HHS Public Access

Author manuscript

Am J Prev Med. Author manuscript; available in PMC 2018 March 01.

Published in final edited form as:

Am J Prev Med. 2017 March ; 52(3): 300–310. doi:10.1016/j.amepre.2016.08.042.

Beyond supermarkets: Food outlet location selection in four U.S. cities over time

Pasquale E. Rummo,

Department of Nutrition, University of North Carolina at Chapel Hill, Carolina Population Center, 137 East Franklin Street, CB #8120, Chapel Hill, NC 27514

David K. Guilkey, PhD,

Department of Economics, University of North Carolina at Chapel Hill, Carolina Population Center, 137 East Franklin Street, Room 6107, CB #8120, Chapel Hill, NC 27514, Tel.: (919) 962-6140

Shu Wen Ng, PhD,

Department of Nutrition, University of North Carolina at Chapel Hill, Carolina Population Center, 137 East Franklin Street, Room 6506, CB #8120, Chapel Hill, NC 27514, Tel.: 919-962-6188

Barry M. Popkin, PhD,

Department of Nutrition, University of North Carolina at Chapel Hill, Carolina Population Center, 137 East Franklin Street, Room 6305, CB #8120, Chapel Hill, NC 27514, Tel.: 919-962-6139

Kelly R. Evenson, PhD, and

Department of Epidemiology, University of North Carolina at Chapel Hill, 137 East Franklin Street, Suite 306, CB #8050, Chapel Hill, NC 27514, Tel. (919)-966-4187

P. Gordon-Larsen, PhD

Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; Carolina Population Center, Chapel Hill, North Carolina

Abstract

Introduction—Understanding what influences where food outlets locate is important for mitigating disparities in access to healthy food outlets. However, few studies have examined how neighborhood characteristics influence the neighborhood food environment over time, and whether these relationships differ by neighborhood-level income.

Corresponding author and last author: Penny Gordon-Larsen, PhD, Department of Nutrition, University of North Carolina at Chapel Hill, Carolina Population Center, 137 East Franklin Street, Room 6124, CB #8120, Chapel Hill, NC 27514, Tel.: (919) 962-6110, Fax: (919) 966-9159, pglarsen@email.unc.edu.

Conflict of interest statement: The authors have no conflicts of interest.

Financial disclosures:

Pasquale E. Rummo has no financial disclosures.
David K. Guilkey has no financial disclosures.
Shu Wen Ng has no financial disclosures.
Barry M. Popkin has no financial disclosures.
Kelly R. Evenson has no financial disclosures.
Penny Gordon-Larsen has no financial disclosures.

Methods—Neighborhood-level data from four U.S. cities (Birmingham, AL; Chicago, IL; Minneapolis, MN; Oakland, CA) from 1986, 1993, 1996, 2001, 2006, and 2011 were used with two-step econometric models to estimate longitudinal associations between neighborhood-level characteristics (*z*-scores) and the log-transformed count/km^2 (density) of food outlets within real-estate-derived neighborhoods. We examined associations with lagged neighborhood-level sociodemographics and lagged density of food outlets, with interaction terms for neighborhood-level income. Data were analyzed in 2016.

Results—Neighborhood-level income at earlier years was negatively associated with the current density of convenience stores (β= −0.27; 95% CI: −0.16, −0.38; p<0.001). The percentage of neighborhood white population was negatively associated with fast food restaurant density in low income neighborhoods (10th percentile of income: β= −0.17; 95% CI: −0.34, −0.002; p=0.05), and the density of smaller grocery stores across all income levels (β= −0.27; 95% CI: −0.45, −0.09; p=0.003). There was a lack of policy-relevant associations between the pre-existing food environment and the current density of food outlet types, including supermarkets.

Conclusions—Socioeconomically-disadvantaged populations and minority populations may attract ‘unhealthy’ food outlets over time. To support equal access to healthy food outlets, the availability of ‘less healthy’ food outlets types may be relatively more important than the potential lack of supermarkets or full-service restaurants.

INTRODUCTION

Between 1970 and 2000, the number of convenience stores and fast food restaurants increased, while availability of supermarkets remained stable in the U.S.[1] There has also been greater attention to ‘food deserts’ (lack of healthy food options) and ‘food swamps’ (abundance of unhealthy food options) in the literature in the past decade.[2–4] The unequal distribution of food outlets (i.e., convenience and grocery stores, supermarkets, and fast food and full-service restaurants) may negatively impact individuals’ cardiometabolic health. For example, some researchers have found that fewer supermarkets and more convenience stores and fast food restaurants are associated with lower quality diet and obesity,[5–10] but these results do not always hold.[1,11–13] Therefore, it is important to understand which factors influence the availability of food outlets in neighborhoods, and whether it differs across type of food outlet.

According to a rich evidence base in urban economics, factors that influence food outlet site selection include competition, zoning ordinances,[14–18] proximity to the central business district,[19] socioeconomic status (SES), and neighborhood racial composition,[19,20] but the literature is mixed [21] and results differ by type of outlet.[22] The current literature is limited by cross-sectional study designs and a lack of sophisticated modeling techniques. For example, the majority of previous studies either focus on one type of food outlet (without accounting for the presence of other food outlets) or use crude measures of retail density. In addition, previous studies have not explicitly examined how the availability of one type of food outlet in a given neighborhood influences the availability of other types of food outlets in later years, or whether relationships differ by neighborhood-level income.

To address these gaps, unique data related to neighborhood sociodemographics, food outlets, and the built environment within four U.S. cities (Birmingham, AL; Chicago, IL; Minneapolis, MN; Oakland, CA) from six time-points (1986, 1993, 1996, 2001, 2006, and 2011) were used. Using two-step econometric models, longitudinal associations between neighborhood-level characteristics and the presence and density of fast food restaurants, full-service restaurants, convenience stores, grocery stores, and supermarkets were separately analyzed. Associations between the availability of food outlets in earlier years and the current availability of food outlets were also examined, as well as interactions by neighborhood-level income. The results may inform policies and interventions to mitigate disparities in food outlet availability and access to healthy food options.

METHODS

Study population

We capitalized on existing data from the Coronary Artery Risk Development in Young Adults (CARDIA) study, including detailed community-level measures corresponding to participants' residential addresses in four U.S. cities (Birmingham, AL; Chicago, IL; Minneapolis, MN; Oakland, CA) across 6 exam periods in 1986, 1993, 1996, 2001, 2006, and 2011 (although individual-level data were not used in this study). These four cities are emblematic of different types of metropolitan areas in the U.S., with broad geographic and sociodemographic diversity.

The goal was to ask questions about whether neighborhood characteristics were associated with the food environment over time. To do this necessitated an understanding of salient neighborhood boundaries. Since real-estate-oriented boundaries reflected such an understanding, they were deemed theoretically appropriate for the research question[23,24]. Real-estate-oriented neighborhood boundary definitions from Zillow[25] (Chicago, IL; Oakland, CA; Minneapolis, MN) and the Regional Planning Commission of Greater Birmingham (Birmingham, AL) were used for a total of 392 neighborhoods at six time points; observations=2,352). Neighborhood observations missing values for any exposure variables were excluded (n=105; 25, 17, 19, 21, 12, and 11 in 1986, 1993, 1996, 2001, 2006, and 2011, respectively), with a final analytic sample of 2,247 neighborhood-observations across follow-up.

Measures

Since our source data were produced at varying geographic levels, we harmonized all source data to fit real-estate oriented neighborhood boundaries using geographic proportional weighting and temporally aligned the data according to the six CARDIA exam periods (Supplemental File 2). Data from Dun & Bradstreet, a commercial data set of US businesses[26] with moderate reliability and validity,[27–29] was used to characterize the food environment in 1986, 1993, 1996, 2001, 2006, and 2011. Eight-digit Standard Industrial Classification (SIC) codes were used to classify food outlets; only 4-digit codes were available in 1986, which were used along with matched business names and prediction model for full-service restaurants (Supplemental File 2). The count of each type of food outlet within each neighborhood unit (km^2) was calculated for each year of analysis.

Census tract-level age structure, racial composition, income, and population density were obtained from the U.S. Census 1980, 1990, and 2000 and American Community Survey 5-year estimates from 2005–2009 and 2007–2011. To match the years of analysis, a continuous change in sociodemographic characteristics across all decennial and quinquennial censuses was estimated using linear interpolation. Regions of significant employment density with any contiguous group of Traffic Analysis Zones meeting a previously-defined threshold criterion[30] were identified using data from the Census Transportation Planning Package corresponding to the years 1990, 2000, 2005, and 2010.

Statistical analysis

Exposures in the analysis included: population density (per km^2); median household income ($); percentage neighborhood white population; percentage foreign-born population; percentage population with less than a high school (HS) education; percentage population ≤18 years of age; percentage of vacant housing units; distance to employment subcenter (km^2); and except when modeled as an outcome, the count per km^2 of fast food restaurants, full-service restaurants, convenience stores, grocery stores, supermarkets, and all other types of food outlets (separately). To account for inflation, we divided median household income by the Consumer Price Index at each year of analysis. The exposures were modeled as *z*-scores to obtain comparable estimates of effect. The goal was to specify a causal model that allowed for exposures at one time point to influence outcomes at a later time point (i.e., latency period). Therefore, exposures were lagged by the period of time between years of analysis (***t-1***); latency periods were unevenly distributed due to unequal intervals between 1985 and 1995.

The count of each type of food outlet per km^2 (described as “density” henceforth) was right-skewed, so log-transformed values were used (outcome variables only). To maximize the interpretability of results, beta coefficients and confidence intervals from two-step models were reported as percentage values. City and year of analysis (1986, 1993, 1996, 2001, 2006, and 2011) were controlled for in all models.

According to the theory of entry thresholds, an area must be a certain market size to support a new business, but once the entry threshold is crossed, competition does not meaningfully impact entry of additional businesses.[31–33] Therefore, it was hypothesized that an increase from zero to one food outlet was meaningfully different than any other increase in the count of food outlets. Consequently, two-step econometric models (versus Poisson models) were utilized (Stata 14.0),[34–36] which account for the threshold effect of zero versus any neighborhood food outlets.

In the first equation, a probit regression model was used to estimate the longitudinal associations between neighborhood-level characteristics and the probability of the presence of each type of food outlet (yes/no). In the second equation, a linear regression model was used to estimate the longitudinal associations between neighborhood-level characteristics and the log-transformed density of each type of food outlet (separately), conditional on the presence of any food outlet of that type. Random effects (versus fixed effects) regression was used for all outcomes because we assumed that unobserved heterogeneity was not correlated with our explanatory variables. For example, unmeasured factors known to affect store placement, such as traffic patterns,[16] are likely not correlated with our exposures.

Using the two-step model,[34–36] the coefficients from the second equation (linear model among neighborhoods with any food store or restaurant; unconditional model) were multiplied by the predicted probability from the first equation (probit model among full sample) to generate conditional estimates; thus, estimates from the two-step model represent the estimated effects of neighborhood-level characteristics (including density of each type of food outlet, except when modeled as an outcome) at year ***t-1*** on the density of each type of food outlet at year ***t***, unconditional on the presence of the same type of food outlet in the neighborhood (across the full sample). It was hypothesized that the same set of covariates would affect the outcome in both equations; thus, separate identification was not necessary.[35]

Given literature showing modification of associations between neighborhood racial composition and accessibility of supermarkets by poverty level,[37] we hypothesized that the effect of each exposure differed by neighborhood-level income. Based on preliminary analyses, we added statistically significant interaction terms to each two-step model, and predicted marginal effects at the 10th and 90th percentile of income level.

Observations with studentized residuals >|4| were excluded from all analyses to address outliers (n=1 and 3 neighborhood-observations excluded from the convenience store and grocery store models, respectively).[38] Standard errors for the combined coefficients of the probit and linear regression models were obtained by bootstrapping with 1000 replications. We assessed spatial autocorrelation by calculating the Moran's I for each outcome variable at baseline and at the end of follow-up[39]. Due to spatial clustering in the outcome variables (Supplemental Table 2), we included polynomial terms for latitude and longitude in each model.[40]

In sensitivity analyses, exposure and outcome data from the same year of analysis were used, which were compared to estimates from the central analysis with lagged exposures. All data were analyzed in 2016.

RESULTS

Across the four cities, the counts of each type of food outlet increased over time (Table 1). Across the follow-up period, the mean (SD) count of neighborhood fast food restaurants, full-service restaurants, convenience stores, grocery stores, and supermarkets per km^2 was 1.8 (4.2), 1.9 (5.3), 1.3 (1.7), 2.0 (3.6), 0.2 (1.0), and 4.7 (13.7), respectively. Including polynomial terms for latitude and longitude coordinates resulted in negligible differences in estimates compared to models without these terms.

Model-based estimates of lagged sociodemographics on density of food outlets at follow-up

The estimated coefficients for the probability of the presence of each food outlet type (equation 1) and the coefficients unconditional on the presence of each food outlet type are shown (two-step model) in Table 2. Across follow-up, median household income was negatively associated with the density of convenience stores (unconditional on the presence of convenience stores). We also observed a statistically significant interaction between the percentage of white population and median household income with the density of fast food restaurants (p-interaction, 0.001). For example, a 1% increase in the percentage of white population was associated with a 17% (95% CI: 34, −0.2; p=0.05) *decrease* in the density of fast food restaurants (unconditional on the presence of fast food restaurants) in low-income neighborhoods (10th percentile); whereas, a 22% (95% CI: 2, 43; p=0.04) *increase* in the density of fast food restaurants in high-income neighborhoods (90th percentile) over time.

Across all income levels, the percentage of vacant housing units and the percentage of population with less than a HS education were positively associated with the density of grocery stores (unconditional on the presence of grocery stores) over time; while the percentage of white population was negatively associated with the density of grocery stores (unconditional on the presence of grocery stores). The percentage of foreign-born population was also negatively associated with the density of grocery stores and supermarkets (unconditional on the presence of the same type of food stores). Adjusting for other factors, no other sociodemographic measures were statistically significantly associated with the density of supermarkets (unconditional on the presence of supermarkets) across the study period.

Model-based estimates of lagged food outlets on density of food outlets at follow-up

A higher density of fast food restaurants at earlier years was positively associated with the current density of full-service restaurants (unconditional on the presence of the same type of food outlets), and vice versa. Additionally, a higher density of convenience stores at earlier years was positively associated with the current density of grocery stores (unconditional on the presence of grocery stores).

The density of food outlet types at earlier years did not influence the current density of grocery stores or supermarkets (unconditional on the presence of those types of food outlets). We also did not observe any statistically significant interactions between food environment exposures at earlier years and neighborhood-level income.

Sensitivity analyses

The direction and magnitude of the predicted marginal effects from models with contemporaneous exposure and outcome data were similar to results from the central analysis, with negligible differences in statistical significance (Table S1).

DISCUSSION

A wide range of time-varying neighborhood-level data related to the food environment in four U.S. metropolitan areas was used to determine how neighborhood characteristics influenced the density of food outlets over time. Neighborhood age and racial composition, educational attainment, and vacancy were found to be associated with food outlet density, with differences by neighborhood-level income level. The findings also suggest that the density of food outlets at earlier years was associated with the current presence or density of fast food and full-service restaurants, smaller grocery stores, and supermarkets, but not convenience stores.

Overall, the results suggest that minority populations and socioeconomically-disadvantaged individuals may live in neighborhoods that attract a higher availability of ‘less healthy’ food outlets’ (i.e., fast food restaurants, convenience stores, and smaller grocery stores) over time, compared to other more advantaged segments of the population. The results also showed that relationships might be stronger in low-income neighborhoods; for example, the percentage of non-Hispanic White population at earlier years was negatively associated with the current density of fast food restaurants only in low-income neighborhoods. Given previously documented negative associations between convenience and grocery stores, and fast food restaurants with in-store food measures, diet, and obesity,[5,7,27,41] individuals living in these areas may be at a heightened risk for less healthy diet behaviors. Thus, future initiatives to modify the food environment should reduce the relative availability of ‘less healthy’ food outlets types rather than the potential lack of supermarkets or full-service restaurants, especially in low-income areas.

In this study, the density of convenience stores at earlier years was positively associated with the current density of smaller grocery stores. We also observed a positive association between the density of fast food restaurants at earlier years and the current density of full-service restaurants, and vice-versa. These findings may be due to unmeasured factors that cause similar types of food outlets to locate together (e.g., proximity to complementary businesses)[16], especially in neighborhoods that can support additional businesses.[31]

Differences in the associations between neighborhood characteristics with the presence versus density of different types of food outlets were also observed. For example, the percentage of foreign-born population at earlier years was negatively associated with the current density of supermarkets, but positively associated with the current presence (yes/no) of supermarkets. These findings suggest that neighborhood factors may influence the growth of additional food outlets, but not initial density (and vice-versa). Inconsistencies between probit and two-step models may be due to zoning ordinances and land-use policies that determine whether food outlets to locate in certain neighborhoods,[15] but do not necessarily influence the density of food outlets in areas that permit food retail.

Limitations

While a wealth of data on neighborhoods was used in this study, data related to crime, within-store food measures, and other factors related to store location choice was lacking,[14,16,18,21,42] as well as employment density data prior to 1990. Furthermore, it was not possible to capture the openings and closings of unique food outlets, and thus not possible to examine whether the availability of each type of food outlet caused new food outlets (or closings) to occur over time. Although polynomial terms for latitude and longitude were used to account for potential spatial autocorrelation, the use of two-step econometric models precluded the use of complex spatial regression analysis to address the modifiable areal unit problem.[40] Despite a lack of continuous annual data and having unequal latency periods between 1980 and 1995, this study captured longitudinal associations between our exposures and outcomes using lagged exposures. Data on zoning ordinances was also unavailable, but the use of two-step models accounted for the probability that a food outlet was present in a neighborhood before estimating associations between exposures and the density of food outlets. Although the sample of neighborhoods is not generalizable to less urban areas, the four metropolitan areas in this study are emblematic of different types of US cities. We also used a unique neighborhood boundary data source in Birmingham, which may not be analogous to Zillow-defined boundaries in the other cities; however, both sources use the same theoretical approach to creating neighborhood boundaries with real estate data. We observed small classification errors in D&B, but previous research shows that misclassification of food stores in secondary data sources is not systematic by neighborhood characteristics.[28]

Conclusion

The results of this study underscore the complex relationships between neighborhood sociodemographics and food outlets over time, and show that minority populations and socioeconomically-disadvantaged individuals may live in areas with greater access to 'less healthy food outlets'. Given the lack of success of interventions introducing supermarkets or banning new fast food establishments in underserved areas,[43–48] the results of this study suggest that new strategies are necessary for promoting equitable food environments. Understanding which neighborhood characteristics influence the availability of different types of neighborhoods food outlets allows policy-makers to identify salient targets for stimulating changes in the distribution of food outlets across social and geographic space.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to acknowledge Marc Peterson of the UNC, Carolina Population Center (CPC); and the CPC Spatial Analysis Unit for the creation of the environmental variables. The authors have no financial or other conflicts of interest to disclose. Pasquale E. Rummo had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of data analysis.

FUNDING:

This work was supported by the National Heart, Lung, and Blood Institute (NHLBI) R01HL104580. The authors are grateful to the Carolina Population Center, University of North Carolina at Chapel Hill (grant P2C HD050924 from the National Institute of Health [NIH] for general support and grant T32 HD007168 for training support; the Nutrition Obesity Research Center (NORC), University of North Carolina (grant P30DK56350 from the National Institute for Diabetes and Digestive and Kidney Diseases [NIDDK]); and the UNC Center for Environmental Health and Susceptibility (CEHS), University of North Carolina (grant P30ES010126 from the National Institute of Environmental Health Sciences [NIEHS]).

References

1. Richardson AS, Meyer KA, Howard AG, et al. Neighborhood socioeconomic status and food environment: a 20-year longitudinal latent class analysis among CARDIA participants. Health & place. Nov.2014 30:145–153. [PubMed: 25280107]
2. Fielding JE, Simon PA. Food Deserts or Food Swamps?: Comment on "Fast Food Restaurants and Food Stores". Archives of internal medicine. 2011; 171(13):1171–1172. [PubMed: 21747012]

3. Saunders P, Saunders A, Middleton J. Living in a 'fat swamp': exposure to multiple sources of accessible, cheap, energy-dense fast foods in a deprived community. British Journal of Nutrition. 2015:1–7.
4. Walker RE, Keane CR, Burke JG. Disparities and access to healthy food in the United States: a review of food deserts literature. Health & place. 2010; 16(5):876–884. [PubMed: 20462784]
5. Duran AC, Diez Roux AV, do Latorre MR, Jaime PC. Neighborhood socioeconomic characteristics and differences in the availability of healthy food stores and restaurants in Sao Paulo, Brazil. Health & place. Sep.2013 23:39–47. [PubMed: 23747923]
6. Hollands S, Campbell MK, Gilliland J, Sarma S. Association between neighbourhood fast-food and full-service restaurant density and body mass index: a cross-sectional study of Canadian adults. Canadian journal of public health = Revue canadienne de sante publique. May-Jun;2014 105(3):e172–178. [PubMed: 25165835]
7. Larson NI, Story MT, Nelson MC. Neighborhood environments: disparities in access to healthy foods in the U.S. American journal of preventive medicine. Jan; 2009 36(1):74–81. [PubMed: 18977112]
8. Morland K, Wing S, Roux AD. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. American journal of public health. 2002; 92(11):1761–1768. [PubMed: 12406805]
9. Powell LM, Han E, Chaloupka FJ. Economic contextual factors, food consumption, and obesity among U.S. adolescents. The Journal of nutrition. Jun; 2010 140(6):1175–1180. [PubMed: 20392882]
10. Powell LM, Nguyen BT. Fast-food and full-service restaurant consumption among children and adolescents: effect on energy, beverage, and nutrient intake. JAMA pediatrics. Jan; 2013 167(1): 14–20. [PubMed: 23128151]
11. Duffey KJ, Gordon-Larsen P, Steffen LM, Jacobs DR Jr, Popkin BM. Regular consumption from fast food establishments relative to other restaurants is differentially associated with metabolic outcomes in young adults. The Journal of nutrition. Nov; 2009 139(11):2113–2118. [PubMed: 19776183]
12. Pearce J, Hiscock R, Blakely T, Witten K. The contextual effects of neighbourhood access to supermarkets and convenience stores on individual fruit and vegetable consumption. Journal of epidemiology and community health. Mar; 2008 62(3):198–201. [PubMed: 18272733]
13. Pearce J, Hiscock R, Blakely T, Witten K. A national study of the association between neighbourhood access to fast-food outlets and the diet and weight of local residents. Health & place. Mar; 2009 15(1):193–197. [PubMed: 18499502]
14. Litz RA, Rajaguru G. Does small store location matter? A test of three classic theories of retail location. Journal of Small Business & Entrepreneurship. 2008; 21(4):477–492.
15. Mayo ML, Pitts SBJ, Chriqui JF. Peer Reviewed: Associations Between County and Municipality Zoning Ordinances and Access to Fruit And Vegetable Outlets in Rural North Carolina, 2012. Preventing chronic disease. 2013:10.
16. Park K, Khan MA. An exploratory study to identify the site selection factors for US franchise restaurants. Journal of Foodservice Business Research. 2006; 8(1):97–114.
17. Prayag G, Landré M, Ryan C. Restaurant location in Hamilton, New Zealand: clustering patterns from 1996 to 2008. International Journal of Contemporary Hospitality Management. 2012; 24(3): 430–450.
18. Tzeng G-H, Teng M-H, Chen J-J, Opricovic S. Multicriteria selection for a restaurant location in Taipei. International Journal of Hospitality Management. 2002; 21(2):171–187.
19. Schuetz J, Kolko J, Meltzer R. Are poor neighborhoods "retail deserts"? Regional Science and Urban Economics. 2012; 42(1):269–285.
20. Meltzer R, Schuetz J. Bodegas or bagel shops? Neighborhood differences in retail and household services. Economic Development Quarterly. 2011 0891242411430328.
21. Hurvitz P, Moudon A, Rehm C, Streichert L, Drewnowski A. Arterial roads and area socioeconomic status are predictors of fast food restaurant density in King County, WA. International Journal of Behavioral Nutrition and Physical Activity. 2009; 6(1):46. [PubMed: 19630979]

22. Morland K, Wing S, Roux AD, Poole C. Neighborhood characteristics associated with the location of food stores and food service places. American journal of preventive medicine. 2002; 22(1):23–29. [PubMed: 11777675]

23. Arribas-Bel D. Accidental, open and everywhere: Emerging data sources for the understanding of cities. Applied Geography. 2014; 49:45–53.

24. Hirsch JA, Meyer KA, Peterson M, et al. obtaining Longitudinal Built environment data Retrospectively across 25 years in Four Us Cities. Frontiers in public health. 2016:4. [PubMed: 26870722]

25. Zillow. [Accessed 5/15/2015] Zillow Neighborhood Boundaries. http://www.zillow.com/howto/api/neighborhood-boundaries.htm

26. Laxer RE, Janssen I. The proportion of excessive fast-food consumption attributable to the neighbourhood food environment among youth living within 1 km of their school. Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme. Apr; 2014 39(4):480–486.

27. Fleischhacker SE, Evenson KR, Rodriguez DA, Ammerman AS. A systematic review of fast food access studies. Obesity reviews : an official journal of the International Association for the Study of Obesity. May; 2011 12(5):e460–471. [PubMed: 20149118]

28. Han E, Powell LM, Zenk SN, Rimkus L, Ohri-Vachaspati P, Chaloupka FJ. Classification bias in commercial business lists for retail food stores in the US. Int J Behav Nutr Phys Act. 2012; 9(1):46. [PubMed: 22512874]

29. Liese AD, Colabianchi N, Lamichhane AP, et al. Validation of 3 food outlet databases: completeness and geospatial accuracy in rural and urban food environments. American Journal of Epidemiology. 2010; 172(11):1324–1333. [PubMed: 20961970]

30. Cho EJ, Rodriguez DA, Song Y. The role of employment subcenters in residential location decisions. Journal of Transport and Land Use. 2008

31. Bresnahan TF, Reiss PC. Entry and competition in concentrated markets. Journal of Political Economy. 1991:977–1009.

32. Chevalier JA. Capital structure and product-market competition: Empirical evidence from the supermarket industry. The American Economic Review. 1995:415–435.

33. Morton FMS. Entry decisions in the generic pharmaceutical industry. The Rand journal of economics. 1999:421–440. [PubMed: 10558596]

34. Haines PS, Popkin BM, Guilkey DK. Modeling food consumption decisions as a two-step process. American Journal of Agricultural Economics. 1988; 70(3):543–552.

35. Madden D. Sample selection versus two-part models revisited: The case of female smoking and drinking. Journal of health economics. 2008; 27(2):300–307. [PubMed: 18180064]

36. Manning WG, Duan N, Rogers WH. Monte Carlo evidence on the choice between sample selection and two-part models. Journal of econometrics. 1987; 35(1):59–82.

37. Zenk SN, Schulz AJ, Israel BA, James SA, Bao S, Wilson ML. Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in metropolitan Detroit. American journal of public health. 2005; 95(4):660–667. [PubMed: 15798127]

38. Donald SG, Maddala G. Identifying outliers and influential observations in econometric models. Handbook of statistics. 1993; 11:663–701.

39. Getis A, Ord JK. The analysis of spatial association by use of distance statistics. Geographical analysis. 1992; 24(3):189–206.

40. Auchincloss AH, Gebreab SY, Mair C, Roux AVD. A review of spatial methods in epidemiology, 2000–2010. Annual review of public health. 2012; 33:107.

41. Gibson DM. The neighborhood food environment and adult weight status: estimates from longitudinal data. American journal of public health. 2011; 101(1):71–78. [PubMed: 21088263]

42. Kincaid C, Baloglu S, Mao Z, Busser J. What really brings them back? The impact of tangible quality on affect and intention for casual dining restaurant patrons. International Journal of Contemporary Hospitality Management. 2010; 22(2):209–220.

43. Cummins S, Flint E, Matthews SA. New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity. Health Affairs. 2014; 33(2):283–291. [PubMed: 24493772]

44. Cummins S, Petticrew M, Higgins C, Findlay A, Sparks L. Large scale food retailing as an intervention for diet and health: quasi-experimental evaluation of a natural experiment. Journal of epidemiology and community health. 2005; 59(12):1035–1040. [PubMed: 16286490]

45. Sadler RC, Gilliland JA, Arku G. A food retail-based intervention on food security and consumption. International journal of environmental research and public health. 2013; 10(8):3325–3346. [PubMed: 23921626]

46. Wrigley N, Warm D, Margetts B. Deprivation, diet, and food-retail access: findings from the Leedsfood deserts' study. Environment and Planning A. 2003; 35(1):151–188.

47. Elbel B, Moran A, Dixon LB, et al. Assessment of a government-subsidized supermarket in a high-need area on household food availability and children's dietary intakes. Public health nutrition. 2015; 18(15):2881–2890. [PubMed: 25714993]

48. Sturm R, Hattori A. Diet and obesity in Los Angeles County 2007–2012: Is there a measurable effect of the 2008 "Fast-Food Ban"? Social science & medicine. 2015; 133:205–211. [PubMed: 25779774]

TABLE 1

Descriptive statistics of food stores and restaurants and other neighborhood-level characteristics [a]

	1986 (n=367 neighborhoods)	1993 (n=375 neighborhoods)	1996 (n=373 neighborhoods)	2001 (n=371 neighborhoods)	2006 (n=380 neighborhoods)	2011 (n=381 neighborhoods)	1986–2011 (n=2247 neighborhood-observations)
Food stores and restaurants – count [mean (SD)]							
Fast food restaurants	2.8 (6.1)	4.2 (9.2)	4.2 (9.6)	4.3 (10.0)	6.9 (15.9)	7.3 (17.7)	1.8 (4.2)
Full-service restaurants	1.9 (7. 7)	3.8 (11.0)	3.6 (11.1)	4.1 (11.2)	6.8 (17.2)	8.3 (20.8)	1.9 (5.3)
Convenience stores	2. 3 (3.6)	4.3 (7.4)	3.5 (5.8)	3.3 (5.3)	4.2 (7.2)	4.0 (6.7)	1.3 (1.7)
Grocery stores	4. 4 (9.0)	7.4 (15.8)	5.3 (10.6)	4.7 (8.2)	5.4 10.0)	5.2 (10.1)	2.0 (3.6)
Supermarkets	0.1 (0.3)	0.4 (0.9)	0.3 (0.7)	0.7 (0.9)	0.7 (1.4)	0.91 (1.8)	0.2 (1.0)
Other types of food outlets [b]	10.3 (22.2)	17.7 (37.2)	16. 7 (37.2)	14.9 (33.2)	15.8 (35.2)	15.8 (36.1)	4.7 (13.7)
Land area (km^2) [c]	3.2 (3.9)	3.2 (4.0)	3.3 (4.2)	3.3 (4.2)	3.2 (4.1)	3.2 (4.1)	3.2 (4.1)
Neighborhood-level characteristics [mean (SD)]							
Population density (per km^2)	3,292 (2,592)	3,300 (2,671)	3,300 (2,671)	3,448 (2,804)	3,379 (2,718)	3,309 (2,639)	3,453 (2,676)
Median household income ($) [d]	14,676 (6,456)	27,208 (14,051)	27,208 (14,051)	40,147 (21,404)	49,280 (28,636)	51,015 (29,904)	35,218 (24,741)
Percentage of white population	52.7 (35.9)	46.2 (34.5)	46.2 (34.5)	39.6 (30.1)	41.0 (29.0)	41.8 (28.9)	45.3 (32.5)
Percentage of residents ≤18 years	25.3 (8.7)	24.4 (7.9)	24.4 (7.9)	24.5 (8.2)	23.0 (7.4)	22.3 (7.1)	23.9 (8.0)
Percentage of vacant housing units	5.4 (3.3)	7.8 (4.8)	7.8 (4.8)	6.7 (5.1)	12.5 (7.1)	13.2 (7.2)	8.9 (6.3)
Distance to nearest employment sub-center (km) [e]	5,903 (3,804)	5,903 (3804)	5,903 (3804)	6,181 (3,584)	6,181 (3,584)	6,066 (3,794)	5,823 (3,636)

[a] Pooled cities (Birmingham, AL; Chicago, IL; Minneapolis, MN; Oakland, CA)

[b] Other types of food outlets includes specialty shops, markets, food stands/cafeterias, co-ops, and warehouse stores

[c] Neighborhood boundaries were fixed over time. Land area appears to change because the number of missing values for each variable differs by year of analysis, so the number of neighborhood units differs slightly over time.

[d] Divided by the Consumer Price Index at each year of analysis to account for inflation.

[e]Distance from neighborhood centroid to nearest employment center centroid (km)

TABLE 2

Estimated model coefficients [a] of associations between lagged neighborhood-level characteristics [b] and count of food outlets [c]

	Equation 1: Estimated coefficients for probability of presence of each type of food outlet (95% CI) [d]	p-value	Two-step model: Estimated coefficients for count per km^2 unconditional on presence of each type of food outlet (95% CI) [e]	p-value
(a)				
Fast food restaurants (log-transformed count per km^2)				
Population density (per km^2) at *t-1*	−0.04 (−0.35, 0.26)	0.78	0.19 (0.07, 0.31)	**0.002**
Median household income ($) at *t-1* [f]	−0.13 (−0.45, 0.18)	0.40	−0.12 (−0.27, 0.03)	0.13
Percentage white population at *t-1*	0.09 (−0.24, 0.42)	0.60	0.0005 (−0.15, 0.15)	0.99
Percentage population ≤18 years at *t-1*	−0.06 (−0.31, 0.19)	0.66	−0.17 (−0.28, −0.06)	**0.003**
Percentage vacant housing units at *t-1*	−0.004 (−0.22, 0.21)	0.97	0.03 (−0.06, 0.12)	0.56
Distance to employment subcenter (km) at *t-1* [g]	−0.29 (−0.59, 0.01)	0.06	−0.15 (−0.28, −0.03)	**0.02**
Percentage foreign-born population at *t-1*	−0.19 (−0.79, 0.39)	0.53	−0.001 (−0.19, 0.19)	0.99
Percentage population with <HS education at *t-1*	−0.15 (−0.43, 0.13)	0.28	−0.08 (−0.20, 0.04)	0.18
Median household income*percentage white population at *t-1*	0.27 (0.02, 0.52)	**0.04**	0.18 (0.08, 0.29)	**0.001**
Full-service restaurants (count/km^2) at *t-1*	0.49 (0.12, 0.85)	**0.01**	0.21 (0.04, 0.37)	**0.01**
Convenience stores (count/km^2) at *t-1*	−0.002 (−0.17, 0.17)	0.98	0.03 (−0.04, 0.10)	0.39
Grocery stores (count/km^2) at *t-1*	0.12 (−0.08, 0.34)	0.22	0.04 (−0.06, 0.14)	0.43
Supermarkets (count/km^2) at *t-1*	0.16 (−0.11, 0.42)	0.25	0.08 (−0.04, 0.20)	0.18
Other stores & restaurants (count/km^2) at *t-1*	0.13 (−0.19, 0.46)	0.42	0.07 (−0.05, 0.19)	0.26
(b)				
Full-service restaurants (log-transformed count per km^2)				
Population density (per km^2) at *t-1*	0.22 (−0.07, 0.51)	0.14	0.28 (0.18, 0.39)	**<0.001**
Median household income ($) at *t-1* [f]	−0.05 (−0.34, 0.23)	0.73	−0.02 (−0.12, 0.08)	0.75
Percentage white population at *t-1*	0.41 (0.11, 0.71)	**0.01**	0.07 (−0.05, 0.19)	0.27
Percentage population ≤18 years at *t-1*	−0.14 (−0.38, 0.10)	0.26	−0.15 (−0.24, −0.07)	**0.001**
Percentage vacant housing units at *t-1*	0.05 (−0.12, 0.22)	0.56	0.04 (−0.02, 0.10)	0.16
Distance to employment subcenter (km) at *t-1* [g]	−0.34 (−0.60, −0.08)	**0.01**	−0.17 (−0.28, −0.06)	**0.002**
Percentage foreign-born population at *t-1*	0.01 (−0.52, 0.55)	0.96	0.04 (−0.10, 0.18)	0.61

	Equation 1: Estimated coefficients for <u>probability</u> of presence of each type of food outlet (95% CI) [d]	p-value	Two-step model: Estimated coefficients for count per km² <u>unconditional</u> on presence of each type of food outlet (95% CI) [e]	p-value
Percentage population with <HS education at *t-1*	−0.25 (−0.51, 0.02)	0.07	−0.08 (−0.17, 0.02)	0.10
Median household income*percentage population ≤18 years at *t-1*	0.10 (−0.07, 0.28)	0.26	0.09 (0.02, 0.17)	**<0.001**
Fast food restaurants (count/km²) at *t-1*	0.79 (0.41, 1.18)	**<0.001**	0.23 (0.07, 0.39)	**0.004**
Convenience stores (count/km²) at *t-1*	0.05 (−0.13, 0.23)	0.59	0.05 (−0.01, 0.10)	0.09
Grocery stores (count/km²) at *t-1*	−0.06 (−0.28, 0.17)	0.62	−0.02 (−0.08, 0.05)	0.59
Supermarkets (count/km²) at *t-1*	0.28 (−0.06, 0.62)	0.10	0.08 (−0.02, 0.18)	0.11
Other stores & restaurants (count/km²) at *t-1*	0.28 (0.0001, 0.56)	**0.05**	0.13 (0.04, 0.21)	**0.01**
(c)				
Convenience stores (log-transformed count per km²)				
Population density (per km²) at *t-1*	0.01 (−0.31, 0.33)	0.94	0.27 (0.16, 0.38)	**<0.001**
Median household income ($) at *t-1* [f]	−0.36 (−0.74, 0.02)	0.06	−0.14 (−0.26, −0.01)	**0.03**
Percentage white population at *t-1*	0.29 (−0.08, 0.67)	0.12	−0.003 (−0.14, 0.13)	0.96
Percentage population ≤18 years at *t-1*	0.14 (−0.12, 0.40)	0.31	−0.03 (−0.13, 0.06)	0.51
Percentage vacant housing units at *t-1*	−0.02 (−0.20, 0.16)	0.79	0.05 (−0.02, 0.12)	0.18
Distance to employment subcenter (km) at *t-1* [g]	−0.07 (−0.40, 0.26)	0.68	−0.13 (−0.26, −0.002)	**0.05**
Percentage foreign-born population at *t-1*	0.04 (−0.54, 0.63)	0.89	0.02 (−0.14, 0.18)	0.82
Percentage population with <HS education at *t-1*	−0.21 (−0.52, 0.09)	0.17	−0.07 (−0.18, 0.04)	0.19
Fast food restaurants (count/km²) at *t-1*	−0.05 (−0.32, 0.22)	0.71	0.05 (−0.05, 0.15)	0.30
Full-service restaurants (count/km²) at *t-1*	0.14 (−0.23, 0.52)	0.45	0.04 (−0.09, 0.16)	0.56
Grocery stores (count/km²) at *t-1*	0.15 (−0.09, 0.38)	0.22	0.03 (−0.06, 0.11)	0.53
Supermarkets (count/km²) at *t-1*	0.20 (−0.06, 0.46)	0.12	0.08 (−0.06, 0.21)	0.27
Other stores & restaurants (count/km²) at *t-1*	0.23 (−0.13, 0.60)	0.21	0.06 (−0.04, 0.17)	0.23
(d)				
Grocery stores (log-transformed count per km²)				
Population density (per km²) at *t-1*	0.35 (0.02, 0.68)	**0.04**	0.49 (0.33, 0.66)	**<0.001**
Median household income ($) at *t-1* [f]	0.04 (−0.28, 0.36)	0.80	−0.05 (−0.21, 0.10)	0.51
Percentage white population at *t-1*	−0.26 (−0.60, 0.08)	0.14	−0.27 (−0.45, −0.09)	**0.003**
Percentage population ≤18 years at *t-1*	−0.12 (−0.36, 0.13)	0.35	−0.18 (−0.32, −0.04)	**0.01**

	Equation 1: Estimated coefficients for <u>probability</u> of presence of each type of food outlet (95% CI) [d]	p-value	Two-step model: Estimated coefficients for count per km^2 <u>unconditional</u> on presence of each type of food outlet (95% CI) [e]	p-value
Percentage vacant housing units at *t-1*	0.03 (−0.16, 0.22)	0.74	0.09 (0.004, 0.18)	**0.04**
Distance to employment subcenter (km) at *t-1* [g]	−0.14 (−0.42, 0.15)	0.35	−0.14 (−0.29, −0.0004)	**0.05**
Percentage foreign-born population at *t-1*	−0.85 (−1.45, −0.24)	**0.01**	−0.32 (−0.58, −0.06)	**0.02**
Percentage population with <HS education at *t-1*	0.31 (0.02, 0.60)	**0.04**	0.23 (0.07, 0.38)	**0.01**
Median household income*distance to employment subcenter at *t-1*	−0.26 (−0.46, −0.05)	**0.01**	−0.15 (−0.29, −0.02)	**0.02**
Fast food restaurants (count/km^2) at *t-1*	0.04 (−0.28, 0.35)	0.83	0.02 (−0.14, 0.18)	0.84
Full-service restaurants (count/km^2) at *t-1*	0.63 (0.22, 1.04)	**0.003**	0.30 (0.10, 0.50)	**0.004**
Convenience stores (count/km^2) at *t-1*	0.34 (0.15, 0.53)	**<0.001**	0.18 (0.06, 0.29)	**0.003**
Supermarkets (count/km^2) at *t-1*	−0.41 (−0.68, −0.14)	**0.003**	−0.14 (−0.30, 0.02)	0.09
Other stores & restaurants (count/km^2) at *t-1*	0.17 (−0.26, 0.59)	0.44	0.07 (−0.19, 0.33)	0.59
(e)				
Supermarkets (log-transformed count per km^2)				
Population density (per km^2) at *t-1*	0.16 (−0.09, 0.41)	0.20	−0.02 (−0.17, 0.13)	0.80
Median household income ($) at *t-1* [f]	−0.09 (−0.35, 0.17)	0.50	0.04 (−0.12, 0.19)	0.64
Percentage white population at *t-1*	0.16 (−0.13, 0.45)	0.27	−0.09 (−0.28, 0.10)	0.33
Percentage population ≤18 years at *t-1*	0.04 (−0.18, 0.26)	0.71	−0.03 (−0.14, 0.09)	0.64
Percentage vacant housing units at *t-1*	0.10 (−0.06, 0.26)	0.24	−0.05 (−0.16, 0.06)	0.39
Distance to employment subcenter (km) at *t-1* [g]	0.14 (−0.11, 0.40)	0.27	−0.09 (−0.23, 0.06)	0.26
Percentage foreign-born population at *t-1*	0.54 (0.04, 1.04)	**0.03**	−0.26 (−0.51, −0.02)	**0.03**
Percentage population with <HS education at *t-1*	0.02 (−0.25, 0.29)	0.88	−0.01 (−0.17, 0.15)	0.91
Fast food restaurants (count/km^2) at *t-1*	0.29 (0.06, 0.53)	**0.01**	−0.16 (−0.32, 0.01)	0.06
Full-service restaurants (count/km^2) at *t-1*	0.15 (−0.11, 0.40)	0.26	−0.07 (−0.22, 0.09)	0.39
Convenience stores (count/km^2) at *t-1*	0.16 (0.01, 0.32)	**0.04**	−0.07 (−0.16, 0.03)	0.16
Grocery stores (count/km^2) at *t-1*	0.04 (−0.16, 0.24)	0.70	−0.02 (−0.13, 0.10)	0.79
Other stores & restaurants (count/km^2) at *t-1*	−0.09 (−0.27, 0.09)	0.35	0.06 (−0.05, 0.18)	0.25

Boldface indicates statistical significance (p<0.05).

[a] All models were adjusted for year of analysis, city, latitude, longitude, latitude*longitude, latitude-squared, and longitude-squared. Year corresponds to the year food outlet data was obtained from Dun & Bradstreet; 1986 was dropped due to exposures being lagged one time period.

[b] All neighborhood-level characteristic values converted to z-scores by city (Birmingham, AL; Chicago, IL; Minneapolis, MN; Oakland, CA) and year (1980–2011) and lagged by one year of analysis.

[c] Food stores and restaurants calculated as the log-transformed count (per km^2) within each neighborhood at each year.

[d] Estimated beta coefficients (SE) from equation 1 are derived from population-averaged probit models of neighborhood-level characteristics on the probability of a fast food restaurant, full-service restaurant, convenience store, grocery store, or supermarket (separately) being present in a neighborhood (yes/no).

[e] Estimated beta coefficients (SE) are derived from two-step models (equation 1 by equation 2) of neighborhood-level characteristics on weighted means of changes in the count of fast food restaurants, full-service restaurants, convenience stores, grocery stores, and supermarkets (separately) per km^2 for all neighborhoods [unconditional on presence of corresponding food outlet].

[f] Divided by the Consumer Price Index at each year of analysis to account for inflation.

[g] Distance from neighborhood centroid to nearest employment center centroid (km).