Selected Presentation at the 2020 Agricultural & Applied Economics Association Annual Meeting, Kansas City, Missouri, July 26-28

Copyright 2020 by authors. All rights reserved.
Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

The Impact of Food Environment on Fresh Vegetable Consumption in North Carolina

June 30, 2020

Haoran Miao[1i], Chyi-Lyi Liang[1], Manoj Jha[1], Leila H. Beni[1], Lyubov Kurkalova[1], Timothy Mulrooney[2], Greg Monty[1]

[1] North Carolina Agricultural and Technical State University, Greensboro, NC

[2] Department of Environmental, Earth and Geospatial Sciences, North Carolina Central University, Durham, NC

Abstract

As one of the top agricultural-producing States, North Carolina can supply enough fresh produce to meet consumer demand. However, food deserts, areas with limited access to available and affordable fresh produce, are common in North Carolina. For example, Guilford County has about 900 farms, and nearly 20% of its census tracts are defined as food deserts. Literature has documented that some chronic diseases, such as diabetes and obesity (e.g., Flórez et al., 2015), and nutritional inequality are associated with food deserts (e.g., Allot et al., 2019). This paper investigates the impact of the food environment in North Carolina, represented by three types of measurement, on fresh vegetable purchasing behavior using Nielsen Homescan data. About 3,000 households in North Carolina reported their purchased grocery items, including packaged and unpackaged vegetables, to the Nielsen database after they made about one million grocery trips between 2012 and 2017. They also reported their geographic locations (zip codes) and their socioeconomic characteristics. We use HUD-USPS ZIP crosswalk files data to identify whether one zip code has any overlap with any food desert area and the portion of residents within a zip code living in a food desert census tract. We also use average driving distance/time from households to the nearest full-service grocery stores, calculated in ArcGIS, in a zip code to approximate the food environment in that zip code.

[i] hmiao@ncat.edu

At the single-trip level, we use a series of weighted logistic regression models to find the important factors of buying fresh vegetables. We find that living in a food desert zip code is associated with a lower probability of buying fresh vegetables during a grocery-shopping trip. We also find that owning a vehicle increase shopping trips, diversify the grocery shopping locations, and reduces the probability of going to a full-service grocery store.

We then aggregate total trip spending and total fresh vegetable spending for each household each year and build models to identify the food environment and other factors that affect these two indicators. In the aggregate year-level data, as expected, we find that household characteristics, such as household income, age, and household size, are correlated with both yearly food spending and yearly total fresh vegetable spending. We also find that in general, households do not have stark differences in yearly food spending and yearly fresh vegetable spending, no matter where they live, ceteris paribus. Households living in food deserts do not visit full-service grocery stores less than their counterparts and visit other types of stores more often. Food deserts may only impact low-income households living in food deserts.

Keywords: Fresh vegetable consumption, food environment, consumption pattern

JEL: Q11, Q15, Q18

1 Introduction

Fresh produce is one important slice of a healthy eating pattern. Healthy U.S.-Style Eating Pattern, one of three United States Department of Agriculture (USDA) eating patterns, includes two and a half cup-equivalent vegetables and two cup-equivalent fruits at the 2,000-Calorie level. Healthy diets with proper vegetable and fruit intakes can have protective effects of stroke (He et al., 2006), and chronic diseases, such as cardiovascular disease (Hung et al., 2004), and obesity (Epstein et al., 2001). However, based on the 2018 State Indicator Report on Fruits and Vegetables, each day, only one in nine adult Americans eat sufficient fruits, and one in ten adult Americans eats sufficient vegetables. To make the statistic number even worse, only seven percent of adults who have incomes below the poverty level meet the vegetable intake standard.

The benefit of increased vegetable and fruit consumption has been documented widely in literature (e.g., Ford and Mokdad, 2001; John et al., 2002; Wang et al., 2014). American governments at the federal, state and local levels advocate the increase of fruit and vegetable consumption in many ways. For example, the Action Guide on Fruits and Vegetables from Centers for Disease Control and Prevention explicitly points out, "Fruit and vegetable consumption in the United States is too low," and vegetable consumption in North Carolina is even lower, compared with the national level. Although a lot of effort has been placed to encourage fruit and vegetable consumption, the overall condition is not improved both at the national and regional levels, even in a lot of agricultural states with plenty of vegetable/fruit production capability (CDC 2013, CDC 2018). In some states, such as North Carolina, the percentage of adults eating enough vegetables goes down to eight percent (CDC, 2018). As one of the top ten agricultural producing states, North Carolina can grow a large variety of vegetables, however. Food deserts, as a metaphor for limited access to fresh food defined by USDA, prevail in many low-income communities in those agricultural states (e.g., Guilford County; see Figure 1). In North Carolina, about fourteen percent

of its residents living in 349 food deserts (PATHS, 2013). The gap between vegetable production capability and vegetable consumption and the stark comparison between food deserts and oasis of fresh vegetables motivates us to investigate the fundamental vegetable consumption pattern of residents. We are particularly interested in finding factors that may contribute to the consumption patterns of fresh vegetables, such as food environment, social and economic status.

Factors contributing to food consumption disparities among different groups of people have been widely researched recently. Internal or individual-level factors include race, ethnicity, income, education, occupation, marriage status, presence of children at home, ownership of automobiles, beliefs about healthy food, and even psychosocial determinants. External factors, such as access to supermarkets, or availability of healthy food, are heavily investigated in the literature because places offering affordable healthy food is seemingly important to a subset of people. Food consumption is tabulated into different groups according to the listed factors, and we can often see different types of disparities among different groups. For example, Morland et al. (2002) find that communities in a few states in the U.S. with large portions of the poor and minority may not have the same opportunities to healthy food choice, compared with nonminority and rich communities. Smith and Morthon (2009) discovered at several focus groups in Minnesota and Iowa that physical food environments constrain the nonminority group's food consumption. In contrast, Allcott et al. (2019) use nationwide data to investigate the nutritional inequality problem, "why the wealthy eat more healthfully than the poor in the United States." Their findings suggest that about 90% of the inequality arises from the internal or individual factors, that is, people in the low-income categories, have a low demand for healthy nutrition, and neighborhood environments do not contribute meaningfully to mitigate the nutritional inequality.

Research exploring the factors contributing to food consumption gives us mixed results. Vegetable consumption varies person by person and varies at the global level (Hall et al., 2009),

however. Research investigating the factors contributing to vegetable consumption provides us even more mixed results. Often, vegetable and fruit consumption are studied together in a certain group of people by previous research papers. For example, in Reinaerts et al. (2006), questionnaire results show that significant factors explaining fruit and vegetable consumption of 4 -12-year-old children in the Netherland are psychosocial factors, parental consumption, habit, and ethnicity. Riediger & Moghadasian (2013) find that fresh vegetable and fruit consumption of the Canadian elderly has a strong correlation with total household income and household education years in a questionnaire. Ethnicity does not affect vegetable and fruit intakes in their study, however. In data from China Chronic Disease Surveillance survey, Li et al. (2017) find that low vegetable and fruit consumption is associated with low socioeconomic statuses, such as low income and low educational level, and minority ethnicity.

The above results indicate that fresh vegetable consumption patterns may be different spatially. Thus, we use a subset of a national-level dataset to estimate the patterns in North Carolina. We aim to find the meaningful factors that associated with fresh vegetable consumption in North Carolina. We contribute to the literature in three ways. First, we demonstrate how to extract fresh vegetable spending, including packaged and unpackaged fresh vegetables, from the Nielsen Homescan data. Often, authors neglect unpackaged vegetables, fruits, and meat consumption (e.g., Alloctt 2019). Second, we match USDA food desert information with Neilsen data using HUD-USPS ZIP crosswalk data and estimate food desert status at the zip code level. Third, we use average driving distance/time from household addresses to the nearest full-service grocery stores in a zip code as the approximation of the food environment. We believe it is a more accurate way to measure the food environment compared with the traditional way using the centroids as the starting points.

Based on our analysis, we find that, in general, the food environment changes the grocery shopping locations, the probability of buying vegetables, and other fresh vegetable consumption patterns at the trip level. However, the food environment does not alter the yearly fresh vegetable spending and yearly total food spending meaningfully. It seems most households living in food deserts have mobility and can visit full-service grocery stores almost at the same frequency as their counterparts.

In Section Two, we will describe the datasets we are using in this study and how we extract and match variables from the datasets. In Section Three, we show how we analyze the data we collect. We report the results in Section Four. We conclude the results and discuss our next steps in Section 5.

2 Data

2.1 Nielsen Homescan Data

The main data we have for our analysis come from the Nielsen Homescan dataset. Nielsen (www.nielsen.com) is a company that offers retail measurement services, including sales information from stores and home-purchasing information from its panelists. The stores they collect information from vary in types (66 types in total), including grocery stores, discount stores, warehouse clubs, military stores, and so on. They ask their panelists to record all take-home items purchased by all household members. The price of each item they record is either from the internal database if the store where they made the purchases is partnering with Nielsen or from the panelists themselves if the store where they made the purchases is not within the Nielsen's database.

We purchased the Nielsen Homescan data from the Chicago Booth Kilts Center, who additionally administrate the collection of the dataset and further clean the dataset. They make efforts to make sure the panelists in the dataset are active, and the categorization of the data is accurate. Since we mainly investigate fresh vegetable consumption, we process items that fall into

the categories of fresh vegetables. In the Nielsen Homescan data, fresh vegetables have about fifteen categories [fifteen standard Universal Product Codes (UPCs)] of packaged fresh vegetables, including fresh sprouts, fresh carrots, fresh lettuce, fresh tomatoes, etc., and one category (one UPC) of unpackaged fresh vegetables. Each item categorized in the packaged fresh vegetables has a vegetable name, price, quantity, and unit. Each item categorized in the unpackaged fresh vegetables, just have one UPC, and does not have its vegetable name and unit. Before the year 2012, Nielsen does not differentiate unpackaged vegetables with other items without standard UPC, such as fruits, and meats, and all of them are under the same category (magnet data). From the Homescan dataset after the year 2012, we find the consumption of unpackaged vegetables consists of about 42% of overall fresh vegetable consumption, which is not a negligible part. Thus, we include Nielsen Homescan data from 2012 to 2017, in which a portion of the panelists reports unpackaged vegetables in North Carolina in our analysis. Also, because we cannot obtain accurate data of quantities of unpackaged fresh vegetables, we use the dollar spending as the measure of fresh vegetable consumption.

In the Nielsen Homescan dataset, one trip is represented by one trip code, and one trip code is corresponding to one shopping location and numerous shopping items. For example, if one household went for grocery shopping in two food stores in one day, and in each food store, the household bought five items, the household would have records of two grocery shopping trips, and ten items. We first clean the Nielsen Homescan data at the item level. Because we focus on fresh vegetable consumption and related food consumption, we sieve non-food items out from the dataset, such as cooking utensils, books, and pesticides. We further exclude items that are from non-food outlets, such as optical stores and toy stores (limited food items come from those outlets, and we treated these items as record errors, and we delete them from the database). Because there is no standard UPC associated with food in full-service restaurants, we measure take-home

food/vegetable only in this study. We further define grocery stores, discount stores, warehouse clubs, and military stores as full-service grocery stores, because they supply various categories of food, including fresh vegetables and fruits, all the year. We then aggregate vegetable/food consumption and other related variables at the trip level. Not every household scans unpackaged vegetable items, however. Including those who only scan packed vegetables would bias our data (e.g., we will underestimate the average vegetable consumption of households). Thus, we identify those who do not scan unpacked vegetables by their household codes and eliminate those trips made by those households, and we obtain 1,107,754 grocery trips in total. At the end of data processing, we aggregate the dataset to the year level, and we obtain 8,640 household-by-year observations. However, we find that one household record that they go to grocery shopping for one time in a year (less than one trip a month) and spend about sixteen dollars. We consider this observation an outlier and delete it from our dataset. Using the consumer price index (CPI) over the years, we deflate incomes and various spending at different levels to 2010 dollars.

In addition to the information of the purchased items of each household, the Nielsen Homescan data have household-characteristic data, such as the zip code, income (in 16 brackets), composition, race, members' ages, members' education level, the status of employment, and weekly working hours for the heads of the households. Based on the selection of time (2012 - 2017) and space (North Carolina), we obtain 8,639 (year × Household) observations, with about 3,008 distinct households. We present the descriptive statistics of household characteristics of those panelist households in Table 1.

We have the average age/education years of two household heads as the age/education years of one household if the household has two household heads. We use the average working hours of two household heads if there are two household heads in the household. If no one in the household is employed, we record the employment status of the household as the value of zero.

We record the employment status as one if no household head is unemployed. However, if only one of two household head is employed, the employment status is recorded as 0.5.

We hypothesize the ownership of a vehicle is an important factor contributing to vegetable consumption, especially for those who live in a food desert. However, Nielsen Homescan data do not have information about households' ownership of vehicles. Thus, we use the spending of gasoline as the proximity of whether one household has a vehicle. If there is no cost in gasoline for one household, we assume the household does not own a vehicle.

From Table 1, we know that the average household income of our samples is about $59,000 (2010 dollars). The average household size is 2.44, and 71 percent of households in our sample own a vehicle. The average age of household heads is 54.21, and the average years of education of household heads are 14.58 years. The average working hours of household heads are 22.22 hours per week. Among all the households, 75 percent are white, 20 percent are African American, and 43 percent of them have children at home.

In our samples, households spend $3,283 on take-home food on average per year, and $125 is spent on fresh vegetables, and 97 percent of vegetables are purchased from full-service grocery stores. Besides, it takes about 127 trips on average to fulfill grocery needs, and 92 trips end in a full-service grocery store.

2.2 HUD-USPS ZIP crosswalk

The smallest geographic location that can be identified for each panelist is at the zip code level in the Nielsen Homescan dataset. Although in some literature, such as Allcott et al. (2019), household codes can be matched to the census tracts level, Nielsen sales representative affirmed us that they could not provide us the census tract information of the panelist households. Nevertheless, the food desert definition by USDA is at the census tract level. While there are about

74 thousand census tracts and 42 thousand zip codes in the U.S., almost no census tracts and zip codes are perfectly matched.

The geographic information of panelist households gives us challenges to estimate the physical food environment of each household; that is, we cannot identify whether one household lives in a USDA-identified food desert. The approach we take to solve this problem is to apply a HUD-USPS ZIP crosswalk to our dataset so that we can estimate the portion of residents of one specific zip code living in a food desert. The HUD-USPS ZIP crosswalk is developed by the U.S. Department of Housing and Urban Development's (HUD's) Office of Policy Development and Research (PD&R). This dataset gives an estimation of portions of residents living in a certain census tract within any zip code. For example, from the HUD-USPS ZIP code crosswalk file, we find that the area of zip code 27055 overlaps with census tracts 37197050501, 37197050400, 37197050102, 37059080100, 37097060801,37197050502, 37197050200, and 37197050101. Only census tract 37197050501is defined as a food desert by USDA, and the resident's ratio in the census tract is about 47 percent. So, we claim that 47 percent of residents of zip code 27055 live in a food desert. Because there is no explicit food desert definition for one zip code, we use the following way to define one zip code as a food desert. We define one zip code overlapping with any food-desert census tract as a food-desert zip code. We match the food-desert variable and the resident's ratio in food deserts with trip-level and year-level data using zip codes.

2.3 Travel time/distance

In addition to creating food environment variables by matching USDA and HUD-USPS ZIP code crosswalk data, we also take another router to measure the food environment of a zip code. We apply geospatial networking tools in ArcGIS to estimate the average driving distance and average driving time to the nearest full-service grocery stores in a certain zip code by optimizing source locations. Compared with traditional ways that use centroids as the source

locations, in our study, we use grocery store locations and locations of resident parcels as the source locations. In this way, the driving distance and time are more accurately measured compared with traditional methods. We match the average travel time/distance data with our original dataset using zip codes. We want to use travel time/distance as additional measures of the food environment.

3 Empirical Strategies

In this study, we investigate the fresh vegetable consumption patterns of grocery shoppers in North Carolina, and we pay special attention to how the food environment, such as food deserts, can affect the fresh vegetable consumption patterns. In our empirical analysis, first, we study how household characteristics may be different from each other between food-desert and non-food-desert households in our sample. We test the balances of household characteristics between non-food-desert and food-desert households.

Second, we investigate how food deserts and household characteristics can affect consumers' behavior at the trip level. For example, we want to understand whether the food environment and household characteristics influence where consumers shop for groceries and the probability of buying vegetables during each trip. We use a series of weighted logistic regressions and weighted least square regressions to test those associations.

Third, we also investigate whether food environment and household characteristics can affect vegetable shopping patterns at the aggregated year level. For example, we want to know how food deserts affect consumers' shopping frequency in different types of stores during a year, whether food deserts affect the total fresh vegetable/food consumption. We implement a series of weighted least square regressions to test those correlations.

After we identify those associations, we try to understand the causal effect of food deserts on fresh vegetable consumption using propensity score matching methods. We treat the food-

desert status as an experimental variable and match observations in pairs to reduce the effect of confounders. In the end, we use a subset of our dataset to further identify the causal effect of food deserts by limiting our data to only movers. That is, we only look at those households moving between the food-desert area and the non-food-desert area. By this approach, we assume we control for unobserved confounders, such as habits or psychological factors that also may contribute to fresh vegetable consumption. We include known control variables in those weighted least square regressions.

Finally, we use other environmental variables, the resident's ratio in food deserts and travel distance/time, to further estimate how the food environment may affect fresh vegetable consumption.

4. Results

4.1 Spatial differences in characteristics between non-food-desert and food-desert households

The unbalanced socioeconomic status between food-desert and non-food-desert households has been well documented in the literature (e.g., see Dutko, Ploeg, & Farrigan, 2012). The household characteristics are also mostly unbalanced in our sample (see Table 2). The average income of households in the non-food-desert area is about $64,000, which is $7,000 higher than that of households in the food-desert area. The education level of households living in the non-food area is also statistically significantly higher than that of households living in the non-food desert area. Besides, household size, working hours, ethnicity, marital status, presence of children, and employment status are also statistically significantly different. However, the differences in household age and the portion of people owning a vehicle are not statistically significant.

Grocery shopping trips are widely different from households to households. The distribution of grocery shopping trips is presented in Figure 2. We find that mode of trip frequency

is around 100, and the average trip frequency is about 92 (Panel B, Table 1). However, we did not observe significant differences in total trips for grocery shopping and even trips to full-service grocery stores between food-desert and non-food-desert households. Intuitively, we believe living in food deserts would decrease the frequency to full-service grocery stores and increase the frequency to other stores, such as convenience stores and dollar stores, where the supply of fresh vegetables is limited. However, the trip frequency to full-service grocery stores is around 92 on average for food desert households and their counterparts. We further test the hypothesis in the rest of the section by controlling for other cofounders.

4.2 The effect of food deserts and household characteristics on fresh vegetable consumption at the trip level

We understand that our sample cannot represent the population. Thus, we do not treat all households equally and use a projection factor provided by the Nielsen dataset to represent demographic composition better. Regression models are all weighted by the projection factor (magnet project factor). To test the associations between food deserts/other factors and fresh vegetable consumption patterns at the trip level, we run three weighted logistic regressions to estimate the marginal probabilities of those factors of going to full-service grocery stores, buying vegetables at full-service grocery stores, and buying vegetables at non-full-service grocery stores.

Because we hypothesize that households living in food deserts and owning no vehicle have the most limits to access to fresh vegetables, we include the interaction of the food desert variable and the ownership of vehicles variable in the regression to capture interaction effects, if any. Table 3 presents the results of the three weighted logit models. Model 1 shows that the marginal probabilities of factors that affect the likelihood of visiting a full-service grocery store. Owning a vehicle has the largest marginal effect and, counterintuitively, decreases the probability of going

to full-service grocery stores by about 13 percent. We conjecture that owning a vehicle diversify shopping locations and increase the shopping frequency. We test this conjecture in Section 4.3. Living in food deserts decreases the probability of going to full-service grocery stores by about two percent. The growth of household income, year education, working hours, presence of children also increases households' probability of visiting full-service grocery stores. However, the employment status and the increase of household size and age decreases the probability of going to full-service grocery stores. Ethnic groups also play important roles in affecting the probability of going to full-service grocery stores. For example, compared with the group other than white, African American, Asian, and Hispanic, Asian people are nine percent more likely to visit full-service grocery stores. In comparison, African Americans are five percent less likely to visit full-service grocery stores.

After arriving full-service grocery stores, households are less likely to buy fresh vegetables in one trip if they live in the food-desert area, own a vehicle, have any child at home, or they are black/Hispanic (Column 2, Table 3). The increase in income and years of education increases the probability of buying fresh vegetables. Households have similar consumption patterns if they visit other types of stores. Still, the magnitude of marginal effects is small since the probability of buying fresh vegetables in other stores is low (Column 3, Table 3).

4.3 The effect of food deserts and household characteristics on fresh vegetable consumption at the aggregate-year level

We find food-desert status and other factors, such as owning vehicles, income, education, and age, are associated with probability variations of going to full-service grocery stores, buying fresh vegetables at full-service grocery stores at the trip level. In this part, we investigate how those factors are associated with aggregate data at the year level.

We run four weighted least square regressions to figure out how those factors affect four long-term indicators, yearly fresh vegetable consumption, yearly food consumption, trip frequency to all food stores, trip frequency to full-service grocery stores. Table 4 presents the results. In the long run, living in food deserts does not decrease trip frequency to full-service grocery stores (Column 4, Table 4). Instead, it increases the overall grocery trip frequency (Column 3, Table 4). This finding aligns with findings from Shannon (2016), which states that households living in food deserts are very mobile in their trips to grocery shopping. Food desert status does not decrease food spending (Column 2, Table 4), and has a marginal impact of ten dollars on yearly fresh vegetable spending.

Not surprisingly, owning a vehicle has a significant impact on four indicators. As household income increases, total food spending and fresh vegetable spending increases. Household income does not have a big impact on grocery shopping frequency, however. Minority groups have less impact on the long-term indicators, compared with their impact on a single grocery trip.

4.4 Causal effect of food deserts on fresh vegetable consumption at the aggregate-year level

We further use two approaches to estimate the causal effect of food deserts on yearly fresh vegetable consumption patterns. In the first approach, we use the Propensity Score Matching (PSM) method to identify the causal effect. In this design, we use the food-desert variable as a treatment variable in an experiment. Then we match treated observations with the most similar control observations by balancing household income, household size, and other household characteristics. We find that overall, living in food deserts has a limited effect on fresh vegetable consumption, and it reduces yearly fresh vegetable spending by $2.89 (Column 1, Table 5), which is not statistically significant.

We further divide our sample into four income quartiles to investigate whether there is any heterogeneous impact of food deserts on households with different income levels. We use income levels $32K, $50K, and 78K to divide our samples into four quartiles, with almost the same observations in each part. We run four additional PSM models to test the heterogeneous impact. Not surprisingly, food desert status has the biggest impact on fresh vegetable spending of households in the first quartile with household income less than $32K. It decreases their yearly fresh vegetable consumption by $22.49 (Column 2, Table 5), which is statistically significant. However, when household income is over 32K, the impact decreases and is not statistically significant (Column 3, 4, and 5, Table 5).

In the second approach, we restrict our sample to only movers who move between the food-desert area and the non-food-desert area. We find 54 distinct households with 254-year × household observations. We assume this approach can find better counterparts than the PSM method. Using four weighted least square regressions, we find that food desert status does not change movers' grocery shopping frequency, total food spending, and fresh vegetable spending (Table 6).

4.5 Effects of other food environment variables on fresh vegetable consumption at the aggregate-year level

Essentially, the food desert status is a binary variable that measures the food environment developed by the USDA. In this part, we use additional measurements to measure the food environment. The first measurement is built upon USDA's food desert definition. However, we use a continuous variable, the portion of residents living in food deserts, which ranges from 0 to 1. By combining USDA food desert data, and HUD-USPS ZIP crosswalk data, we create this continuous variable that better describe the food environment of one zip code than the food desert variable. The second measurement is driving distance and driving time. With ArcGIS software,

we calculate the driving distances/time from households to the nearest full-service grocery stores and use the average driving distance/time within a zip code to represent the food environment.

Tables 7 and 8 present the results. We find that the food environment, represented by three continuous variables, does not meaningfully change grocery shopping frequency (Columns 3 and 4). Its impact on fresh vegetable spending and total food spending is also limited (Columns 1 and 2).

3 Conclusion and next steps

In this study, we investigate the fresh vegetable consumption patterns in North Carolina, with interests in finding how the food environment impacts households' fresh vegetable spending. We extract sample observations of fresh vegetable spending and related variables from Nielsen Homescan data and match them with food environment data, which are generated from USDA food desert data, HUD-USPS ZIP crosswalk data, and ArcGIS calculation. The data extracted from Nielsen Homescan data has a geographic location at the zip code level and limits our ability to estimate the food environment. We propose three ways to measure the food environment at the zip code level.

We find that households buy fresh vegetables mostly from full-service grocery stores, despite different food environment. Households living in food deserts are quite mobile, and their shopping frequency at full-service grocery stores are not significantly lower than their counterparts. The correlations between the food environment and household characteristics are strong at the trip level. The food desert status reduces the probability of going to full-service grocery stores and buying vegetables when visiting full-service stores. Owning a vehicle diversify food shopping locations. Household characteristics, such as income, age, and ethnic groups, play important roles in choosing grocery shopping locations and buying vegetables during a grocery shopping trip. However, at the aggregate level, the food environment does not have strong

associations with shopping frequency in full-service grocery stores, fresh vegetable spending, and total food spending. The food environment may have a significant impact on low-income households.

Our next step is to estimate local fresh vegetable demand and supply at the local level, such as Guilford County, using the technique of agent-based modeling. We will utilize fresh vegetable consumption patterns and household characteristics data to calculate the local fresh vegetable demand over time. Besides, we will use local farm-level data to estimate the fresh vegetable supply. We are going to build an integrated model so that we can simulate fresh vegetable demand and supply at the local level. Ultimately, we can use the model simulate the scenarios that we use local food systems to 1) mitigate the impact of the underserved food environment, such as food deserts, and 2) increase overall vegetable consumption.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1824949. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Researcher(s) own analyses calculated (or derived) based in part on data from The Nielsen Company (U.S.), LLC, and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

References

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. "Food Deserts and the Causes of Nutritional Inequality*." *The Quarterly Journal of Economics* 134, no. 4 (November 1, 2019): 1793–1844. https://doi.org/10.1093/qje/qjz015.

Boynton-Jarrett, Reneé, Tracy N. Thomas, Karen E. Peterson, Jean Wiecha, Arthur M. Sobol, and Steven L. Gortmaker. "Impact of Television Viewing Patterns on Fruit and Vegetable Consumption Among Adolescents." *Pediatrics* 112, no. 6 (December 1, 2003): 1321–26.

Brug, Johannes, Lilian Lechner, and Hein De Vries. "Psychosocial Determinants of Fruit and Vegetable Consumption." *Appetite* 25, no. 3 (1995): 285–96. https://doi.org/10.1006/appe.1995.0062.

Centers for Disease Control and Prevention, "2018 State Indicator Report on Fruits and Vegetables,": 2018 https://www.cdc.gov/nutrition/data-statistics/2018-state-indicator-report-fruits-vegetables.html

Dehghan, M., N. Akhtar-Danesh, and A. T. Merchant. "Factors Associated with Fruit and Vegetable Consumption among Adults." *Journal of Human Nutrition and Dietetics* 24, no. 2 (2011): 128–34. https://doi.org/10.1111/j.1365-277X.2010.01142.x.

"Disparities and Access to Healthy Food in the United States: A Review of Food Deserts Literature." *Health & Place* 16, no. 5 (September 1, 2010): 876–84. https://doi.org/10.1016/j.healthplace.2010.04.013.

Dutko, Paula, Michele Ver Ploeg, and Tracey Farrigan. "Characteristics and Influential Factors of Food Deserts," n.d., 36.

Epstein, Leonard H., Constance C. Gordy, Hollie A. Raynor, Marlene Beddome, Colleen K. Kilanowski, and Rocco Paluch. "Increasing Fruit and Vegetable Intake and Decreasing Fat and Sugar Intake in Families at Risk for Childhood Obesity." *Obesity Research* 9, no. 3 (2001): 171–78. https://doi.org/10.1038/oby.2001.18.

Fleischhacker, Sheila E., Daniel A. Rodriguez, Kelly R. Evenson, Amanda Henley, Ziya Gizlice, Dolly Soto, and Gowri Ramachandran. "Evidence for Validity of Five Secondary Data Sources for Enumerating Retail Food Outlets in Seven American Indian Communities in North Carolina." *International Journal of Behavioral Nutrition and Physical Activity* 9, no. 1 (November 22, 2012): 137. https://doi.org/10.1186/1479-5868-9-137.

Ford, Earl S., and Ali H. Mokdad. "Fruit and Vegetable Consumption and Diabetes Mellitus Incidence among U.S. Adults." *Preventive Medicine* 32, no. 1 (January 1, 2001): 33–39. https://doi.org/10.1006/pmed.2000.0772.

Hall, Justin N., Spencer Moore, Sam B. Harper, and John W. Lynch. "Global Variability in Fruit and Vegetable Consumption." *American Journal of Preventive Medicine* 36, no. 5 (May 1, 2009): 402-409.e5. https://doi.org/10.1016/j.amepre.2009.01.029.

Haynes-Maslow, Lindsey, Sarah E. Parsons, Stephanie B. Wheeler, and Lucia A. Leone. "A Qualitative Study of Perceived Barriers to Fruit and Vegetable Consumption Among Low-Income Populations, North Carolina, 2011." *Preventing Chronic Disease* 10 (March 14, 2013). https://doi.org/10.5888/pcd10.120206.

He, Feng J., Caryl A. Nowson, and Graham A. MacGregor. "Fruit and Vegetable Consumption and Stroke: Meta-Analysis of Cohort Studies." *The Lancet; London* 367, no. 9507 (February 28, 2006): 320–26. http://dx.doi.org.ncat.idm.oclc.org/10.1016/S0140-6736(06)68069-0.

Hung, Hsin-Chia, Kaumudi J. Joshipura, Rui Jiang, Frank B. Hu, David Hunter, Stephanie A. Smith-Warner, Graham A. Colditz, Bernard Rosner, Donna Spiegelman, and Walter C. Willett. "Fruit and Vegetable Intake and Risk of Major Chronic Disease." *JNCI: Journal of the National Cancer Institute* 96, no. 21 (November 3, 2004): 1577–84. https://doi.org/10.1093/jnci/djh296.

John, JH, S Ziebland, P Yudkin, LS Roe, and HAW Neil. "Effects of Fruit and Vegetable Consumption on Plasma Antioxidant Concentrations and Blood Pressure: A Randomised Controlled Trial." *The Lancet* 359, no. 9322 (June 8, 2002): 1969–74. https://doi.org/10.1016/S0140-6736(02)98858-6.

Lee-Kwan, Seung Hee, Latetia V. Moore, Heidi M. Blanck, Diane M. Harris, and Deb Galuska. "Disparities in State-Specific Adult Fruit and Vegetable Consumption — United States, 2015." *MMWR. Morbidity and Mortality Weekly Report* 66, no. 45 (November 17, 2017): 1241–47. https://doi.org/10.15585/mmwr.mm6645a1.

Leone, Lucia A., Diane Beth, Scott B. Ickes, Kathleen MacGuire, Erica Nelson, Robert Andrew Smith, Deborah F. Tate, and Alice S. Ammerman. "Attitudes Toward Fruit and Vegetable Consumption and Farmers' Market Usage Among Low-Income North Carolinians." *Journal of Hunger & Environmental Nutrition* 7, no. 1 (January 1, 2012): 64–76. https://doi.org/10.1080/19320248.2012.651386.

Li, Yi Chong, Bo Jiang, Mei Zhang, Zheng Jing Huang, Qian Deng, Mai Geng Zhou, Zhen Ping Zhao, You Fa Wang, and Li Min Wang. "Vegetable and Fruit Consumption among Chinese Adults and Associated Factors: A Nationally Representative Study of 170,847 Adults." *Biomedical and Environmental Sciences* 30, no. 12 (December 1, 2017): 863–74. https://doi.org/10.3967/bes2017.117.

McGuire, Shelley. "State Indicator Report on Fruits and Vegetables, 2013, Centers for Disease Control and Prevention, Atlanta, GA." *Advances in Nutrition* 4, no. 6 (November 6, 2013): 665–66. https://doi.org/10.3945/an.113.004598.

McGuirt, Jared T., Stephanie B. Jilcott, Haiyong Liu, and Alice S. Ammerman. "Produce Price Savings for Consumers at Farmers' Markets Compared to Supermarkets in North Carolina." *Journal of Hunger & Environmental Nutrition* 6, no. 1 (February 28, 2011): 86–98. https://doi.org/10.1080/19320248.2010.551031.

"PATHS_NC_Healthy-Food-Access-Fact-Sheet-Final.Pdf." Accessed June 14, 2020. https://www.chlpi.org/wp-content/uploads/2013/12/PATHS_NC_Healthy-Food-Access-Fact-Sheet-Final.pdf.

Reinaerts, Evelien, Jascha de Nooijer, Math Candel, and Nanne de Vries. "Explaining School Children's Fruit and Vegetable Consumption: The Contributions of Availability, Accessibility, Exposure, Parental Consumption and Habit in Addition to Psychosocial Factors." *Appetite* 48, no. 2 (March 1, 2007): 248–58. https://doi.org/10.1016/j.appet.2006.09.007.

Riediger, Natalie D., and Mohammed H. Moghadasian. "Patterns of Fruit and Vegetable Consumption and the Influence of Sex, Age and Socio-Demographic Factors among Canadian Elderly." *Journal of the American College of Nutrition* 27, no. 2 (April 1, 2008): 306–13. https://doi.org/10.1080/07315724.2008.10719704.

Shannon, Jerry. "Beyond the Supermarket Solution: Linking Food Deserts, Neighborhood Context, and Everyday Mobility." *Annals of the American Association of Geographers* 106, no. 1 (January 2, 2016): 186–202. https://doi.org/10.1080/00045608.2015.1095059.

Smith, Chery, and Lois W. Morton. "Rural Food Deserts: Low-Income Perspectives on Food Access in Minnesota and Iowa." *Journal of Nutrition Education and Behavior* 41, no. 3 (May 1, 2009): 176–87. https://doi.org/10.1016/j.jneb.2008.06.008.

United States Department of Agriculture, "2015-2020 Dietary Guidelines for Americans,": 2015 https://health.gov/sites/default/files/2019-09/2015-2020_Dietary_Guidelines.pdf

Van duyn, MARY ANN S, and ELIZABETH Pivonka. "Overview of the Health Benefits of Fruit and Vegetable Consumption for the Dietetics Professional: Selected Literature." *Journal of the American Dietetic Association* 100, no. 12 (December 1, 2000): 1511–21. https://doi.org/10.1016/S0002-8223(00)00420-X.

Walker, Renee E., Christopher R. Keane, and Jessica G. Burke. "Disparities and Access to Healthy Food in the United States: A Review of Food Deserts Literature." *Health & Place* 16, no. 5 (September 2010): 876–84. https://doi.org/10.1016/j.healthplace.2010.04.013.

Wang, May C., Kara E. MacLeod, Cathy Steadman, Laura Williams, Shene L. Bowie, Denise Herd, Mia Luluquisen, and Mark Woo. "Is the Opening of a Neighborhood Full-Service Grocery Store Followed by a Change in the Food Behavior of Residents?" *Journal of Hunger & Environmental Nutrition* 2, no. 1 (May 2007): 3–18. https://doi.org/10.1080/19320240802077789.

Wang, Xia, Yingying Ouyang, Jun Liu, Minmin Zhu, Gang Zhao, Wei Bao, and Frank B. Hu. "Fruit and Vegetable Consumption and Mortality from All Causes, Cardiovascular Disease, and Cancer: Systematic Review and Dose-Response Meta-Analysis of Prospective Cohort Studies." *BMJ* 349 (July 29, 2014). https://doi.org/10.1136/bmj.g4490.

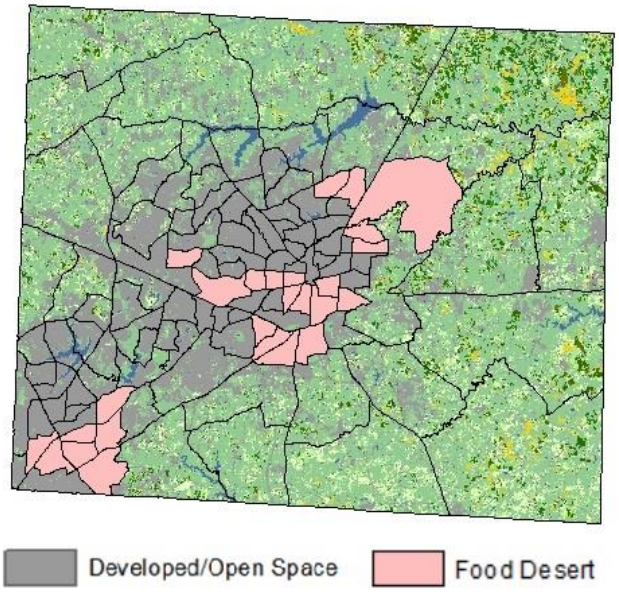

Figure 1: Food deserts surrounded by undeveloped area/farms (represented by colors other than gray and pink) in Guilford County, NC

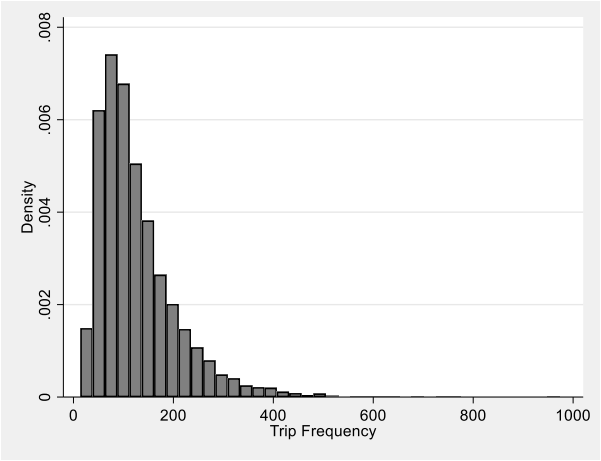

Figure 2: Distribution of yearly take-home-food trip frequency

Table 1. Descriptive Statistics

Variables	Description	Median	Mean	Std. Dev.
Panel A. yearly take-home food consumption (Year 2010 $)				
Vege-all	Consumption of fresh vegetables	89.75	124.68	122.09
Vege-full	Consumption of fresh vegetables in a full-service grocery store	87.43	121.30	119.25
All-food	Consumption of all food from grocery shopping trips	2937.85	3283.11	1785.76
Panel B. Yearly grocery shopping trip characteristics				
Trip-all	Grocery shopping trip numbers	106	127.33	80.66
Trip-full	Grocery shopping trip numbers to a full-service grocery store	80	92.29	53.69
Panel C. Household characteristics				
Food Desert	= 1 if living in a zip code intersecting at least one food desert, = 0 otherwise	1	0.68	0.470
Vehicle	= 1 if owning a vehicle, = 0 otherwise	1	0.71	0.450
Income	Household income ($1000s)	50.60	59.28	34.54
Education years	Average of education years of household head(s)	14	14.58	1.880
Household size	Member numbers of a household	2	2.44	1.250
Age	Average of ages of household head(s)	55	54.22	12.11
Hours	Average of weekly working hours of household head(s)	20	22.22	15.79
White	= 1 if white households, = 0 otherwise	1	0.75	0.430
Black	= 1 if black households, = 0 otherwise	0	0.20	0.400
Asian	= 1 if Asian households, = 0 otherwise	0	0.02	0.130
Hispanic	= 1 if Hispanic households, = 0 otherwise	0	0.03	0.170
Married	= 1 if married households, = 0 otherwise	1	0.68	0.470
Child	= 1 if any child under 18 years old in the households, = 0 otherwise	0	0.24	0.430
Employed	= 0 if no head employed; = 0.5 if one of the heads employed in a two heads household; = 1 if head(s) employed.	0.5	0.60	0.420

Notes: The sample includes 8,639 household-by-year observations for 2012-2017 in North Carolina.

Table 2. Descriptive statistics of household characteristics by food-desert status

Variables	**Non-food-desert**			**Food-desert**			Balance test (χ^2)
	Median	Mean	S.D.	Median	Mean	S.D.	*p*-value
Vehicle	1	0.70	0.45	1	0.71	0.45	0.45
Income	55.52	64.24	35.89	49.50	56.97	33.65	0.00***
Education years	15	14.70	1.86	14	14.53	1.89	0.00***
Household size	2	2.50	1.24	2	2.42	1.25	0.00***
Age	55	53.98	12.18	55	54.33	12.07	0.108
Hours	20	22.81	15.40	20	21.94	15.97	0.00***
White	1	0.83	0.37	1	0.71	0.45	0.00***
Black	0	0.10	0.30	0	0.24	0.43	0.00***
Asian	0	0.02	0.14	0	0.01	0.12	0.00***
Hispanic	0	0.04	0.19	0	0.03	0.16	0.03**
Married	1	0.73	0.44	1	0.66	0.47	0.02**
Child	0	0.25	0.44	0	0.23	0.42	0.00***
Employed	0.50	0.62	0.41	0.50	0.59	0.42	0.00***

Notes: Sample size for non-food-desert households is 2,748, sample size for food-desert households is 5,891. Total sample includes 8,639 household-by-year observations for 2012-2017 in North Carolina. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3. Association between factors and store choices/probability of buying fresh vegetables

Independent variables	Going to full-service stores	Buy fresh vegetable at full-service stores	Buy fresh vegetable at non-full-service stores
	(1)	(2)	(3)
Food desert	-0.0220***	-0.0147***	-0.00135*
	(0.00137)	(0.00161)	(0.000726)
Vehicle	-0.127***	-0.00343**	0.000411
	(0.00136)	(0.00166)	(0.000846)
Income	0.000590***	0.000289***	5.41e-05***
	(2.29e-05)	(2.47e-05)	(1.05e-05)
Education years	0.0152***	0.00777***	0.000690***
	(0.000354)	(0.000424)	(0.000201)
Household size	-0.00144*	-0.00417***	-0.00153***
	(0.000757)	(0.000890)	(0.000438)
Age	-0.000178***	0.000382***	0.000379***
	(6.74e-05)	(7.97e-05)	(3.52e-05)
Hours	0.00138***	0.00211***	0.000124
	(0.000163)	(0.000203)	(8.31e-05)
White	-0.00543	0.00765**	-0.00177
	(0.00344)	(0.00390)	(0.00163)
Black	-0.0492***	-0.0511***	-0.00545***
	(0.00354)	(0.00409)	(0.00172)
Asian	0.0877***	0.0221***	0.0126***
	(0.00651)	(0.00636)	(0.00244)
Married	-0.00643**	-0.0216***	-0.00133
	(0.00305)	(0.00364)	(0.00161)
Hispanic	0.0503***	0.0506***	0.00607***
	(0.00157)	(0.00190)	(0.000801)
Child	0.0489***	-0.00797***	0.00908***
	(0.00217)	(0.00251)	(0.00123)
Employed	-0.0952***	-0.0775***	-0.00686**
	(0.00587)	(0.00736)	(0.00303)
N	1,107,754	801,972	305,782

Notes: Weighted logit regressions for model 1-3. Marginal effects are reported. Delta-method standard error in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4**.** Determinants of yearly food consumption patterns using all the year × household observations

Independent variables	Fresh Vegetable spending ($)	All food spending ($)	Trip frequency to all food stores	Trip frequency to full-service stores
	(1)	(2)	(3)	(4)
Food desert	-10.42*	22.60	8.847**	3.266
	(5.391)	(83.77)	(3.933)	(2.628)
Vehicle	10.45**	207.7***	30.72***	8.498***
	(4.473)	(72.08)	(3.440)	(2.427)
Income	0.318***	3.973***	-0.0327	0.0372
	(0.0771)	(1.352)	(0.0862)	(0.0503)
Education years	5.075***	-61.26***	-0.569	1.358*
	(1.275)	(18.77)	(1.110)	(0.708)
Household size	7.414***	318.5***	5.144***	3.470***
	(2.526)	(41.12)	(1.957)	(1.316)
Age	0.449**	5.928*	0.835***	0.603***
	(0.226)	(3.351)	(0.165)	(0.117)
Hours	-0.102	-18.29**	-0.873*	-0.558*
	(0.530)	(8.751)	(0.500)	(0.291)
White	2.458	154.2	-3.191	-3.111
	(10.60)	(188.2)	(8.821)	(6.361)
Black	-19.67*	-225.1	15.87*	4.019
	(10.68)	(198.9)	(9.420)	(6.734)
Asian	9.974	-584.7**	-5.981	2.912
	(24.21)	(271.3)	(13.39)	(12.43)
Married	12.08	291.3*	23.74**	15.77**
	(13.32)	(159.8)	(9.456)	(6.440)
Hispanic	41.18***	635.5***	9.274*	12.05***
	(5.701)	(90.08)	(4.967)	(3.146)
Child	-21.81***	-25.86	-15.74***	-5.246
	(7.185)	(117.7)	(5.546)	(3.730)
Employed	-5.349	577.9*	30.74*	14.63
	(19.95)	(312.1)	(16.92)	(10.27)
N	8,639	8,639	8,639	8,639

Notes: Weighted least square regressions for all models. Robust standard errors (clustered at the household level) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5**.** Effect of food desert on vegetable consumption (propensity score matching method)

	Consumption of fresh vegetables ($)				
	Full sample	Income Q1	Income Q2	Income Q3	Income Q4
Food desert	-2.89	-22.49***	2.87	-11.74	24.43
	(5.48)	(9.46)	(9.86)	(12.90)	(16.90)
N	8,639	2,161	2,185	2,288	1,903

Notes: propensity score matching method for all models. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6. Effect of food desert on vegetable consumption (movers' response)

Independent variables	Fresh Vegetable spending ($)	All food spending ($)	Trip frequency to all food stores	Trip frequency to full-service stores
	(1)	(2)	(3)	(4)
Food desert	9.58	475.42	-3.26	7.26
	(15.66)	(324.39)	(10.70)	(6.08)
Control of Household characteristics	X	X	X	X
N	254	254	254	254

Notes: Weighted least square for all models. Robust standard errors (clustered at the household level) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7. Effect of the portion of residents living in food deserts on consumption patterns

Independent variables	Fresh Vegetable spending ($)	All food spending ($)	Trip frequency to all food stores	Trip frequency to full-service stores
	(1)	(2)	(3)	(4)
Portion of residents living in food deserts	-10.55	25.75	9.56	3.70
	(11.44)	(195.25)	(10.59)	(6.37)
Control of Household characteristics	X	X	X	X
N	8,639	8,639	8,639	8,639

Notes: Weighted least square for all models. Robust standard errors (clustered at the household level) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8. Effect of driving distance and driving time on consumption patterns

Independent variables	Fresh Vegetable spending ($)	All food spending ($)	Trip frequency to all food stores	Trip frequency to full-service stores
	(1)	(2)	(3)	(4)
Driving distance	-2.42	-2.84	0.76	-1.86*
	(1.58)	(31.49)	(1.73)	(0.99)
Driving time	1.40	3.41	-1.04	0.40
	(0.92)	(20.42)	(0.96)	(0.57)
Control of Household characteristics	X	X	X	X
N	8,639	8,639	8,639	8,639

Notes: Weighted least square for all models. Robust standard errors (clustered at the household level) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.