Spatial differences in fresh vegetable spending: A case study in Guilford County, North Carolina

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

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

[2] North Carolina Central University, Durham, NC

Abstract

This paper investigates the spatial differences in fresh vegetable spending in Guilford County, North Carolina. We create a geo-coded spatial-temporal database for both human factors and natural factors to understand why food deserts have become a serious issue in a county with many farming activities. Using an agent-based toy model, we find that the formation of food deserts may root from the demand side. Social-economic factors are most sensitive and are important determinants of fresh food demand based on the simulation results. Food deserts are more likely to be equilibrium responses to consumers' demand. We find that residents living in food deserts in Guilford County do not buy enough fresh food compared with their counterparts, even when they are shopping at supermarkets far away from home.

Keywords: Fresh vegetable consumption, food desert, agent-based model

[i] hmiao@ncat.edu

1 Introduction

Vegetables are important in a healthy diet. However, only ten percent adult Americans meet the vegetable consumption standard recommended by Dietary Guidelines for Americans and vegetable is the most under-consumed nutritional food in the States (2018 State Indicator Report on Fruits and Vegetables). There is growing literature in enhancing food security, especially providing sufficient fresh vegetables to communities through local food systems (e.g., Torjusen, Lieblein &Vittersø 2008; Wilkins, Farrell & Rangarajan 2015). Local food systems can improve food access (Kantor 2001), decrease carbon footprint (Kaplin 2011), reduce consumers' energy intake (Rose et al. 2008), ease the food desert problem (McKinney and Kato, 2017), one type of geographic disparities in the food supply. We propose to mitigate the food desert problem by using local food systems from a coupled human and natural system perspective. We are particularly focusing on improving the availability and accessibility of fresh vegetables in the food desert area by making good use of local food systems. This research takes the first step to investigate spatial differences of fresh vegetable demand and use Guilford County, North Carolina (NC), as a case study.

A food desert is a metaphor for neighborhood health food deprivation. It has many versions of definition. In general, it refers to areas with low access to affordable fresh vegetables and fruits, usually in the units of census tracts. The United States Department of Agriculture (USDA) defines one census tract as a food desert if it meets two thresholds. 1) "a poverty rate of 20 percent or greater, or a median family income at or below 80 percent of the statewide or metropolitan area median family income; 2) "at least 500 persons and/or at least 33 percent of the population lives more than 1 mile from a supermarket or large grocery store (10 miles, in the case of rural census tracts)". In other words, one measurement of food deserts by USDA is whether most residents in one census tract have nearby access to a full-service grocery store, which serves a variety of fresh vegetables and fruits, such as a supermarket or a wholesale market. Based on recent data (USDA ERS 2010, 2015; US Census Bureau 2017), we find the food deserts in North Carolina are expanding, and the number of full-service grocery stores is declining. Some solutions were taken by policymakers to ease the food desert problem. But very few cases had meaningful effects. For example, In Greensboro, NC, Renaissance Co-op opened as a food desert rescuer, but they were out of business in two years due to a lack of enough demand.

Guilford County, NC, has a population of more than half a million. It is divided into about 120 census tracts, and 21 of them are defined as food deserts by USDA. In Figure 1a, we find those deserts are surrounded by the undeveloped area where exists about 854 farms (USDA 2017, Census of Agriculture), and those farms are capable of producing a large variety of fresh vegetables. If we further overlay individual household addresses in the map (see Figure 1b), we may find that a significant portion of households living in food desert areas. USDA estimates that about one-fifth of the county's population is underserved by full-service grocery stores. One important difference between one full-service grocery store and one corner store is the different supply of fresh vegetables and fruits. As we know, supply and demand always come side by side, and most literature argues that the food desert area (e.g. Allcott et al. 2019) is underserved mostly because of the low demand for healthy food. However, we find that literature mostly focuses on finding the association between food environment/food deserts and dietary behavior/healthy food at the individual level. There is a lack of literature that appropriately calculate the spatial differences in vegetable demand between food-desert and nonfood-desert area. The gap in the literature leads to our first step research question: what the spatial differences in fresh vegetable demand between food-desert and nonfood-desert census tracts in Guilford County are. We propose to use a combination of a private dataset and several public datasets to solve this problem. The technique to calculate the aggregated fresh vegetable consumption is agent-based modeling, which is a from-bottom-to-up approach. In general, we start from household level and use agents in the software NETLOGO to represent about 200k household of Guilford County. Their fresh vegetable purchasing behavior is affected by the food environment and their household features, which are programmed as one set of fresh vegetable purchasing rules in the NETLOGO. We use the software to simulate each household behavior for one year, and then we aggregate fresh vegetable demand for each census tract.

2 Data

The dataset that we first look at in our study area is the Cropland Data Layer (CDL) from the USDA National Agricultural Statistics Service (NASS). Using images remotely sensed by satellites and national agricultural statistics, USDA NASS creates the CDL dataset annually to illustrate how the US continents are covered by specific crops, includes different types of vegetables. CDL data contain one layer of raster data that show how specific crops are distributed over space and over time. To fulfill the purpose of the paper, we mainly utilize the CDL dataset in the year 2019 to find out how the current spatial relationships among vegetable growing parcels, household addresses, and food desert areas. For next steps, we would apply CDL data in about ten years to estimate how land use, especially vegetable growing area, changed over space and time, and whether they are correlated with farm-level characteristics and uncertain scenarios, such as a flood.

The second dataset that we use for our study area is the RTI international's US synthetic household population dataset (Wheaton et al. 2009). Because we use a bottom-up strategy to estimate how fresh vegetable demand is spatially different using an agent-based model and socioeconomic differences matter in the correlation between food deserts and dietary behavior (Mackenbach et al. 2019), we need data containing the spatial distribution of households and the characteristics of each household in our study area. Different from aggregated data at census tracts or zip code levels, The RTI international's US synthetic household population dataset represents an accurate and complete set of household addresses and household (member) characteristics, such as household income, member ages, race, household size, etc. Therefore, in the agent-based model, one household, represented by one agent, can follow specific rules and make fresh fruit purchasing decisions based on their household-specific data (e.g., whether one household locates in food desert area). Also, we can intuitively understand the spatial relationships between the food desert area/fresh vegetable production area and household locations.

The third dataset, which we use for extracting behavior patterns, is Nielsen Homescan dataset. Nilsen Homescan data are one type of national-level dataset provided by Nielsen company. The company has a balanced sample all over the States, and each panelist in the sample reports all purchased items, including all kinds of fresh vegetables. We can observe the types of vegetables, unit price, and total spending of one vegetable item from the dataset. Also, we can observe the types of stores where those vegetables are bought from, such as a wholesale club, a supermarket, or a convenience store.

Furthermore, each panelist reports the location of his/her household, and household characteristics, such as household income, household member ages, employment status, race, etc. We use the subset of North Carolina so that we can better mimic the fresh vegetable purchasing behavior in our study area. The dataset records panelists' each grocery shopping trip, and each item they bought during one trip. We also combine the Nielsen Homescan data and USDA food desert data to estimate the food environment.

3 Method

We aim to build an integrated agent-based model to solve our research question in this paper. The overarching goal of the agent-based model is to build a food system that includes vegetable production and consumption, and related environmental impact. Thus, the agent-based model includes three types of agents, households, farms, and environmental agents. In Table 1, we identify the rules of the agents in the second column. Households make one to four decisions each tick or day based on food environment and household characteristics: whether they are going to do grocery shopping, which store they are going to do grocery shopping, whether they are going to buy fresh vegetables, and how much they are going to spend on fresh vegetables. By analyzing Nielsen Homesman data, we can find the parameters for fresh vegetable purchasing behavior. As farm agents, they decide whether they change the use of land at the beginning of one season, either turn vegetable planting into other uses, or vice versa. Based on the previous years' CDL data, we model the change as functions of changes in previous years' consumption and changes in vegetable prices (Second Column, Table 1). The environmental agents would take variables from Farm actions, such as land-use change, to simulate environment changes, such as water quality, using Agricultural Policy/Environmental eXtender (APEX model). The three parts, production, consumption, and environment, would be linked and integrated by variables proposed in the third column, Table 1.

4 Results

Table 2 presents a simple description of our sample data. We can see that vehicle ownership and age are not statistically different between nonfood-desert and food-desert households. Other variables, such as income, education year, household size, are sharply different. And we can see a large portion of minorities living in a food desert area.

After looking into the data more systematically, we can conclude a few consumption patterns. Firstly, the numbers of total trips in a year and the numbers of trips to a full-service grocery store are not different from each other between food desert households and their counterpart. In a year, both food desert and nonfood desert households visit full-service grocery stores 92 times on average.

If we take other confounders together and try to understand how they affect the probabilities of households going to a full-service grocery store, we can see the ownership of a vehicle is the most powerful indicator (Table 3). It decreases the probability of choosing a full-service grocery store by about 15%. Intuitively, one household will have many choices if they own vehicles. The food environment (food desert status) matters, but the magnitude seems small. The ethnic groups also play an important role in choosing grocery stores.

Secondly, we find that most vegetables are bought in a full-service grocery store. After arriving at a full-service grocery store (Table 4), the logistic regression shows that owning a vehicle does affect the probabilities of buying vegetables. Ethnic group variables still play a big role.

Thirdly, we find that the fresh vegetable spending during each trip is different if we compare the two groups. Nonfood desert households buy 4.35-dollar fresh vegetables on average in each trip, while the number for their counterparts is 3.95 dollars. If we add impacts of other confounders by running a weighted least square regression (Table 5), we can find that the impact of food deserts is still significant. Minorities groups do not have a big difference in consumption if they arrive at a full-service grocery store and decide to buy vegetables.

After we implement those patterns in the agent-based model and simulate agent's behavior with other environmental data, such as household food environment, and household characteristics, we can find the spatial differences in yearly fresh vegetable consumption in Guilford County (Table 6). We can see that during a year, a nonfood desert census tract consumes about 184k dollars of fresh vegetables on average, while a food desert census tract consumes about 124k dollars. Nonfood census tracts consumers about 50% more than their counterparts.

5 Conclusion and future steps

In this paper, we propose a way to build an integrated model of a local food system using the agent-based model technique. We take the first step to estimate the spatial differences in fresh vegetable spending in Guilford County, NC, using the model. Combing a few private and public datasets, we can estimate aggregated fresh vegetable spending at the census tract level. We find that census tracts defined as food deserts by USDA have a much lower fresh vegetable demand, which indicates Food deserts may be equilibrium responses to consumers' demand.

Soon, we will develop one more integrated model, which includes consumers, farmers, and environmental agents in the model as agents. More data, such as remote sensing data, GIS info, land use data, will also be included in the model as environment data. Ultimately, we want to find out how the local food system can help alleviate the food desert problem, and any environmental consequences would happen.

Acknowlegement

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 (US), 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

"2014-Economics-of-Local-Food-Systems.Pdf." Accessed July 11, 2020. https://conservancy.umn.edu/bitstream/handle/11299/171637/2014-Economics-of-Local-Food-Systems.pdf?sequence=1.

"2018 State Indicator Report on Fruits and Vegetables," n.d., 18.

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.

Dong, Diansheng, and Biing-Hwan Lin. "Fruit and Vegetable Consumption by Low-Income Americans: Would a Price Reduction Make a Difference?," n.d., 23.

Kantor, Linda Scott. "Community Food Security Programs Improve Food Access." Food Review/ National Food Review, January 1, 2001. https://doi.org/10.22004/ag.econ.266234.

Kaplin, Lauren. "Energy (In)Efficiency of the Local Food Movement: Food for Thought" 23 (n.d.): 24.

Mackenbach, Joreintje D., Kyra G. M. Nelissen, S. Coosje Dijkstra, Maartje P. Poelman, Joost G. Daams, Julianna B. Leijssen, and Mary Nicolaou. "A Systematic Review on Socioeconomic Differences in the Association between the Food Environment and Dietary Behaviors." *Nutrients* 11, no. 9 (September 2019): 2215. https://doi.org/10.3390/nu11092215.

Markowitz, Lisa. "Produce(Ing) Equity: Creating Fresh Markets in a Food Desert." In *Hidden Hands in the Market: Ethnographies of Fair Trade, Ethical Consumption, and Corporate Social Responsibility*, edited by Geert De Neve, Luetchford Peter, Jeffrey Pratt, and Donald C. Wood, 28:195–211. Research in Economic Anthropology. Emerald Group Publishing Limited, 2008. https://doi.org/10.1016/S0190-1281(08)28009-1.

McClintock, Nathan, Jenny Cooper, and Snehee Khandeshi. "Assessing the Potential Contribution of Vacant Land to Urban Vegetable Production and Consumption in Oakland, California." *Landscape and Urban Planning* 111 (March 1, 2013): 46–58. https://doi.org/10.1016/j.landurbplan.2012.12.009.

McKinney, Laura, Yuki Kato, 1 Department of Sociology, Tulane University, USA, and 2 Department of Sociology, Georgetown University, USA. "Community Context of Food Justice: Reflections on a Free Local Produce Program in a New Orleans Food Desert." *AIMS Agriculture and Food* 2, no. 2 (2017): 183–200. https://doi.org/10.3934/agrfood.2017.2.183.

Morrison, Kathryn T., Trisalyn A. Nelson, and Aleck S. Ostry. "Mapping Spatial Variation in Food Consumption." *Applied Geography*, Local Food Systems and the Applied Geography of Food, 31, no. 4 (October 1, 2011): 1262–67. https://doi.org/10.1016/j.apgeog.2010.11.020.

MS, Nick Rose, Elena Serrano PhD, Kathy Hosig PhD RD MPH, Carola Haas PhD, Dixie Reaves PhD, and Sharon M. Nickols-Richardson PhD RD. "The 100-Mile Diet: A Community Approach to Promote Sustainable Food Systems Impacts Dietary Quality." *Journal of Hunger & Environmental Nutrition* 3, no. 2–3 (August 25, 2008): 270–85. https://doi.org/10.1080/19320240802244082.

Nolasco, Camille, Luciana Soler, Marcos Freitas, Myanna Lahsen, and Jean Ometto. "Scenarios of Vegetable Demand vs. Production in Brazil: The Links between Nutritional Security and Small Farming." *Land* 6, no. 3 (July 25, 2017): 49. https://doi.org/10.3390/land6030049.

Noy, Sue, Rebecca Patrick, Claire Henderson-Wilson, Sonia Nuttman, and Isabella Ryan. "New Frontiers in Community Initiatives to Increase Vegetable Consumption." *Health Promotion Journal of Australia* 30, no. S1 (2019): 52–61. https://doi.org/10.1002/hpja.207.

Osvald, Ana, and Lidija Zadnik Stirn. "A Vehicle Routing Algorithm for the Distribution of Fresh Vegetables and Similar Perishable Food." *Journal of Food Engineering* 85, no. 2 (March 1, 2008): 285–95. https://doi.org/10.1016/j.jfoodeng.2007.07.008.
Peters, Christian, Nelson L. Bills, Jennifer Wilkins, and R. David Smith. "Vegetable Consumption, Dietary Guidelines and Agricultural Production in New York State—Implications for Local Food Economies." AgEcon Search, 2002. https://doi.org/10.22004/ag.econ.122636.
Torjusen, Hanne, Geir Lieblein, and Gunnar Vittersø. "Learning, Communicating and Eating in Local Food-Systems: The Case of Organic Box Schemes in Denmark and Norway." *Local Environment* 13, no. 3 (April 15, 2008): 219–34. https://doi.org/10.1080/13549830701669252.
Wheaton, William D., James C. Cajka, Bernadette M. Chasteen, Diane K. Wagener, Philip C. Cooley, Laxminarayana Ganapathi, Douglas J. Roberts, and Justine L. Allpress. "Synthesized Population Databases: A US Geospatial Database for Agent-Based Models." *Methods Report (RTI Press)* 2009, no. 10 (May 1, 2009): 905. https://doi.org/10.3768/rtipress.2009.mr.0010.0905.
Wilkins, Jennifer L., Tracy J. Farrell, and Anusuya Rangarajan. "Linking Vegetable Preferences, Health and Local Food Systems through Community-Supported Agriculture." *Public Health Nutrition* 18, no. 13 (September 2015): 2392–2401. https://doi.org/10.1017/S1368980015000713.

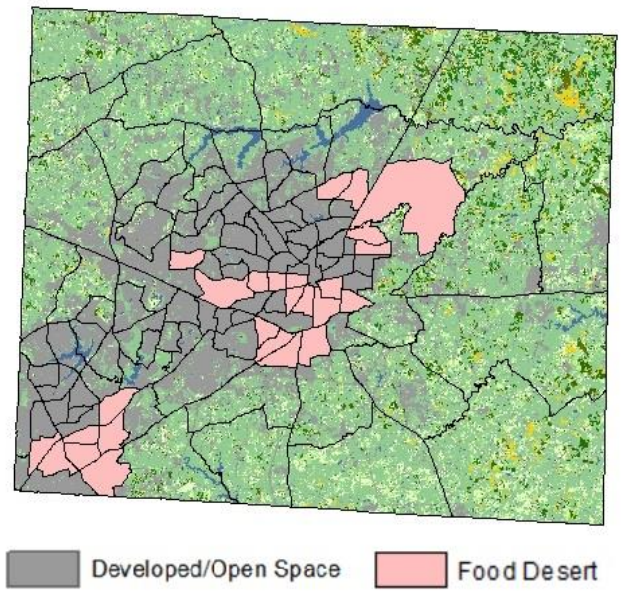

Figure 1a: Food deserts surrounded by undeveloped area/farms (represented by colors other than gray and pink) in Guilford County, NC

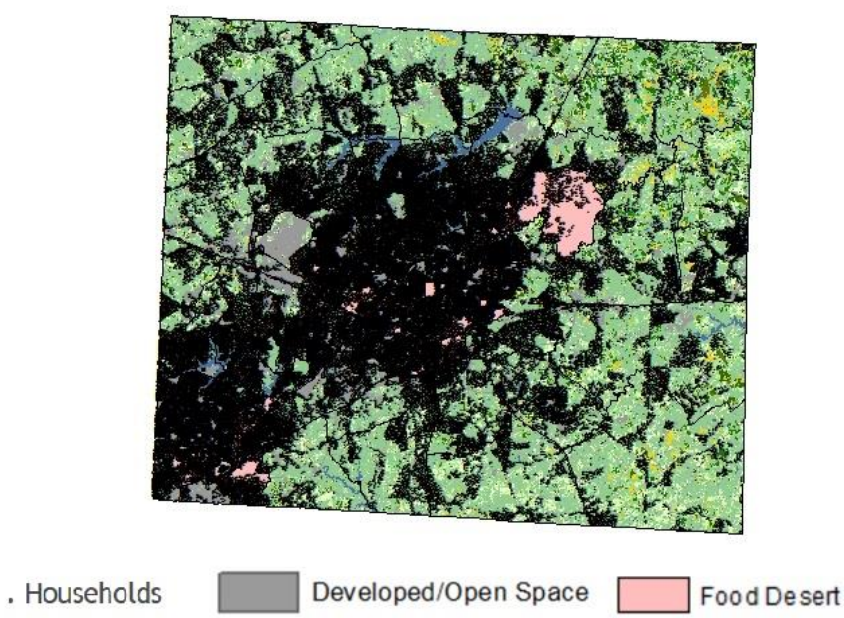

Figure 1b: Food deserts surrounded by undeveloped area/farms (represented by colors other than gray and pink) and households scattered in Guilford County, NC

Table 1. Architecture for the agent-based model

<table>
<tr><th>Agent types</th><th>Rules</th><th>Parameters and variables</th></tr>
<tr><td>Households</td><td>Decision rules:
1) Decrease food stock (FS) every day based on household size (hs), gender(g), age (a), etc. every tick (every day).
2) Go for shopping based on FS:
a. Decide which store(s) to go based on mobility (m), income (in), living in a food desert (fd), etc. (In the first-stage model, we do not have food vendor agents, so we will skip this step)
b. Decide what items (prob. of one item to be one vegetable $prob._v$) and how many (q, q_v)to buy based on household size, income, occupation (oc), income, etc.
c. Total vegetable consumption (Q) equals the summation of vegetables bought by all households</td><td>Parameters corresponding to hs, g, m, in, oc, fd
Parameters: $prob._v$
Variables: FS, q, q_v</td></tr>
<tr><td>Farms</td><td>1) Decide what to plant on farms at the beginning of every year based on previous experience, previous year demand (Q), previous years' price(p)
<table><tr><th></th><th>Vegetable</th><th>Non-Vegetable</th></tr><tr><td>Vegetable</td><td>$0.8*(1+\frac{Q_t-Q_{t-1}}{Q_t}+\frac{p_t-p_{t-1}}{p_t})$</td><td>$1 - 0.8*(1+\frac{Q_t-Q_{t-1}}{Q_t}+\frac{p_t-p_{t-1}}{p_t})$</td></tr><tr><td>Non-Vegetable</td><td>$0.1*(1+\frac{Q_t-Q_{t-1}}{Q_t}+\frac{p_t-p_{t-1}}{p_t})$</td><td>$1 - 0.1*(1+\frac{Q_t-Q_{t-1}}{Q_t}+\frac{p_t-p_{t-1}}{p_t})$</td></tr></table>2) Total vegetable supply must be greater than or equal Q.
3) In the second stage of the model, farms take the yield of vegetables (Y) and production costs into consideration (C).</td><td>Parameter: probabilities
Variables: Q, p, Y, C</td></tr>
<tr><td>Environmental Agent</td><td>The environmental agent has input such as precipitation, soil condition, land use., etc., and has output such as yield, water quality, soil erosion, etc.</td><td>Parameters: APEX parameters
Variables: Yield, environmental indicators</td></tr>
</table>

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

Variables	**Nonfood deserts**			**Food deserts**			Balance test (χ^2)
	Median	Mean	SD.	Median	Mean	SD.	*p*-value
Vehicle	1	0.70	0.45	1	0.71	0.45	0.45
Income	55.52	64.24	35.89	49.50	56.98	33.66	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	54.32	12.07	0.108
Hours	20	22.81	15.40	20	21.95	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 residents is 2,748, sample size for residents living in or close to a food desert is 5,891. Total sample includes 8,640 household-by-year observations for 2012-2017 in North Carolina. Vehicle = 1 if owning a vehicle, = 0 otherwise; Income: household income ($1000s); Education years: Average of education years of household head(s); Household size: member numbers of a household; Age: Average of ages of household head(s); Hours: Average of weekly working hours of household head(s); White = 1 if white households, = 0 otherwise; Black = 1 if black households, = 0 otherwise; Asian = 1 if Asian households, = 0 otherwise; Hispanic = 1 if Hispanic households, = 0 otherwise; Married = 1 if married households, = 0 otherwise; Child = 1 if any child under 18 years old in the households, = 0 otherwise; Employed = 0 if no head employed; = 0.5 if one of the heads employed in a two heads household; = 1 if head(s) employed. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3. Determinants of choosing full-service grocery stores

Food Deserts	**Vehicle**	**Income**	**Education**	**Household size**	**Age**	**Black**	**Asian**	**Child**
-1.54%* **	-15.50%***	0.06%***	0.99%***	0.46%***	0.06%***	-5.50%***	4.90%***	3.63%***
(0.00093)	(0.0011)	(0.000015)	(0.00024)	(0.00055)	(0.000004)	(0.002)	(0.0045)	(0.0015)

Notes: Logit regression results for the prob. of choosing a full-service grocery store, N = 1,107,754. First row: Marginal probability. Second row: Delta-method standard errors. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Determinants of buying vegetables

Food Deserts	**Vehicle**	**Income**	**Education**	**Household size**	**Age**	**Black**	**Asian**	**Child**
-2.17%* **	-0.11%	0.034%***	0.78%***	0.55%***	0.094%***	-4.12%***	3.24%***	0.11%
(0.00112)	(0.0012)	(0.000018)	(0.00031)	(0.00069)	(0.000059)	(0.0031)	(0.0047)	(0.0019)

Notes: Logit regression results for prob. of buying vegetable at a full-service grocery store, N = 801,972. First row: Marginal probability. Second row: Delta-method standard errors. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Determinants of the spending on fresh vegetables

Food Deserts	Vehicle	Income	Education	Household size	Age	Black	Asian	Child
-0.32* **	0.0045	0.0058***	0.050*	0.19***	-0.018***	-0.27	0.07	3.63%***
(0.12)	(0.10)	(0.0018)	(0.029)	(0.058)	(0.0051)	(0.26)	(0.26)	(.0015)

Notes: Weighted least square regression results, N = 264,582. First row: Marginal fresh vegetable spending. Second row: robust standard errors clustered at household level. *** *p*<0.01, ** *p*<0.05, * *p*<0.1.

Table 6: Spatial differences in yearly fresh vegetable consumption in Guilford County

	Nonfood desert census tracts				Food desert census tracts			T-test
	Median ($)	Mean ($)	SD.		Median ($)	Mean ($)	SD.	p-value
Fresh vegetable spending	177K	184K	84K		106K	124K	75K	0.00***