Applied Economic Perspectives and Policy (2020) volume 00, number 00, pp. 1–25.
doi:10.1002/aepp.13098

Submitted Article

If You Build Them… Will it Matter? Food Stores' Presence and Perceived Barriers to Purchasing Healthy Foods in the Northeastern U.S.

Lauren Chenarides*, Alessandro Bonanno, and Anne Palmer

Lauren Chenarides is an assistant professor at the Morrison School of Agribusiness at Arizona State University. Alessandro Bonanno is an associate professor at the Department of Agricultural and Resource Economics at Colorado State University. Anne Palmer is a senior research Associate at the Bloomberg School of Public Health at Johns Hopkins University.
Editor in charge: Craig Gundersen
*Correspondence may be sent to: E-mail: lauren.chenarides@asu.edu

Submitted 4 September 2020; editorial decision 11 September 2020.

Abstract *Policies incentivizing store entry or store improvements are aimed at increasing access to healthy foods; however, findings about their effectiveness to improve diets are mixed. Similarly, little is known about whether food stores' presence affects consumers' perceived barriers to purchasing healthy foods, which reflect the subjective hardships experienced by shoppers to purchase and consume healthier foods. In this study, we assess the relationship between the two most widely studied perceived barriers to purchasing healthy foods (price and availability) and the local retail food environment using individual-level survey data collected across the northeastern US and census data on the numbers of grocery stores and warehouse clubs and supercenters. Our results indicate that unobserved heterogeneity plays an important role in determining the sign and magnitude of the relationship between store presence and perceptions. The likelihood that an individual cites price or availability as a perceived barrier depends upon the barrier considered, whether respondents live in the zip code where they shop, and the method of controlling for unobserved heterogeneity. Thus, policies focusing on improving access to a given store type may only mitigate some of the negative perceptions associated with one's food environment.*

Key words: Food access, Food environment, Perceived barriers to healthy eating.

JEL codes: I14, I18, Q18.

Introduction

Because of the increasing incidence of diet-related noncommunicable diseases, improving individuals' diets has been at the forefront of policymakers' agendas across the globe (Kearney and McElhone 1999; Ver Ploeg, Dutko, and Breneman 2015; Ghosh-Dastidar et al. 2017; Allcott et al. 2019). The exposure

© 2020 Agricultural & Applied Economics Association

to certain food outlets has long been believed to influence households' food choices and subsequently the ability to follow a healthy diet (Wrigley et al. 2002; Moore and Diez Roux 2006; Black and Macinko 2008; Cobb et al. 2015; Ver Ploeg and Wilde 2018). Some studies suggest the importance of supermarkets and grocery stores' presence for the purchase of healthy,[1] affordable foods (Glanz et al. 2007; Zenk et al. 2014; Chenarides and Jaenicke 2018). Thus, policies encouraging the entry of new supermarkets are proposed to improve access to nutritious and affordable foods. The rationale behind these policies is that ease of access to food stores may lead to increased purchases of healthier foods, the adoption of healthier diets, and improved health outcomes (Larson, Story, and Nelson 2009; Sharkey 2009; Richardson et al. 2017).

However, the evidence that interventions modifying the food environment[2] lead to increased purchases and/or consumption of healthy foods is dubious,[3] whereas individuals' *perceptions* about the food environment seem to be affected by store entry. For example, Cummins, Flint, and Matthews (2014) found that the opening of a supermarket did not lead to increases in fruit and vegetable intake or a reduction in BMI; however, perceptions of access to healthy foods improved following the supermarket opening. Similarly, two studies examining the effect of supermarket entry in a neighborhood in Pittsburgh (Dubowitz et al. 2015; Ghosh-Dastidar et al. 2017) found that even though households in the neighborhood where the new store opened showed improved diet quality overall (compared to those in a neighborhood with no store entry),[4] the only outcome positively associated with using the new supermarket was an improved perception of access to healthy foods.

Surprisingly, the applied economics literature has mostly focused on measuring the relationship between objective measures of access to food and food choices (Kyureghian and Nayga 2013; Handbury, Rahkovsky, and Schnell 2015; Ver Ploeg, Dutko, and Breneman 2015), leaving a gap in our understanding of what may affect consumers' *perceived* constraints in acquiring healthy foods.[5] Ver Ploeg, Dutko, and Breneman (2015) argue that

[1]*We acknowledge the difference between* purchasing *healthy foods and* consuming *healthy foods, although existing literature indicates that purchases are a relatively accurate representation of consumption (Appelhans et al. 2017). This distinction is relevant to our study, as the data we analyze consider food purchases, and not consumption.*

[2]*Extrapolating from Cummins and Macintyre (2006), we consider the retail "food environment" to include stores where food can be purchased and consumed at home* (e.g., *supermarkets and grocery stores).*

[3]*Multiple studies find several other factors affecting shopping behaviors, purchasing patterns and food choices more than food stores per se. For example, Rahkovsky and Snyder (2015) show that demographic factors* (i.e. *presence of young children, race, education, and rurality) explain the demand for a series of food groups, more than supermarket access. Ver Ploeg and Wilde (2018) indicate that variation in the proximity of supermarkets plays a limited role to explaining where people choose to shop and households' variation in food security and dietary quality. Allcott et al. (2019) find that exposing low-income households to the same prices and product availability of better served communities, reduced nutritional differences (expressed in Healthy Eating Index - HEI - points) by less than 10%.*

[4]*The authors note in this study that the improvement in diet quality was not due to new supermarket use. They note that the "changes in diet (among the respondents in the intervention neighborhood)… occurred regardless of frequency of supermarket use" (Dubowitz et al. 2015, 1864).*

[5]*Interestingly, a strong consensus exists (in disciplines other than economics) that* perceptions *about the food environment have a stronger relationship with purchasing/dietary behaviors than objective measures of it (Kearney and McElhone 1999; Gustafson et al. 2011; Traill, Chambers, and Butler 2012; Davison et al. 2015; Farahmand et al. 2015; de Mestral, Stringhini, and Marques-Vidal 2016; de Mestral et al. 2017; McMorrow et al. 2017).*

focusing on neighborhood-level, objective measures of access to food (e.g., store presence, price levels) may underestimate the extent of the barriers faced by some individuals while overestimating the problem for others. As intervention policies can target the structural features of the food environment (i.e., number and types of stores, services offered), analyzing the relationship between store numbers and the perceived barriers to purchase healthy foods may help policymakers take more appropriate actions to mitigate issues of perceived food access hardships.

This article examines the relationship between different types of food stores in the zip code where an individual shops and his/hers stated perceived barriers to buying more healthy foods. We assess the relationship between the two most cited barriers to purchasing healthy foods—price and availability—and the presence of supermarkets and grocery stores, and wholesale clubs and supercenters. We combine survey data collected across eight locations in the US Northeast with zip code-level data on the presence of food retail outlets collected from the U.S. Census Bureau.

We focus on price and availability for three reasons. First, even though studies on individuals' perceptions and barriers to healthy eating vary in terms of geographic contexts/scope, methodology, and the definition of healthy eating used, price and availability are consistently among the most cited barriers to purchasing healthy foods (Eikenberry and Smith 2004; Inglis, Ball, and Crawford 2008; Davison et al. 2015; Farahmand et al. 2015).[6] Second, by focusing on price and availability, we complement other literature on the food environment and food purchases: In their systematic review, Glanz et al. (2016) find most of the studies reviewed focused on the availability (85.6%) and prices (64.0%) of healthful food options. Third, price levels and assortment are heavily influenced by both the number and types of stores in a given area (Hausman and Liebtag 2007; Volpe and Lavoie 2008; Basker and Noel 2009; Matsa 2011; Courtemanche and Carden 2014; Chenarides and Jaenicke 2018; Bauner and Wang 2019), making prices and availability two barriers that are likely to be affected by the number of food stores.

We believe, except for Ver Ploeg, Dutko, and Breneman (2015), to be the first applied economic study to consider how objective food environment measures are related to perceived hardships in acquiring healthy foods. Whether, in fact, a relationship exists between perceived barriers to healthy foods and its direction is an empirical question. While some small-scale studies find an association between the presence of food stores and the perception of healthy food availability (Gustafson et al. 2011), others highlight a dissonance between perceived and objective access (Caspi et al. 2012).[7] Our expectations are unclear, and because of the presence of confounding factors affecting both perceived and objective measures, unobserved heterogeneity may affect our estimates. Therefore, in our approach, we control for different sets of zip code-level fixed effects, survey-year indicators, and their interactions. Further, we allow the relationship between the number of stores and perceived barriers to purchasing healthy foods to vary conditionally on

[6] *Additionally, perceptions regarding food prices and availability are associated with fruits and vegetables consumption in ways consistent with economic intuition: An inverse relationship between perceived high prices and consumption; a positive association between perceived availability of fresh produce and consumption (Kearney and McElhone 1999; Inglis, Ball, and Crawford 2008; McMorrow et al. 2017).*

[7] *Caspi et al. (2012) found that respondents who declared they did not have access to a store within walking distance (but actually had one) had lower fruits and vegetables intake than those who correctly identified the presence of a store.*

whether shoppers live in the same zip code where they shop. As willingness to travel and shop elsewhere is more important than store proximity to determine dietary behavior (Cummins, Flint, and Matthews 2014; Dubowitz et al. 2015), it is possible that consumers *choosing* to shop in a specific area may have different perceptions about the features of the area where they shop than those who shop in the area where they reside.

Our findings indicate a generally robust result that the number of supermarkets and grocery stores is a predictor of higher (lower) perceptions of price (availability) as a barrier to purchasing more healthy foods. On the other hand, warehouse clubs and supercenters are negatively related to price as a barrier, and positively associated with availability. These estimated relationships vary in magnitude and statistical significance across models including different sets of fixed effects, and on whether the relationships are assessed on average (for all respondents) or conditionally on respondents who live in the same zip code where they shop for food. Thus, we do not find unequivocal support to the notion that, in general, the presence of more food stores is related to a reduction of perceived barriers to purchasing healthy foods.

The remainder of the article is organized as follows. First, we discuss the empirical methods used in our analysis, including estimation and identification strategy. Then, we discuss the data used. Empirical results and robustness checks are next, followed by an illustration of the policy implications of our results. Closing remarks and avenues for future research conclude.

Methods

Empirical Framework

We model an individual's perception of a given barrier to purchase more healthy foods as a function of the number and types of stores in the area where they choose to shop, as well as a set of other factors.

An individual's declared perceived barriers to purchasing more healthy foods are treated as binary variables; $B_{il}^{k}=1$ indicates that an individual declared to have experienced barrier k, $B_{il}^{k}=0$ otherwise. The determinants of the probability that individual i in zip code l experienced the k-th perceived barrier to purchasing more healthy foods, B_{il}^{k}, is:

$$P\left(B_{il}^{k}|\mathbf{NStores}_{l},\mathbf{X}_{il},\mathbf{SH}_{il},\mathbf{d}_{l}\right)=f\left(\mathbf{NStores}_{l},\mathbf{X}_{il},\mathbf{SH}_{il},\mathbf{d}_{l}|\boldsymbol{\delta}^{k},\boldsymbol{\beta}^{k},\boldsymbol{\gamma}^{k},\kappa^{k}\right)+e_{i}^{k} \quad (1)$$

where NStores_{l} is a vector of variables capturing the number and types of different food stores available in zip code l; X_{il} are individual i's characteristics and SH_{il} are individual i's shopping habits; d_{l} are time-invariant location factors as well as time-varying ones that may influence perceived barriers. δ^{k}, β^{k}, γ^{k} and κ^{k} are vectors of parameters conformable to NStores_{l}, X_{il}, SH_{il}, and d_{l}, respectively, and e_{i}^{k} is a stochastic term.

There are two issues in estimating the relationship in equation (1). First, the presence of food stores in an area is likely correlated with unobserved factors, which can also be related to an individual's perceived barrier to purchasing more healthy foods. In other words, endogeneity bias may affect the parameter estimates. Although similar issues have been resolved in the literature using instrumental variables methods (Bonanno and Li 2015; Bonanno et al. 2017; Zeng et al. 2019), given the limited geographic scope of our data, finding variables correlated with NStores and uncorrelated to unobserved variation in B^k (that is, suitable instruments that guarantee identification) would be a challenge. Alternatively, one can use a combination of time-varying and time-invariant, neighborhood-level controls to limit omitted variable bias and other unobservables that may confound results (Barrientos-Gutierrez et al. 2017; Mui, Gittelsohn, and Jones-Smith 2017). Therefore, because unobserved factors are likely to be correlated with both objective (store counts) and perceived access (the barrier outcome variables), we follow a similar approach, wherein we use indicator variables to account for time-varying and time-invariant, location-specific unobservables. Year indicators are used for the former. As for the latter, because the main variables of interest (store counts) are measured at the zip code level, we include zip code-level fixed effects to control for unobserved zip code characteristics. By including these fixed effects,[8] we control for any time-invariant or location-invariant unobservables that might bias the parameter estimates on our variable of interest (NStores).

We estimate four different specifications of the empirical model, progressively adding layers of time-varying and time-invariant fixed effects. The first specification includes both year and zip code-level fixed effects for the zip codes where the individual's primary food store is located (Shopping Zip FE). The second model includes year dummies, Shopping Zip FE, and an indicator variable for the zip code where individuals live (Residence Zip FE). These two variations do not control for changes in zip code-specific unobservables over time, so the third and fourth specifications include the same controls as the first two specifications, as well as interactions of year and zip code-level fixed effects. With that, our identification assumption is as follows: After controlling for time-varying, time-invariant, and unobserved variation in area characteristics over time and space (in the last two specifications), the remaining variation in the probability of experiencing a perceived barrier to purchasing more healthy foods will be captured by the observed individual characteristics, their shopping habits, and the presence of different types of food stores.

Before proceeding, we make another important consideration. Fixed-effects limited dependent variables models estimated using nonlinear estimators, such as probit or logit, results in the well-known "incidental parameter problem" leading to biased estimates. As discussed in Greene (2004), fixed-effects probit and logit estimates obtained using an asymptotic variance–covariance matrix are likely biased upwards (away from zero), whereas the standard errors may be biased downwards. In other words, estimates will likely be larger and show a higher level of statistical significance, than the "true" underlying relationship. Thus, we use Ordinary Least Squares (OLS) and treat equation (1) as a Linear Probability Model (LPM). Besides being free from the incidental parameters problem, LPM estimates also have the advantage of being easy to interpret and reliably represent marginal effects (Angrist and Pischke 2008). However, even though examples of analyses using LPMs can be found in the recent literature (Courtemanche et al. 2019), LPM estimates by OLS can be affected by both bias and inconsistency and their predictions may fall outside the plausible range (0–1) of probability values (Greene 2003). Thus, we also compare estimates obtained using LPM and all

[8] *It should be noted that the data at our disposal are not a panel, but repeated cross sections. Instead of using prepackaged panel estimator commands, we include fixed-effects directly in our model by means of dummy variables.*

the data at our disposal to those obtained with trimmed samples including only observations leading to predictions in the 0–1 range using the "trimmed sample estimator" (TSE), as proposed by Horrace and Oaxaca (2006). Horrace and Oaxaca's (2006) TSE is based on the notion that, if a researcher knew (and used) only the subsample of observations resulting in predicted values laying in the 0–1 interval, an LPM will likely produce consistent estimates. Estimates obtained using such subsample (whose observations are isolated using an iterative process), will show reduced finite sample bias which may be present in the full sample estimates.

Data, Variables, and Model Specification

To characterize the relationship between the number of food stores and perceived barriers to buying more healthy foods, we use individual-level survey data collected across the Northeastern US and publicly available data on the numbers of grocery stores and warehouse clubs and supercenters from the U.S. Census Zip Code Business Patterns (ZBP), described in detail below.

Survey Instrument and Sample Size. Two rounds of shoppers' intercept surveys were conducted by a team of community partners as part of the United States Department of Agriculture (USDA) Agriculture and Food Research Institute (AFRI)-GIS research project "Enhancing Food Security in the Northeast with Regional Food Systems" (EFSNE). A total of 1,997 individual surveys were collected across thirteen food retailers. Study stores met the Food Marketing Institute's definition of a supermarket and stocked a market basket of food items (see Clancy, Bonanno, and Canning 2017 for more details). Each store was located in one of nine locations across the Northeastern US. Nearly half of the stores were in zip codes with poverty rates above 20% and all were in counties with Supplemental Nutrition Assistance Program (SNAP) participation rate greater than 10%.

Individuals surveyed were adult shoppers, the primary or frequent food buyer of the household, English speakers, and able to provide informed consent. Persons who did not meet these criteria were excluded from participating.

The survey included fifteen questions,[9] collecting information on barriers to purchasing more healthy foods, respondents' food shopping habits (including information on the store where they primarily shop for food), and demographic characteristics. Respondents were also asked about their zip code of residence, which allows us to identify respondents residing in the same zip code where they habitually shop for food and those who do not.

Of the original 1,997 survey observations collected, 158 were dropped because of missing demographic information, 107 because of missing information regarding shopping habits, 16 for missing zip code-level food environment information, and two for missing responses. Of the remaining observations, we identified and retained only respondents who declared their primary store to be one of our study stores. Of the 1,082 respondents retained in the estimation sample, about 52.4% shopped for food primarily within the

[9] *The survey took between five and seven minutes to administer. We recruited shoppers as they were exiting the store and screen them as part of the consent process. We do not have a response rate per se because we did not ask people to record how many people refused to take the survey. They were given a $5 gift card to the store from which they exited.*

zip code of residence, whereas 47.6% shopped for food primarily outside their zip code of residence.

Perceived Barriers. To measure perceived barriers to purchasing more healthy foods, respondents were asked the following question: "Is there anything that might prevent you from buying more healthy foods?"[10] and were offered a list of items to choose from, with the option of selecting multiple items. The list included some of the most commonly cited perceived barriers to healthy eating reported in the literature: Price; Not available/limited availability; Taste; Family preferences; Preparation time; Need more knowledge on how to prepare healthy foods; Others; and No barriers. In our analysis, we consider only the first two barriers, Price and Not available/limited availability (herein referred to as "Availability"). As indicated in the top row of Table 1, within the estimation sample, 41.6% of the respondents cited *Price* as a barrier to purchasing more healthy foods whereas 10.4% cited *Availability*. [11]

Retail Food Environment. To capture the retail food environment of each respondent, the survey data were augmented with data on the number of food stores available in the zip code where the respondent's primary food store is located. Thus, we consider a respondent's retail food environment to be comprised of the stores present in the zip code where the respondent shops for most of their food. We construct measures of food store presence using data from the U.S. Census Bureau ZBP. Specifically, we collect the total number of Supermarkets and Other Grocery Stores - NAICS 445110 (*NGrocery*) and the number of Warehouse Clubs and Supercenters - NAICS 452910 (*NWCSC*).[12]

Additionally, we create two binary indicator variables, *Resident Shopper* and *Nonresident Shopper*, to capture whether a respondent's residence is or is not, respectively, in the zip code of their primary food store. These variables are interacted with *NGrocery* and *NWCSC* to determine whether the relationship between store presence and perceived barriers to purchasing healthy foods varies conditional on the residence status of the respondent, to account for differences in the behavior of shoppers *choosing* to shop in a specific area *versus* those shopping in the area where they reside.

Values in the middle panel of Table 1 show that there are, on average, about 13.5 supermarkets and grocery stores in the zip code where respondents shop. While every zip code has at least one grocery store, only one in twenty zip codes has a WCSC (average of about 0.04 stores per zip code). Respondents who shop in the same zip code of residence have access to a very similar number of grocery stores and supermarkets (about 7.1 stores), on average, as those who shop outside their zip code of residence (about 6.4); whereas,

[10]*To ensure a common interpretation of "healthy foods," the following statement was read to the respondents before asking the "barriers" question:* "In the next question, I am asking about healthy food, which, for this survey, refers to foods like fruits and vegetables, lean meat, low fat dairy, and whole grain breads."

[11]*Thirty-eight percent of respondents indicated that they did not face any barriers to purchasing healthy foods; whereas,* Taste *(10%),* Family Preferences *(8%),* Lack of Knowledge *(5%), and* Preparation Time *(4%), while 3% of respondents indicated "other" as a barrier.*

[12]*We follow the USDA Economic Research Service's Food Environment Atlas and Glanz et al. (2007) to restrict the store formats to outlets where food is predominantly purchased and consumed at home (FAH).*

Table 1 Summary Statistics

Variable Group	Variable	Mean	St. Dev.	Min	Max
Perceived	Price	0.416	0.493	0	1
Barrier [a]	Availability	0.104	0.306	0	1
Food	NGrocery	13.513	12.126	1	34
Environment [b]	NGrocery * Resident Shopper	7.088	11.021	0	34
	NGrocery * Nonresident Shopper	6.425	10.806	0	34
	NWCSC	0.039	0.193	0	1
	NWCSC * Resident Shopper	0.026	0.159	0	1
	NWCSC * Nonresident Shopper	0.013	0.113	0	1
	NGrocery | NWCSC = 1	5.429	0.501	5	6
	NGrocery | NWCSC = 0	13.839	12.257	1	34
Demographics [a]	Gender (Female)	0.645	0.479	0	1
	Age of Respondent	48.329	15.130	18	91
	Household Size	3.062	1.817	0	12
	Number of Children in HH	0.359	0.821	0	8
	Years of Education	12.888	2.673	0	24.5
	Participates in Food Assistance Program	0.495	0.500	0	1
	Rural	0.185	0.388	0	1
Shopping Habits [a]	Shops at Farmers Market	0.518	0.500	0	1
	Data Collected on a SNAP Day	0.403	0.491	0	1
	End of the Month	0.158	0.365	0	1
	Monthly Food Expenditure ($100 s)	2.502	2.866	0.05	34.667
N[+]	1,082				

[a]*Source of data is the shopper intercept survey (unit of analysis: individual).*
[b]*Source of data is the Zip code Business Patterns (unit of analysis: zip code).*
[+]*The number of observations for which there is no missing data.*
NGrocery*: Number of grocery stores in the zip code of the respondent's primary store.*
NWCSC*: Number of warehouse clubs and supercenters in the zip code of the respondent's primary store.*
Resident Shopper*: Indicator variable equal to one for respondents living in the same zip code as the one where their primary store is located.*
Nonresident Shopper*: Indicator variable equal to one for respondents living in a zip code different than the one where their primary store is located.*

respondents who shop in the zip code where they reside have access to more WCSC's compared to those of a nonresident shopper ($0.05*52.4\% \approx 0.026$ against $0.027*47.6\% \approx 0.013$).[13]

Shopper Characteristics and Shopping Habits. To control for heterogeneity among respondents' shopping behaviors, we rely on survey questions capturing demographic characteristics (X) and shopping habits (SH). Demographic characteristics include the respondent's gender (*Female*), age (*Age*), household size (*Household Size*), number of children under the age of five living in the house (*Child*), and maximum number of years of education achieved (*Education*). Respondents were also asked whether he/she or any member of his/her household participated in any federally funded food assistance programs, such as SNAP, the

[13]*Recall that 52.4% of the estimation sample resides in the same zip code as their primary store; 47.6% reside in a different zip code; multiplying these percentages times* NGrocery *and* NWCSC *gives the sample averages of the interactions in Table 1.*

Women, Infant, Children's Program (WIC), or National School Lunch Program (indicator variable *Program Participation*). Additionally, the metropolitan status of the site location was collected and used to create the binary variable *Rural* to characterize respondents interviewed in rural areas.

Respondents' shopping habits are captured by several variables. First, we include the monthly shopping frequency to their primary store (*Monthly Frequency*). Average dollars spent per trip is multiplied by monthly shopping frequency to calculate average monthly food expenditure (*Monthly Expenditure*). Shoppers were also asked whether they shop at farmers markets (binary variable, *Farmers Market Shopper*). In addition, the day of the month the shopper was interviewed was used to construct two binary variables indicating whether the interview took place on a day when SNAP benefits were issued (*SNAP Day*)[14] and whether the interview was conducted during the last five days of the month (*End of the Month*). These variables control for, respectively, differences in food expenditures during the SNAP-cycle (Damon, King, and Leibtag 2013) and differences in income for food purchases which may vary in different times of the month.

The respondents included in the estimation sample are predominantly female (64.5%), with an average age of approximately 48.3 years. The average household size was about three, and 35.9% of the respondents had children under the age of five living in their home. The average number of years of schooling was slightly below 13, equivalent to an associate degree. Forty-nine and a half percent of the respondents in the estimation sample participated in a food assistance program, and slightly below one-fifth (18.5%) reside in a rural area. Fifty-two percent of shoppers indicated to have shopped at farmers markets during the last month; 40.3% were interviewed on a day during the SNAP period (i.e., when monthly SNAP benefits were issued) and 15.8% were interviewed at the end of the month. On average, respondents' average monthly expenditure on groceries at their primary store of about $250.

All data manipulation and model estimation are performed using STATA V.14 (StataCorp., College Station, TX, USA).

Empirical Results

The estimated OLS food environment coefficients for the four model specifications are reported in Table 3 and Table 4 for *Price* and *Availability*, respectively. The estimates in Panel A are for the model without interactions (Baseline model), those in Panel B for a model where we interact *NGrocery* and *NWCSC* with *Resident Shopper* and *Nonresident Shopper*, referred to as the Conditional model.

For each specification, we report model fit statistics (R-squared, Adjusted R-Squared and Bayesian Information Criterion) as well as the P-value of joint F-tests for the significance of the fixed effects, year effects, and their interactions. Estimated coefficients for the demographic and

[14] *Varying by state, SNAP benefits can become available on different days of the month. We created a binary variable called "SNAP Day" which indicates the days when monthly benefits were eligible to have been deposited. More information by state can be found at https://www.freshebt.com/, and searching for "food stamp deposit schedule."*

shopping habits variables are omitted for brevity but available in the Appendix (Table A1-A4).[15]

Focusing first on the results in Table 2 (Dependent variable: "Price is a barrier to purchasing more healthy foods"), we find that the R-Squared roughly doubles in each specification that includes both year and zip code fixed effect interactions (comparing columns (1) and (2); (3) and (4), respectively). Between the baseline and conditional models, the values of the Adjusted R-Squared and BIC are largest in column (4), which includes year fixed effects, zip code fixed effects at both the primary shopping zip code and residence of the respondents, and their interactions, indicating that this model best fits the data. These interactions appear to be jointly significant, along with shopping zip code fixed effects, suggesting that, when time-varying unobservables are taken into account, the relevance of year fixed effects and time-invariant controls for the zip code of residence lose in their joint explanatory power.

Next, we discuss the estimated coefficients of the food store variables. In the baseline model, we find that overall, the magnitude and statistical significance of the coefficients increases as we control for more sources of unobserved variation and their signs remains unchanged, with two exceptions.[16] In particular, for the models where we control for both time-invariant and time-varying heterogeneity (Table 2, columns (3) and (4)), we find a positive (negative) relationship between *NGrocery* (*NWCSC*) and the probability of declaring *Price* to be a barrier to purchasing more healthy food. As indicated above, the difference in estimated parameters is quite large across models, and in those where fixed effects are interacted with year indicator variables, they vary from 0.021 to 0.181 for *NGrocery*, and from −0.078 to −0.778 for *NWCS*. Thus, we find some evidence that, in the model with fewer controls, the parameters measuring the relationship between store numbers and price as a perceived barrier for purchasing healthy foods may be substantially biased towards zero.

In the conditional model, where the presence of food stores is allowed to vary depending upon respondents living in (or out of) the zip code where they shop, we find that there is no statistically significant difference between the estimated parameters on *NGrocery* for resident and nonresident shoppers

[15]*Estimated coefficients for demographics and shopping habits variables are highly stable across specifications. Female respondents are more likely to declare* Price *as a barrier, but gender has no relationship with* Availability *being perceived as a barrier. The older the respondents, the lower is the probability of indicating* Price *and/or* Availability *as a barrier to purchasing more healthy foods. Household size is inversely related to both* Price *and* Availability *as barriers; the same pattern emerges for number of children in the household (for* Price*). Years of education are not associated with* Price *as a barrier; however, for more than half specifications, it has a positive and statistically significant relationship with* Availability. *"Rural" shows heterogenous signs for* Price *as a barrier depending upon the FE used, whereas the effect is negative (and statistically significant in most cases) for* Availability. *Moving on to the shopping habits variables, we find no relationship between shopping at farmers markets and the probability of indicating* Price *as a barrier, although it has a positive and statistically significant relationship with* Availability. *Monthly expenditure at the primary store is not related to the probability of perceiving price as a barrier, and weakly and positively related with the probability of perceiving availability as a barrier. We find some, evidence that respondents interviewed in a SNAP day were more (less) likely to indicate* Price *(*Availability*) as a barrier. Respondents interviewed at the end of the month were less likely to declare* Price *as a barrier, and in some specifications, more likely to indicate* Availability. *Respondents receiving federal assistance are more likely to indicate* Price *as a barrier; we find no relationship between this variable and* Availability.

[16]*The two exceptions are the coefficient on* NGrocery *for nonresident shoppers, and the coefficient on* NWCSC *for in-residence shoppers.*

Table 2 LPM Regression Results (Price)

Dep. Var. = Price	**(1)**	**(2)**	**(3)**	**(4)**
(A) Baseline - Stores in shopping zip code				
NGrocery	0.001	−0.002	0.021***	0.181***
	(0.014)	(0.006)	(0.002)	(0.010)
NWCSC	−0.038	−0.088**	−0.078***	−0.778***
	(0.069)	(0.041)	(0.022)	(0.057)
R-Squared	0.104	0.218	0.157	0.321
Adj. R-Squared	0.081	0.086	0.127	0.151
BIC	1470.93	1323.13	1404.6	1170.55
P-val F Test Year	*0.052*	*0.140*	*0.000*	*0.466*
P-val F Test Shopping Zip	*0.000*	*0.090*	*0.000*	*0.000*
P-val F Test Residence Zip		*0.197*		*0.150*
P-val F Test Shopping Zip * Year			*0.000*	*0.000*
P-val F Test Residence Zip * Year				*0.004*
(B) Conditional - Stores in shopping zip code by respondents' residence				
NGrocery *	0.002	0.001	0.021***	0.189***
Resident Shopper	(0.015)	(0.005)	(0.005)	(0.010)
NGrocery *	0.001	−0.009	0.021***	0.179***
Nonresident Shopper	(0.014)	(0.010)	(0.002)	(0.010)
NWCSC *	−0.059	0.277***	−0.117***	−0.572***
Resident Shopper	(0.077)	(0.059)	(0.029)	(0.073)
NWCSC *	−0.008	−0.236***	−0.025	−0.910***
Nonresident Shopper	(0.074)	(0.062)	(0.027)	(0.051)
P-val test Ho: β(resident) = β(nonresident)				
NGrocery	0.727	0.345	0.954	0.000
NWCSC	0.037	0.000	0.000	0.000
R-Squared	0.104	0.231	0.157	0.322
Adj. R-Squared	0.080	0.095	0.126	0.151
BIC	1470.47	1304.79	1404.23	1168.05
P-val F Test Year	*0.050*	*0.150*	*0.000*	*0.520*
P-val F Test Shopping Zip	*0.000*	*0.075*	*0.000*	*0.000*
P-val F Test Residence Zip		*0.155*		*0.137*
P-val F Test Shopping Zip * Year			*0.000*	*0.000*
P-val F Test Residence Zip * Year				*0.004*
Fixed Effects				
Year	Y	Y	Y	Y
Shopping Zip	Y	Y	Y	Y
Residence Zip	N	Y	N	Y
Shopping Zip * Year	N	N	Y	Y
Residence Zip * Year	N	N	N	Y
N	1,082	1,082	1,082	1,082

*Note: Standard errors in parenthesis clustered by site location. ***, ** and * indicate statistical significance at 10%, 5%, and 1% level. Other coefficients, omitted for brevity. A full set of results is available in Tables A1 and A2.*
The dependent variable "Price" is equal to one if a respondent indicated that "Price" was a barrier to purchasing more healthy food, zero otherwise.
Shopping Zip*: Zip code where the respondent's primary store is located.*
Residence Zip*: Zip code where the respondent resides (may or may not be the same as* Shopping Zip*).*

in three out of four models, whereas we find statistically significant differences between the estimated parameters on *NWCSC* for residents and nonresident shoppers in all four models.[17] We find that the relationship of *NWCSC*

[17] *Refer to the results of the* P-val *test Ho: β(resident) = β(nonresident) in Table 2, panel B.*

Table 3 LPM Regression Results (Availability)

Dep. Var. = Availability	**(1)**	**(2)**	**(3)**	**(4)**
(A) Baseline - Stores in shopping zip code				
NGrocery	−0.018***	−0.016***	−0.009***	−0.042***
	(0.005)	(0.002)	(0.001)	(0.009)
NWCSC	0.096***	0.165***	0.057***	0.177***
	(0.024)	(0.018)	(0.012)	(0.036)
R-Squared	0.090	0.196	0.107	0.242
Adj. R-Squared	0.068	0.055	0.075	0.052
BIC	453.63	320.17	433.53	256.43
P-val F Test Year	*0.000*	*0.011*	*0.002*	*0.465*
P-val F Test Shopping Zip	*0.000*	*0.000*	*0.000*	*0.001*
P-val F Test Residence Zip		*0.782*		*0.946*
P-val F Test Shopping Zip * Year			*0.054*	*0.027*
P-val F Test Residence Zip * Year				*0.521*
(B) Conditional - Stores in shopping zip code by respondents' residence				
NGrocery *	−0.018***	−0.009***	−0.010***	−0.027***
Resident Shopper	(0.006)	(0.002)	(0.001)	(0.009)
NGrocery *	−0.018***	−0.027***	−0.009***	−0.046***
Nonresident Shopper	(0.005)	(0.004)	(0.001)	(0.010)
NWCSC *	0.063**	0.028*	0.018	0.022
Resident Shopper	(0.026)	(0.016)	(0.014)	(0.042)
NWCSC *	0.161***	0.264***	0.113***	0.240***
Nonresident Shopper	(0.025)	(0.030)	(0.013)	(0.034)
P-val test Ho: β(resident) = β(nonresident)				
NGrocery	0.997	0.000	0.801	0.000
NWCSC	0.000	0.000	0.000	0.000
R-Squared	0.091	0.201	0.108	0.246
Adj. R-Squared	0.067	0.060	0.074	0.055
BIC	452.59	312.41	432.51	250.37
P-val F Test Year	*0.000*	*0.010*	*0.001*	*0.428*
P-val F Test Shopping Zip	*0.000*	*0.000*	*0.000*	*0.000*
P-val F Test Residence Zip		*0.658*		*0.915*
P-val F Test Shopping Zip * Year			*0.055*	*0.025*
P-val F Test Residence Zip * Year				*0.589*
Fixed Effects				
Year	Y	Y	Y	Y
Shopping Zip	Y	Y	Y	Y
Residence Zip	N	Y	N	Y
Shopping Zip * Year	N	N	Y	Y
Residence Zip * Year	N	N	N	Y
N	1,082	1,082	1,082	1,082

*Note: Standard errors in parenthesis clustered by site location. ***, ** and * indicate statistical significance at 10%, 5%, and 1% level. Other coefficients, omitted for brevity. A full set of results is available in Tables A3-A4.*
The dependent variable "Availability" is equal to one if a respondent indicated that "Not available/limited availability" was a barrier to purchasing more healthy food, zero otherwise.
Shopping Zip: Zip code where the respondent's primary store is located.
Residence Zip: Zip code where the respondent resides (may or may not be the same as Shopping Zip*).*

and *Price* as perceived barriers to purchase more healthy food differs across models in both size and magnitude. In Table 2, column (4), the specification with store and residence fixed effects interacted with year dummies, we find a 60% stronger relationship (−0.572 *vs.* -0.910) for respondents who live in a

Table 4 Comparison of LPM Estimates Obtained Using Entire Sample and Truncated Sample Estimates (TSE) Producing Predictions in the 0–1 Range (Price)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Dep. Var. = Price	OLS	TSE	OLS	TSE	OLS	TSE	OLS	TSE
(A) Baseline - Stores in shopping zip code								
NGrocery	0.001	0.001	0.021***	0.021***	−0.002	−0.002	0.181***	0.216***
	(0.014)	(0.014)	(0.002)	(0.002)	(0.006)	(0.006)	(0.010)	(0.015)
NWCSC	−0.038	−0.039	−0.078***	−0.084***	−0.088**	−0.102**	−0.778***	−0.977***
	(0.069)	(0.071)	(0.022)	(0.024)	(0.041)	(0.043)	(0.057)	(0.105)
Adj. R-Squared	0.081	0.070	0.127	0.112	0.093	0.072	0.151	0.071
BIC	1470.93	1463.61	1404.61	1388.47	1308.80	1290.62	1170.55	1186.97
N	1,082	1,063	1,082	1,049	1,082	1,005	1,082	940
% difference in N		1.76%		3.05%		7.12%		13.12%
(B) Conditional - Stores in shopping zip code by respondents' residence								
NGrocery *	0.002	0.002	0.021***	0.022***	0.001	0.001	0.189***	0.217***
Resident Shopper	(0.015)	(0.016)	(0.005)	(0.005)	(0.005)	(0.005)	(0.010)	(0.016)
NGrocery *	0.001	0.001	0.021***	0.021***	−0.009	−0.009	0.179***	0.212***
Nonresident Shopper	(0.014)	(0.014)	(0.002)	(0.002)	(0.010)	(0.011)	(0.010)	(0.016)
NWCSC *	−0.059	−0.060	−0.117***	−0.122***	0.277***	−0.546***	−0.572***	−1.268***
Resident Shopper	(0.077)	(0.078)	(0.029)	(0.029)	(0.059)	(0.070)	(0.073)	(0.078)
NWCSC *	−0.008	−0.008	−0.025	−0.026	−0.236***	−0.290**	−0.910***	−1.159***
Nonresident Shopper	(0.074)	(0.075)	(0.027)	(0.027)	(0.062)	(0.093)	(0.051)	(0.094)
Adj. R-Squared	0.080	0.069	0.126	0.111	0.095	0.069	0.151	0.069
BIC	1470.47	1463.13	1404.23	1388.65	1304.79	1286.78	1168.05	1190.96
Fixed Effects								
Year	Y	Y	Y	Y	Y	Y	Y	Y
Shopping Zip	Y	Y	Y	Y	Y	Y	Y	Y
Residence Zip	N	N	N	N	Y	Y	Y	Y

(Continues)

Table 4 Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Shopping Zip * Year	N	N	Y	Y	N	N	Y	Y
Residence Zip * Year	N	N	N	N	N	N	Y	Y
N	1,082	1,063	1,082	1,051	1,082	1,006	1,082	942
% difference in N		1.76%		2.87%		7.02%		12.94%

*Note: Standard errors in parenthesis clustered by site location. ***, ** and * indicate statistical significance at 10%, 5%, and 1% level. Other coefficients are available upon request. The dependent variable "Price" is equal to one if a respondent indicated that "Price" was a barrier to purchasing more healthy food, zero otherwise.*

different zip code than where they shop. Note that a unitary increase in *NGrocery* corresponds to (about) a 7.4% increase in the number of these stores (on average); as *NWCSC* is, essentially, an indicator variable, given that there are no observations with more than one Warehouse Club/Supercenter, the presence of one of these stores results in shoppers 57% to 91% less likely to indicate price as a barrier. These results suggest that improving access to large food outlets or making access to these outlets less prohibitive would be more marked for those who shop outside the zip code where they live.

Next, we turn our attention to the barrier "Availability," the dependent variable whose results are presented in Table 3. The model diagnostics and goodness of fit measures do not provide a clear indication as to which model specification (in terms of zip code fixed effects and their interaction with year indicators) is preferred. In every model including residence zip code fixed effects, the F-test for the joint significance of the parameters show them to be jointly statistically nonsignificant. Additionally, the Adjusted R-squared values are lower (in columns (2) and (4)) than the specification that does not include them (columns (1) and (3)), suggesting that the inclusion of these fixed effects may not be necessary. Yet, these models show lower BICs, which would indicate a better fit. Also, the bias of the estimated parameters follows opposite directions for the two types of food stores, and their patterns change depending upon whether the store count variables interacted with the residence dummies are included in the model.

Even if we cannot clearly point to a "best" model specification or a clear direction of the bias, in most cases the sign and statistical significance of the estimated coefficients are consistent across specifications. Overall, we find a negative (positive) and statistically significant relationship between *NGrocery* (*NWCSC*) and the perception of *Availability* as a barrier for purchasing more healthy food. The difference in estimated parameters is smaller in models where the zip code fixed effects are not interacted with year indicators (the parameters on *NGrocery* do not vary considerably, whereas those on *NWCSC* vary from 0.096 to 0.177), and are larger in those specifications where zip code fixed effects are interacted with year dummies (from −0.009 to −0.042 for *NGrocery*; from 0.057 to 0.177 for *NWCSC*).

The models where we condition on whether respondents live (or not) in the zip code where they shop show that *NGrocery*'s estimated parameters are statistically different for resident and nonresident shoppers in models where fixed effects are interacted with year indicator variables. In these specifications, the estimated parameters are larger for nonresidents than for residents, whereas they are not statistically different from one another in the other specifications. We find that the relationship of *NWCSC* and *Availability* as a perceived barrier to purchase more healthy food to be positive and statistically significant in all model specifications for respondents who live in a different zip code than that where they shop, with coefficients varying between 0.113 and 0.264. For respondents shopping in the zip code where they live, the relationship is smaller in magnitude, and only statistically different from zero in models where zip code fixed effects are not interacted with year indicators.

Additional Estimations and Robustness Checks

In this section, we discuss results of additional estimation and robustness checks.

The first set of additional estimates compares the results presented in Tables 2 and 3 with those obtained truncating the sample so to obtain only predictions in the 0–1 range, which, as illustrated by Horrace and Oaxaca (2006), should limit bias. The estimated coefficients for these robustness checks are reported in Tables 4 and 5 for *Price* and *Availability*, respectively. The odd columns include the same coefficients from Tables 2 and 3, whereas the even columns include estimates obtained using trimmed samples. Overall, we find that dropping observations generating out-of-range predictions leads to small reductions in sample size (from about 1.8% to 13%, depending on the specification), and small differences in estimated parameters. The results in Table 4 point to small, if any, bias in the estimates discussed above.

First, we discuss the results obtained using the TSE samples for the barrier *Price*. In the baseline model (top panel of Table 4), the estimated coefficients for the number of stores appear almost identical across subsamples where only shopping zip code FE are used. For the model specifications that also include fixed effects for the respondents' zip codes of residence, the estimates using the TSE sample are larger in magnitude yet the signs are the same. In the conditional model (bottom panel of Table 4), the largest differences in estimates are for the coefficients on *NWCSC*. The negative sign on the coefficients of *NWCSC* interacted with residence status indicates an inverse relationship with the probability of perceiving *Price* as a barrier to purchasing more healthy foods. Moreover, the coefficients show that the presence of a WCSC has a stronger relationship for respondents residing in the zip code where they shop than for those who do not.

The results in Table 5 compare the estimates in Table 3, with those obtained using the TSE samples for the barrier *Availability*. Noticeably, the trimming estimation yields considerably fewer observations, ranging from 22% to 58.6% of the full sample. The TSE parameter estimates confirm the negative and statistically significant relationship between *NGrocery* and the probability of perceiving *Availability* as a barrier to purchasing healthy food. This result holds in all but one specification (column (8)), where including zip code-level controls interacted with year indicators produces an estimation sample less than half the size of the original sample and may show issues of sample selection. The magnitude of the *NGrocery* TSE estimates are in line or larger than those obtained using the full sample, similar to what we observe for *Price*. The positive and statistically significant relationship of *NWCSC* and *Availability* also persists in all the TSE models where residence status is not interacted with the store variables, and it shows either similar or larger magnitude than those from the full sample (almost three times as large when zip code fixed effects are interacted with year indicators). Due to the large number of observations dropped, the interaction of *NWCSC* with "resident shopper" is dropped in three out of four specifications, which leaves us unable to detect any possible bias in this coefficient's estimates.

As a final robustness check, we estimated models including an interaction of *NGrocery* and *NWCSC* and the results are shown in Table 6. The aim of this additional model specification is to capture how the relationship between *Price* and *Availability* as perceived barriers to purchasing more healthy food and the number of stores may change when shoppers have access to multiple store formats. Because different retail formats interact strategically in ways that may affect prices and assortment levels (e.g. Hausman and Liebtag 2007; Volpe and Lavoie 2008; Basker and Noel 2009; Matsa 2011; Cleary & Lopez 2014;

Table 5 Comparison of LPM Estimates Obtained Using Entire Sample and Truncated Sample Estimates (TSE) Producing Predictions in the 0–1 Range (Availability)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var. = Availability	OLS	TSE	OLS	TSE	OLS	TSE	OLS	TSE
(A) Baseline - Stores in shopping zip code								
NGrocery	−0.018***	−0.039***	−0.009***	−0.010***	−0.016***	−0.043***	−0.042***	0.006
	(0.005)	(0.005)	(0.001)	(0.001)	(0.002)	(0.010)	(0.009)	(0.003)
NWCSC	0.096***	0.198***	0.057***	0.058***	0.165***	0.458***	0.177***	0.536***
	(0.024)	(0.028)	(0.012)	(0.014)	(0.018)	(0.071)	(0.036)	(0.100)
Adj. R-Square	0.068	0.041	0.075	0.044	0.055	0.083	0.052	0.089
BIC	453.63	551.41	433.53	553.54	320.17	451.16	256.43	425.66
N	1,082	805	1,082	844	1,082	504	1,082	452
% difference in N		25.60%		22.00%		53.42%		58.23%
(B) Conditional - Stores in shopping zip code by respondents' residence								
NGrocery*	−0.018***	−0.038***	−0.010***	−0.010***	−0.009***	−0.029**	−0.027***	0.022***
Resident Shopper	(0.006)	(0.006)	(0.001)	(0.001)	(0.002)	(0.012)	(0.009)	(0.003)
NGrocery*	−0.018***	−0.039***	−0.009***	−0.010***	−0.027***	−0.084***	−0.046***	−0.037***
Nonresident Shopper	(0.005)	(0.005)	(0.001)	(0.001)	(0.004)	(0.021)	(0.010)	(0.006)
NWCSC*	0.063**	0.095***	0.018		0.028*		0.022	
Resident Shopper	(0.026)	(0.033)	(0.014)		(0.016)		(0.042)	
NWCSC*	0.161***	0.313***	0.113***	0.415***	0.264***	0.708***	0.240***	0.807***
Nonresident Shopper	(0.025)	(0.035)	(0.013)	(0.086)	(0.030)	(0.136)	(0.034)	(0.105)
Adj. R-Square	0.080	0.069	0.126	0.111	0.095	0.069	0.151	0.069
BIC	1470.47	1463.13	1404.23	1388.65	1304.79	1286.78	1168.05	1190.96
Fixed Effects								
Year	Y	Y	Y	Y	Y	Y	Y	Y
Shopping Zip	Y	Y	Y	Y	Y	Y	Y	Y

(Continues)

Table 5 Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Residence Zip	N	N	N	N	Y	Y	Y	Y
Shopping Zip * Year	N	N	Y	Y	N	N	Y	Y
Residence Zip * Year	N	N	N	N	N	N	Y	Y
N	1,082	799	1,082	833	1,082	509	1,082	448
% difference in N		26.16%		23.01%		52.96%		58.60%

*Note: Standard errors in parenthesis clustered by site location. ***, ** and * indicate statistical significance at 10%, 5%, and 1% level. Other coefficients are available upon request. The dependent variable "Availability" is equal to one if a respondent indicated that "Not available/limited availability" was a barrier to purchasing more healthy food, zero otherwise.*

Table 6 Estimated Food Store Coefficients and Interaction Marginal Effects

	(1)	(2)
	Price	**Availability**
(A) Baseline - Stores in shopping zip code		
NGrocery	0.181***	−0.042***
	(0.010)	(0.009)
NWCSC	−2.008***	−2.424***
	(0.405)	(0.182)
NGrocery * NWCSC	0.246**	0.520***
	(0.088)	(0.043)
Interaction Marginal Effects		
NGrocery | NWCSC = 1	0.427***	0.478***
	(0.080)	(0.035)
NWCSC | NGrocery = mean	−0.673***	0.400***
	(0.088)	(0.053)
NWCSC | NGrocery = max	−0.532***	0.698***
	(0.134)	(0.077)
(B) Conditional - Stores in shopping zip code by respondents' residence		
NGrocery * Resident Shopper	0.189***	−0.027**
	(0.010)	(0.009)
NGrocery * Nonresident Shopper	0.179***	−0.046***
	(0.010)	(0.010)
NWCSC * Resident Shopper	−1.492**	−1.036***
	(0.571)	(0.246)
NWCSC * Nonresident Shopper	−1.983***	−3.360***
	(0.454)	(0.257)
NGrocery * NWCSC * Resident Shopper	0.185	0.232***
	(0.124)	(0.053)
NGrocery * NWCSC * Nonresident Shopper	0.214*	0.708***
	(0.092)	(0.056)
Resident Shoppers Interaction Marginal Effects		
NGrocery | NWCSC = 1	0.375***	0.206***
	(0.115)	(0.048)
NWCSC | NGrocery = mean	−0.486***	0.224***
	(0.104)	(0.060)
NWCSC | NGrocery = max	−0.380**	0.357***
	(0.173)	(0.084)
Nonresident Shoppers Interaction ME		
NGrocery | NWCSC = 1	0.392***	0.662***
	(0.086)	(0.047)
NWCSC | NGrocery = mean	−0.822***	0.484***
	(0.078)	(0.053)
NWCSC | NGrocery = max	−0.700***	0.888***
	(0.118)	(0.082)

*Note: Standard errors in parenthesis clustered by site location. ***, ** and * indicate statistical significance at 10%, 5%, and 1% level. Both regressions include fixed effects for year, shopping zip, residence zip, and their interactions.*
Dependent variables: "Price" and "Not available/limited availability" as barriers to purchasing more healthy foods. Estimates from LPM with interactions of NGrocery *and* NWCSC*. All models include interactions of fixed effects (both shopping zip codes and residence zip codes) and year dummies.*

Courtemanche and Carden 2014; Bauner and Wang 2019), it is at least possible that changes in price and assortment strategy may shape consumers perceptions, and the relationships discussed above may be, in fact, more nuanced.

The values included in Table 6 present estimates of specifications including the interaction of *NGrocery* and *NWCSC* (panel A), also allowing for the estimates to vary by residence status (panel B). Both *Price* (column (1)) and *Availability* (column (2)) are considered, separately, as dependent variables. All models include interactions of year dummies with both shopping and residence zip code fixed effects. Before proceeding, it should be noted that in our sample, all the zip codes with a WCSC have a positive number of supermarkets and grocery stores, but the average value of *NGrocery* conditional on *NWCSC* = 1 is smaller than the sample average (about 5.43 versus 13.5). The results in the top panel suggest that, for both *Price* and *Availability*, the simultaneous presence of supermarkets and other grocery stores and WCSC (*NGrocery**NWCSC*) is positively associated with perceiving barriers to purchasing more healthy food. Furthermore, in zip codes where a WCSC is located, the positive relationship between *NGrocery* and *Price* is magnified, whereas the negative relationship between *NGrocery* and *Availability* becomes positive. These patterns also emerge in the model where relationships are estimated conditional on residence (Table 6, panel B). In the absence of traditional stores, we find that WCSCs would be related to lower probabilities of declaring *Price* and *Availability* as barriers, a relationship that is more marked for nonresident shoppers than for resident shoppers.[18] Finally, we consider the marginal effects of the interacted terms. As *NGrocery* grows, the relationship between the probability to declare barriers to healthy eating and the presence of a WCSCs becomes smaller in magnitude, remaining negative and statistically significant for *Price* (and again, larger for nonresident shoppers), whereas it reaches positive values for *Availability*.

Overall, these results indicate that the "average" relationships illustrated in Tables 2 and 3 may mask more complex relationships as shoppers' perceptions may change when retailers act strategically in response to other stores' presence. The changes in the estimated relationships may also hint to consumers' perception shifting when they switch between different types of stores. For example, the price-decreasing effect that is usually associated with some WCSCs (e.g. Hausman and Liebtag 2007; Volpe and Lavoie 2008; Basker and Noel 2009) may lead to lower probabilities of perceiving price as a barrier to purchasing healthy foods; however, in areas with more traditional food retailers such an effect may be less marked as retailers differentiate and the effect dissipates.

Discussion and Policy Implications

Our findings report some relatively robust results that the presence of supermarkets and grocery stores may not mitigate individuals' perceptions of price being a barrier to purchasing more healthy foods. Traditional stores do, however, show a negative and statistically significant association with availability being a barrier. These findings are overall consistent with the literature on food retail competition. Ample empirical evidence exists that some large stores, such as Walmart Supercenters, have a mitigating effect on competitor's prices (Hausman and Liebtag 2007; Volpe and Lavoie 2008; Basker and Noel 2009). It should be noted, however, that other WCSCs are generally associated with higher prices (Courtemanche and Carden 2014) or with

[18] *We do not observe any cases where NWCSC = 1 and NGrocery = 0, so a coefficient on NWCSC cannot be interpreted as isolated effect; i.e, the estimated coefficient is considered "out of sample."*

mixed effects on price changes (as well as product assortment) depending upon the characteristics of the product considered (Bauner and Wang 2019).

Our results also suggest that availability as a barrier to purchasing more healthy foods is related positively to the presence of warehouse clubs and supercenters, and negatively to the number of grocery stores and supermarkets, although these results are less robust than those for *Price*. Even so, some findings from the economic literature can explain this result. For example, Matsa (2011) finds that supermarket firms have fewer inventory shortfalls (that is, less disruption in product availability) when they face more local competition. In other words, as the number of supermarkets increases and product availability becomes less variable, shoppers may become more consistently exposed to deeper assortment levels. Also, Bauner and Wang (2019) find changes in assortments as a strategic response to warehouse clubs depending upon the perishability and stockpilability of the products considered.

Even though we can contextualize our results in the existing economics literature, in terms of policy implications, our results do not show definitive evidence that the probability of experiencing perceived barriers to purchasing more healthy foods is lower in presence of more food stores.[19] While the presence of stores of a given type may be related to a lower probability to declare a given barrier, the same store-type may be related to a higher probability of experiencing another. Furthermore, as we progressively include more controls to limit bias due to unobserved heterogeneity, some of the estimates change in significance level, size, and, at times, sign. Moreover, restricting the samples to obtain only predicted values between 0 and 1, we note changes in parameters' sign and significance and, in the case of *Availability*, one of the variables of interest is dropped. Last, we find that perceptions may also change when shoppers are exposed to multiple types of stores.

Thus, we suggest caution to make definitive statements about the relationship between stated hardships to purchase healthy foods and store presence. In spite of the general patterns indicated above, our results vary enough across specifications to be used as a cautionary tale for researchers in this area. Given the strong public support for store intervention policies, and the lack of evidence regarding some aspects of their effectiveness, researchers should be open to discuss the possibility that they may be in the presence of a false-positive result (i.e., a result supporting a policy which may have no effect) or a false negative one (i.e., a null result for a policy which may instead be effective).

Study Limitations

Our study has several limitations. First, using a respondent's zip codes of residence to define the scope of the food environment limits the number of stores to which respondents are exposed. Second, this study focuses on the relationship between physical food environment and perceived barriers to purchasing healthy foods, and not diet quality, directly. Third, as our sample is limited to individuals living in a limited number of communities across the US Northeast, the validity and generalizability of our results are also limited. Fourth, and last, our data does not allow us to

[19] *An example of such a policy where the presence of food stores is central to its design is the Healthy Food Financing Initiative (HFFI), which, through tax incentives and subsidies, encourages food retailer entry in areas without access to supermarkets, where the availability of healthy food options was limited.*

consider the counterfactual to measure whether nonresident shoppers will choose to shop in their zip code of residence if the food environment changes.

Conclusion

Households' perceived hardships to purchase healthy foods is hypothesized to be related to lack of access to supermarkets and large grocery stores. Several studies document a "null" or limited relationship between supermarket density/proximity and purchases of healthy foods. Additionally, policies aimed at promoting healthy food choices by ameliorating the food environment (e.g., financing programs to subsidize store entry) find little empirical support.

The applied economics literature has focused on the relationship between objective measures of access to healthy foods and food choices, sidestepping the importance of assessing consumers perceived constrains in acquiring more healthy foods as a more refined measure of what factors, if any, of a food environment, should be modified to promote healthier eating. In this study, we assess the relationship between individuals' perceived barriers to buying more healthy foods and the number of food stores in the zip code where they shop. We use shoppers' intercept survey data collected across the Northeastern US augmented with secondary zip code-level data on food store numbers.

Our findings indicate that the number of supermarkets and grocery stores, as well as wholesale clubs and supercenters, have mixed relationships with perceived barriers to purchasing healthy foods. Overall, our findings suggest that the number of traditional food stores' presence is related to a lower (higher) probability of stating availability (prices) as a barrier; the presence of WCSC is instead related to a lower (higher) probability of perceiving prices (availability) as a barrier. Additionally, we find that the estimated relationships are larger for respondents who do not live in the zip code where they primarily shop for food. These results indicate that shoppers' *choosing* to shop in a given zip code may be more likely to be affected by changes in the food environment than those who actually live in the place where they shop. Also, because of the opposite directions of the relationship between the number of stores of a given type and different perceived barriers, our results indicate that investigating research in this area should examine whether, and under what circumstances, modifications to the food environment would serve to benefit residents.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Acknowledgments

The authors gratefully acknowledge funding from USDA/NIFA under Global Food Security Grant No. 2011-68004-30057 (Enhancing the Food Security of Underserved Populations in the Northeast US through Sustainable Regional Food Systems Development) and the Agricultural Experiment Station of Colorado State University. We would like to thank Kate Clancy and Katherine Ramirez for their valuable input in developing this manuscript, and for comments received on earlier versions of this work.

REFERENCES

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. 2019. Food Deserts and the Causes of Nutritional Inequality. *The Quarterly Journal of Economics* 134(4): 1793–1844.

Angrist, Joshua, and Jörn–steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Appelhans, Bradley M., Simone A. French, Christy C. Tangney, Lissa M. Powell, and Yamin Wang. 2017. To What Extent Do Food Purchases Reflect Shoppers' Diet Quality and Nutrient Intake? *International Journal of Behavioral Nutrition and Physical Activity* 14(1): 46.

Barrientos-Gutierrez, Tonatiuh, Kari A.B. Moore, Amy H. Auchincloss, Mahasin S. Mujahid, Carmella August, Brisa N. Sanchez, and Ana V. Diez Roux. 2017. Neighborhood Physical Environment and Changes in Body Mass Index: Results from the Multi-Ethnic Study of Atherosclerosis. *American Journal of Epidemiology* 186(11): 1237–1245.

Basker, Emek, and Michael Noel. 2009. The Evolving Food Chain: Competitive Effects of Wal-mart's Entry into the Supermarket Industry. *Journal of Economics and Management Strategy* 18(4): 977–1009.

Bauner, Christoph, and Emily Wang. 2019. The Effect of Competition on Pricing and Product Positioning: Evidence from Wholesale Club Entry. *International Journal of Industrial Organization* 67.

Black, Jennifer L., and James Macinko. 2008. Neighborhoods and Obesity. *Nutrition Reviews* 66(1): 2–20.

Bonanno, Alessandro, Francesco Bimbo, Elena Castellari, and Paolo Sckokai. 2017. Five-a-Day, Fruit and Vegetables Portions, and the Food Environment: The Italian Case. *Applied Economic Perspectives and Policy* 39(4): 682–709.

Bonanno, Alessandro, and Jing Li. 2015. Food Insecurity and Food Access in U.S. Metropolitan Areas. *Applied Economic Perspectives and Policy* 37(2): 177–204.

Caspi, Caitlin E., Ichiro Kawachi, S.V. Subramanian, Gary Adamkiewicz, and Glorian Sorensen. 2012. The Relationship between Diet and Perceived and Objective Access to Supermarkets among Low-Income Housing Residents. *Social Science & Medicine* 75(7): 1254–1262.

Chenarides, Lauren, and Edward C. Jaenicke. 2018. Documenting the Link between Poor Food Access and less Healthy Product Assortment across the U.S. *Applied Economic Perspectives and Policy*: 1–41.

Clancy, Kate, Alessandro Bonanno, Patrick Canning, et al. 2017. Using a Market Basket to Explore Regional Food Systems. *Journal of Agriculture, Food Systems, and Community Development* 7(4): 163–178.

Cleary Rebecca, L.O, and Lopez Rigoberto, A. 2014. Supermarket responses to Wal-Mart Supercenter expansion: a structural approach. *Empircal Econ* 47: 905–925. https://doi.org/10.1007/s00181-013-0767-5.

Cobb, Laura K., Lawrence J. Appel, Manuel Franco, Jessica C. Jones-Smith, Alana Nur, and Cheryl A.M. Anderson. 2015. The Relationship of the Local Food Environment with Obesity: A Systematic Review of Methods, Study Quality, and Results. *Obesity* 23(7): 1331–1344.

Courtemanche, Charles, and Art Carden. 2014. Competing with Costco and Sam's Club: Warehouse Club Entry and Grocery Prices. *Southern Economic Journal* 80(3): 565–585.

Courtemanche, Charles, Art Carden, Xilin Zhou, and Murugi Ndirangu. 2019. Do Walmart Supercenters Improve Food Security? *Applied Economic Perspectives and Policy* 41(2): 177–198.

Cummins, Steven, Ellen Flint, and Stephen A. Matthews. 2014. New Neighborhood Grocery Store Increased Awareness of Food Access but Did Not Alter Dietary Habits or Obesity. *Health Affairs* 33(2): 283–291.

Cummins, Steven, and Sally Macintyre. 2006. Food Environments and Obesity—Neighbourhood or Nation? *International Journal of Epidemiology* 35(1): 100–104.
Damon, Amy L., Robert P. King, and Ephraim Leibtag. 2013. First of the Month Effect: Does it Apply across Food Retail Channels? *Food Policy* 41: 18–27.
Davison, Jenny, Michelle Share, Marita Hennessy, and Barbara S. Knox. 2015. Caught in a 'Spiral'. Barriers to Healthy Eating and Dietary Health Promotion Needs from the Perspective of Unemployed Young People and their Service Providers. *Appetite* 85: 146–154.
de Mestral, Carlos, Saman Khalatbari-Soltani, Silvia Stringhini, and Pedro Marques-Vidal. 2017. Fifteen-Year Trends in the Prevalence of Barriers to Healthy Eating in a High-Income Country. *American Journal of Clinical Nutrition* 105(3): 660–668.
de Mestral, Carlos, Silvia Stringhini, and Pedro Marques-Vidal. 2016. Barriers to Healthy Eating in Switzerland: A Nationwide Study. *Clinical Nutrition* 35(6): 1490–1498.
Dubowitz, Tamara, Madhumita Ghosh-Dastidar, Deborah A. Cohen, et al. 2015. Diet and Perceptions Change with Supermarket Introduction in A Food Desert, but Not because of Supermarket Use. *Health Affairs* 34(11): 1858–1868.
Eikenberry, Nicole, and Chery Smith. 2004. Healthful Eating: Perceptions, Motivations, Barriers, and Promoters in Low-Income Minnesota Communities. *Journal of the American Dietetic Association* 104 (7): 1158–1161.
Farahmand, Maryam, Parisa Amiri, Fahimeh Ramezani Tehrani, Amir A. Momenan, Parvin Mirmiran, and Fereidoun Azizi. 2015. What Are the Main Barriers to Healthy Eating among Families? A Qualitative Exploration of Perceptions and Experiences of Tehranian Men. *Appetite* 89: 291–297.
Ghosh-Dastidar, Madhumita, Gerald Hunter, Rebecca L. Collins, Shannon N. Zenk, Steven Cummins, Robin Beckman, Alvin K. Nugroho, Jennifer C. Sloan, Lavette Wagner, and Tamara Dubowitz. 2017. Does Opening a Supermarket in a Food Desert Change the Food Environment? *Health & Place* 46: 249–256.
Glanz, Karen, Lauren Johnson, Amy L. Yaroch, Matthew Phillips, Guadalupe X. Ayala, and Erica L. Davis. 2016. Measures of Retail Food Store Environments and Sales: Review and Implications for Healthy Eating Initiatives. *Journal of Nutrition Education and Behavior* 48(4).
Glanz, Karen, James F. Sallis, Brian E. Saelens, and Lawrence D. Frank. 2007. Nutrition Environment Measures Survey in Stores (NEMS-S). Development and Evaluation. *American Journal of Preventive Medicine* 32(4): 282–289.
Greene, William. 2004. The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Modesl in the Presence of Fixed Effects. *Econometrics Journal* 7: 98–119.
——. 2003. *Econometric Analysis*, 5th ed. Pearson Education India.
Gustafson, Alison, Joseph Sharkey, Carmen D. Samuel-Hodge, Jesse Jones-Smith, Mary C. Folds, Jianwen Cai, and Alice S. Ammerman. 2011. Perceived and Objective Measures of the Food Store Environment and the Association with Weight and Diet among Low-Income Women in North Carolina. *Public Health Nutrition* 14(6): 1032–1038.
Handbury, Jessie, Ilya Rahkovsky, and Molly Schnell. 2015. What Drives Nutritional Disparities? Retail Access and Food Purchases across the Socioeconomic Spectrum. *NBER Working Paper*.
Hausman, Jerry, and Ephraim Liebtag. 2007. Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart. *Journal of Applied Econometrics* 22: 1157–1177.
Horrace, Wiiliam C., and Ronald L. Oaxaca. 2006. Results on the Bias and Inconsistency of Ordinary Least Squares for the Linear Probability Model. *Economics Letters* 90(3): 321–327.
Inglis, Victoria, Kylie Ball, and David Crawford. 2008. Socioeconomic Variations in women's Diets: What Is the Role of Perceptions of the Local Food Environment? *Journal of Epidemiology and Community Health* 62(3): 191–197.

Kearney, John M., and Sinead McElhone. 1999. Perceived Barriers in Trying to Eat Healthier – Results of a Pan-EU Consumer Attitudinal Survey. *British Journal of Nutrition* 81(S1): S133.

Kyureghian, Gayaneh, and Rodolfo M. Nayga. 2013. Food Store Access, Availability, and Choice when Purchasing Fruits and Vegetables. *American Journal of Agricultural Economics* 95(5): 1280–1286.

Larson, Nicole I., Mary T. Story, and Melissa C. Nelson. 2009. Neighborhood Environments: Disparities in Access to Healthy Foods in the U.S. *American Journal of Preventive Medicine* 36(1): 74–81.

Matsa, David A. 2011. Competition and Product Quality in the Supermarket Industry. *Quarterly Journal of Economics* 126(3): 1539–1591.

McMorrow, Liam, A. Ludbrook, Jennie I. Macdiarmid, and Damilola Olajide. 2017. Perceived Barriers towards Healthy Eating and their Association with Fruit and Vegetable Consumption. *Journal of Public Health* 39(2): 330–338.

Moore, Latetia V., and Ana V. Diez Roux. 2006. Associations of Neighborhood Characteristics with the Location and Type of Food Stores. *American Journal of Public Health* 96(2): 325–331.

Mui, Yeeli, Joel Gittelsohn, and Jessica C. Jones-Smith. 2017. Longitudinal Associations between Change in Neighborhood Social Disorder and Change in Food Swamps in an Urban Setting. *Journal of Urban Health* 94(1): 75–86.

Rahkovsky, Ilya, and Samantha Snyder. 2015. Food Choices and Store Proximity. Washington DC: U.S. Department of Agriculture, Economic Research Report ERR-195.

Richardson, Andrea S., Madhumita Ghosh-Dastidar, Robin Beckman, Karen R. Flórez, Amy DeSantis, Rebecca L. Collins, and Tamara Dubowitz. 2017. Can the Introduction of a Full-Service Supermarket in a Food Desert Improve residents' Economic Status and Health? *Annals of Epidemiology* 27(12): 771–776.

Sharkey, Joseph R. 2009. Measuring Potential Access to Food Stores and Food-Service Places in Rural Areas in the U.S. *American Journal of Preventive Medicine* 36(4): S151–S155.

Traill, W. Bruce, Stephanie A. Chambers, and Laurie Butler. 2012. Attitudinal and Demographic Determinants of Diet Quality and Implications for Policy Targeting. *Journal of Human Nutrition and Dietetics* 25(1): 87–94.

Ver Ploeg, Michele, Paulo Dutko, and Vince Breneman. 2015. Measuring Food Access and Food Deserts for Policy Purposes. *Applied Economic Perspectives and Policy* 37(2): 205–225.

Ver Ploeg, Michele, and Parke E. Wilde. 2018. How Do Food Retail Choices Vary within and between Food Retail Environments? *Food Policy* 79(August): 300–308.

Volpe, Richard J., and Nathalie Lavoie. 2008. The Effect of Wal-Mart Supercenters on Grocery Prices in New England. *Review of Agricultural Economics* 30(1): 4–26.

Wrigley, Neil, Daniel Warm, Barrie Margetts, and Amanda Whelan. 2002. Assessing the Impact of Improved Retail Access on Diet in a 'Food Desert': A Preliminary Report. *Urban Studies* 39(11): 2061–2982.

Zeng, Di, Michael R. Thomsen, Rodolfo M. Nayga, and Heather L. Rouse. 2019. Neighbourhood Convenience Stores and Childhood Weight Outcomes: An Instrumental Variable Approach. *Applied Economics* 51(3): 288–302.

Zenk, Shannon N., Lisa M. Powell, Leah Rimkus, Zeynep Isgor, Dianne C. Barker, Punam Ohri-Vachaspati, and Frank Chaloupka. 2014. Relative and Absolute Availability of Healthier Food and Beverage Alternatives across Communities in the United States. *American Journal of Public Health* 104(11): 2170–2178.