(0.031)	(0.051)	(0.031)	(0.082)	(0.095)	(0.082)
asian	0.024	0.047	0.024	-0.202	-0.162	-0.202
	(0.023)	(0.024)	(0.023)	(0.047)**	(0.050)**	(0.047)**
black	-0.011	-0.011	-0.011	0.078	0.122	0.078
	(0.015)	(0.021)	(0.015)	(0.040)	(0.054)*	(0.040)
female	0.129	0.167	0.129	-1.378	-1.811	-1.378
	(0.131)	(0.212)	(0.131)	(0.352)**	(0.398)**	(0.352)**
college	-0.051	-0.030	-0.051	-0.064	-0.044	-0.064
	(0.024)*	(0.032)	(0.024)*	(0.074)	(0.081)	(0.074)
job	-0.156	-0.195	-0.156	0.163	0.204	0.163
	(0.029)**	(0.040)**	(0.029)**	(0.097)	(0.106)	(0.097)
native	-0.089	-0.057	-0.089	0.213	0.063	0.213
	(0.031)**	(0.036)	(0.031)**	(0.084)*	(0.088)	(0.084)*
obese	0.248	0.292	0.248			
	(0.031)**	(0.045)**	(0.031)**			
smoke	-0.001	-0.016	-0.001	-0.094	-0.107	-0.094
	(0.030)	(0.036)	(0.030)	(0.076)	(0.082)	(0.076)
white	-0.044	-0.034	-0.044	0.019	0.020	0.019
	(0.011)**	(0.012)**	(0.011)**	(0.027)	(0.030)	(0.027)
_diabete3				1.789	1.693	1.789
				(0.227)**	(0.239)**	(0.227)**
_cons	0.111	0.091	0.111	0.731	0.930	0.731
	(0.085)	(0.133)	(0.085)	(0.223)**	(0.273)**	(0.223)**
SER	0.01	0.00	0.01	0.02	0.01	0.02
R^2	.	0.965	.	.	0.937	.
Adjusted R^2	.	0.960	.	.	0.927	.
N	102	102	102	102	102	102

* $p<0.05$; ** $p<0.01$

Discussion

In this study, the soda taxes did not have any impact on the diabetes or obesity outcomes with coefficients close to zero and no statistical significance. Conversely, other variables had stronger predictive powers such as having a job leading to a lessening of diabetes incidence and the same for being female with obesity. Since the primary focus of this study was the soda tax, the findings of this study did not support junk food taxes as an effective intervention. The results mirror what other studies said about it not having enough impact on its own. Also, this type of intervention uses regressive taxation that limits the purchasing power of US constituents, especially those of lower socioeconomic status. If the public health impacts do not improve, it might not be worth it to limit the public in such a way because of the risk of adverse effects like increased food insecurity. That said, the scope of this study was not enough to make any conclusive statements on the issue.

One of the limitations of this study was the small sample size, the OLS regressions only had 51 observations per year and with the combined years in the panel data, it only aggregated to 102 observations. In the future it would be beneficial to attempt this study at the county level although it might be difficult to get BRFSS county level data. Similarly, it would be good to expand the range of years. Initially, this study had chosen to focus on food prices rather than a food price-related intervention. However, there was not such available data at the state level. It would be interesting to see that type of study. Similarly, it would be good to get a more thorough update with more available years of data to compare. There can be a focus on other types of junk food taxes as well. Finally, there was not as much data supporting the efficacy of healthy food subsidies so an inverse of this study would be an interesting direction.

References

Anekwe T. & Rahkovsky I. (2018) “The Association Between Food Prices and the Blood Glucose Level of US Adults With Type 2 Diabetes.” American Journal of Public Health, 108, S475_S482. https://doi.org/10.2105/AJPH.2013.301661r

Christiansen P., Dickson J., Lorenzetti V., Hardman C., & Spinosa J. (2019). From Socioeconomic Disadvantage to Obesity: The Mediating Role of Psychological Distress and Emotional Eating. https://doi.org/10.1002/oby.22402

Finkelstein, E. A., Strombotne, K. L., Zhen, C., & Epstein, L. H. (2014). Food prices and obesity: a review. Advances in nutrition, 5(6), 818-821.

Ghosh-Dastidar, B., Cohen, D., Hunter, G., Zenk, S. N., Huang, C., Beckman, R., & Dubowitz, T. (2014). Distance to store, food prices, and obesity in urban food deserts. American journal of preventive medicine, 47(5), 587–595. https://doi.org/10.1016/j.amepre.2014.07.005

Goldman D. & Lu Y. (2010) "The Effects of Relative Food Prices on Obesity — Evidence from China: 1991-2006.” NBER Working Paper No. 15720 JEL No. D01,I1,J88

Leicester, A., & Windmeijer, F. (2004). The 'fat tax': economic incentives to reduce obesity.

Lin, B. H., Smith, T. A., Lee, J. Y., & Hall, K. D. (2011). Measuring weight outcomes for obesity intervention strategies: the case of a sugar-sweetened beverage tax. Economics & Human Biology, 9(4), 329-341.

Peñalvo, J.L., Cudhea, F., Micha, R. et al. The potential impact of food taxes and subsidies on cardiovascular disease and diabetes burden and disparities in the United States. BMC Med 15, 208 (2017). https://doi.org/10.1186/s12916-017-0971-9

Powell, L. M., & Chaloupka, F. J. (2009). Food prices and obesity: evidence and policy implications for taxes and subsidies. The Milbank quarterly, 87(1), 229–257. https://doi.org/10.1111/j.1468-0009.2009.00554.x

"Products - Data Briefs - Number 360 - February 2020." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 27 Feb. 2020, www.cdc.gov/nchs/products/databriefs/db360.htm.

Sacks, G., Veerman, J. L., Moodie, M., & Swinburn, B. (2011). 'Traffic-light'nutrition labelling and 'junk-food'tax: a modelled comparison of cost-effectiveness for obesity prevention. International Journal of Obesity, 35(7), 1001-1009.

Wang, Y. C., Coxson, P., Shen, Y. M., Goldman, L., & Bibbins-Domingo, K. (2012). A penny-per-ounce tax on sugar-sweetened beverages would cut health and cost burdens of diabetes. Health Affairs, 31(1), 199-207.

Appendix

Figure 4. Race and Ethnicity by Year

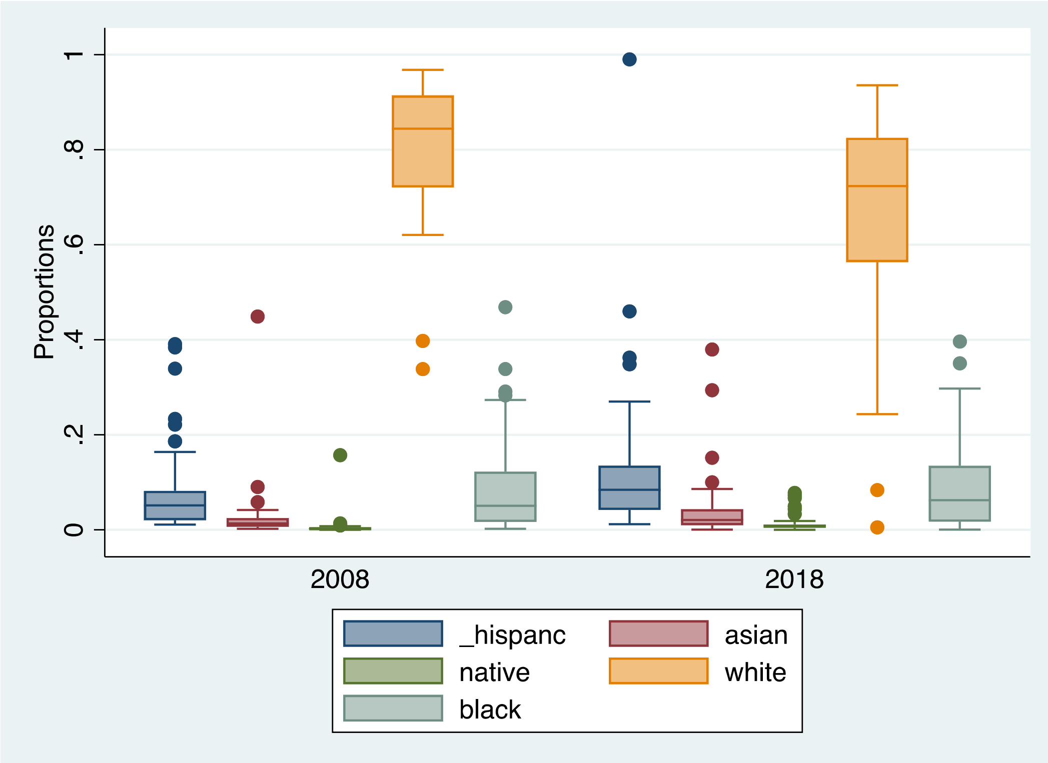

Figure 5. Diabetes by State in Percent

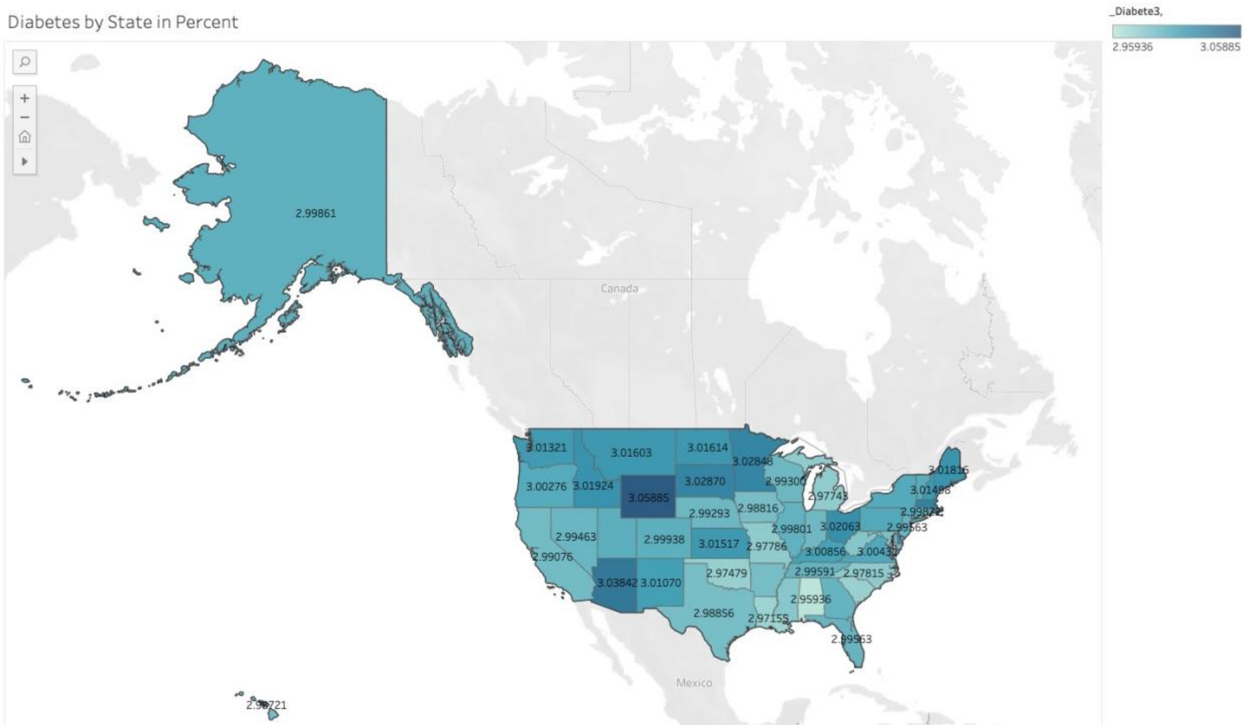

Figure 6. Obesity by State in Percent

Figure 7. Correlations

```
. correl Sodatax _35kplus _age65yr _diabete3 _hispanc active asian black college female job native obese smoke white
(obs=102)
```

	Sodatax	_35kplus	_age65yr	_diabe~3	_hispanc	active	asian	black	college	female	job	native	obese	smoke	white
Sodatax	1.0000														
_35kplus	0.0621	1.0000													
_age65yr	0.0131	-0.1203	1.0000												
_diabete3	-0.0800	-0.1142	0.5415	1.0000											
_hispanc	-0.1133	-0.0493	0.0526	0.2013	1.0000										
active	0.0475	0.1169	-0.0901	-0.5635	-0.2718	1.0000									
asian	0.0535	0.0837	-0.0847	0.0777	0.0860	0.1788	1.0000								
black	-0.1921	0.0160	-0.0956	0.3207	-0.1157	-0.2512	-0.0908	1.0000							
college	0.0096	0.3041	-0.1624	-0.6614	-0.1084	0.7613	0.0677	-0.0735	1.0000						
female	-0.0941	0.0883	0.1971	0.2929	0.0602	-0.3672	-0.2045	0.6399	-0.0872	1.0000					
job	0.1374	0.2506	-0.5134	-0.8070	-0.3570	0.5502	0.0966	-0.2070	0.6537	-0.3443	1.0000				
native	-0.0135	0.0877	0.1976	0.0805	0.0644	0.1569	0.4242	-0.1654	0.0785	-0.2376	-0.0520	1.0000			
obese	0.0228	-0.0629	0.5449	0.8293	-0.0498	-0.4418	-0.1806	0.2272	-0.5720	0.1044	-0.6105	0.0879	1.0000		
smoke	0.1283	-0.1818	-0.1401	0.1637	-0.4352	-0.4780	-0.2442	0.1018	-0.5947	-0.0032	-0.1647	-0.0792	0.2777	1.0000	
white	0.1748	-0.0435	-0.0033	-0.4816	-0.5793	0.2507	-0.5738	-0.3918	0.1493	-0.2031	0.3222	-0.2128	-0.1388	0.3334	1.0000

Figure 9. 2008 Descriptive Statistics

```
. su
```

Variable	Obs	Mean	Std. Dev.	Min	Max
stname	0				
_state	51	28.96078	15.83283	1	56
stusps	0				
binary	51	1	0	1	1
iyear	51	2008	0	2008	2008
_35kplus	51	.6526961	.0634714	.5271931	.7846367
_age65yr	51	.1711504	.0196677	.1010357	.2207026
_diabete3	51	.0852372	.0151232	.0605871	.1218161
_hispanc	51	.0822966	.0913116	.0108792	.390826
active	51	.7516522	.0356495	.6745294	.8191778
asian	51	.0269255	.0624506	.0021465	.4489994
black	51	.0910855	.1036164	.0022353	.4686771
college	51	.6060582	.0545051	.4530715	.715755
female	51	.5126979	.0095609	.483311	.5395824
job	51	.6144651	.040739	.5211908	.6765587
native	51	.0061374	.02176	.0000743	.1571495
obese	51	.2559558	.0295805	.1825567	.3236736
smoke	51	.1895066	.0335836	.0931179	.2653217
white	51	.8042911	.1361699	.3384055	.9679431
Foodandbev~d	51	15080.11	16706.54	1634.8	93758.7
Sodatax	51	3.549216	2.790379	0	7

Figure 10. 2018 Descriptive Statistics

```
. su
```

Variable	Obs	Mean	Std. Dev.	Min	Max
stname	0				
_state	51	28.96078	15.83283	1	56
stusps	0				
iyear	51	2018	0	2018	2018
_35kplus	51	.6492835	.0620502	.4857976	.7474124
_age65yr	51	.2136859	.026403	.0990259	.2643275
_diabete3	51	.1157099	.0204033	.0722547	.1661484
_hispanc	51	.1227412	.1559286	.011905	.9901778
active	51	.7563297	.0500234	.5279035	.8380194
asian	51	.041451	.067707	.0006496	.3796058
black	51	.096641	.0973022	.0007059	.396428
college	51	.594658	.0559084	.4425462	.7280772
female	51	.5119822	.0095474	.4902741	.5332796
job	51	.5752441	.0424759	.4405498	.6506476
native	51	.0134217	.0175732	0	.0777769
obese	51	.313249	.0385229	.2315836	.3953986
smoke	51	.164659	.0348288	.0898411	.2524688
white	51	.6761125	.202438	.0052172	.9355184
Sodatax	51	3.549216	2.790379	0	7

Figure 11. Hausman Obesity

```
. hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
     _35kplus |    .1155513      .050888        .0646633        .0903516
     _age65yr |    .3451894      .200212        .1449775        .0811564
     _hispanc |   -.0154664    -.0412217        .0257552        .0276908
       active |   -.1670062     -.06715        -.0998563        .0803679
        asian |   -.1615821    -.2018296        .0402475        .0488577
        black |    .1224094     .0778155        .0445939         .038754
      college |   -.0436911    -.0642084        .0205172        .0554792
       female |   -1.810841    -1.377534       -.4333067        .3152027
          job |    .2042608       .16322        .0410408        .0845672
       native |    .0630971     .2134242        -.150327        .0807332
    _diabete3 |    1.692559     1.788788       -.0962293        .1824652
        smoke |   -.1071566    -.0935453       -.0136113        .0841251
        white |    .0203099     .0193485        .0009614        .0242261
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(13) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                           =       19.69
                Prob>chi2 =      0.1031
```

Figure 12. Hausman Diabetes

```
. hausman fixed1 random1
```

	Coefficients (b) fixed1	(B) random1	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
obese	.291593	.2481495	.0434435	.0375353
_35kplus	.003798	.0085094	-.0047115	.0389037
_age65yr	.0233083	.0901187	-.0668103	.0440707
_hispanc	-.0234488	-.0221685	-.0012804	.0123424
active	-.0019148	-.0134876	.0115728	.0380378
asian	.0470196	.0238548	.0231648	.022353
black	-.0108931	-.0106099	-.0002832	.0190665
college	-.0296792	-.0514685	.0217893	.0260146
female	.1665387	.128674	.0378647	.1777069
job	-.1951499	-.1556688	-.0394812	.0298986
native	-.0573073	-.0886106	.0313033	.0368192
smoke	-.0161018	-.0008139	-.0152879	.0389852
white	-.0343215	-.0444694	.0101479	.011014

```
                       b = consistent under Ho and Ha; obtained from xtreg
        B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                 chi2(13) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        10.48
                Prob>chi2 =       0.6545
```

Website

https://bakeralexan.github.io/Pepperdine-2020-Econometrics/

https://github.com/bakeralexan/Pepperdine-2020-Econometrics/blob/main/README.md