Claremont Colleges

Scholarship @ Claremont

2018

The Effect of Fast Food Restaurants on Type 2 Diabetes Rates

Grace Bailey
Claremont McKenna College

Claremont McKenna College

The Effect of Fast Food Restaurants on Type 2 Diabetes Rates

submitted to

Professor David Bjerk

by

Grace Bailey

for

Senior Thesis

Fall 2017

December 4, 2017

Abstract

This paper conducts an analysis of county level data to determine the effect of fast food restaurants on type 2 diabetes rates. Due to endogeneity concerns with respect to the location of fast food restaurants, this paper follows the work of Dunn (2010) and uses the number of interstate exits in a given county to serve as an instrument for fast food restaurants. The strength of the instrument, which is theoretically and empirically tested in this paper, imposes some restraints on the interpretation of the findings. Using the Two-Stage Least Squares estimation method, I find that the presence of fast food restaurants has a positive and statistically significant effect on type 2 diabetes rates at the county level.

Introduction

A national epidemic in healthcare has begun to unfold. According to the Centers for Disease Control and Prevention, in the United States today, 30.3 million adults have diabetes. 95% of this population suffers from type 2 diabetes, the form that is not purely based on genetics and can thus be prevented or mitigated. Though typically developing in adulthood, an increasing number of children and young adults are being diagnosed with type 2 diabetes, further fueling the moral and economic concerns of the chronic disease.

Diabetes is the seventh leading cause of death in the U.S. and is the number one cause of kidney failure, lower-limb amputations and adult-onset blindness. In the last 20 years, the number of adults diagnosed with diabetes has more than tripled as the American population has aged and become more overweight or obese. Beyond the ethical considerations and lifestyle challenges this disease poses, the increasing rate of type 2

diabetes has yielded growing economic costs. The Centers for Disease Control and Prevention estimated that diabetes costs an annual amount of $245 billion in the United States, consisting of both medical costs and costs associated with lost work and wages for people with diagnosed diabetes.

Research and statistics around this national challenge are now prevalent, but work to understand the actual drivers, particularly in the field of economics, remains somewhat sparse. Understanding the source and cause of type 2 diabetes will enable legislators to develop effective policies to halt or slow the rising prevalence of this chronic disease.

This paper seeks to provide useful research about one potential driver of type 2 diabetes through an examination of the effect of fast food restaurants on rates of this chronic disease. Specifically, this study will attempt to answer the following question: Does an increase in the number of fast food restaurants per 10,000 people increase the prevalence of type 2 diabetes? To do so, I conduct an econometric analysis using cross-sectional data at the United States county level and an instrumental variable to account for endogeneity in the model. The hypothesis tested in this study is that a rise in the presence of fast food restaurants will increase the prevalence of type 2 diabetes at the county level. This hypothesis is indeed validated by the statistically significant findings in my paper, though the overall strength of the instrument may pose some limitations on the interpretation of the results.

Modeled off of Dunn's (2010) approach, I will account for the issue of endogeneity bias using an instrumental variable and the Two-Stage Least Squares method. Endogeneity in the model primarily derives from the fact that a demand for unhealthy food may affect 1) the location of fast food restaurants, because restaurants are

not placed randomly but rather strategically in a profit-maximizing location, and 2) type 2 diabetes rates, because a steady consumption of unhealthy food will yield negative long-term health outcomes. To address "the possibility that unobserved characteristics associated with the probability of becoming obese may also affect the location decision of restaurant owners," Dunn employs an instrumental variable – the number of interstate exits in a county. Using highway exits, Dunn argues, serves as a valid instrumental variable because the location of highway exits is correlated with the number of fast food restaurants in an area, but unrelated to someone's body mass index, which is the dependent variable in his study. Given these relationships, using highway exits to instrument for the endogenous covariate should remove the issue of endogeneity from the model. Because highway exits are correlated with the endogenous variable in this paper (fast food restaurants) and unrelated to type 2 diabetes rates, this paper will employ the same instrumental variable, following Dunn's approach. I will discuss the instrument's exact use and validity in more detail in the coming sections.
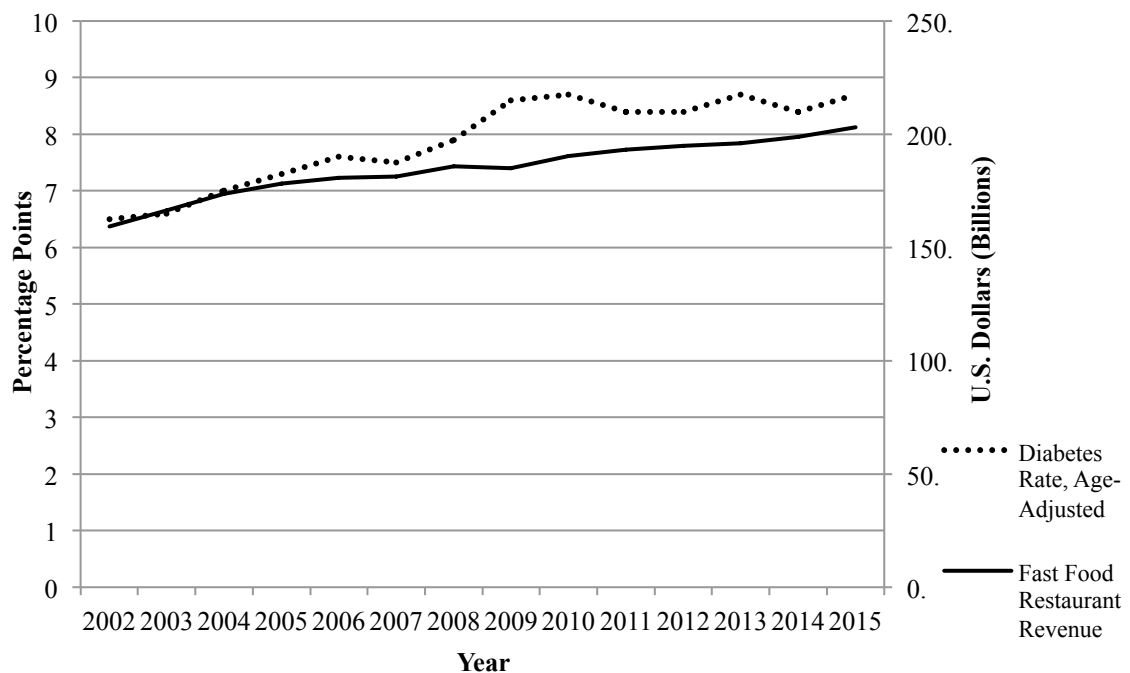
Two economic theories, which may work concurrently or separately, provide the theoretical underpinnings of this empirical analysis. One theory stems from a supply-based argument: As the number of fast food restaurants relative to the population of a county increases, the presence of this unhealthy food has a crowding out effect on other, likely healthier options, namely grocery stores, fast casual restaurants, food trucks, and casual dining restaurants. If an individual in a county chooses to eat out, he or she will have fewer substitutes for fast food restaurants, which may induce the individual to dine at the fast food restaurant.

A second theory covering why the increased presence of fast food restaurants has a causal and positive impact on type 2 diabetes rates is cost-related and can be understood by examining the breakdown of the cost of obtaining a fast food meal. The cost of obtaining a meal has two main components – the actual price paid for the meal and the opportunity cost of obtaining the food. The opportunity cost of obtaining the food is primarily related to the ease of accessibility of the restaurant, namely how much time it takes to drive to the restaurant. Jekanowski, Binkley and Eales (2001) provide a useful theoretical framework for thinking about this economic understanding of food choice. As more fast food restaurants become available to consumers in a given county, the "increase in the number of fast food outlets in a market directly increases quantity consumed by decreasing the costs of obtaining a fast food meal." In other words, more availability equates to, on average, less time driving to the restaurant and potentially less time waiting in line at the restaurant, depending on the demand in the given market. This lower opportunity cost will decrease the overall cost per fast food meal. Additionally, as more fast food restaurants flood a market, competition is induced and actual menu prices will lower as well.

The implication of these two effects, which may occur independently or simultaneously, is a lower overall cost of consuming an unhealthy meal. Because it is common knowledge that the food one eats affects one's health, a rational consumer should factor the health consequences of eating a particular meal into his or her overall cost equation. However, given a lower overall cost of consuming an unhealthy meal due to an increase in fast food restaurants, a discounting effect of the negative health costs associated with eating an unhealthy meal will occur.

Another interesting nuance to this discussion, presented by Jekanowski, Binkley and Eales (2001), is that increased availability and accessibility to fast food restaurants will yield greater sensitivity to menu prices because "menu prices will account for a larger portion of the 'full' price of" a food choice. This means that within the cost equation of consuming a fast food meal, the menu price will be of greater consideration as opportunity costs of obtaining the meal, due to more availability, will decrease. According to this theory, as consumers consider the various price ranges of different types of restaurants, the prices offered by fast food restaurants, which tend to be the least expensive restaurant option, will be more appealing in comparison to other types of restaurants, further perpetuating the increase in consumption of fast food.

Figure 1. Growth in Diabetes Rate and Fast Food Restaurant Revenue in the United States, 2002-2015

Regardless of the exact drivers, unhealthy eating choices, like those at a fast food restaurant, will yield negative health outcomes over the long-term, such as type 2 diabetes, hypertension and other chronic conditions. As Figure 1 demonstrates, both type 2 diabetes rates and fast food restaurant industry revenue in the United States are trending upward, motivating an analysis to determine the relationship between these two factors. This paper focuses on the rising prevalence of type 2 diabetes, specifically exploring the impact that the presence of fast food restaurants has on these increasing rates.

Literature

Given the multidisciplinary nature of the topic, a relevant literature review spans the fields of economics, medicine, public health, psychology, and public policy. The topic of my empirical study fills an important gap in the literature: Causal relationships affecting the increase in prevalence of type 2 diabetes have not been explored extensively in an econometric setting. The majority of health-related studies are focused on what people actually eat, without necessarily considering the underlying drivers of those decisions. Therefore, more research is needed in the realm of how a built environment, especially the various food choices one faces on a daily basis, may impact his or her health outcomes.

The primary model for this study is Dunn's (2010) analysis of the effect of fast food restaurants on obesity. Dunn examines the relationship between the concurrent increase in fast food restaurants and obesity rates using cross-sectional, individual-level data. Specifically, Dunn explores the effects of fast food restaurant availability on obesity rates in the context of gender (male or female) and race (white, black or Hispanic) across

three different population density categories (low, medium and high; or rural, suburban and urban). Dunn employs the instrumental variable of interstate exits to account for the issue of endogeneity, running both Two-Stage Least Squares and Ordinary Least Squares regressions and comparing the results. Dunn finds that fast food restaurant availability increases the body mass index (BMI) amongst females and non-white racial groups in medium-density counties. This implies that the presence of fast food restaurants has a causal relationship with higher BMIs, specifically for women and people who identify as African American or Hispanic.

Other economic literature explores the relationship between one's income and one's choice to purchase fast food, which is relevant as it relates to the theory underlying my empirical study. Drewnowski and Darmon (2005) note that low-income consumers are more likely to purchase and eat fast food, as opposed to going to a full-service restaurant. The hypothesis that poverty and obesity are linked is driven by the fact that unhealthy food, like "refined grains, added sugars and added fats have… a low energy cost." Additionally, Drewnowski and Darmon point to low-income consumers living in "areas with less physical access to healthier foods," which further increases the cost of healthy food and lowers the cost of unhealthy food. Because options at fast food restaurants are less healthy, the logical next step is a less healthy food choice, and ultimately a less healthy diet. Additionally, the fact that low-income individuals are more likely to live in areas with less physical access to healthy food provides a partial explanation for the disproportionate negative health impact of fast food restaurant presence on non-white racial groups demonstrated in Dunn's study.

Powell, Han, and Chaloupka (2010) explore the socioeconomic implications of adolescent dietary behavior and obesity prevalence. As of 2006, fast food restaurants made up 30% of the entire restaurant industry. Given the rising trend of fast food chains' overall representation in the restaurant industry (up from 17% in 1997), their share of the overall restaurant market, following this historical upward trend, has likely risen above 30% today. Interestingly, Powell, Han, and Chaloupka find no significant relationship between the availability of fast food and teen BMI. However, they did find that the price of fast food had a significant effect on teen BMI, with an estimated price elasticity of – 0.08. Their findings suggest that it is "the low cost of fast food and not its widespread availability that affects youth diet and obesity, particularly for youths at risk of becoming overweight and those in lower to middle-SES families who are most price sensitive." However, the authors of the study also note a significant association between higher income, grocery store availability and access, and lower BMI, suggesting that place and choice do have a significant impact on people's food choices, and consequently their health outcomes.

Other studies that examine the relationship between food choice and income support the findings of Powell, Han, and Chaloupka. Powell et al. (2001) reported that low-income neighborhoods, when compared with high-income neighborhoods, have 1.25-1.3 times more fast food restaurants. This study also shows that although restaurants of all types are less available in predominantly racial or ethnic minority communities, a significantly higher proportion of restaurants in those communities are those classified as fast food restaurants. Because income and race are highly correlated in the United States,

living in predominately black versus mainly white neighborhoods may contribute to racial differences in obesity rates.

The psychology literature offers useful research on the underlying effects of fast food restaurants on rates of type 2 diabetes. Burgoine et al. (2014) found evidence of a dose-response association, meaning increased exposure to fast food restaurants drives an increase in consumption at those restaurants. According to the study, a greater presence of fast food restaurants creates an unhealthy "food environment," meaning the people living in areas of high exposure to fast food are more likely to over consume "energy dense, nutrient poor foods" and be overweight or obese. Additionally, studies by Harris, Bargh, and Brownell (2009) and Reisch et al. (2013) have demonstrated that exposure to unhealthy food advertising will induce people to make worse health choices. Harris, Bargh, and Brownell (2009) note that "advertising for food and beverages communicates potentially powerful food consumption cues, including images of attractive models eating, snacking at non-meal times, and positive emotions linked to food consumption," which in turn drives people to consume more of the typically unhealthy products that are advertised in response to these cues. Thus, applying this framework, greater exposure to fast food restaurants could have a similar, though more indirect, form of food advertising, which could yield an increase in overall consumption of unhealthy food amongst those being exposed to the restaurants.

Public health studies offer additional insight into the question of fast food restaurants' impact on rates of type 2 diabetes. Though health is an inherently complex research topic, deconstructing the many factors that yield unhealthy outcomes, and understanding their significance, is useful for guiding individual health choices and

public policy. A 2015 report published by the California Department of Public Health's Office of Health Equity presented a breakdown of the different factors that account for a person's health status. According to the report, the medical care one receives only accounts for 20% of a person's total health, suggesting that the bigger picture of health lies elsewhere. This same estimate suggested that the rest of a person's health is accounted for by the following: 30% by their health behaviors, 40% by social and economic factors, and 10% by their physical environment.

Of these drivers, this paper will primarily focus on the physical or built environment in which people live, in particular the local "food environments." The Centers for Disease Control and Prevention define a food environment as the "physical presence of food that affects a person's diet; a person's proximity to food store locations; the distribution of food stores, food service, and any physical entity by which food may be obtained; or a connected system that allows access to food." The term food environment is a useful catch-all term relating to the ways in which social and economic factors influence how people make choices about food, given their built surroundings and options.

From the public policy and law literature, Freeman (2007) emphasizes the impact of fast food restaurants on people's health through the lens of income and race. She notes the "prevalence of fast food in low-income urban neighborhoods across the United States, combined with the lack of access to fresh, healthy food" as the primary contributor to the "disproportionate incidence of food-related death and disease among African Americans and Latinos as compared to whites." For example, West Oakland, California had one supermarket and 36 liquor and convenience stores for their population of 30,000 people,

who are primarily African American and Latino. Studies like Freeman's point to the importance of controlling for race and income, given the extent to which these factors may have an effect on fast food restaurant location and resulting health outcomes.

Inagami et al. (2006) conduct an empirical study to explore their hypothesis that people living in lower income neighborhoods have a higher BMI. They report: "The better predictor of BMI was not the individual's specific choice of grocery store but the location of where the average resident shopped. This suggests a group-level influence that might be related to pressures to conform to local norms." The theory behind this analysis is that people with fewer healthy options will eat less healthily, which is propelled by the social environment and expectations of where one should shop for food. Given the logic that proximity of grocery stores has an effect on people's weight, it is plausible to examine the effects of proximity of unhealthy food sources, namely fast food restaurants, on a longer term health outcome – type 2 diabetes.

Currie et al. (2010) examine the relationship between the increase in the number of local fast food restaurants and obesity, specifically for young teens and pregnant women. They found that the "presence of a fast food restaurant within one-tenth of a mile of a school is associated with at least a 5.2 percent increase in the obesity rate in that school." For pregnant women, a fast-food restaurant within 0.5 miles of their home yields a 1.6 percent increase in the probability of gaining over 20 kilos. Their paper provides a useful description of the effects that the proximity of a fast food restaurant may have on one's health choices. One could argue that fast food restaurants in closer proximity merely provide the potential for substitution away from unhealthy food prepared at home, but does not affect the aggregate amount of unhealthy food consumed. However, a more

probable understanding of the relationship between distance and health choices is that "proximity to a fast food restaurant could lower the monetary and nonmonetary costs of accessing unhealthy food," thus causing people to make less healthy decisions than they would typically. Over time, this effect of fast food restaurant presence on health decisions could have an impact on long-term health, which is the relationship examined in this paper.

Jeffery et al. (2006) did not find an association between the proximity of fast food restaurants and eating at fast food restaurants or higher BMIs, contrary to the central findings of Dunn. Though this may suggest conflicting evidence to the research question explored in this article, Jeffery et al. only consider consumption of fast food, not whether or not the location may have an effect on long-term health outcomes, like type 2 diabetes. Since these rates tend to develop in the long run, and given that food environments 10 years ago and 20 years ago were vastly different than those today, the hypothesis of this paper is that the long-term and more widespread effects of an unhealthy food environment are now beginning to reach their full level of impact. Additionally, Jeffery et al. acknowledge that their study may have a key methodological fault: the homogeneity of their restaurant proximity metric, which may explain their insignificant results of the effects a food environment has on one's likelihood of eating at a fast food restaurant or having a higher BMI. Finally, their data is gathered from a survey of 1,033 individuals. This format, given that the survey participants were asked questions about their health, may face issues of response bias, in that participants would subconsciously think more highly of their health when responding to the survey, which could skew the results of the study.

In order to extend the analysis that Dunn explores (fast food restaurants' effect on obesity) to that which is examined in this paper (fast food restaurants' impact on type 2 diabetes rates), it is useful to inspect the medical literature as a means of understanding the relationship between obesity and type 2 diabetes. Ginter and Simko (2013) highlight the increasing prevalence of type 2 diabetes as a function of rising obesity rates. While the link between obesity and type 2 diabetes has been strongly established, the exact biochemical relationship is still being explored, though strong hypotheses point to biochemical factors such as abnormalities in free fatty acids, adipokines, and leptin. Obesity is predicted to be the explanatory factor for 70-90% of individuals in the United States and the European Union who suffer from type 2 diabetes. Furthermore, backing up Dunn's findings, Burgoine et al. (2014) report that the more individuals are exposed to unhealthy food options, like fast food restaurants, the more likely they are to consume marginally more takeout food, have a higher body mass index, and face likelier odds of obesity.

A key innovation of this paper is to employ a different outcome variable, type 2 diabetes rates rather than BMI (the proxy for obesity rates). Rothman (2008) provides evidence of inherent flaws of the BMI metric. The BMI has problems with sensitivity, as it does not reflect the physiological body fat and muscle mass changes that naturally occur with age, and specificity, in terms of misclassification problems, which, in cross-sectional studies, can yield misinterpretations of effects related to obesity. Despite its shortcomings, the majority of obesity data is measured using BMI, and the literature examining the impact of food environments on health has primarily focused on using BMI as the key measure for one's health. Thus, a gap exists in the literature, which can

be filled by using a health metric that is not BMI. In the case of this paper, the different measure for one's health is rates of type 2 diabetes.

Given this medical literature, my extension of Dunn's analysis, by changing the outcome variable from obesity to type 2 diabetes rates, is worth consideration. Additionally, the existing literature and focus on the rising prevalence of type 2 diabetes presents both a space and motivation for this paper to add to the research surrounding food environments', specifically fast food restaurants', impact on people's health.

Data

The analysis in this paper is conducted using cross-sectional, county-level data from four sources: the United States Department of Agriculture's Food Environment Atlas, the Centers for Disease Control and Prevention, Dunn (2010), and the United States Census Bureau. While the Food Environment Atlas, Centers for Disease Control and Prevention, and United States Census Bureau data sets provide information on 3,142 counties and county-equivalents – namely parishes, organized boroughs, census areas, and independent cities – in the United States, the instrumental variable data set has 1,147 observations. Thus, the main regressions in this paper are limited to 1,147 counties, but this amount of observations still provides a strong sample size and the opportunity for a robust analysis. Summary statistics of the variables discussed in this section and employed in the analysis for this paper are presented in Table 1.

Table 1. Summary Statistics of 2011 County-Level Data from the Food Environment Atlas, the Centers for Disease Control and Prevention, the U.S. Census Bureau, and Dunn (2010)

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Diabetes Percentage	3,138	10.98384	2.311213	3.9	21.6
Diabetes Number	3,142	7,048.808	20,072.16	9	580,158
Fast Food Restaurants	3,106	69.32357	238.3487	0	7211
Population	3,142	9,8263.02	312,946.4	82	9,818,605
Population Density	3,141	259.4872	1,724.708	0	69,468.4
Median Income	3,108	43,059.98	10719	20577	119,075
White	3,142	78.31531	19.85771	2.667918	99.16318
African American	3,142	8.74928	14.42369	0	85.43878
Hispanic	3,142	8.282409	13.19277	0	95.74477
Male	3,142	49.96903	2.200891	43.2	72.1
Female	3,142	50.03097	2.200891	27.9	56.8
Poverty Rate	3,133	18.35822	6.747441	2.9	48.1
Interstate Exits	1,179	10.3503	18.97524	0	229
Fast Food Restaurants per 10,000 people	3,106	37.05848	199.4385	0	7,561.375

The outcome variable of this study is the percentage of people in a county who had type 2 diabetes in 2011, which is data obtained from the Centers for Disease Control and Prevention. Following the topic of this paper, the independent variable of interest is

the number of fast food restaurants in a given county in 2011. This data comes from the United States Department of Agriculture's Food Environment Atlas, an expansive data set with county-level information on population, socioeconomic status, stores and restaurants, prices and taxes, food insecurity, local food sources, health metrics, provisions of federal assistance, and more. For the purpose of scaling the number of fast food restaurants to the population of each county for the regressions, I created a new variable: fast food restaurants per 10,000 people. I generated this new variable by multiplying the number of fast food restaurants by 10,000 and dividing this product by the 2011 population.

The third key variable in this study is the instrumental variable – highway exits in each county in 2005. I thank Richard Dunn for kindly providing me with the interstate exits data. The definition of the interstate exits data used in Dunn's study is: "any exit from any roadway with an Eisenhower Interstate designation except exits for nonservice rest stops, tolls, or private ranch roads." While the year of the highway exits data is six years prior to the rest of the data in this study, the relatively permanent nature of highway exits due to limited changes in this type of roadway infrastructure implies that the values in 2011 are likely quite similar to those in 2005.

Given the variation in factors that could influence a long-term health outcome like type 2 diabetes, I employ a number of controls to refine the model, including: population density per square mile of land (which is data obtained from the Food Environment Atlas), race data, specifically the percentage of people who identify as African American or Hispanic in each county (which is also from the Food Environment Atlas), and the poverty rate (which I acquired through the Census Bureau). Poverty rates from this data

set are defined as the percentage of families or individuals in the county whose combined pre-tax income is less than their relevant poverty threshold, of which there are 48 possibilities. The weighted average poverty threshold for a family of four in 2011 was $23,021.

Examining the dependent variable, the type 2 diabetes percentage at the county level, reveals significant variation across counties, with the minimum percent in a county of 3.6% and a maximum value of 23.5%. Many of the control variables display variation as well: The smallest population size is 82 people and the largest is 9,818,605 people. Population density, which is a key variable in both this study and Dunn's paper, ranges from zero people per square mile of land to 69,468.4 people per square mile of land, reflecting the significant geographical and demographic variation of the United States, which demonstrates the usefulness of employing county level data in an econometric setting.

Additionally, the minimum poverty rate is 2.9%, which is a significant contrast to the maximum of 48.1%. Racial and gender demographics also have a wide range, with some counties having a racial breakdown of 0% African American and/or Hispanic population and others consisting of a large majority of the population represented by a single racial minority (maximums of 85.43% and 95.74% for African Americans and Hispanic/Latino people, respectively). As made evident by the review of the literature in the last section, controlling for race and income variation is key given their significant impact on fast food restaurant location and resulting health outcomes.

Methods

The goal of this paper is to establish a causal relationship between fast food restaurants and type 2 diabetes rates. The theoretical multiple linear regression model that I will estimate using Ordinary Least Squares (OLS) is:

$$(1)\ Type\ 2\ Diabetes\ Rate$$

$$= \alpha + \beta_1 \times Fast\ Food\ Restaurant$$

$$+ \beta_2 \times Low\ population\ dummy\ variable$$

$$+ \beta_3 \times Medium\ population\ dummy\ variable + \beta_4 \times Poverty\ Rate$$

$$+ \beta_5 \times \%\ African\ American + \beta_6 \times \%\ Hispanic$$

$$+ \beta_7\ Population\ Density + \mathcal{E}$$

In this equation, the βs represent the coefficients derived from the regression and the α is the intercept in the model. The $\mathcal{E}$ in Equation (**1**) is the error term, which captures unobservable factors that are not controlled for in the model, or in other words anything that is not a covariate. The specific reasons for creating dummy variables to classify populations by size will be addressed later on in this section. As is evident in the equation, I control for population (both categorically and in terms of population density), the poverty rate, and race (both African American and Hispanic population percentages at the county level) in this model.

Given the issue of endogeneity, which I touched upon earlier, this OLS regression will not provide sufficient and valid results. Thus, I will employ an instrumental variable to account for the issue of endogeneity bias in the model. As touched upon in previous sections, endogeneity is a problem in this analysis because what causes a fast food chain to open a restaurant in a certain location (i.e. a given county) is partially explained by a

demand for unhealthy eating, which the fast food organization may track by observing rates of overweight and obese individuals in an area. This same factor, demand for unhealthy eating, is also very likely related to higher rates of type 2 diabetes, a plausible argument given that type 2 diabetes is a condition largely stemming from a poor diet. Given that both the location of fast food restaurants and rates of type 2 diabetes are strongly correlated with a demand for unhealthy eating, which is captured in the error term of the model, an estimation of the model would produce invalid results. However, using an instrumental variable, which is interstate exits in a county in this paper, can remove the issue of endogeneity from the model.

Several conditions, laid out by Wooldridge (2016), must be met to ensure the use of a valid instrumental variable. The first requirement is that the instrumental variable must be relevant, meaning it is correlated with the independent variable of interest, in this case fast food restaurants. To establish this correlation, I regress fast food restaurants per 10,000 people on the instrument, interstate exits per 10,000 people, and the control variables, as shown in Table 2.

As we see from the results, the instrumental variable of highway exits is indeed positively correlated with the explanatory variable of interest, fast food restaurants per 10,000 people. The coefficient on interstate exits in Table 2 demonstrates a positive relationship between the instrument and the endogenous covariate, which is needed to satisfy one of the conditions of a valid instrument. However, the p-value of 0.110 is not significant at the 5% or 10% level, as expected.

Table 2. Fast Food Restaurants per 10,000 People Regressed on Control Variables and Instrumental Variable (OLS)

Variable	Coefficient	P-value
Interstate Exits per 10,000 People	0.8326046	0.110
Poverty Rate	0.0328625	0.846
Hispanic	– 0.0285634	0.771
African American	0.082043	0.355
Constant	10.64742	0.001

Number of observations = 1,151

The fact that the relationship between the instrument and the endogenous explanatory variable is not of the desired statistical significance (a p-value of 0.05) may be explained by several factors. The presence of interstate exits in a county may affect other health behaviors, like physical activity, which would in turn impact type 2 diabetes rates. For example, areas with more interstate exits, and therefore more traffic and pollution, likely make exercising outside more difficult. If this is indeed the case, then interstate exits would affect unobserved factors captured in the error term of the model, yielding slightly biased results.

There are additional explanations for the empirical results in Table 2 indicating that interstate exits serve as a weaker instrument than expected. Dunn cites the fact that interstate exits may draw other unhealthy food providers beyond just fast food restaurants, such as small local marts that primarily sell junk food. Because unhealthy food is a primary cause of type 2 diabetes, not including data from these other places of business may misreport the effect of fast food restaurants on type 2 diabetes rates.

Additionally, counties with more interstate exits may attract people who have different health preferences than those living in areas without interstate exits. If this is the case, interstate exits would have an effect on type 2 diabetes rates, through this location preference based on health habits. Both of these criticisms are worth noting and provide further insight into why the instrument's correlation with the variable it will instrument for, fast food restaurants, is not one of statistical significance.

Despite these concerns, the fact that the p-value on interstate exits per 10,000 people is 0.110, relatively close to a p-value representing statistical significance, and the coefficient on interstate exits demonstrates that there is a positive relationship between highway exits and fast food restaurants per 10,000 people, the use of the instrument is still of valid consideration. However, when interpreting the results from the Two-Stage Least Squares regressions that employ the instrumental variable in both this study and Dunn's paper, the statistical insignificance of the instrument is an important nuance to consider.

Beyond the first condition, correlation with the independent variable of interest, Wooldridge cites two other parameters that an instrumental variable must satisfy in order to be considered valid. The second condition that the instrumental variable must meet is excludability, which implies that the instrument should not causally affect the outcome variable. Though there may be some relationship between interstate exits and type 2 diabetes rates, as touched upon, a thought experiment would suggest that interstate exits alone is not a strong enough factor to have a causal effect on rates of type 2 diabetes. According to the third condition laid out by Wooldridge, the instrument must be exogenous, implying that interstate exits cannot be correlated with the unobservable

factor causing the endogeneity problem, which in the case of this paper is primarily a demand for unhealthy eating. While this relationship cannot be proven, as demand for unhealthy eating is not a measured variable in the data set, it is theoretically valid to presume that the presence of interstate exits and a demand for unhealthy food are likely unrelated.

The instrument of interstate exits has flaws, given that it does not perfectly meet all three criteria, particularly only loosely meeting the first condition. However, the fact that interstate exits to a certain extent satisfy the three conditions required for a valid instrumental variable implies that it is a useful technique to employ the highway exits data in the model as a means of accounting for endogeneity bias. Moving forward under the assumption that the instrument is considered adequate in addressing the issue of endogeneity, I will employ the interstate exits data to conduct Two-Stage Least Squares regressions in contrast to the results of the Ordinary Least Squares regression.

Because county population is implicitly controlled for through the denominator of the independent variable – fast food restaurants per 10,000 people in the population – it is useful to conduct two different estimations of the model using Two-Stage Least Squares: one with only an indirect population control and one with both an indirect and direct population control. In this case, the indirect control refers to implicitly controlling for county population through the denominator of the fast food restaurants independent variable and through county population size categories, but omitting an actual population or population density covariate from the model. The more direct method of controlling for population size is by both implicitly controlling for county population through the denominator of the fast food restaurants variable and controlling for population density.

Of the two key regressions with the instrumental variable, this more direct approach will be examined first.

The first multiple linear regression model that employs the instrumental variable and is estimated using the Two-Stage Least Squares method is:

(2) **Stage 1**: $Fast\ Food\ Restaurant\ (\widehat{X}\ )$

$$= \alpha + \beta_1 \times Interstate\ Exit + \beta_2 \times Population\ Density$$

$$+ \beta_3 \times \%\ African\ American + \beta_4 \times \%\ Hispanic + \beta_5 \times Povery\ Rate$$

$$+ \mathcal{E}$$

$$\widehat{X} = Z(Z'Z)^{-1}\ Z'X$$

Stage 2: $Type\ 2\ Diabetes\ Rate$

$$= \alpha + \beta_1 \times \widehat{X}_{Fast\ food\ restaurant} + \beta_2 \times Population\ Denisty$$

$$+ \beta_3 \times \%\ African\ American + \beta_4 \times \%\ Hispanic$$

$$+ \beta_5 \times Poverty\ Rate + \mathcal{E}$$

As we see in Equation (2), the first model using the instrumental variable of interstate exits employs four control variables, the instrumental variable and the independent variable of interest, fast food restaurants per 10,000 people. The poverty rate, African American population percentage and Hispanic population percentage covariates are important to include given the extensive literature demonstrating the effect of income and race on health outcomes. Thus, controlling for these factors is essential in all regressions examining the influence of any factor on a health outcome. In Equation (2), α is the intercept, the βs represent the coefficients derived from each stage, and $\mathcal{E}$ is the error term. The first stage of the Two-Stage Least Squares estimation is to regress the endogenous explanatory variable on the instrumental variable and the other covariates.

This regression yields $\widehat{X}$, which is the predicted value of X (fast food restaurants) from the regression of X on Z, where Z is the predicted values of the instrumental variable and the control variables. These predicted values are then used to conduct the second stage regression, in which type 2 diabetes rates are regressed on $\widehat{X}$ and the other independent variables. The results of this second stage will yield the coefficient on fast food restaurants per 10,000 people without endogeneity to the extent that the instrument is considered valid.

In order to conduct a regression with a slightly less direct approach to controlling for county population size to avoid conflicting with the implicit control of population via the fast food restaurants variable, I create three dummy variables based on the population distribution of United States counties. Table 3 presents the population distribution.

Table 3. United States County Population Distribution, 2011

Percentiles	Corresponding Values
1%	968
5%	2,908
10%	5,208
25%	11,113
50%	25,872
75%	66,861
90%	197,465
95%	422,718
99%	1,202,362

Number of observations = 3,142

The county population dummy variable classifications are as follows: Low population counties are counties with populations of 15,000 people or less; medium population counties are those with between 15,001 and 70,000 residents; large population counties are classified as counties with a population between 70,001 and 9,818,605 people. Clearly, these population breakdowns appear heavily skewed toward low population counties. This category bias toward low population counties may be partially explained by the data limitations from having fewer counties to work with, due to the more limited number of observations of interstate exits, the instrumental variable. However, the more illuminating reason for this seemingly heavy concentration of low population counties is that, according to an analysis by Business Insider, half of the United States population was clustered in just 146 counties in 2013. This was likely the case two years prior, which explains why the 2011 data in this study has this type of population distribution. The important takeaway of the statistic is that only a small number of counties have extremely large populations, while the large majority of United States counties have smaller populations, capturing just a small portion of the other half of Americans not living in the 146 densely populated counties. This explanation is further validated by an analysis of the data. The counties classified as low population counties have 1,032 observations; medium population counties have 1,351 observations; and high population counties have just 120 observations.

Creating dummy variables, as opposed to just controlling via the continuous variable of population in 2011, is important because population is already implicitly controlled for in the denominator of the fast food restaurants per 10,000 people independent variable. Thus, creating categorizations of county population removes the

possibility of having two independent variables in the model that directly relate. Additionally, using these population categorizations, as opposed to just relying on the manipulation of the fast food restaurants dependent variable, is essential to fully account for the ways in which population size will affect type 2 diabetes – the most obvious being that more people living in a county statistically increases the likelihood of a greater incidence of type 2 diabetes.

    For the second Two-Stage Least Squares regression, I control for population size via both the denominator of the independent variable of interest and through two dummy variables – low population and medium population counties (with the third dummy variable, the high population county classification, not included in the regression).

    The second multiple linear regression model with the instrumental variable that is estimated using Two-Stage Least Squares is:

(3) **Stage 1**: $Fast\ Food\ Restaurant\ (\widehat{X})$

$$= \alpha + \beta_1 \times Interstate\ Exit + \beta_2 \times \%\ African\ American$$
$$+ \beta_3 \times \%\ Hispanic + \beta_4 \times Poverty\ Rate$$
$$+ \beta_5 \times Low\ Population\ dummy\ variable$$
$$+ \beta_6 \times Medium\ Population\ dummy\ variable + \mathcal{E}$$

$$\widehat{X} = Z(Z'Z)^{-1}\ Z'X$$

Stage 2: $Type\ 2\ Diabetes\ Rate$

$$= \alpha + \beta_1 \times \widehat{X}_{Fast\ food\ restaurant} + \beta_2 \times \%\ African\ American$$
$$+ \beta_3 \times \%\ Hispanic + \beta_4 \times Poverty\ Rate$$
$$+ \beta_5 \times Low\ Population\ dummy\ variable$$
$$+ \beta_6 \times Medium\ Population\ dummy\ variable + \mathcal{E}$$

As we see in Equation (**3**), the second model using the instrumental variable of interstate exits employs three control variables, two dummy variables, the instrument, and the explanatory variable of interest (fast food restaurants per 10,000 people). Like Equation (**2**), in this equation, α is the intercept, the βs represent the coefficients derived from each stage, $\mathcal{E}$ is the error term, and race and income controls are included. $\widehat{x}$ is the predicted value of X from the regression of X on Z, where Z is the predicted values of the instrumental variable and the control variables. These predicted values are then used to conduct the second stage regression, in which the type 2 diabetes rates are regressed on these variables to yield the key results demonstrating the relationship between fast food restaurants and rates of type 2 diabetes.

The results of the OLS and the Two-Stage Least Squares estimates and analysis of the findings are presented in the following section, including a comparison of the findings of this paper to Dunn's results, which will provide useful context for the discussion.

Results

The empirical analysis in this paper is multi-staged, beginning with an OLS regression and building up to two final Two-Stage Least Squares estimations, as highlighted in the former section. Before delving into the key regressions used to examine the causal relationship between fast food restaurants and type 2 diabetes rates, I conduct several simple regressions to empirically justify the theoretical model, particularly the control variables. Specifically, I regress the variable measuring type 2 diabetes rates on all of the controls that will be employed in the various forthcoming OLS and Two-Stage Least Squares regressions to confirm the expected correlation and

statistical significance. The results from regressing the type 2 diabetes values separately on all of the controls are shown in Table 4. As we see, all of the control variables are positively correlated with type 2 diabetes. Additionally, each of these relationships is statistically significant, at either the 1% level or the 5% level.

Table 4. Simple Regressions of Dependent Variable on Control Variables (OLS)

Regression ("Y on X")	Coefficient	P-value
Diabetes Number on Population Density per Square Mile of Land	4.11244	0.000
Diabetes Number on Number of Fast Food Restaurants	5.609231	0.000
Diabetes Rate on Physical Inactivity Value	0.3215169	0.000
Diabetes Rate on Poverty Rate	0.0739191	0.000
Diabetes Number on Diabetes Screening Value	7287.983	0.020

Number of observations = 3,142

Upon confirming that the control variables align with the theoretical expectations, I conducted more comprehensive regressions. I first estimated the model from Equation **(1)** by conducting an OLS regression. These results are presented in Table 5.

This regression yields a positive, though small, coefficient for the explanatory variable of interest, fast food restaurants per 10,000 people. Despite this positive coefficient, the explanatory variable has a high p-value and thus does not prove a statistically significant, and therefore causal, relationship between fast food restaurants and type 2 diabetes rates. The statistical insignificance on the fast food restaurants per 10,000 people is further empirical justification for the presence of endogeneity in the model that must be addressed to have valid results.

Table 5. Effect of Fast Food Restaurants on Type 2 Diabetes Rates (OLS)

Variable	Coefficient	Standard Error	P-value
Fast Food Restaurants per 10,000 People	0.0000433	0.0001667	0.795
Low Population Dummy Variable	1.364253	0.0911905	0.000
Medium Population Dummy Variable	1.315634	0.0857393	0.000
Poverty Rate	0.0363538	0.0049946	0.000
African American	0.0756322	0.0023554	0.000
Hispanic	− 0.0312302	0.0025265	0.000
Population Density per Square Mile of Land	− 0.0000872	0.0000194	0.000
Constant	8.935008	0.114233	0.000

Number of observations = 3,106

The empirical evidence from the OLS regression suggests that an instrumental variable is indeed a necessary component to remove the issue of endogeneity and yield a useful estimation of the model. Following this need, I examined the relationship between the instrument, interstate exits, and the endogenous explanatory variable, fast food restaurants. As discussed in the last section, I conduct an OLS regression to explore this relationship, regressing fast food restaurants per 10,000 people on the control variables and interstate exits per 10,000 people. These results, as discussed and presented in Table 2 from the last section, yield a positive relationship between fast food restaurants and interstate exits, with a coefficient of 0.8326046, but an insignificant p-value of 0.110.

As previously touched upon, the instrument is not perfect, in that the first condition is not met with statistical significance. However, there is indeed a positive correlation between the instrument and the key explanatory variable and the other two

conditions required for a valid instrument are plausibly satisfied at the theoretical level. Therefore, proceeding with the use of the instrument, bearing in mind that it is weaker than the theoretical expectation predicted, remains a useful and valid step for the purpose of examining the effect of fast food restaurants on type 2 diabetes with the full or partial removal of the issue of endogeneity from the estimated model.

Having established a baseline comparison via an OLS regression, I employ the instrumental variable to conduct regressions with results that are both valid, assuming the instrument is considered adequate in addressing the issue of endogeneity, and significant. As discussed in the Methods section, I conduct two different Two-Stage Least Square regressions, one with a more direct approach to controlling for population and another with an indirect approach. The direct method, in which population size is controlled both implicitly through the denominator of the fast food restaurants variable and directly by including population density in the regression, is the first set of results that are presented.

As we saw in Equation (**2**), the first model using the instrumental variable of interstate exits employs four control variables and the independent variable of interest, fast food restaurants per 10,000 people. Results of this regression are presented in Table 6.

Hausman (1978) suggested comparing the OLS and Two-Stage Least Squares estimations to determine whether or not endogeneity is present in the estimated model. According to Hausman, OLS and Two-Stage Least Squares estimations should yield consistent results if all variables in the model are exogenous. If the two estimations are different in terms of the statistical significance of the explanatory variable of interest, then endogeneity is an issue in the model. Thus, the statistical insignificance from the

OLS estimation, and the statistical significance from the Two-Stage Least Squares regression as seen in Table 6, provide useful empirical backing for the theoretical argument laid out as to why an instrumental variable is necessary to account for the issue of endogeneity in this model.

Table 6. Effect of Fast Food Restaurants on Type 2 Diabetes Rates (IV, with Population Density)

Variable	Coefficient	Standard Error	P-value
Fast Food Restaurants per 10,000 People	0.0947206	0.0288613	0.001
Population Density per Square Mile of Land	− 0.0000158	0.0000475	0.739
African American	0.0575237	0.0096078	0.000
Hispanic	− 0.0429862	0.0103179	0.000
Poverty Rate	0.0391373	0.0175948	0.026
Constant	8.456304	0.4856723	0.000

Instrumental variables (2SLS) regression
Instrumented: Fast Food Restaurants per 10,000 people

Examining the main results from the first Two-Stage Least Squares estimation reveals that the key relationship – fast food restaurants and type 2 diabetes – is positive and now statistically significant, which validates the main hypothesis of the study. The coefficient of 0.0947 implies that for every one-unit increase in fast food restaurants per 10,000 people in a given county, the type 2 diabetes rate is expected to increase by 0.0947%, holding all else equal. A more tangible interpretation is as follows: Given a ten-unit increase in fast food restaurants per 10,000 people, type 2 diabetes rates would rise by 0.947%, or almost 1%.

Interestingly, the only statistically insignificant relationship in the first Two-Stage Least Squares regression is that between population density per square mile of land and the dependent variable, type 2 diabetes rates. As previously discussed, this high p-value may be a consequence of an overly direct approach to controlling for population in the regression, given the fact that population data already exists within the denominator of the fast food restaurants per 10,000 people covariate.

Thus, for the purpose of comparison to examine how different population controls may impact the relationship of interest (fast food restaurants and type 2 diabetes rates), I conduct a second Two-Stage Least Squares regression, estimating the model from Equation (**3**). Using the indirect approach to control for population, I include two dummy variables representing county population size categories – low population and medium population counties – in the regression and omit the third dummy variable, the high population county classification. Additionally, like the last regression, population is implicitly controlled for through the denominator of the fast food restaurants independent variable. The results of this second Two-Stage Least Squares estimation are presented in Table 7.

The key results of this paper, which we see in Table 7, is type 2 diabetes rates regressed on fast food restaurants per 10,000 people, which is instrumented with interstate exits, and on several controls. The first noticeable difference between the two regressions conducted using the instrumental variable of interstate exits is the coefficient on fast food restaurants per 10,000 people: In Table 7, we see that this regression yields a coefficient of greater magnitude, 0.1161854 rather than 0.0947206, compared to that of the first Two-Stage Least Squares regression. Though the main result, the beta coefficient

indicating the relationship between fast food restaurants and type 2 diabetes rates, is not significant at the 5% level, it is significant at the 10% level, as it yields a p-value of 0.096.

Table 7. Effect of Fast Food Restaurants on Type 2 Diabetes Rates (IV, with Population Categories)

Variable	Coefficient	Standard Error	P-value
Fast Food Restaurants per 10,000 People	0.1161854	0.0698947	0.096
African American	0.0468112	0.0178319	0.009
Hispanic	− 0.0503276	0.0137666	0.000
Poverty Rate	0.0414499	0.0212078	0.051
Low Population Dummy Variable	− 1.797853	1.688307	0.287
Medium Population Dummy Variable	− 0.3317857	0.9396232	0.724
Constant	8.596852	0.504799	0.000

Instrumental variables (2SLS) regression
Instrumented: Fast Food Restaurants per 10,000 people

Outcomes from this second Two-Stage Least Squares estimation imply that a one-unit increase in fast food restaurants per 10,000 people would yield a 0.1162 percentage point increase in type 2 diabetes rates, holding all else equal. Similarly, a ten-unit increase in fast food restaurants per 10,000 people would yield a 1.162 percentage point increase in type 2 diabetes rates. Though the fast food restaurant variable has a higher p-value in this second regression with the instrumental variable, the larger value of the coefficient is noteworthy, as it implies that fast food restaurants have a greater impact, in terms of magnitude, on type 2 diabetes rates.

For comparison purposes, all three of the key regressions in this paper are presented side by side in Table 8, which provides both the coefficient and the p-value for each covariate. As touched upon, the significant difference between the p-values from the OLS regression and the Two-Stage Least Squares regressions is notable, with the key explanatory factor for the difference being the use of the instrumental variable. The statistical significance of the fast food restaurants per 10,000 people explanatory variable in the regressions using the instrument, as compared to the statistical insignificance from the regression without the instrument, provides empirical justification for the need of an instrumental variable to account for the issue of endogeneity. The need to account for the factor creating the endogeneity problem – demand for unhealthy food – is evident theoretically, but the empirical backing through a comparison of the OLS and Two-Stage Least Squares regressions is essential to confirm the theoretical expectation and to conduct a valid study.

Another interesting note when comparing these three regressions is the statistical significance of the race covariates. Both the Hispanic and African American independent variables are statistically significant at the 1% level for the OLS and Two-Stage Least Squares regressions. In all three estimations, the positive coefficient on the African Americans covariate indicates a positive relationship, on average, between African Americans and type 2 diabetes rates at the county level. This finding is expected, as studies have demonstrated that racial minorities tend to bear a disproportionate burden of chronic diseases. However, the coefficient on the Hispanic covariate across all three estimations is negative, indicating that changing the racial makeup of a county through an increase in the number of Hispanic individuals in that county would lead to a decrease in

rates of type 2 diabetes rates, on average and holding all else equal. This result is unexpected, as American Hispanics/Latinos are also a minority in the United States and public health research has suggested that this minority group, like African Americans, experiences a disproportionate number of cases of type 2 diabetes.

Table 8. Effect of Fast Food Restaurants on Type 2 Diabetes Rates (Two-Stage Least Squares and OLS)

Variable	IV *population categories*	IV *population density*	OLS
Fast Food Restaurants per 10,000 People	0.1161854 (0.096)	0.0947206 (0.001)	0.0000433 (0.795)
Low Population Dummy Variable	− 1.797853 (0.287)		1.364253 (0.000)
Medium Population Dummy Variable	− 0.3317857 (0.724)		1.315634 (0.000)
Poverty Rate	0.0414499 (0.051)	0.0391373 (0.026)	0.0363538 (0.000)
African American	0.0468112 (0.009)	0.0575237 (0.000)	0.0756322 (0.000)
Hispanic	− 0.0503276 (0.000)	− 0.0429862 (0.000)	− 0.0312302 (0.000)
Population Density per Square Mile of Land		− 0.0000158 (0.739)	− 0.0000872 (0.000)
Constant	8.596852 (0.000)	8.456304 (0.000)	8.935008 (0.000)

Notes: The results are present in the following format – coefficient (p-value). Some blank spaces appear in this table because I did not employ all covariates for all regressions.

A useful extension of these results is to analyze them in relation to Dunn's findings. Dunn, who explores fast food restaurants' effect on obesity via the BMI measure, took a similar approach to the second of the two estimations I conducted using

Two-Stage Least Squares in that he created three different county categories based on low, medium, and high population density. Dunn found that in medium-density counties, females, African Americans and Hispanics all displayed a positive and statistically significant, at the 5% level, relationship between the number of fast food restaurants and BMI. Unlike Dunn, I keep the data aggregated to analyze the overall impact of fast food restaurants on type 2 diabetes rates, which I find to be positive and statistically significant at the county level. Another key difference is that Dunn combined African American and Hispanic individuals into a single group, whereas I kept data on the percentage of African American and Hispanic people at the county level separate and included the two as separate variables in the regressions in this paper.

Like the results of this study, Dunn's results must be examined with some skepticism, as he employs the same instrumental variable as I do in this paper and does not justify the interstate exit instrumental variable empirically (through the OLS regression of fast food restaurants on the instrument and the control variables). Therefore, both the results of this paper and Dunn's must be examined bearing in mind the weaker nature of the instrumental variable. Despite this, the results of this paper, which demonstrate a statistically significant and positive relationship between fast food restaurants and type 2 diabetes rates, may have important implications for policy and further economic research.

<u>Conclusion</u>

This paper provides additional insight into the realm of health economics and policy discussions surrounding the exponentially growing rates of type 2 diabetes and obesity. The results indicate that fast food restaurants have a positive and statistically

significant effect on type 2 diabetes rates, confirming the underlying theory and hypothesis. While the findings themselves are intuitive, the model and analysis in this paper is distinct from existing literature in that it examines the effect of the mere presence of fast food restaurants on rates of type 2 diabetes. Given the fact that the discussion is typically centered around the actual consumption of unhealthy food from a place of business like a fast food restaurant, exploring other factors that influence long-term health outcomes can provide useful insights for policymakers and health professionals.

The relationship between fast food restaurants and type 2 diabetes rates has been and continues to be relevant to policymakers considering the profound economic and social costs of the exponentially rising rates of chronic diseases. For example, the New York Times reported in 2011 that Los Angeles placed a temporary ban on the opening of new fast food restaurant branches in the low-income area of South Los Angeles. The action stemmed from the fact that, according to the Los Angeles County Department of Public Health, this area of the county has far higher rates of type 2 diabetes. This step is not typical for most counties, however, and fast food restaurant prevalence remains a persisting issue, as justified by the findings of this paper.

While my final results are both robust and statistically significant, there are limitations to the interpretation of the findings. There are important theoretical and empirical criticisms about the validity of the instrument that are worth considering. The strongest of these critiques is the fact that areas with more interstate exits, and thus more fast food restaurants, may attract people who have different health preferences than those living in areas with fewer interstate exits. Additionally, the fact that the correlation between fast food restaurants and the interstate exits was not statistically significant,

though close with a p-value of 0.110, imposes a need for skepticism when interpreting the results of this paper. Therefore, a useful extension of this study could employ a different instrumental variable that is of greater validity and can thus better account for the issue of endogeneity in the model.

Additional research could examine individual-level data and consider how the actual distance of a house to a fast food restaurant may affect rates of type 2 diabetes. A study of this sort could be especially interesting in a state like California, which has significant variation in terms of population and socioeconomic status, with distinct pockets of wealth and poverty. Furthermore, the negative coefficient on the covariate capturing the rate of Hispanic people at the county level suggests that further research is needed in exploring specific burdens of diseases on racial minorities relative to white people in the United States to determine which groups, and to what extent, bear a disproportionate share of type 2 diabetes. Lastly, a time series analysis would provide useful context to the discussions around type 2 diabetes, as it would explore long-term changes to fast food restaurant availability and the corresponding resulting changes in rates of type 2 diabetes rates.

References

"A Snapshot: Diabetes in the United States." 2017. *Centers for Disease Control and Prevention*.

"About Diabetes." 2017. *Centers for Disease Control and Prevention*.

"County Data Indicators." 2016. *Centers for Disease Control and Prevention*.

"Diagnosed Diabetes Percentage." 2017. *Centers for Disease Control and Prevention*.

"Food Environment Atlas." 2017. *Economic Research Service, Department of Agriculture*.

"General Food Environment Resources." 2014. *Centers for Disease Control and Prevention*.

"Income, Poverty and Health Insurance Coverage in the United States: 2011." 2012. *United States Census Bureau*.

"Population, Housing Units, Area, and Density." 2010. *American FactFinder, United States Census Bureau*.

"Portrait of Promise: The California Statewide Plan to Promote Health and Mental Health Equity." 2015. *Office of Health Equity, California Department of Public Health*.

"State and County Estimates for 2011." 2017. *United States Census Bureau*.

Burgoine, Thomas, Forouhi, Nita, Griffin, Simon, Wareham, Nicholas, Monsivais, Pablo. 2014. "Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: population based, cross sectional study." *BMJ* 348: 1-10.

Currie, Janet, DellaVigna, Stefano, Moretti, Enrico, Pathania, Vikram. 2010. "The Effect of Fast Food Restaurants on Obesity and Weight Gain." *American Economic Journal: Economic Policy* 2 (3): 32-63.

Drenowski, Adam, Darmon, Nicole. 2005. "The economics of obesity: dietary energy density and energy cost." *The American Journal of Clinical Nutrition* 82 (suppl): 265S-73S.

Dunn, Richard. 2010. "The Effect of Fast-food Availability on Obesity: An Analysis by Gender, Race, and Residential Location." *American Journal of Agricultural Economics* 92 (4): 1149-1164.

Freeman, Andrea. 2007. "Fast Food: Oppression through Poor Nutrition." *California Law Review* 95 (6): 2221-2260.

Ginter, Emil, Simko, Vlado. 2013. "Type 2 Diabetes Mellitus, Pandemic in 21st Century." *Diabetes. Advances in Experimental Medicine and Biology*: 42-50. Springer, New York, NY.

Harris, Jennifer, Bargh, John, Brownell, Kelly. 2009. "Priming Effects of Television Food Advertising on Eating Behavior." *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *28*(4), 404–413.

Hausman, Jerry. 1978. "Specification Tests in Econometrics." *Econometrica* 46 (6): 1251-1271.

Hickey, Walter, Weisenthal, Joe. 2013. "Half Of The United States Lives In These Counties." *Business Insider*.

Inagami, S, Cohen, Deborah, Finch, Brian, Asch, Steven. 2006. "You are where you shop: grocery store locations, weight, and neighborhoods." *American Journal of Preventive Medicine* 31 (1): 10-17.

Jeffery, Robert, Baxter, Judy, McGuire, Maureen, Linde, Jennifer. 2006. "Are fast food restaurants an environmental risk factor for obesity?" *The International Journal of Behavioral Nutrition and Physical Activity* 3 (2).

Jekanowski, Mark, Binkley, James, Eales, James. 2001. "Convenience, Accessibility, and the Demand for Fast Food." *Journal of Agricultural and Resource Economics* 26 (1): 58-74.

Medina, Jennifer. 2011. "In South Los Angeles, New Fast-Food Spots Get a 'No, Thanks.'" *New York Times*.

Powell, Lisa, Chaloupka, Fank, Bao, Yanjun. 2007. "The availability of fast-food and full-service restaurants in the United States: associations with neighborhood characteristics." *American Journal of Preventive Medicine* 33 (4): S240-S245.

Powell, Lisa, Han, Euna, Chaloupka, Frank. 2010. "Economic Contextual Factors, Food Consumption, and Obesity among U.S. Adolescents" *Journal of Nutrition* 140: 1175-1180.

Reisch, Lucia, Gwozdz, Wencke, Barba, Gianvincenzo, De Henauq, Stefaan, Lascorz, Natalia, Pigeot, Iris. 2013. "Experimental Evidence on the Impact of Food Advertising on Children's Knowledge about and Preferences for Healthful Food." *Journal of Obesity* vol. 2013: 1-13.

Rothman, Kenneth. 2008. "BMI-related errors in the measurement of obesity." *International Journal of Obesity* 32 (S3): S56-S59.

Wooldridge, Jeffrey. *Introductory Econometrics: A Modern Approach*. Boston, MA: Cengage Learning, 2016.