Steven Espinoza

Professor Liu

ECON 970: Big Data in Economics

30 April 2018

Obesity, Food Deserts, and Minorities: A Case Study of Colorado

Access to healthy food is something that many take for granted. Especially when considering that most people in the United States today live in urban areas where supermarket stores like Wal-Mart and Whole Foods, which have a dedicated section of their floor plan towards fresh produce, are abundant, some cannot imagine how certain segments of society can lead healthy lives without supermarkets that are located within reach. Unfortunately, however, this is the reality for a sizeable portion of the United States population who live in "food deserts," which is defined by the United States Department of Agriculture (USDA) simply as "neighborhoods that lack healthy food sources." Anywhere between 17.3 million to 54.4 million people live in such areas, depending on the method of measurement.

Intuitively, one would assume that food deserts are linked with worse health outcomes. The goal of this paper is to test the hypothesis that census tracts within the state of Colorado with food deserts are associated with higher rates of obesity. Additionally, this paper tests the hypothesis that some interaction exists between food deserts and the level of minorities in predicting obesity outcomes within a census tract. To test this hypothesis, this paper will use data on food deserts by census tract provided by the USDA, as well as predicted levels of obesity provided by the State of Colorado's open data website.

With the lowest rate of obesity among all states, Colorado is generally accepted to be the healthiest state in the country. Moreover, access to food in Colorado is, on average, slightly better than that of other states as well. The Center for Disease Control (CDC) calculates an index, the mRFEI score, which indicates the number of healthy food retailers among all food retailers in the state. A higher mRFEI score indicates a healthier food environment, while a lower mRFEI score indicates the opposite. According to a 2011 report, Colorado's mRFEI score is 11, while the national average is 10. Despite these indicators that speak to the healthiness of Colorado, significant variation exists within the state with regards to obesity outcomes and food deserts. My goal in this paper is to determine if food deserts and the population of minorities can explain this variation.

Literature Review

Previous literature has examined the relationship among food deserts, socioeconomic status, and obesity outcomes. Arterburn et al. (2014) studied obesity outcomes in the census tracts that make up King County, Washington, within which lies much of the Seattle metropolitan area. Arterburn and his colleagues found that census-tract level home values and college education "were more strongly associated with obesity than household incomes." With regards to food deserts within the county, Arterburn and his colleagues used a sample of around 60,000 insured adults and found that "there was no evidence that the mRFEI score was associated with the prevalence of obesity, either independently or after accounting for demographic covariates."

Another study conducted by Brownell et al. (2017) comes across a different conclusion. In their research, Brownell and her colleagues find from data across the entire United States that "food swamps have a positive, statistically significant effect on adult obesity rates." In their research, Brownell and her colleagues differentiate "food swamps" from "food deserts" as "a spatial metaphor to describe neighborhoods where fast food and junk food inundate healthy alternatives" versus a "residential area with limited access to affordable and nutritious food." According to their research, "minorities are more likely than Whites to live near unhealthy food retailers," though the model they use does not consider how the presence of minorities alone could impact obesity rates. The variables they use instead are "food swamp," "food desert," "recreation/fitness centers," "natural amenities," and "neighborhood characteristics."

Arcury et al. (2006) consider the experience of Hispanic/Latino immigrant families in North Carolina. Instead of focusing on food deserts specifically, they consider the implications that food insecurity and hunger (FIH) have on the health of immigrants in North Carolina. They argue that "Recent evidence linking FIH to negative health and social outcomes, like obesity, through economic limits on food choice (28) raised the prospect of long-term consequences of FIH. This suggests that changes in state and national policies are needed to improve the situation for new immigrants." Some of the policies they consider include considering how immigrants can gain access to federal programs like the food stamp program, or stressing the importance of adequately funding school lunch programs on FIH outcomes.

Methods and Data

The different hypotheses, results, and approaches taken by the researchers in this existing literature within health economics have forced me to consider the best possible method in determining the relationship between food accessibility and minority populations with obesity outcomes per census tract in Colorado. The model I have formulated to consider this hypothesis takes the form of

$$y_i = \alpha_i + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 X_2) + \psi + \lambda + \varepsilon_i$$

In this model, y_i measures the level of overweightness and obesity per census tract in Colorado. This data comes from the State of Colorado's Department of Public Health and

Environment (CDPHE) website, last updated in February 2018, which gives the predicted prevalence of overweight or obese adults per census tract in Colorado. Overweight is defined as a Body Mass Index (BMI) of 25 or more; obese is defined as a BMI of 30 or more. Access to obesity data in general is difficult; to the best of my knowledge, Colorado is the only state to provide this census-tract level data publicly and for free. The research conducted by Arterburn et al. (2014) on obesity takes data from a private insurance company, while Brownell et al.'s (2017) data on obesity is taken from the Food Environment Atlas, but at a county-level. By analyzing data at the census-tract level, I can obtain a clearer picture of how food deserts on the micro-level can influence obesity outcomes, for there may be more food desert variation within any county.

X_1 is the continuous independent variable measuring the extent to which a certain census tract is a food desert. Using data from the USDA's Economic Research Service (ERS), the way I calculated this is by determining the proportion of people per census tract who live beyond one mile from a supermarket in urban areas, or ten miles away from a supermarket in rural areas.

$$"food\ desertness" = \frac{\#\ of\ people\ sufficiently\ far\ enough\ from\ healthy\ food}{\#\ total\ of\ people\ within\ census\ tract\ (2010\ census)}$$

It should be noted that the dataset provided by the USDA's ERS has three measures of determining low access to food: counting the number of people per census tract beyond 1 mile for urban areas (or 10 miles for rural areas) from a supermarket, the number of people per census tract beyond ½ mile for urban areas (or 10 miles for rural areas) from a supermarket, or the number of people per census tract beyond 1 mile for urban areas or 20 miles for rural areas from a supermarket. Moreover, the way the USDA determines how many people live in food deserts depends on either of these three measures combined with whether a census tract is designated as "low-income." Thus, the USDA counts that 54.4 million people live in food deserts by the first measure, 19 million people by the second measure, and 17.3 million people by the third measure respectively. For simplicity, I will use the method in my research that gives the median indicator of people who live in food deserts. This explains the numerator of my equation measuring "food desertness." This value takes on a range of 0 < X_1< 1. A value closer to 0 would indicate a "food oasis," where everyone in the census tract has access to healthy food based on their location. A value closer to 1 would indicate a "food desert," where no one in the census tract has access to healthy food.

X_2 measures the proportion of minorities per census tract. This data is also taken from the CDPHE website on a census tract level. The way this is measured is by calculating the number of non-White people per census tract. In the regression analyses that follow, however, X_2 will also come to mean the percentage of each minority (African Americans, Hispanic/Latino, Asian, and Native American/Alaskan) that resides within a census tract to analyze if the relationship between minority census tracts and obesity is stronger within certain minorities than others.

The interaction term, X_1X_2, evaluates whether the relationship of food deserts and minority populations on obesity is stronger when considering how they relate to one another as opposed to considering their relationship with obesity on an individual basis. Additionally, it evaluates the interaction between food deserts and the presence of a specific minority. This would mean that five different interaction terms would be generated, since X_2 would consider the presence of minorities (as a whole), African Americans, Hispanic/Latinos, Asians, and Native Americans/Alaskans.

Finally, ψ is a control variable to measure the relationship among obesity, minority populations, and obesity holding income constant. This data (median income) comes from the CDPHE's website and, like the other variables, is recorded on a census tract level. λ is a dummy control variable to measure this relationship holding a census tract's designation as "urban" or "rural" constant. This data is taken from the USDA's ERS dataset on food access. "1" indicates an urban census tract while "0" indicates a rural one. According to the USDA, "a census tract is considered rural if the population-weighted centroid of that tract is located in an area with a population of less than 2,500." ε is an error term to account for the difference between predicted obesity outcomes and actual obesity outcomes per census tract.

In total, there are 1242 observations in the dataset containing these values. Eight observations have been dropped from the original dataset. Six of these omissions were due to the very small population of the census tracts (adult population was less than 50), while two of these were omitted because the census tracts exclusively contained two federal correctional institutions. Additionally, missing values for income have been imputed with the median income for Colorado, which according to the U.S. Census website was $62,520 in 2016 dollars.

Results

Table 1 shows the first regression where X_2 is the percentage of minorities per census tract, which is defined by the number of non-White people in each census tract. The coefficients suggest that there is a statistically significant relationship between food deserts and overweightness/obesity, as well as a statistically significant relationship between the percentage of minorities within a given census tract and the level of overweightness/obesity in a census tract. This is true for when there are no income or urban controls, as well as for when they are held constant. The interaction terms are a quite surprising; the fact that they have negative coefficients left me puzzled, since I initially assumed that food deserts with a high percentage of minorities would lead to an increased presence of overweightness and obesity within a census tract. In these regressions from table 1, however, the most significant variable that stood out to me was the "food desertness" variable. When controlling for urban and income, and when considering the interaction taking place between minorities and food deserts, a one unit increase in "food desertness" predicts a 6.076 percentage point increase in the proportion of people within a census tract who are overweight and obese. Figures 1 and 2 illustrates this relationship using the `binscatter` command on Stata, where a positive slope is shown with the percentage of overweight and obese adults per census tract as the dependent variable and the proportion of the total population with limited access to healthy food per census tract as the independent variable.

To obtain a deeper understanding in the role that minorities play in predicting overweight and obesity outcomes in Colorado, the X_2 variable had to take on different meanings. Instead of merely taking on the value of the percentage of all minorities within each census tract, I thought it would be best to see how the presence of each minority within each census tract can lead to predictive outcomes of obesity and overweightness in each census tract. Tables 2–6 show what this looked like in practice.

With regards to African Americans, the model suggested a statistically significant positive relationship between overweight/obesity outcomes and food deserts holding all else constant, as well as a statistically significant positive relationship between overweight/obesity outcomes and the percentage of African Americans in a census tract. However, as shown in Table 2, the interaction term is negative and significant at the 0.01 level, suggesting that the presence of food deserts in census tracts where African Americans are more prevalent can lead to a lower predicted outcome of overweight/obesity levels than one would predict if one was merely considering only the presence of African Americans, or only the proportion of people with limited access to healthy food per census tract.

Table 3 shows X_2 as representative of the percentage of Hispanic/Latinos within a census tract. The results are similar to that of Table 2: When considering the interaction between Latinos and food deserts in Colorado, the result is that predicted obesity outcomes are less than what one would expect if they were only considering either the level of Latinos in a census tract or the level of “food desertness.” However, unlike the results from Table 2, Table 3 suggests a statistically insignificant outcome when not controlling for income and urban; when controlling for these factors, however, the coefficient for the interaction is still only significant at the 0.1 significance level. This suggests that food deserts and the percentage of Hispanic/Latino population within a census tract may be independent of one another. Table 3 also shows the coefficients for X_1 and X_2 are positive and statistically significant holding all else constant. This is true for Latinos in Colorado even when controlling for income and urban. Figure 5 shows the relationship between overweight/obesity outcomes and percentage of Hispanic/Latinos per census tract controlling for income and urban.

Table 4 shows X_2 as representative of the percentage of Asians within a census tract. The table shows a statistically significant negative relationship between the percentage of Asians within a census tract and outcomes of overweightness/obesity. However, the statistically insignificant interaction term suggests that the percentage of Asians in a census tract and the “food desertness” are independent of one another.

Table 5 shows a statistically significant between the interaction of “food desertness” with the presence of Native Americans within a census tract, thus suggesting that there is a higher presence of obesity in food deserts where there are also a lot of Native Americans holding all other factors constant. However, the coefficient for the interaction term here is negative once again, this time at a statistically significant level.

To provide greater context, I provided a sixth table for non-minority White people, which I calculated as (percent white) = 100–(percent of minorities). This table shows a subversion of what was revealed within the other tables with minorities; X_1 and X_2 are both negative in this case, while the interaction term is positive. Given this information, then, we can conclude that the higher percentage of non-minority White people within a census tract, the more positive the impact of food deserts on obesity.

Figures 1–2 visualize the relationship between overweight/obesity outcomes and "food desertness," while figures 3-7 visualize the relationship between overweight/obesity outcomes and the level of minorities within a census tract. The strong positive relationship showed in figure 3 seems to be most heavily influenced by the relationship shown in figure 5, suggesting that most of the minority observations in the data are of Hispanic/Latino descent. Figure 4 (African Americans) and Figure 7 (Native Americans) show a weaker positive relationship between overweight/obesity outcomes, while Figure 6 (Asians) shows a negative relationship. Nonetheless, when considering minorities as a whole, the relationship is strongly positive.

Finally, figures 9–12 visualize the relationship between the percentage of a specific minority and "food desertness," a relationship that was not considered in the model but is otherwise useful and interesting to understand. It shows a generally negative relationship between "food desertness" and every minority, with Native American/Alaskan as the only exception (Figure 12).

Discussion and Conclusion

The results for the interaction terms suggest is a slight divergence from what I initially expected to find. Though I assumed that the relationship between food deserts on obesity would be positive (and the results show strong evidence for that), I should acknowledge that this relationship only came about because of the way I defined food deserts. By calculating my own formula for “food desertness,” I could obtain a clearer picture as to how many people exactly within each census tract were lacking the access to healthy food. This measurement could thus explain why the mRFEI data used in Arterburn et al. (2014) was found to have a statistically insignificant impact on obesity outcomes. Nonetheless, the fact that my results show a significant positive relationship between the proportion of people who lack access to healthy food and overweight/obesity outcomes is enough to go against the findings made by Arterburn and his colleagues.

Additionally, breaking down the X_2 variable into different categories was key to my research. Though the results from table 1 show a significant positive relationship between the percentage of minorities per census tract and overweight/obesity outcomes, it fails to show to what extent this is true for all minorities. As further data showed, for example, this was not the case for Asians in the data, who showed a negative relationship between the percentage of Asians in a census tract and outcomes of obesity.

Explaining the negative coefficient behind the interaction terms for African-Americans and Native Americans/Alaskans was initially difficult for me to do, since I really expected the coefficients to be significant and positive for all the different minorities. Nonetheless, there are a few ways to explain this relationship. The first explanation might simply be that the interaction term is negative because that is how it is–after all, this study was done in Colorado, which is probably the healthiest state in the nation. The extent to which this explanation is externally valid is tough to determine; this would require looking at obesity outcomes on a census-tract level basis, and the data for this is simply not available to the best of my knowledge. Nonetheless, it really might just be that some of Colorado's policies (perhaps, for example, its zoning laws on grocery stores and food availability) make it easy for minorities anywhere within the state to have access to healthy food.

The second explanation as to why this interaction between the two variables might be negative is because more controls need to be added. Though I added controls for income and whether a census tract was considered urban, other controls that could have been considered could be the amount of physical activity in a census tract, or the level of parks and recreational facilities located within a census tract. After all, the beautiful scenery of the Rockies and the outdoors is what comes to many people's minds when they think about Colorado. This might have been a more interesting factor to look at and something that further research might benefit from.

Finally, a third explanation as to why this interaction between X_1 and X_2 was positive may have to do not with the independent variables, but the dependent variable. While the interaction between food deserts and minorities may have a limited impact on overweightness/obesity outcomes within a census tract, perhaps this may not be the case for other indicators of poor health, such as the prevalence of diabetes or heart disease. Heart disease especially might show a different outcome; after all, the data obtained on obesity was only among adults, and conditions like heart disease are age-related. Perhaps this could also be a fourth explanation for the negative interaction between X_1 and X_2; maybe clustering by age in the future could show different results.

All in all, the results from my research confirm my hypothesis that the more a census tract is a "food desert," the more likely it is that the census tract has higher outcomes of overweightness and obesity. The same positive relationship is shown between the level of minorities within a census tract and overweight/obesity outcomes; however, when considering each minority on an individual basis, one finds that Asians in Colorado do not follow this trend. Despite this confirmation of my hypothesis, the negative coefficient on the interaction terms go against my hypothesis that a positive interaction existed between food deserts, the level of minorities, and overweight/obesity outcomes. The explanations for this result are varied, though given Colorado's level of healthiness it should not be as surprising as one may expect.

Given these results, perhaps more states should look at Colorado's policies with regards to food accessibility. Practices such as considerate zoning laws with regards to food stores, urban food gardens, or open-air farmers markets are only some of the potential practices that policymakers as well as private individuals could consider in order to encourage a healthy lifestyle in one's community.

Works Cited

Arterburn, David; Drewnowski, Adam and Rehm, Colin D. "The geographic distribution of obesity by census tract among 59 767 insured adults in King County, WA." *International Journal for Obesity*, June 2014, 38 (6), 833-839.

Brownell, Kelly D.; Cooksey-Stowers, Kristen and Schwartz, Marlene B. "Food Swamps Predict Obesity Rates Better Than Food Deserts in the United States." *International Journal of Environmental Research and Public Health*, November 2017, 14(1366).

Arcury, Thomas A.; Clark, Heather M.; Hernandez-Pelletier, Mercedes; Quandt, Sara A.; Shoaf, John I. and Tapia, Janeth. "Experiences of Latino Immigrant Families in North Carolina Help Explain Elevated Levels of Food Insecurity and Hunger." *Journal of Nutrition*, August 2006, 136 (10), 2638-2644.

Tables and Figures

Table 1: Regression Results (X_2 = Minority)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
"Food desertness" (X_1)	2.624	3.045	2.929	5.298	6.212	6.076
	(0.559)	(0.556)	(0.556)	(0.973)	(0.969)	(0.968)
Minority (X_2)	0.115	0.129	0.144	0.134	0.153	0.168
	(0.001)	(0.009)	(0.010)	(0.011)	(0.011)	(0.012)
Interaction (X_1X_2)				–0.087	–0.103	–0.102
				(0.026)	(0.026)	(0.026)
Constant	53.009	54.576	52.929	52.365	53.917	52.297
	(0.351)	(0.435)	(0.730)	(0.399)	(0.463)	(0.743)
Observations	1,242	1,242	1,242	1,242	1,242	1,242
Income Control	No	No	Yes	No	No	Yes
Urban Control	No	Yes	Yes	No	Yes	Yes

Note: The dependent variable is the percentage of overweight and obese adults per census tract in Colorado according to the Colorado Department of Public Health and Environment. Parentheses show standard errors. All results in Table 1 were significant at the 0.1, 0.05, and 0.01 significance levels.

Table 2: Regression Results (X_2 = African American)

Variables	(1)	(2)
"Food desertness" (X_1)	3.774***	4.212***
	(0.701)	(0.703)
African American (X_2)	0.100***	0.093***
	(0.033)	(0.034)
Interaction (X_1X_2)	–0.279***	–0.285***
	(0.074)	(0.073)
Constant	56.027	58.753
	(0.284)	(0.634)
Observations	1,242	1,242
Income Control	No	Yes
Urban Control	No	Yes

Note: Dependent variable is the percentage of overweight and obese adults per census tract in Colorado according to the Colorado Department of Public Health and Environment. Parentheses show standard errors. * p<0.1, ** p<0.05, *** p<0.01

Table 3: Regression Results (X_2 = Hispanic/Latino)

Variables	(1)	(2)
"Food desertness" (X_1)	3.685***	3.940***
	(0.826)	(0.820)
Hispanic/Latino (X_2)	0.165***	0.201***
	(0.012)	(0.013)
Interaction (X_1X_2)	–0.049	–0.053*
	(0.030)	(0.024)
Constant	53.012	51.980
	(0.333)	(0.697)
Observations	1,242	1,242
Income Control	No	Yes
Urban Control	No	Yes

Note: Dependent variable is the percentage of overweight and obese adults per census tract in Colorado according to the Colorado Department of Public Health and Environment. Parentheses show standard errors. * p<0.1, ** p<0.05, *** p<0.01

Table 4: Regression Results (X_2 = Asian)

Variables	(1)	(2)
"Food desertness" (X_1)	2.513***	2.552***
	(0.863)	(0.864)
Asian (X_2)	–0.175***	–0.118*
	(0.064)	(0.066)
Interaction (X_1X_2)	–0.075	0.013
	(0.174)	(0.175)
Constant	57.213	59.160
	(0.340)	(0.644)
Observations	1,242	1,242
Income Control	No	Yes
Urban Control	No	Yes

Note: Dependent variable is the percentage of overweight and obese adults per census tract in Colorado according to the Colorado Department of Public Health and Environment. Parentheses show standard errors. * p<0.1, ** p<0.05, *** p<0.01

Table 5: Regression Results (X_2 = Native American/Alaskan)

Variables	(1)	(2)
"Food desertness" (X_1)	3.433***	3.705***
	(0.676)	(0.947)
Native American/Alaskan (X_2)	0.492***	0.405***
	(0.109)	(0.112)
Interaction (X_1X_2)	–0.538***	–0.487***
	(0.148)	(0.149)
Constant	55.542	58.016
	(0.321)	(0.700)
Observations	1,242	1,242
Income Control	No	Yes
Urban Control	No	Yes

Note: Dependent variable is the percentage of overweight and obese adults per census tract in Colorado according to the Colorado Department of Public Health and Environment. Parentheses show standard errors. * p<0.1, ** p<0.05, *** p<0.01

Table 6: Regression Results (X_2 = Non-minority white)

Variables	(1)	(2)
"Food desertness" (X_1)	–3.445*	–4.122**
	(0.676)	(1.864)
Non-minority White (X_2)	–0.135***	–0.168***
	(0.011)	(0.012)
Interaction (X_1X_2)	0.087***	0.102***
	(0.026)	(0.026)
Constant	65.851	69.146
	(0.770)	(0.938)
Observations	1,242	1,242
Income Control	No	Yes
Urban Control	No	Yes

Note: Dependent variable is the percentage of overweight and obese adults per census tract in Colorado according to the Colorado Department of Public Health and Environment. Parentheses show standard errors. * p<0.1, ** p<0.05, *** p<0.01

Figure 1: Relationship between overweight/obesity outcomes and "food desertness" (no controls)

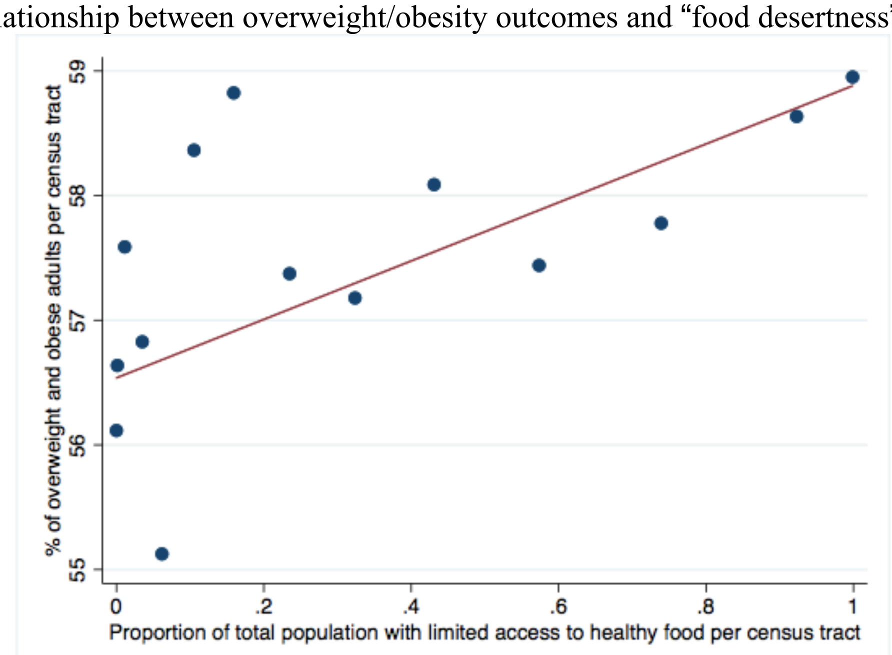

Figure 2: Relationship between overweight/obesity outcomes and "food desertness" (controlling for income and urban)

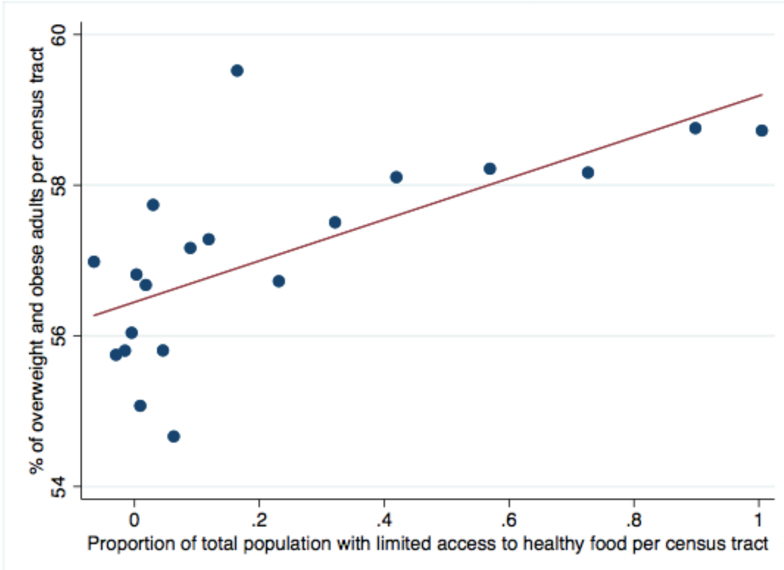

Figure 3: Relationship between overweight/obesity outcomes and percentage of minorities per census tract (controlling for income and urban)

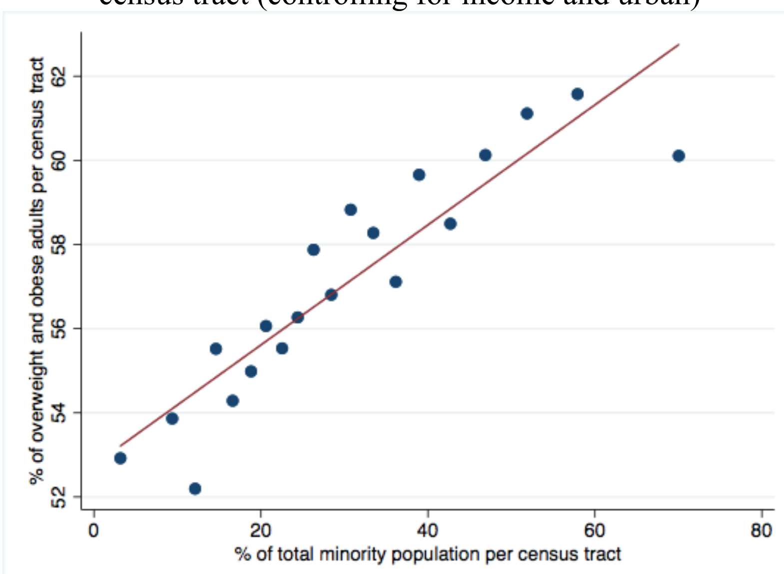

Figure 4: Relationship between overweight/obesity outcomes and percentage of African-Americans per census tract (controlling for income and urban)

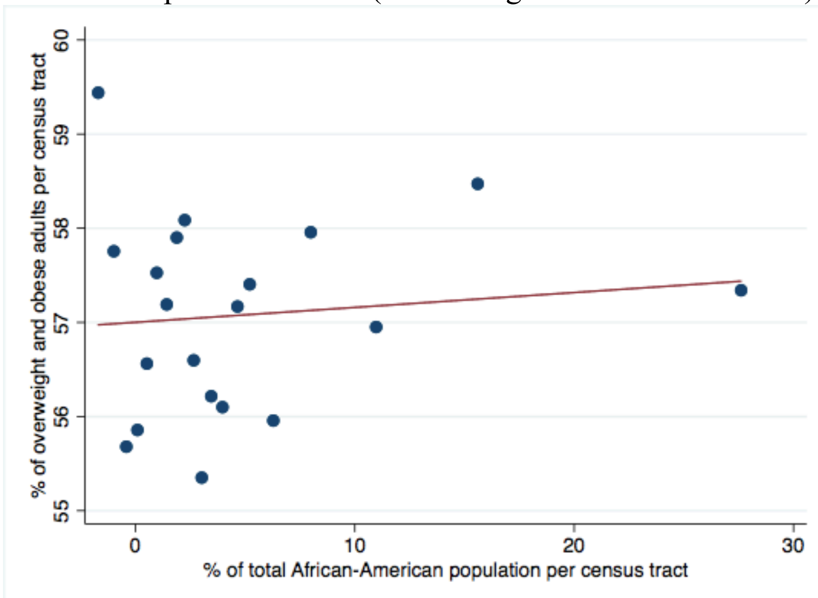

Figure 5: Relationship between overweight/obesity outcomes and percentage of Hispanic/Latinos per census tract (controlling for income and urban)

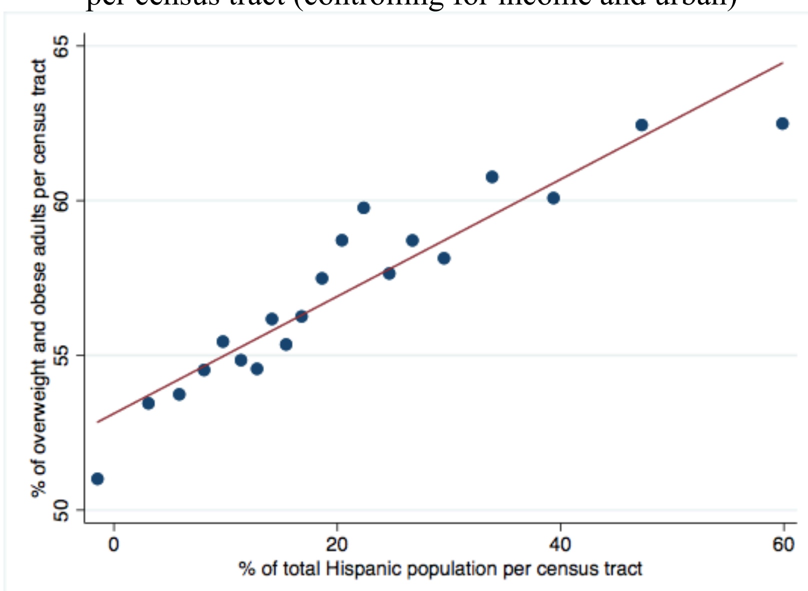

Figure 6: Relationship between overweight/obesity outcomes and percentage of Asians per census tract (controlling for income and urban)

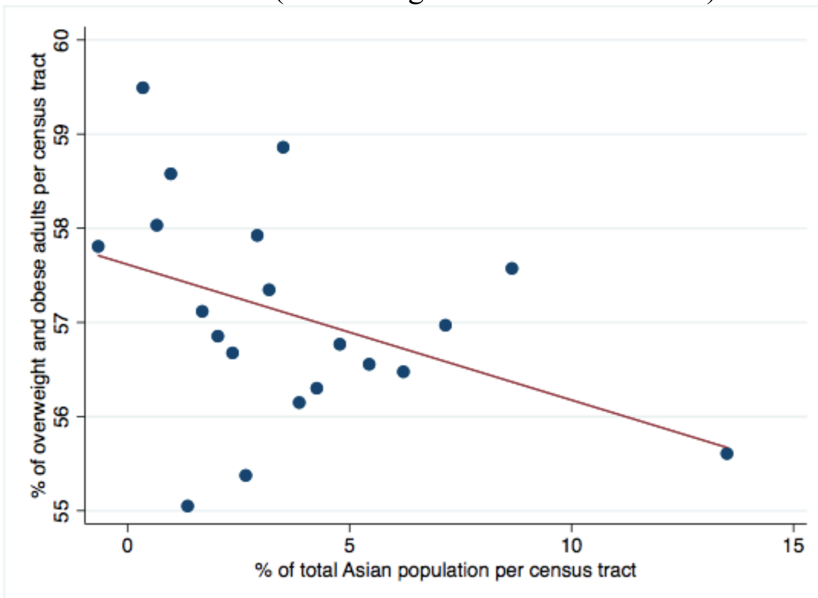

Figure 7: Relationship between overweight/obesity outcomes and percentage of Native American/Alaskans per census tract (controlling for income and urban)

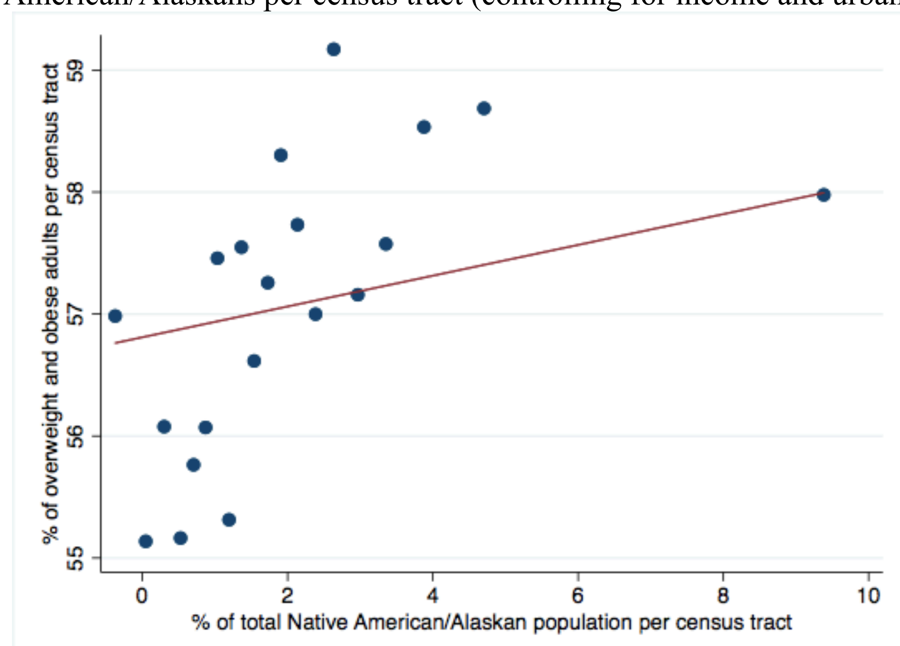

Figure 8: Relationship between "food desertness" and percentage of minorities per census tract (controlling for income and urban)

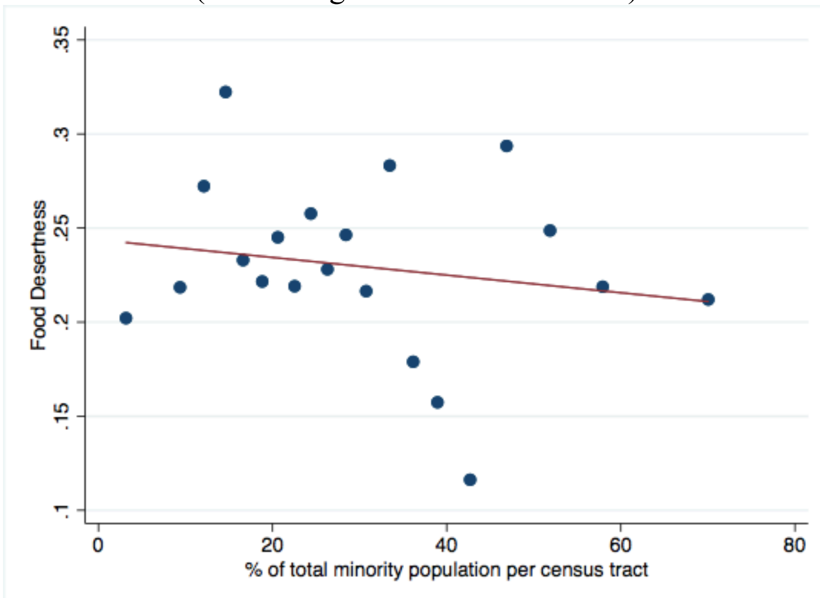

Figure 9: Relationship between "food desertness" and percentage of African-Americans per census tract (controlling for income and urban)

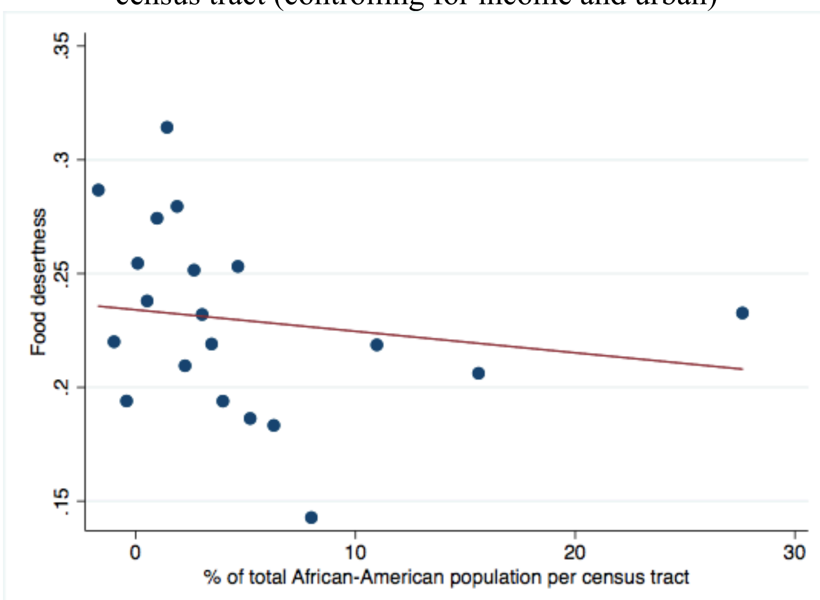

Figure 10: Relationship between "food desertness" and percentage of Hispanic/Latinos per census tract (controlling for income and urban)

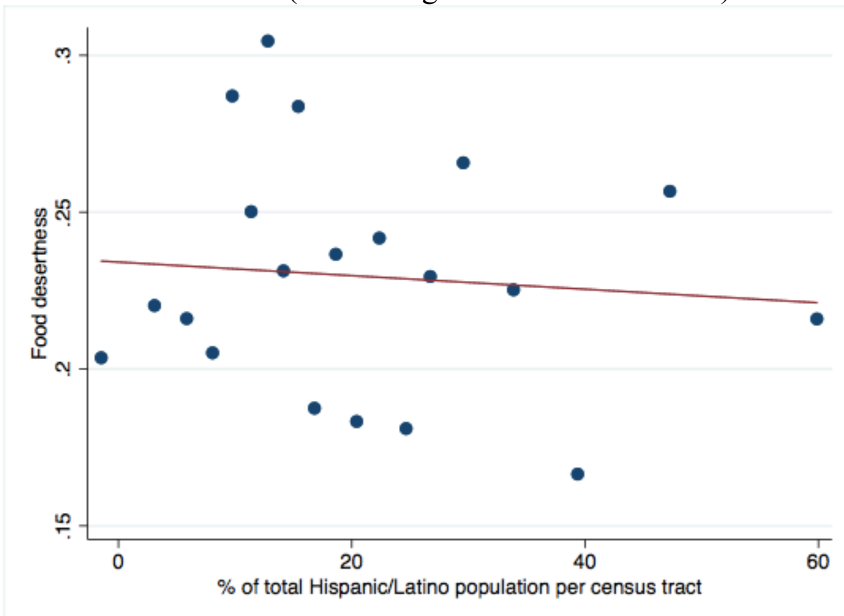

Figure 11: Relationship between "food desertness" and percentage of Asians per census tract (controlling for income and urban)

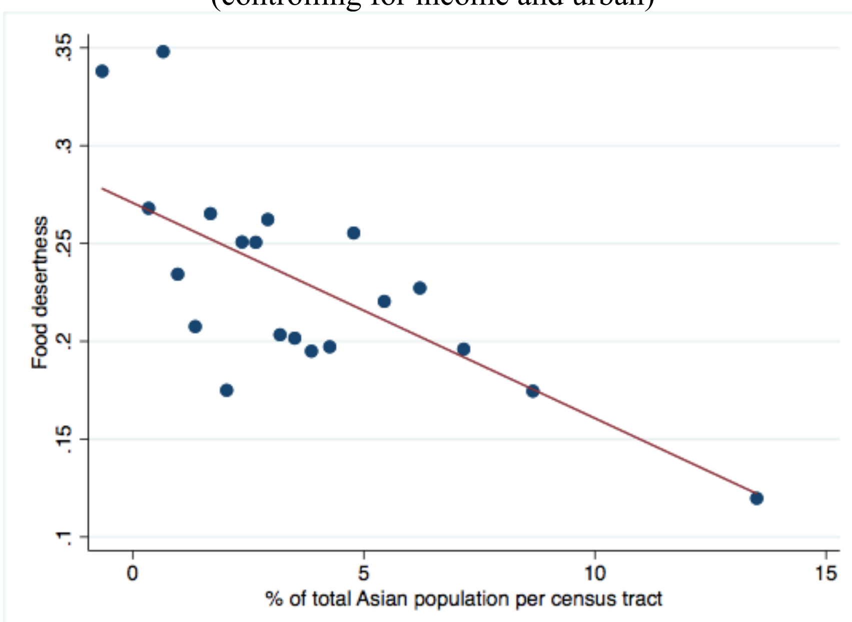

Figure 12: Relationship between "food desertness" and percentage of Native American/Alaskans per census tract (controlling for income and urban)