

# A method for estimating neighborhood characterization in the association with availability of sit-down restaurants and supermarkets

Ke Peng (✉ [kpeng6@hnu.edu.cn](mailto:kpeng6@hnu.edu.cn))

Hunan University <https://orcid.org/0000-0001-6708-3873>

Daniel A. Rodriguez

University of California Berkeley

Jana A. Hirsch

Drexel University

Penny Gordon-Larsen

University of North Carolina at Chapel Hill

---

## Research

**Keywords:** sit-down restaurants, sociodemographic factors, neighborhoods, supermarkets, urban, suburban

**Posted Date:** January 4th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-33989/v3>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on March 25th, 2021. See the published version at <https://doi.org/10.1186/s12942-020-00257-7>.

# Abstract

**Background :** Although neighborhood-level access to food differs by sociodemographic factors, almost all research on neighborhoods and food access has used a single construct of neighborhood context, such as income or race. Neighborhoods possess many interrelated built environment and sociodemographic characteristics, a condition that obscures relationships between neighborhood factors and food access.

**Methods :** The objective of this study is to examine the association between neighborhood characterization and neighborhood food availability while accounting for the effects of many interrelated aspects of neighborhoods associated with food access. We used cluster analyses method that used multiple measures (e.g., population density, mix of land use, and sociodemographic factors) that are easy to obtain to define six neighborhood types in 1993 in the Twin Cities Region, Minnesota. We then used mixed effects regression models to estimate the differences in the relative availability of sit-down restaurants and supermarkets in 1993, 2001, and 2011 across the six neighborhood types.

**Results :** Between 1993 and 2011, inner city neighborhoods experienced a greater increase in the percent of sit-down restaurants compared with urban core, urban, and aging suburb. The differences in the percent of sit-down restaurants between inner city and aging suburb, high-income suburb and suburban edge neighborhoods increased between 1993 and 2011. Using classification of neighborhood type based on sociodemographic and built environment characteristics, we found a complex and increasingly varied distribution of restaurants and food stores.

**Conclusions :** The temporal increase in the relative availability of sit-down restaurants in inner cities after accounting for all restaurants might be associated with a higher proportion of residents who eat-away-from-home, which is associated with higher calorie and fat intake. Keywords: built environment, sociodemographic, food stores, urbanization

## 1 Introduction

Previous studies on access to healthy food have generally characterized neighborhoods by sociodemographic attributes of neighborhood context [1,2], such as income or race. Although low-income and minority-dominant neighborhoods generally have poor access to healthy food [3], findings on this subject are inconclusive in a U.S context. For example, some investigators have observed that, compared with moderate and high-income neighborhoods, low-income neighborhoods tend to have greater availability of fast-food restaurants [4,5], whereas other investigators have not observed a higher prevalence of fast-food restaurants in low-income neighborhoods [6,7]. In accord with the definition of Caspi et al. [8], herein we refer to *availability* as the presence of certain types of restaurants or food stores in the neighborhood. This definition does not include the degree of ease of getting to food outlets, the food prices, or a person's attitude about whether the supply of products meets their standards. There is a new trend of combining a variety of sociodemographic or socio-economic factors as composite indices

[2,9,10] to characterize neighborhoods and relate such indices to food availability. However, a common weakness of previous studies is that they failed to adequately address built environment factors that related to the type and distribution of food outlets, such as population density and land use pattern. For example, purveyors of some types of restaurants and food stores may choose to locate in poor neighborhoods because residential densities are sufficiently high to maintain demand [11]. A few recent studies [9,12] in Spain and western Australia (Perth) have combined built environment with sociodemographic factors in characterizing neighborhoods. Such studies classified the neighborhoods by, for example, predefined level of development stage first and then sub-classified newly-developed neighborhood by income level (high, medium, low). This scheme raises an issue as to whether there is a clear line between, for example, newly-developed neighborhoods and old neighborhoods [13].

In fact, we know little about neighborhoods defined by a multidimensional categorization that acknowledges the patterning of neighborhoods across many interrelated built environment and sociodemographic characteristics [14]. Because neither aggregate indices of sociodemographic factors nor specific aspects of the built environment appear in isolation in neighborhoods [15], we used a grouping technique, namely cluster analysis, to classify neighborhood types by a combination of several domains. Although cluster analysis was widely used by other researchers as a typical approach to classify objects into clusters, it was rarely used to characterize neighborhoods based on multiple interrelated sociodemographic and built environmental variables that related to the types and distribution of food outlets. Cluster analysis can take into account a broad set of neighborhood resource variables to fully capture multiple neighborhood dimensions. Thus, cluster analysis can be used to disentangle the mixed results derived from different neighborhood types. This measurement strategy identifies groups of neighborhoods with shared characteristics (such as population density, mix of land use, and sociodemographic factors) that may entice a restaurant or food store to locate within a neighborhood. Thus, similar to previous studies [14–16], we used cluster analysis as a strategy to define neighborhood types and document their patterns of restaurants and food stores.

The aim of this study was to examine the association between baseline year neighborhood characterization and change in neighborhood food availability while accounting for the effects of many interrelated aspects of neighborhoods associated with food access. Using the baseline-change method of analysis, we examined the distribution of types of restaurants and food stores within each type of neighborhood to determine whether a particular neighborhood type had relatively greater access to a specific type of restaurant or food store compared with other neighborhood types over three observational years. We make two major contributions to the food access literature. First, the study acknowledged that neighborhoods are patterned by interrelated features; thus, we used cluster analysis techniques and sociodemographic and built environmental factors to assemble neighborhoods into homogenous groups instead of relying on a couple of pre-specified factors and cut-off levels (aka, high, medium, low) to find the homogenous neighborhoods. Second, on the basis of this new way of characterizing neighborhoods, we documented and compared the patterns of restaurants and food stores for the three observational years, which provided an approach to single out the neighborhoods that disclosed varied change in food availability.

## 2 Methods

### 2.1 Study area

We analyzed the Twin Cities Region of Minnesota (abbreviated as Twin Cities Region), an area of nearly three million people living in 186 communities across the seven counties of Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington. The Twin Cities Region has developed several distinctive types of neighborhoods (e.g., active downtown, vibrant urban) [15]. In addition, from 1985-2010, the neighborhood environment in the Twin Cities Region became increasingly diverse in social composition and physical form [15]. Therefore, we expected that the Twin Cities Region would be an ideal case in which to observe temporal differences of, and changes in, the distribution of neighborhood food resources. Our study area included 2,083 census block groups defined in 2010 by the U.S. Census Bureau in the Twin Cities Region with diverse built environment and sociodemographic characteristics [17]. We used census block groups to operationalize neighborhoods. The census block group (approximate population of 1,500) is the smallest unit for which data are available on built environment and sociodemographic measures. We excluded only seven census block groups because of missing data.

### 2.2 Relative availability of sit-down restaurants and supermarkets

We obtained food resource data from the D&B Duns Market Identifiers File (restaurant and food store Standard Industrial Classification categories; Dun & Bradstreet, Inc., Short Hills, NJ), a secondary commercial data source widely available in the U.S. We then classified the food resources according to primary eight-digit Standard Industrial Classification codes for data in years 1993, 2001, and 2011 (See Table S1 in the Additional File 1). We had expected to compare the business types from years 1990, 2000, and 2010; however, data for 1993, 2001 and 2011 were the only available Dun & Bradstreet business data.

Recent reports suggest that relative availability, i.e., particular proportions of various types of retail food outlets, may be more important to diet-related behaviors than the total number of outlets because relative availability offers residents competing options [18–20]. We chose to study the relative availability of sit-down restaurants and supermarkets. Sit-down restaurants such as ethnic food restaurants and seafood restaurants provide seating to eat instead of only food-to-go (either inside or drive-through). See Table S1 for the SIC codes we used to identify restaurants and food stores. Although fast food restaurants have been blamed for poor U.S. diet quality, evidence indicates that neither fast food nor sit-down restaurant were consistently more healthful [21–23]. Supermarkets are large food stores that included chained or independent hypermarkets (greater than 100,000 square feet), supermarkets (66,000-99,000 square feet), and superstores (55,000-65,000 square feet) in the current study. In the U.S. context, evidence shows that supermarkets have more or cheaper healthy food options compared with grocery stores and convenience stores, which are ubiquitous, smaller in size, and stocked with fewer or more expensive fresh and healthier food items compared with supermarkets [21–23]. We defined the relative availability of sit-down restaurants as the percent relative to total sit-down and fast food restaurants in a neighborhood (abbreviated below as percent of sit-down restaurants). We defined the relative availability of

supermarkets as the percent relative to total supermarkets, grocery stores, and convenience stores in a neighborhood (abbreviated as percent of supermarkets below). We used a container-based approach to measure the relative availability of sit-down restaurants and supermarkets and defined the Census Block Group as neighborhood. Therefore, our measure of the relative availability was based on the evidence [24] that the types and distribution of food outlets in the neighborhood are associated with diet-related behavior. We used ArcGIS 10.3 to calculate the count of each type of food resource within each neighborhood in each observational year, and then we used the counts to calculate the percent of sit-down restaurants and supermarkets in STATA 14.0. When there was no sit-down restaurant and fast food restaurant, a constant of one was added to that case so that it remained in the analysis [13]. A previous study validated the D&B food resource data and showed that the matched rate of fast food restaurants may differ by various neighborhood characteristics such as income, race, and location (urbanized area, urban cluster and non-urban area as defined by the US Census Bureau) [25]. For example, if sit-down restaurants had a higher matched rate compared with fast food restaurants in low-income neighborhoods versus high-income neighborhoods in the D&B data, we risked exaggerating the gap in the numbers of sit-down restaurants relative to total sit-down restaurants and fast food restaurants between low-income and high-income neighborhoods. By using multiple dimensions to characterize neighborhood, we may partly address the varied matched rate issue because the lower matching rate raised by, for example, income is partly compensated by introducing mix use or population density factors to characterize neighborhoods jointly.

## **2.3 Neighborhood type**

To classify neighborhood type, we used a cluster analysis that included 13 built environment and sociodemographic characteristics in 1990. Because we did not have data for the same factors in 1993, we assumed that the 1990 built environment and sociodemographic data were a valid substitute for the 1993 data. In the following sections 2.3.1 and 2.3.2, we elaborated on the built environment and sociodemographic characteristics that we chose to generate the six types of neighborhoods. In section 2.3.3 we elaborated on the type of cluster analysis we employed to generate neighborhood type and techniques to examine the robustness of type classification. We did not generate the neighborhood type in 2001 and 2011 because our focus was to examine the change in neighborhood food availability over time based on the neighborhood type identified in the baseline year (1990).

### *2.3.1 Neighborhood built environment characteristics*

Neighborhood built environment characteristics included residential population density, employment population density, mix of land use, and percent of single-family housing in the neighborhood. These characteristics were used widely in the characterization of Western built environment [26–29]. We obtained the census population and land area size data in 1990, 2000, and 2006–2009 from the Census 1990, Census 2000, and the 2006–2009 American Community Survey. We used such data from the US Census Longitudinal Tract Database, which normalized the 1990, 2000, and 2006–2009 census data to the boundaries of census tracts in 2010. We interpolated the normalized census population density data

for years from the census tract level to the census block-group level for years 1990, 2000, and 2010. We then measured residential population density as the total residential population divided by the total land area of the block group [30,31], and we measured employment population density as the total employed civilian labor force aged 16 years and above divided by the total land area of the block group. These measures of total land area excluded large bodies of water and parks but included other land uses such as commercial lands and roadways. We obtained data on categories and areas of different types of land uses for the creation of land use mix and percent of single-family housing from the GIS-based current land-use map in 1990, 2000, and 2010 from the Minneapolis Metropolitan Council. We measured the mix of land use by using the 3-tier land use entropy equation (with the denominator set to the static 3 land use types in the block group), which used three land use categories (residential, employment and retail) to calculate mix of land use in the block group [32]. Land use entropy ranges from zero (total homogeneity, with all land use in one category) to 1 (maximum heterogeneity, with an even mixture of land use). We defined the percent of single-family housing as the number of single-family housing units divided by the total number of single-family and multi-family housing units.

### *2.3.2 Neighborhood sociodemographic characteristics*

Neighborhood sociodemographic characteristics included percent of population aged under 14, aged 15-29, 30-44, 45-64, and aged 65 or above according to working age, percent of education of college or above, percent of white race, percent of black race, and median household income. We retrieved all the census sociodemographic characteristics in 1990, 2000, and the 2006–2009 American Community Survey of the U.S. Census Bureau from the US Census Longitudinal Tract Database. We then interpolated the normalized census sociodemographic characteristics data from the census tract level to the census block-group level.

### *2.3.3 Cluster analyses*

Previous work used data reduction techniques such as Principal Component Analysis and factor analysis [10,33] to group variables and generate a composite index; then, previous work used quantile values to classify neighborhoods into different types. Conversely, instead of variables, we used K-means cluster analysis to group observations (i.e., neighborhoods) by data-mining techniques that measured the intrinsic relationship between neighborhood characteristics based on Euclidean k-median clustering algorithm. We first transformed each 1990 built environment and sociodemographic variable into a z-score to achieve more comparable scales and ranges; otherwise, variables with large ranges might have weighed heavier in the analysis than variables with small ranges [34]. We then used the transformed data to perform partition cluster analyses within the 13 built environment and sociodemographic characteristics, using K-means in Stata 14.0. Because a wrong assessment of the number of clusters can lead to sub-optimum allocation of precious resources, we used three statistical approaches, Gap Statistic Method, Average Silhouette Method and Elbow Method [34], to identify the goodness of the number of clusters we chose. These three methods recommended six, seven, and six or seven clusters, respectively

(Figures S1-S3 in the Additional File 1). We finally chose a six-cluster solution based on the associated cluster statistics and the interpretability of the results.

## 2.4 Covariates

Previous studies suggested that sit-down restaurants tend to be located in high density neighborhoods because of walkability and the cozy atmosphere offered by urban environments [35,36]. Owners of sit-down restaurants and supermarkets as basic amenities may be disinclined to locate in Black or poor neighborhoods [36–40] because of uncertainty in investing in such neighborhoods. In addition, highly restrictive land use such as single-family housing may limit the introduction of sit-down restaurants locating near restrictive land use because sit-down restaurants may attract traffic, generates noise, and promote unlawful behavior [41,42]. On the basis of such reports, we incorporated the four variables, residential/employment population density, median household income, percent of white race, and percent of single-family housing as covariates into the models. For those four covariates, to represent the changes in neighborhood characteristics during that period, we added to our models four time-varying variables, which were the changes in residential/employment population density, median household income, percent of white race, and percent of single-family housing from 1990. For example, we calculated the change in employment population density in 1990 as zero. We then calculated the change in employment population density in 2000 as the employment population density in 2000 minus the employment population density in 1990. We then calculated the change in employment population density in 2006-2009 as the employment population in 2006-2009 minus the employment population density in 1990. We calculated changes in residential population density, median household income, percent of white race, and percent of single-family housing by the same method that we used to calculate changes in employment population density. We used the changes in residential population density and employment population density in the sit-down restaurant and supermarket models, respectively. Adding such change variables was necessary because we measured neighborhood type only for 1993, which could not be used to explain the change in percent of sit-down restaurants and supermarkets between 1993 and 2011.

Sit-down restaurant purveyors may prefer to locate their restaurants in neighborhoods that already have a large number of restaurants to draw customers who may seek variety [43,44]. Therefore, we added the total number of sit-down restaurants and fast food restaurants as one of covariates in the sit-down restaurant model. However, supermarket purveyors may not prefer to locate in neighborhoods that already have a large number of different types of food stores because competition may reduce the likelihood of customers who tend to prefer to shop at a specific outlet [45]. Therefore, we added the total number of supermarkets, grocery stores and convenience stores as one of the covariates in the supermarket model.

## 2.5 Statistical analyses

All descriptive analyses and multivariable models were performed using Stata 14.0 (StataCorp, College Station, TX).

### *2.5.1 Descriptive statistics*

We calculated means and standard deviations (for continuous variables) of neighborhood built environment characteristics, neighborhood sociodemographic characteristics, and the relative availability of sit-down restaurants and supermarkets in the neighborhood in 1990/1993, 2001 and 2011. We used one-tailed Student's t-test and Kruskal-Wallis H test to test for statistically significant differences in means and medians for continuous variables.

### *2.5.2 Relationship between neighborhood type and relative availability of sit-down restaurants and supermarkets*

We used multivariable linear mixed effects regression models to estimate the associations between neighborhood type in 1993 and the percent of sit-down restaurants and percent of supermarkets in 1993, 2001, and 2011 (n=2,083). These models appropriately accounted for the clustered data structure of repeated measurements over time within each neighborhood. Specifically, one neighborhood in 1993 had many similarities compared to the same neighborhood in 2001 and 2011, which may have violated the principal of independently and identically distributed observations. To address the "repeated-measurement" feature of the data, we implemented mixed effects regression models for the percent of sit-down restaurants and percent of supermarkets. We modeled the percent of sit-down restaurants/supermarkets in each neighborhood as a function of neighborhood type in 1993, the time elapsed in years from 1993, the term for the interaction of neighborhood type in 1993 with elapsed time, and the time-varying covariates, which we denoted as baseline-change models [46]. We performed baseline-change analysis to assess how neighborhood characteristics (as measured by neighborhood type) at the baseline year modified the effect of time on the relative availability of sit-down restaurants and supermarket. If "neighborhood type at the baseline year" failed to modify the effect of time on the relative availability of sit-down restaurants and supermarkets, then the increase rates in the relative availability of sit-down restaurants and supermarkets should be the same across the baseline-year neighborhood type. Although the results of baseline-change models explicitly disclosed which types of neighborhoods experienced greater increases in the relative availability of sit-down restaurants and supermarkets, we did not stop at that point. Instead, we next employed post-estimated linear contrasts based on the results of same models, which enabled us to compare the relative availability of sit-down restaurants and supermarkets across neighborhood type in each observational year. Approximately 68% of neighborhoods did not change neighborhood type between 1993 and 2011, and the models failed to converge after we incorporated the variable of change in neighborhood type over time. Thus, we added neighborhood-level time-varying variables into the model to address the issue that the neighborhood type at the baseline year failed to account for the change in the relative availability of sit-down restaurants and supermarkets over time. We included random intercepts for each neighborhood in the sit-down restaurant and supermarket models to enable responses to vary within neighborhoods. Because a census block group is a small area in dense areas, we tested whether our results were sensitive with respect to different measures of relative availability of sit-down restaurants and supermarkets based on Census Tract as well as Census Place (i.e., city or town).

# 3 Results

## 3.1 Descriptive statistics

Compared with 1993, in 2011, the percent of sit-down restaurants and supermarkets in the study area increased 10.1 and 3.3 percentage points, respectively (Table 1). Our study area's population in 2011 (compared with 1993) tended to be older (45-64 or 65 or above), more non-white, more college educated or higher, and having higher household incomes. The study area had a greater population density, greater mix of land use, and greater percent of single-family housing in 2011 compared with 1993.

Table 1 Selected characteristics of neighborhoods in years 1993, 2001 and 2011, Twin Cities Region

Neighborhood characteristic	1993 <sup>a</sup>	2001	2011	Change <sup>b</sup>	P value <sup>c</sup>
Number of observations (neighborhoods)	2,083	2,083	2,083	---	---
Relative availability of sit-down restaurants and supermarkets					
Percent of sit-down restaurants <sup>d</sup> , mean (SD)	16.1 ± 33.1	22.7 ± 36.4	26.2 ± 36.8	10.1 ± 41.4	<0.05
Percent of supermarkets <sup>e</sup> , mean (SD)	2.0 ± 11.9	2.4 ± 12.9	5.3 ± 19.6	3.3 ± 20.1	<0.05
Built environment characteristics					
Residential population density, 1,000 person/km <sup>2</sup> , median (IQR)	1.2 (1.6)	1.2 (1.3)	1.3 (1.3)	0.0 (0.5)	<0.05
Employment population density, 1,000 person/km <sup>2</sup> , median (IQR)	0.6 (0.8)	0.7 (0.7)	0.7 (0.7)	0.0 (0.1)	<0.05
Mix of land use <sup>f</sup> , median (IQR)	45.0 (48.0)	51.3 (49.7)	57.0 (46.0)	8.7 (20.4)	<0.05
Percent of single-family housing <sup>g</sup> , median (IQR)	68.5 (65.4)	73.2 (64.9)	94.9 (36.8)	16.3 (27.0)	<0.05
Total sit-down restaurants and fast food restaurants, median (IQR)	0.0 (1.0)	0.0 (2.0)	1.0 (2.0)	1.3 (3.2)	<0.05
Total supermarkets, grocery stores and convenience stores, median (IQR)	0.0 (1.0)	1.0 (1.0)	1.0 (2.0)	0.2 (1.2)	<0.05
Sociodemographic characteristics					
Age, mean (SD)					
Percent of population under 14	22.2 ± 6.4	21.3 ± 6.8	19.9 ± 6.0	-2.3 ± 4.8	<0.05
Percent of population 15–29	23.7 ± 7.3	20.9 ± 7.8	21.6 ± 8.4	-2.1 ± 4.6	<0.05
Percent of population 30–44	27.2 ± 4.7	26.0 ± 4.8	21.9 ± 4.5	-5.3 ± 5.0	<0.05
Percent of population 45–64	16.9 ± 5.0	21.2 ± 5.4	25.8 ± 5.7	8.9 ± 5.8	<0.05
Percent of population 65 or above	9.9 ± 6.6	10.5 ± 7.3	10.8 ± 6.0	0.9 ± 5.7	<0.05
Percent of population with education level of college or above, mean (SD)	57.6 ± 15.2	66.1 ± 14.9	68.1 ± 14.3	10.5 ± 8.6	<0.05
Race, median (IQR)					
Percent of white race	96.0 (5.0)	90.6 (13.0)	87.0 (16.0)	-10.8 (11.6)	<0.05
Percent of black race	1.0 (3.0)	2.9 (5.1)	4.0 (9.0)	4.1 (7.2)	<0.05
Median household income <sup>h</sup> , \$1,000, mean (SD)	38.2 ± 12.5	40.5 ± 15.1	37.1 ± 14.5	0.4 ± 7.5	<0.05
Time elapsed from 1993, year, mean (SD)	0 ± 0	8 ± 0	18 ± 0	18 ± 0	---

IQR, interquartile range; SD, standard deviation.

<sup>a</sup> Because we did not have neighborhood built environment and sociodemographic data for 1993, we assumed that data for 1990 would be valid substitutes for the missing 1993 data

<sup>b</sup> Change in neighborhood characteristics from year 1993 to 2011.

<sup>c</sup> P value for one-tailed Student's t-test of difference in means and Kruskal-Wallis H test of difference in medians from years 1993 and 2011

<sup>d</sup> Percent of sit-down restaurants relative to total sit-down restaurants and fast food restaurants.

<sup>e</sup> Percent of supermarkets relative to total supermarkets, grocery stores and convenience stores.

<sup>f</sup> The mix of land use was measured by 3-tier land use entropy (denominator set to the static 3 land use type in the census block group), which used three land use categories (residential, employment and retail) to calculate mix of land use in the block group.

<sup>g</sup> Percent of single-family housing relative to total single-family and multi-family housings.

<sup>h</sup> The median household incomes in 1993 and 2001 were adjusted for inflation to compare with that in 2011.

### **3.2 Results from cluster analyses: neighborhood type (Year 1993)**

The six robust neighborhood types that we defined by the final cluster solution represented non-overlapping groupings of Twin Cities Region neighborhoods based on built environment and sociodemographic attributes in 1990 (the first observational year). These clusters included: cluster 1 - high-density urban core; cluster 2 - low-income, non-white inner city; cluster 3 - urban; cluster 4 - aging suburb; cluster 5 - high-income suburb; and cluster 6 - suburban edge.

We labeled clusters based on their most prominent built environment and sociodemographic characteristics in 1993 (See Table S2 in the Additional File 1). Compared with most of the other clusters, cluster 1, "high-density urban core", had relatively greater levels of residential and employment population densities, a greater mix of land use, comparatively lower percent single-family housing, comparatively higher percent population aged 15-29, and comparatively lower percent population aged under 14. Cluster 2, "low-income, non-white inner city", had moderate-to-high residential and employment population densities and comparatively higher percent non-white race population, relatively lower level of median household income and comparatively lower percent population with a college education or above. Cluster 5 and Cluster 6, "high-income suburb" and "suburban edge", had relatively lower levels of residential and employment population densities, lower degrees of mix of land use, and relatively greater levels of median household income. Cluster 3 ("urban") and Cluster 4 ("aging suburb") had moderate levels of almost all neighborhood features, except for a greater degree of mix of land use and comparatively higher percent population aged 65 and above.

lower levels of residential and employment population densities, lower degrees of mix of land use, and relatively greater levels of median household income. Cluster 3 ("urban") and Cluster 4 ("aging suburb") had moderate levels of almost all neighborhood features, except for a greater degree of mix of land use and comparatively higher percent population aged 65 and above.

Figure 1 shows that the high-density urban core (abbreviated as urban core) and low-income, non-white inner city (abbreviated as inner city) neighborhoods were tightly clustered in a small segment within the municipal boundaries of the Twin Cities. Urban and aging suburb neighborhoods comprised those

transitional areas located between the urban core or inner city neighborhoods and the suburban areas. Another small grouping of aging suburb and high-income extended into the counties of Carver and Scott and the county of Washington, respectively. The generated clusters reflected comprehensive but distinguishable physical and sociodemographic environments.

### **3.3 Relationship between neighborhood type and relative availability of sit-down restaurants and supermarkets”**

The results of multivariable linear mixed effects regression models showed that urban core ( $p=0.191$ ), urban ( $p=0.063$ ), and aging suburb ( $p=0.071$ ) did not experience a significant increase (at a 0.05 statistical level) in the percent of sit-down restaurants (Figure 2); high-income suburb ( $p=0.091$ ) did not experience a significant increase in the percent of supermarkets (Figure 3) between 1993 and 2011. The coefficients of -1.15, -0.53, and -0.58 in the same sit-down restaurant model showed that inner city neighborhoods experienced a greater increase in the percent of sit-down restaurants compared with urban core, urban, and aging suburb between 1993 and 2011 (Table S5 in the Additional File 1).

The results in Tables 2-4 were derived from the same models as for Figures 2-3 and Tables S5-S6 in the Additional File 1. Tables 2-4 shows the post-estimated linear contrasts of percent of sit-down restaurants and percent of supermarkets in the neighborhood by year and for each neighborhood type pair from the multivariable linear mixed effects regression models. Urban core neighborhoods had a higher percent of sit-down restaurants (by 26.61-35.48 percentage points) compared with the other five types of neighborhoods in 1993 (Table 2). The coefficient of -27.13 in 1993 in the sit-down restaurant model (Table 2) suggested that the percent of sit-down restaurants in inner city neighborhoods was 27.13 percent lower than that of urban core neighborhoods in 1993. And the confidence interval of -28.62 and -25.63 indicated that we had 95% confidence that the actual difference in the percent of sit-down restaurants between inner city and urban core neighborhoods fell between 25.63 and 28.62 in 1993. We did not observe any differences in the percent of supermarkets between urban core and other types of neighborhoods for 1993 (Table 2). For 2001, we observed more differences in percent of sit-down restaurants and supermarkets by neighborhood type. Inner city neighborhoods had a higher percent of sit-down restaurants (by 5.38 percentage points) than did urban neighborhoods; aging suburb neighborhoods had slightly more supermarkets (1.59-1.78 percentage points) compared with the urban and suburban edge neighborhoods (Table 3). In 2011, the differences in the percent of sit-down restaurants between urban core and other neighborhoods decreased to between 8.52 and 23.57, whereas the differences in the percent of sit-down restaurants between inner city and other neighborhoods (aging suburb, high-income suburb and suburban edge neighborhoods) increased to between 8.7 and 15.05 (Table 4), compared to the difference in the percent of sit-down restaurants between 1.69 and 8.36 in 1993 and between 5.38 and 11.72 in 2001. We performed correlation analysis and did not find high correlation between covariates in the mixed effects models.

Table 2 Contrast<sup>a</sup> of percent of sit-down restaurants<sup>b</sup> and supermarkets<sup>c</sup> for each neighborhood type<sup>d</sup> pair in 1993

Sit-down restaurant model: Estimated beta (95% confidence interval)						
	Urban core	Inner city	Urban	Aging suburb	High-income suburb	Suburban edge
Urban core	--	--	--	--	--	--
Inner city	<b>-27.13</b> <b>(-28.62, -25.63)</b>	--	--	--	--	--
Urban	<b>-26.61</b> <b>(-29.45, -23.77)</b>	0.51 (-1.71, 2.74)	--	--	--	--
Aging suburb	<b>-28.82</b> <b>(-30.25, -27.39)</b>	<b>-1.69</b> <b>(-2.82, -0.57)</b>	<b>-2.21</b> <b>(-3.29, -1.11)</b>	--	--	--
High-income suburb	<b>-31.69</b> <b>(-32.90, -30.48)</b>	<b>-4.57</b> <b>(-5.52, -3.61)</b>	<b>-5.08</b> <b>(-6.37, -3.78)</b>	<b>-2.87</b> <b>(-3.56, -2.18)</b>	--	--
Suburban edge	<b>-35.48</b> <b>(-36.50, -34.47)</b>	<b>-8.36</b> <b>(-9.15, -7.56)</b>	<b>-8.87</b> <b>(-9.81, -7.93)</b>	<b>-6.67</b> <b>(-7.18, -6.14)</b>	<b>-3.79</b> <b>(-4.31, -3.27)</b>	--
Supermarket model: Estimated beta (95% confidence interval)						
	Urban core	Inner city	Urban	Aging suburb	High-income suburb	Suburban edge
Urban core	--	--	--	--	--	--
Inner city	-1.87 (-6.24, 2.50)	--	--	--	--	--
Urban	-1.41 (-5.33, 2.51)	0.46 (-2.67, 3.59)	--	--	--	--
Aging suburb	-0.18 (-4.11, 3.75)	1.69 (-1.39, 4.77)	1.23 (-0.66, 3.12)	--	--	--
High-income suburb	-0.58 (-4.78, 3.63)	1.30 (-2.07, 4.66)	0.84 (-1.38, 3.06)	-0.39 (-2.30, 1.51)	--	--
Suburban edge	-1.68 (-5.83, 2.46)	0.19 (-3.08, 3.46)	-0.27 (-2.29, 1.75)	-1.50 (-3.13, 0.13)	-1.11 (-2.94, 0.73)	--

Bold indicates significant difference in percent of sit-down restaurants or percent of supermarkets across neighborhood type at the 0.05 significance level.

<sup>a</sup> Multivariable linear mixed effects regressions modeling the percent of sit-down restaurants relative to total sit-down restaurants and fast food restaurants and percent of supermarkets relative to total supermarkets, grocery stores and convenience stores as functions of neighborhood type in 1993, time elapsed since 1993, interaction between neighborhood type in 1993 and time elapsed, change in employment population density (sit-down restaurant model only), change in residential population density (supermarket model only), median household income, percent of white race and percent of single-family housing since 1993, total sit-down restaurants and fast food restaurants (sit-down restaurant model only), and total supermarkets, grocery stores and convenience stores (supermarket model only) and a random intercept for each neighborhood.

<sup>b</sup> Percent of sit-down restaurants relative to total sit-down restaurants and fast food restaurants in the neighborhood.

<sup>c</sup> Percent of supermarkets relative to total supermarkets, grocery stores and convenience stores in the neighborhood.

<sup>d</sup> Derived from cluster analysis of block-group level data from 1993: percent of population aged under 14, aged 15-29, 30-44, 45-64, and aged above 65, percent of education of college or above, percent of white race, percent of black race, median household income, residential population density, employment population density, mix of land use and percent of single-family housing.

Table 3 Contrast<sup>a</sup> of percent of sit-down restaurants<sup>b</sup> and supermarkets<sup>c</sup> for each neighborhood type<sup>d</sup> pair in 2001

Sit-down restaurant model: Estimated beta (95% confidence interval)						
	Urban core	Inner city	Urban	Aging suburb	High-income suburb	Suburban edge
Urban core	--	--	--	--	--	--
Inner city	<b>-16.66</b> (-18.97, -14.35)	--	--	--	--	--
Urban	<b>-22.04</b> (-25.49, -18.60)	<b>-5.38</b> (-8.05, -2.71)	--	--	--	--
Aging suburb	<b>-24.45</b> (-26.35, -22.55)	<b>-7.79</b> (-9.25, -6.33)	<b>-2.41</b> (-3.74, -1.08)	--	--	--
High-income suburb	<b>-25.10</b> (-26.96, -23.25)	<b>-8.44</b> (-9.85, -7.03)	<b>-3.06</b> (-4.66, -1.46)	-0.65 (-1.58, 0.28)	--	--
Suburban edge	<b>-28.38</b> (-30.00, -26.75)	<b>-11.72</b> (-12.97, -10.46)	<b>-6.33</b> (-7.53, -5.14)	<b>-3.92</b> (-4.65, -3.20)	<b>-3.27</b> (-4.08, -2.46)	--
Supermarket model: Estimated beta (95% confidence interval)						
	Urban core	Inner city	Urban	Aging suburb	High-income suburb	Suburban edge
Urban core	--	--	--	--	--	--
Inner city	-2.02 (-5.51, 1.48)	--	--	--	--	--
Urban	-2.09 (-5.34, 1.15)	-0.78 (-2.58, 2.42)	--	--	--	--
Aging suburb	-0.32 (-3.66, 3.03)	1.70 (-0.82, 4.22)	<b>1.78</b> (0.26, 3.29)	--	--	--
High-income suburb	-1.91 (-5.52, 1.70)	0.11 (-2.66, 2.88)	0.19 (-1.63, 2.00)	0.25 (-1.40, 1.91)	--	--
Suburban edge	-1.84 (-5.41, 1.74)	0.18 (-2.52, 2.88)	0.19 (-1.63, 2.00)	<b>-1.59</b> (-3.13, -0.05)	<b>-1.52</b> (-2.84, -0.21)	--

Refer to the legends in Table 2.

Table 4 Contrast<sup>a</sup> of percent of sit-down restaurants<sup>b</sup> and supermarkets<sup>c</sup> for each neighborhood type<sup>d</sup> pair in 2011

Sit-down restaurant model: Estimated beta (95% confidence interval)						
	Urban core	Inner city	Urban	Aging suburb	High-income suburb	Suburban edge
Urban core	--	--	--	--	--	--
Inner city	<b>-8.52</b> (-11.99, -5.05)	--	--	--	--	--
Urban	<b>-17.22</b> (-23.24, -11.21)	<b>-8.70</b> (-13.35, -4.06)	--	--	--	--
Aging suburb	<b>-22.16</b> (-24.92, -19.40)	<b>-13.64</b> (-15.73, -11.55)	<b>-4.94</b> (-7.14, -2.73)	--	--	--
High-income suburb	<b>-20.52</b> (-23.28, -17.75)	<b>-12.00</b> (-14.03, -9.97)	-3.29 (-6.02, 0.57)	<b>1.64</b> (0.31, 2.98)	--	--
Suburban edge	<b>-23.57</b> (-26.56, -20.58)	<b>-15.05</b> (-17.33, -12.77)	<b>-6.34</b> (-8.50, -4.19)	<b>-1.41</b> (-2.61, -0.21)	<b>-3.05</b> (-4.46, -1.64)	--
Supermarket model: Estimated beta (95% confidence interval)						
	Urban core	Inner city	Urban	Aging suburb	High-income suburb	Suburban edge
Urban core	--	--	--	--	--	--
Inner city	-2.19 (-6.76, 2.37)	--	--	--	--	--
Urban	-2.95 (-7.13, 1.24)	-0.75 (-3.98, 2.48)	--	--	--	--
Aging suburb	-0.48 (-4.72, 3.76)	1.71 (-1.47, 4.90)	<b>2.46</b> (0.50, 4.43)	--	--	--
High-income suburb	-3.57 (-8.04, 0.89)	-1.38 (-4.80, 2.04)	-0.63 (-2.98, 1.72)	<b>-3.09</b> (-5.14, -1.05)	--	--
Suburban edge	-2.03 (-6.39, 2.32)	0.16 (-3.10, 3.42)	0.91 (-1.16, 2.98)	-1.55 (-3.26, 0.15)	1.54 (-0.38, 3.46)	--

Refer to the legends in Table 2.

### 3.4 Sensitivity Testing

Tables S5-S6 in the Additional File 1 contain regression results using the census tract and place to measure food availability. Tract models generated similar results to the major results based on Block Group. But Place models showed inconsistencies, particularly for the sit-down restaurant model. Urban and aging suburb neighborhoods experienced lower increases in the percent of sit-down restaurants than inner city in the Block Group and Tract models, whereas we failed to observe such a difference in the Place model. Similarly, urban core had higher percent of sit-down restaurants than inner city in 1993 in the Block Group and Tract models, but the Place model did not show such a difference. The differences between the Block Group, Tract and Place models suggested that the measure of relative availability was sensitive to spatial unit. Because the size of a census block group was not always small (varying from 0.04 in the urban core to 154.19 km<sup>2</sup> in the suburban edge with median and interquartile range values of 0.88 and 1.63) and actually increased with the distance to urban core increases, we argued that our measure of relative food availability was a reasonable small-area measure.

## 4 Discussion

The objective of this study was to assess the relationship between neighborhood characterization and neighborhood food availability based on the fact that neighborhoods are composed of interrelated sociodemographic and built environment factors, and such factors jointly affect the distribution and type of food outlets. We recognized that analyses may be confounded by correlations among neighborhood features; thus, we used cluster analysis to identify six types of neighborhoods in the Twin Cities Region of Minnesota that reflected distinct combinations of built environment with sociodemographic features. We then examined the association between neighborhood type in the baseline year and neighborhood food availability as measured by the relative competitiveness of one food outlet among its complementary competitors. Our results indicated an increasingly varied distribution of restaurants and food stores by neighborhood type over time. Our findings contribute to a growing literature on the associations between the multifaceted composition of the built environment, sociodemographic features, and the distribution of food resources. Our study demonstrated the need to use cluster analysis to disclose the complexity of assessing homogenous neighborhoods and singled out the neighborhoods that experienced different trends in the change in food availability, compared with other neighborhoods in the region. Simple neighborhood characterization by sociodemographic factors alone may mask these important complexities.

Our neighborhood types were not spatially clustered into homogeneous regions but, instead, were distributed across the Twin Cities Region. For example, the municipal boundaries of the Twin Cities did not contain only urban core and inner city neighborhoods but also included urban and aging suburbs. Similarly, aging suburbs and high-income neighborhoods extended to the boundaries of the region; thus, they were atypically closer to the city center. Therefore, our results support the work of others who noted a

recent blending of built environment and sociodemographic characteristics, resulting in reduced demarcation between the central city and its outlying suburban areas [47,48]. Because both the central cities and the outlying areas in metropolitan U.S. are becoming more diverse in form and social composition [48,49], reliance on single constructs of neighborhoods, such as population density or distance to central business district, may not adequately capture the complexity of neighborhood types.

The inner city neighborhoods had a greater relative availability of sit-down restaurants compared with other neighborhoods in 2001 and 2011; we failed to observe such great differences in 1993. This suggested that inner city neighborhoods became relatively more appealing to proprietors of sit-down restaurants and, perhaps, less appealing to owners of fast food restaurants. Specifically, inner city neighborhoods had only 1.69-8.36 percent more sit-down restaurants than aging suburb, high-income suburb and suburban edge neighborhoods in 1993; however, in 2011, these differences increased to 13.64-15.05. In addition, inner city neighborhoods had greater absolute numbers of sit-down restaurants than other neighborhoods except for urban core in 2011 when compared to that in 1993 (Figure 2 and Table S7 in the Additional File 1). These observations were noteworthy. Although inner city neighborhoods consistently had the lowest household income during the observational period (Tables S2-S4 in the Additional File 1), inner city neighborhoods had greater spatial access to sit-down restaurants than other neighborhoods in 2011. In another of our studies [24], we observed a positive association between the GIS-measured count of neighborhood sit-down restaurants and the frequency of using neighborhood sit-down restaurants in non-rural areas in four metropolitan areas including the Twin Cities Region. If we applied our previous finding to inner cities, possibly, we would have observed that local residents who resided in inner cities used sit-down restaurants more frequently than residents in other neighborhoods because of increasingly larger availability (both relative and absolute) of sit-down restaurants in the immediate neighborhood. Another possibility is that more sit-down restaurants in the inner city does not necessarily mean that inner city people eat at those venues. Inner cities increasingly serve as one part of a social hub in a region, and many people who visit inner cities and use sit-down restaurants are suburbanites or from the surrounding rural area [50]. That is, the desire for cultural amenities, entertainment and other facilities in central cities by an influx of employees and tourists led to the increase in sit-down restaurants in the inner city, as seen in evidence from New York, NY [51], Houston, TX [52], Washington, DC [53], and metropolitan areas [54]. The positive association between change in employment population density and the relative availability of sit-down restaurants (Table S5 in the Additional File 1) supported the second contemplation above (namely, more sit-down restaurants in the inner city does not necessarily mean that residents in inner city eat at those venues) to consider how the increase in “floating population” encouraged a greater availability of sit-down restaurants in the immediate neighborhood. During our study period, U.S. inner cities transitioned from goods production sectors toward relatively place-bound service sector industries [27,33], which includes restaurants [34,35]. Lester et al. [35] observed that, in twenty U.S. inner cities between 1990 and 2000, jobs in retail services replaced jobs lost in goods-producing industries. Retail- and service-dominated neighborhoods may provide a complementary environment for clustering of restaurants, food stores and other retail options [33]. Improvements in transportation [33] and landscaping may create a more spatially accessible and/or

walkable features that attract service and retail options [36,37]. During the study period, the Twin Cities experienced improvements in light rail, the park system, and new sports stadiums [38]. Future research should evaluate who actually used the sit-down restaurants in the inner city and, in turn, drove the increase in sit-down restaurants therein. The health implication of this study is a focus on the neighborhoods with high proportions of sit-down restaurants, such as inner cities and urban core neighborhoods. Mezuk et al. [55] found that neighborhoods with high proportions of health-harming stores were associated with higher prevalence of Type II diabetes. Although sit-down restaurants are not necessarily health-harming outlets, many investigators [21–23] have found that sit-down restaurants are no more proficient in selling healthy foods than are fast food restaurants.

The percent of sit-down restaurants in urban core neighborhoods was stable during the observational period, which indicated the role of the urban core as the social hub of a region [50,56]. Unsurprisingly, urban core did not experience an increase in the percent of sit-down restaurants. This lack of increase occurred because urban cores already had a high percent of sit-down restaurants in 1993 (46.9% in Figure 2). This constant percent implies a “saturated” urban core with respect to the relative availability of sit-down restaurants. The unchanged relative availability of sit-down restaurants paralleled an increase in the well-educated population that predominately lived in urban core neighborhoods. This finding was similar to observations for Houston, TX [52], which also showed an increase in the well-educated population in urban core neighborhoods.

The Twin Cities Region experienced multiple different economic conditions during the period of our study: economic expansion (1993-2007), economic recession (2007-2009), and economic recovery (2009-2011) [57]. Nevertheless, the average percent of sit-down restaurants increased in the Region from 16 percent in 1993 to 23 percent in 2001 to 26 percent in 2011. We had assumed that the percent of sit-down restaurants would stay intact or even decrease during economic recession (2007-2009) because demand for meals from sit-down restaurants was price-elastic [58]. Our results seemed to contradict that assumption. The increase in the percent of sit-down restaurants in the Twin Cities Region, particularly from 2006 to 2011, was perhaps greater than that of the nation at large. We have this hint because Richardson’s study [59] implied a relatively stable percent of sit-down restaurants between 1985 and 2006 for four regions in the U. S. (see Table 5 in her work), Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA. Our results probably corresponded to the argument that convenience, sociodemographic characteristics, and macroeconomic forces such as the business cycle, instead of only relative prices and income, determined why we still saw a significant increase in the percent of sit-down restaurants instead of fast food restaurants in most Twin Cities neighborhoods [60,61].

We also found a more varied distribution of food stores across neighborhoods in 2001 and 2011 that we did not see for 1993. Specifically, aging suburb neighborhoods had a greater percent of supermarkets (i.e., fewer percent of grocery stores and convenience stores) than did the urban and high-income suburb neighborhoods in 2001 and 2011, but not in 1993. Such differences were driven largely by the great increase in the number of grocery stores and convenience stores in the high-income and suburban edge neighborhoods in comparison with increases in numbers of aging suburb supermarkets. The higher

percent of grocery and convenience stores in urban and high-income neighborhoods may compound barriers to accessing healthful foods, if such foods are less available in grocery and convenience stores [62]. Small food stores offer an abundance of less nutritious foods such as sugar-sweetened beverages, salty snacks, and candy, and the prices of fresh foods are more expensive than in supermarkets in U.S context [3–5]. Thus, it is possible that the higher percent of grocery and convenience stores in urban and high-income neighborhoods may encourage accessing less healthful foods if less nutritious foods are abundantly available in such stores [33]. We make this conjecture because two of our recent studies [63,64] suggested that convenience stores discouraged a majority of high-income samples from purchasing healthy foods and encouraged such samples from purchasing snacks.

In the supermarket model, however, we found that an increased percent of supermarkets was associated with a smaller increase (or more rarely a decrease) in the percent of single-family housing units (See Table S6 in the Additional File 1). These largely incompatible land uses—single-family housing and supermarkets – may have opened opportunities for urban planners to use regulatory tools (e.g., zoning) to introduce targeted food stores into the neighborhoods. These regulatory tools could side-step concerns/requirements such as intrusive light [65], sufficient parking [41], or increased traffic, thereby avoiding resistance to introducing a supermarket into neighborhoods with large numbers of single-family housing.

Researchers are increasingly using complex data-rich methods to define and distinguish neighborhoods. Our analysis mingled with this trend because we chose sociodemographic and built environment variables that potentially affect the distribution of food outlets to define neighborhoods. The representative attributes varied by region and research purpose; thus, researchers should conceptualize such factors according to local characteristics. Because of the wide application of Geographic Information System techniques, it is feasible for researchers to generate spatial data and combine abundant location data into cluster analysis. Therefore, although we examined only one large metropolitan (geographical) region, our method to assess associations between this complex group of neighborhood characteristics and food availability is generalizable. We suggest cluster analysis to characterize neighborhoods is particularly useful because metropolitan areas become increasingly complex as urbanization continues [13]. For example, for neighborhoods in which street patterns are dominated by large blocks and residents live a suburban lifestyle, some such neighborhoods (A) will of course be located in core cities, but (B) more are likely to be in suburbs; cluster analysis will likely be able to identify the homogenous neighborhoods that span an extensive region (as long as other features are similar). If we used simple approaches to classify areas, such as distance to the census tract that contains the city hall, they are likely to treat A and B as two different types of neighborhoods. That is, cluster analysis could be powerful in characterizing neighborhoods for which it is increasingly not possible to use a single aspect to define, although a single dimension could be easy to understand by non-experts. Because cluster analysis classifies neighborhoods based on predefined food-patterning-related features, the classification is reasonable provided that a good theory supports the predefined features. Cluster analysis provided clear research variables to classify neighborhoods; thus, by adding or

removing constructs, the process is easy to rerun and readjust if the classification is unreasonable for research purposes.

Our study has several caveats. First, the Twin Cities Region was notably more affordable for housing and transportation and offered more diverse housing choices compared with similar metropolitan areas [17]. Those features may have fostered more convenient access to restaurants and small food stores. Second, the multidimensional class structure we identified by our data-driven approach is difficult to compare with class structure based on single features that other researchers have used. However, because of a lack of consistent association between individual neighborhood characteristics and specific food resource types [66], we elected to use our data-driven approach to characterize the neighborhood environment. Third, the marked undercount of food outlets in the D&B data may have introduced bias [67]. Fourth, the block group was probably too small to reflect the service area of restaurants and food stores, especially in suburban areas. However, census block group level data yield better estimates of the locations of food resources and households [68], compared with data from larger geographic units such as census tracts and zip codes. In addition, we could not obtain some retrospective built environment and sociodemographic data, such as traffic and crime, for the whole region, which have been suggested as relevant factors [69,70]. We interpolated the sociodemographic factors at the block group level from the tract level for year 1993, the results of which may be inaccurate in rural areas and highly developed urban cores [71]. In addition, our models failed to converge when we included a change variable termed “the change in neighborhood type.” We expect to explicitly incorporate the change in neighborhood type to predict the change in neighborhood food availability when we have data for a greater number of observational years.

## 5 Conclusion

We used cluster analysis to characterize food-related urban environments in the Twin Cities Region and examined the relationships between neighborhood type and relative availability of sit-down restaurants and supermarkets. We observed a complex and increasingly varied distribution of restaurants and food stores across six types of neighborhoods with distinctive built environment and sociodemographic characteristics, particularly for inner cities, during an 18-year time span. The composite index generated by cluster analysis and the associated food retailing landscaping provide an analytical tool to support public health policy in monitoring the neighborhoods that experienced great change in food availability. Our results echoed the national trend that the U.S. inner cities have undergone substantial changes in sociodemographic and built environment characteristics, which may impact the types and distribution of restaurants therein. The great change in food outlet type in inner cities may have health implications for people who reside or work in such neighborhoods and who rely on such restaurants.

## List Of Abbreviations

Twin Cities Region of Minnesota    Twin Cities Region

# Declarations

## Ethics approval and consent to participate

Not applicable

## Consent for publication

Not applicable

## Availability of data and material

The data that support the findings of this study are available from Carolina Population Center. Restrictions apply to availability because the data were licensed for the current study, and, therefore, not publicly available.

## Competing interests

The authors have no financial or other conflicts of interest to disclose.

## Funding

This work was funded by the National Heart, Lung, and Blood Institute (NHLBI) R01HL104580. The Coronary Artery Risk Development in Young Adults Study (CARDIA) is supported by contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN26820090-0041C from the NHLBI, the Intramural Research Program of the National Institute on Aging (NIA), and an intra-agency agreement between NIA and NHLBI (AG0005). The authors are grateful to the Carolina Population Center, University of North Carolina at Chapel Hill, for general support (Grant R24 HD050924 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development [NICHD]), 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 to the Center for Environmental Health Sciences (CEHS), University of North Carolina (Grant P30ES010126 from the National Institute for Environmental Health Sciences [NIEHS]). NIH had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data, and preparation, review, or approval of the manuscript.

## Authors' contributions

Ke Peng, PhD had full access to all the data in the study and takes responsibility for its integrity and the accuracy of data analysis. Study concept and design: Peng, Rodriguez, Gordon-Larsen. Acquisition of data: Peng, Gordon-Larsen. Analysis and interpretation of data: Peng, Rodriguez, Hirsch, Gordon-Larsen. Drafting of the manuscript: Peng, Rodriguez, Hirsch, Gordon-Larsen. Critical revision of the manuscript for important intellectual content: Rodriguez, Gordon-Larsen. Statistical analysis: Peng, Hirsch, Rodriguez, Gordon-Larsen. Obtained funding: Gordon-Larsen. Final approval of the version to be published:

Rodriguez, Gordon-Larsen. Study supervision: Rodriguez, Gordon-Larsen. All authors have read and approved the manuscript.

## Acknowledgements

The authors would like to acknowledge Marc Peterson, of the University of North Carolina, Carolina Population Center (CPC) for geocoding the food resource data

## Authors' information (optional)

Not applicable

## References

1. Lytle LA. Measuring the Food Environment. State of the Science. *Am J Prev Med.* 2009;36. doi:10.1016/j.amepre.2009.01.018
2. Bilal U, Jones-Smith J, Diez J, Lawrence RS, Celentano DD, Franco M. Neighborhood social and economic change and retail food environment change in Madrid (Spain): The heart healthy hoods study. *Health Place.* 2018. doi:10.1016/j.healthplace.2018.03.004
3. 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: 876–884. doi:10.1016/j.healthplace.2010.04.013
4. Powell LM, Chaloupka FJ, Bao Y. The Availability of Fast-Food and Full-Service Restaurants in the United States. Associations with Neighborhood Characteristics. *Am J Prev Med.* 2007;33. doi:10.1016/j.amepre.2007.07.005
5. Zenk SN, Powell LM. US secondary schools and food outlets. *Health Place.* 2008;14: 336–346. doi:10.1016/j.healthplace.2007.08.003
6. Lamichhane AP, Warren J, Puett R, Porter DE, Bottai M, Mayer-Davis EJ, et al. Spatial patterning of supermarkets and fast food outlets with respect to neighborhood characteristics. *Health Place.* 2013;23: 157–164. doi:10.1016/j.healthplace.2013.07.002
7. James P, Arcaya MC, Parker DM, Tucker-Seeley RD, Subramanian S V. Do minority and poor neighborhoods have higher access to fast-food restaurants in the United States? *Health Place.* 2014. doi:10.1016/j.healthplace.2014.04.011
8. Caspi CE, Sorensen G, Subramanian S V, Kawachi I. The local food environment and diet: a systematic review. *Health Place.* 2012;18: 1172–87. doi:10.1016/j.healthplace.2012.05.006
9. Bivoltsis A, Trapp G, Knuiman M, Hooper P, Ambrosini GL. The evolution of local food environments within established neighbourhoods and new developments in Perth, Western Australia. *Health Place.* 2019. doi:10.1016/j.healthplace.2019.04.011
10. Apparicio P, Cloutier MS, Shearmur R. The case of Montréal's missing food deserts: Evaluation of accessibility to food supermarkets. *Int J Health Geogr.* 2007. doi:10.1186/1476-072X-6-4

11. Helling A, Sawicki DS. Race and residential accessibility to shopping and services. *Hous Policy Debate*. 2003;14: 69–101. doi:10.1080/10511482.2003.9521469
12. Luan H, Minaker LM, Law J. Do marginalized neighbourhoods have less healthy retail food environments? An analysis using Bayesian spatial latent factor and hurdle models. *Int J Health Geogr*. 2016. doi:10.1186/s12942-016-0060-x
13. Forsyth A. Defining Suburbs. *J Plan Lit*. 2012. doi:10.1177/0885412212448101
14. Jones M, Huh J. Toward a multidimensional understanding of residential neighborhood: A latent profile analysis of Los Angeles neighborhoods and longitudinal adult excess weight. *Health Place*. 2014. doi:10.1016/j.healthplace.2014.01.011
15. Nelson MC, Gordon-Larsen P, Song Y, Popkin BM. Built and Social Environments. Associations with Adolescent Overweight and Activity. *Am J Prev Med*. 2006;31: 109–117. doi:10.1016/j.amepre.2006.03.026
16. Meyer KA, Boone-Heinonen J, Duffey KJ, Rodriguez DA, Kiefe CI, Lewis CE, et al. Combined measure of neighborhood food and physical activity environments and weight-related outcomes: The CARDIA study. *Health Place*. 2015;33: 9–18. doi:10.1016/j.healthplace.2015.01.004
17. Minneapolis Metropolitan Council. Thrive MSP 2040 Plan: housing policy plan. St. Paul, MN: Metropolitan Council; 2015. Available: <https://metrocouncil.org/Planning/Projects/Thrive-2040.aspx>
18. Clary CM, Ramos Y, Shareck M, Kestens Y. Should we use absolute or relative measures when assessing foodscape exposure in relation to fruit and vegetable intake? Evidence from a wide-scale Canadian study. *Prev Med (Baltim)*. 2015. doi:10.1016/j.ypmed.2014.11.023
19. Mercille G, Richard L, Gauvin L, Kestens Y, Shatenstein B, Daniel M, et al. Associations between residential food environment and dietary patterns in urban-dwelling older adults: Results from the VoisiNuAge study. *Public Health Nutr*. 2012. doi:10.1017/S136898001200273X
20. Rummo PE, Meyer KA, Boone-Heinonen J, Jacobs DRJ, Kiefe CI, Lewis CE, et al. Neighborhood availability of convenience stores and diet quality: findings from 20 years of follow-up in the coronary artery risk development in young adults study. *Am J Public Health*. 2015;105: e65–73. doi:10.2105/AJPH.2014.302435
21. Saelens BE, Glanz K, Sallis JF, Frank LD. Nutrition Environment Measures Study in Restaurants (NEMS-R). Development and Evaluation. *Am J Prev Med*. 2007. doi:10.1016/j.amepre.2006.12.022
22. Bruemmer B, Krieger J, Saelens BE, Chan N. Energy, Saturated Fat, and Sodium Were Lower in Entrées at Chain Restaurants at 18 Months Compared with 6 Months Following the Implementation of Mandatory Menu Labeling Regulation in King County, Washington. *J Acad Nutr Diet*. 2012. doi:10.1016/j.jand.2012.04.019
23. Pereira RF, Sidebottom AC, Boucher JL, Lindberg R, Werner R. Peer Reviewed: Assessing the Food Environment of a Rural Community: Baseline Findings From the Heart of New Ulm Project, Minnesota, 2010–2011. *Prev Chronic Dis*. 2014;11.
24. Peng K, Rodríguez DA, Peterson M, Braun LM, Howard AG, Lewis CE, et al. GIS-Based Home Neighborhood Food Outlet Counts, Street Connectivity, and Frequency of Use of Neighborhood

- Restaurants and Food Stores. *J J Urban Heal*. 2020;97: 213–225.
25. Powell LM, Han E, Zenk SN, Khan T, Quinn CM, Gibbs KP, et al. Field validation of secondary commercial data sources on the retail food outlet environment in the U.S. *Health Place*. 2011;17: 1122–1131. doi:10.1016/j.healthplace.2011.05.010
  26. Cervero R, Kockelman K. Travel demand and the 3Ds: Density, diversity, and design. *Transp Res Part D Transp Environ*. 1997. doi:10.1016/S1361-9209(97)00009-6
  27. Saelens BE, Handy SL. Built environment correlates of walking: A review. *Medicine and Science in Sports and Exercise*. 2008. doi:10.1249/MSS.0b013e31817c67a4
  28. Troped PJ, Wilson JS, Matthews CE, Cromley EK, Melly SJ. The Built Environment and Location-Based Physical Activity. *Am J Prev Med*. 2010. doi:10.1016/j.amepre.2009.12.032
  29. Lu Y, Xiao Y, Ye Y. Urban density, diversity and design: Is more always better for walking? A study from Hong Kong. *Prev Med (Baltim)*. 2017. doi:10.1016/j.ypmed.2016.08.042
  30. Bostean G, Sanchez L, Lippert AM. Sociodemographic disparities in e-cigarette retail environment: Vape stores and census tract characteristics in Orange County, CA. *Health Place*. 2018. doi:10.1016/j.healthplace.2017.12.004
  31. Guerra E, Dong X, Kondo M. Do Denser Neighborhoods Have Safer Streets? Population Density and Traffic Safety in the Philadelphia Region. *J Plan Educ Res*. 2019. doi:10.1177/0739456X19845043
  32. Ramsey K, Bell A. Smart location database Version 2.0 User Guide. Washington, DC: United States Environmental Protection Agency; 2014. Available: <https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide>
  33. Rahman AS, Rahman A %J W. Application of Principal Component Analysis and Cluster Analysis in Regional Flood Frequency Analysis: A Case Study in New South Wales, Australia. 2020;12: 781.
  34. Mohamad I Bin, Usman D. Standardization and its effects on K-means clustering algorithm. *Res J Appl Sci Eng Technol*. 2013. doi:10.19026/rjaset.6.3638
  35. Carroll GR, Torfason MT. Restaurant Organizational Forms and Community in the U.S. in 2005. *City Community*. 2011;10: 1–24. doi:10.1111/j.1540-6040.2010.01350.x
  36. Zukin S, Lindeman S, Hurson L. The omnivore’s neighborhood? Online restaurant reviews, race, and gentrification. *J Consum Cult*. 2017. doi:10.1177/1469540515611203
  37. Short A, Guthman J, Raskin S. Food Deserts, Oases, or Mirages? *J Plan Educ Res*. 2007. doi:10.1177/0739456x06297795
  38. Raja S, Ma C, Yadav P. Beyond food deserts: Measuring and mapping racial disparities in neighborhood food environments. *J Plan Educ Res*. 2008. doi:10.1177/0739456X08317461
  39. Schuetz J, Kolko J, Meltzer R. Are poor neighborhoods retail deserts? *Reg Sci Urban Econ*. 2012. doi:10.1016/j.regsciurbeco.2011.09.005
  40. Vojnovic I, Lee J, Kotval-K Z, Podagrosi A, Varnakovida P, Ledoux T, et al. The Burdens of Place: A Socio-economic and Ethnic/Racial Exploration into Urban Form, Accessibility and Travel Behaviour in the Lansing Capital Region, Michigan. *J Urban Des*. 2013. doi:10.1080/13574809.2012.683403

41. Cameron N, Amrhein CG, Smoyer-Tomic KE, Raine KD, Chong LY. Cornering the market: Restriction of retail supermarket locations. *Environ Plan C Gov Policy*. 2010. doi:10.1068/c0915
42. Irvin-Erickson Y. Identifying risky places for crime: An analysis of the criminogenic spatiotemporal influences of landscape features on street robberies. Rutgers University-Graduate School-Newark. 2014.
43. Smith D, Cummins S, Clark C, Stansfeld S. Does the local food environment around schools affect diet? Longitudinal associations in adolescents attending secondary schools in East London. *BMC Public Health*. 2013;13. doi:10.1186/1471-2458-13-70
44. Leslie TF, Frankenfeld CL, Makara MA. The spatial food environment of the DC metropolitan area: Clustering, co-location, and categorical differentiation. *Appl Geogr*. 2012. doi:10.1016/j.apgeog.2012.07.008
45. Krider RE, Putler DS. Which birds of a feather flock together? Clustering and avoidance patterns of similar retail outlets. *Geogr Anal*. 2013. doi:10.1111/gean.12005
46. Singer JD, Willett JB. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. 2009. doi:10.1093/acprof:oso/9780195152968.001.0001
47. Forsyth a. Defining Suburbs. *J Plan Lit*. 2012;27: 270–281. doi:10.1177/0885412212448101
48. Moos M, Mendez P. Suburban ways of living and the geography of income: How homeownership, single-family dwellings and automobile use define the metropolitan social space. *Urban Stud*. 2015. doi:10.1177/0042098014538679
49. Teaford JC. *The American suburb: the basics*. Routledge; 2008.
50. Bereitschaft B. Gentrification and the evolution of commuting behavior within America's urban cores, 2000–2015. *J Transp Geogr*. 2020. doi:10.1016/j.jtrangeo.2019.102559
51. Graziose M, Koch P, Gray H, Contento I. Assessing the Retail Food Environment Surrounding Elementary Schools Across New York City (NYC) Neighborhoods Varying in Their Level of Gentrification. *J Nutr Educ Behav*. 2016. doi:10.1016/j.jneb.2016.04.124
52. Podagrosi A, Vojnovic I, Pigozzi B. The diversity of gentrification in Houston's urban renaissance: From cleansing the urban poor to supergentrification. *Environ Plan A*. 2011. doi:10.1068/a43526
53. Hyra D. The back-to-the-city movement: Neighbourhood redevelopment and processes of political and cultural displacement. *Urban Stud*. 2015. doi:10.1177/0042098014539403
54. Timberlake JM, Johns-Wolfe E. Neighborhood Ethnoracial Composition and Gentrification in Chicago and New York, 1980 to 2010. *Urban Aff Rev*. 2017. doi:10.1177/1078087416636483
55. Mezuk B, Li X, Cederin K, Rice K, Sundquist J, Sundquist K. Beyond access: Characteristics of the food environment and risk of diabetes. *Am J Epidemiol*. 2016. doi:10.1093/aje/kwv318
56. Florida R. *The Creative Class and Economic Development*. *Econ Dev Q*. 2014. doi:10.1177/0891242414541693

57. Minnesota Department of Employment and Economic Development. State of the Economy. St. Paul, MN: Minnesota Department of Employment and Economic Development; 2014. Available: <https://mn.gov/deed/newscenter/publications/trends/june-2014/state-of-the-economy.jsp>
58. Richards TJ, Mancino L. Demand for food-away-from-home: A multiple-discrete-continuous extreme value model. *Eur Rev Agric Econ*. 2014. doi:10.1093/erae/jbt008
59. Richardson AS, Meyer KA, Howard AG, Boone-Heinonen J, Popkin BM, Evenson KR, et al. Multiple pathways from the neighborhood food environment to increased body mass index through dietary behaviors: A structural equation-based analysis in the CARDIA study. *Health Place*. 2015. doi:10.1016/j.healthplace.2015.09.003
60. U.S. Department of Agriculture. Food Environment Atlas. Washington, DC : USDA's Economic Research Service; 2017.
61. Koh Y, Lee S, Choi C. The income elasticity of demand and firm performance of US restaurant companies by restaurant type during recessions. *Tour Econ*. 2013. doi:10.5367/te.2013.0250
62. D'Angelo H, Suratkar S, Song H-J, Stauffer E, Gittelsohn J. Access to food source and food source use are associated with healthy and unhealthy food-purchasing behaviours among low-income African-American adults in Baltimore City. *Public Health Nutr*. 2011;14: 1632–1639. doi:10.1017/S1368980011000498
63. Peng K, Kaza N. Availability of neighbourhood supermarkets and convenience stores, broader built environment context, and the purchase of fruits and vegetables in US households. *Public Health Nutr*. 2019;22: 2436–2447. doi:10.1017/S1368980019000910
64. Peng K, Kaza N. Association between Neighborhood Food Access, Household Income, and Purchase of Snacks and Beverages in the United States. *Int J Environ Res Public Health*. 2020;17.
65. Harvey HM. Civic Engagement In The Age of Devolution: How Anthropological Approaches Can Help Navigate Grassroots Conflicts. 2017.
66. Gustafson A, Hankins S, Jilcott S. Measures of the consumer food store environment: A systematic review of the evidence 2000-2011. *Journal of Community Health*. 2012. pp. 897–911. doi:10.1007/s10900-011-9524-x
67. Liese AD, Colabianchi N, Lamichhane AP, Barnes TL, Hibbert JD, Porter DE, et al. Validation of 3 food outlet databases: Completeness and geospatial accuracy in rural and urban food environments. *Am J Epidemiol*. 2010;172: 1324–1333. doi:10.1093/aje/kwq292
68. Ploeg M ver, Breneman V, Farrigan T, Hamrick K, Hopkins D, Kaufman P, et al. Access to affordable and nutritious food: measuring and understanding food deserts and their consequences. Report to Congress. USDA. 2009; 1–150.
69. Bowes DR. A Two-Stage Model of the Simultaneous Relationship Between Retail Development and Crime. *Econ Dev Q*. 2007;21: 79–90. doi:10.1177/0891242406292465
70. Handy SL, Clifton KJ. Local shopping as a strategy for reducing automobile travel. *Transportation (Amst)*. 2001;28: 317–346. doi:10.1023/A:1011850618753

## Figures

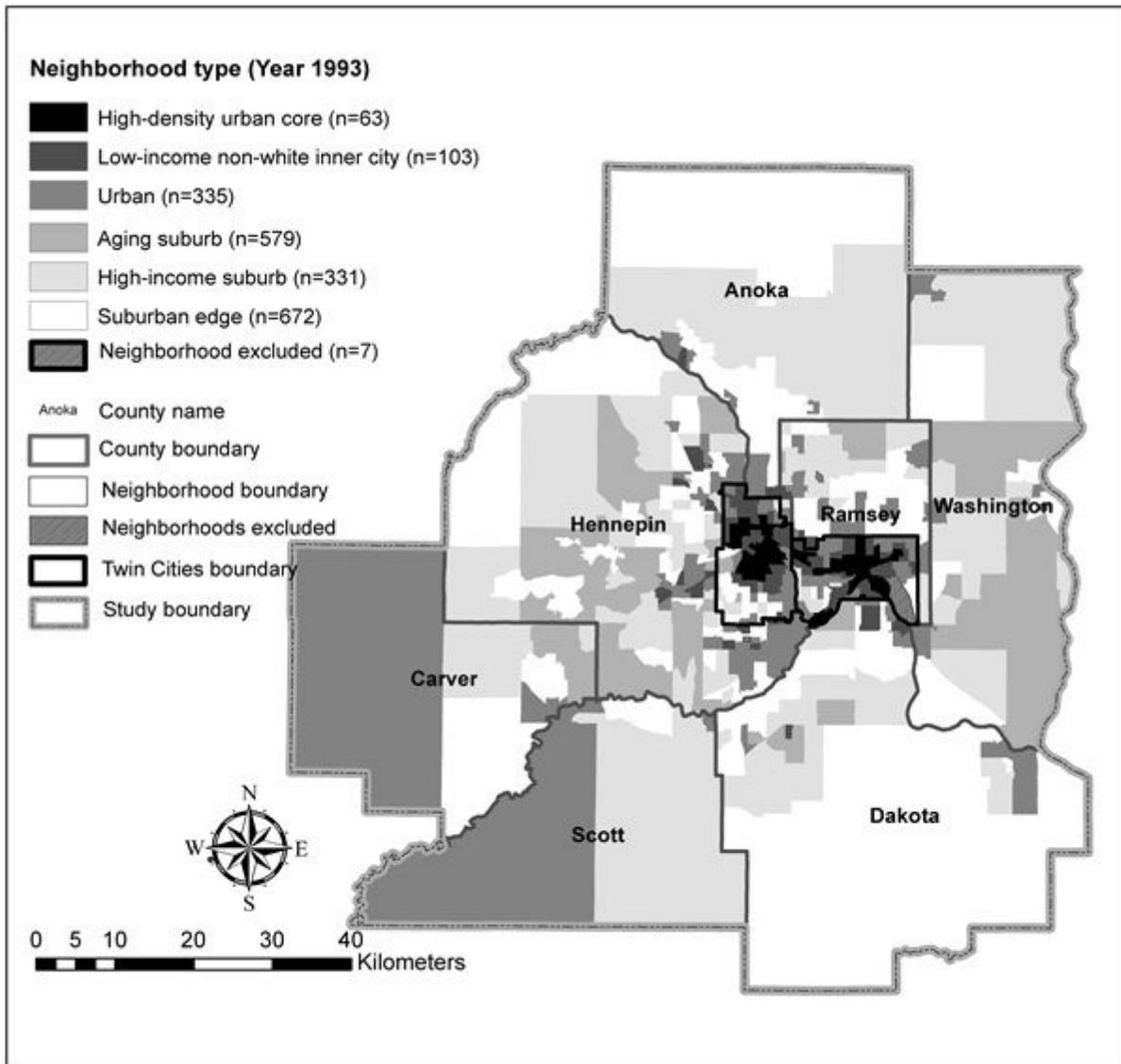
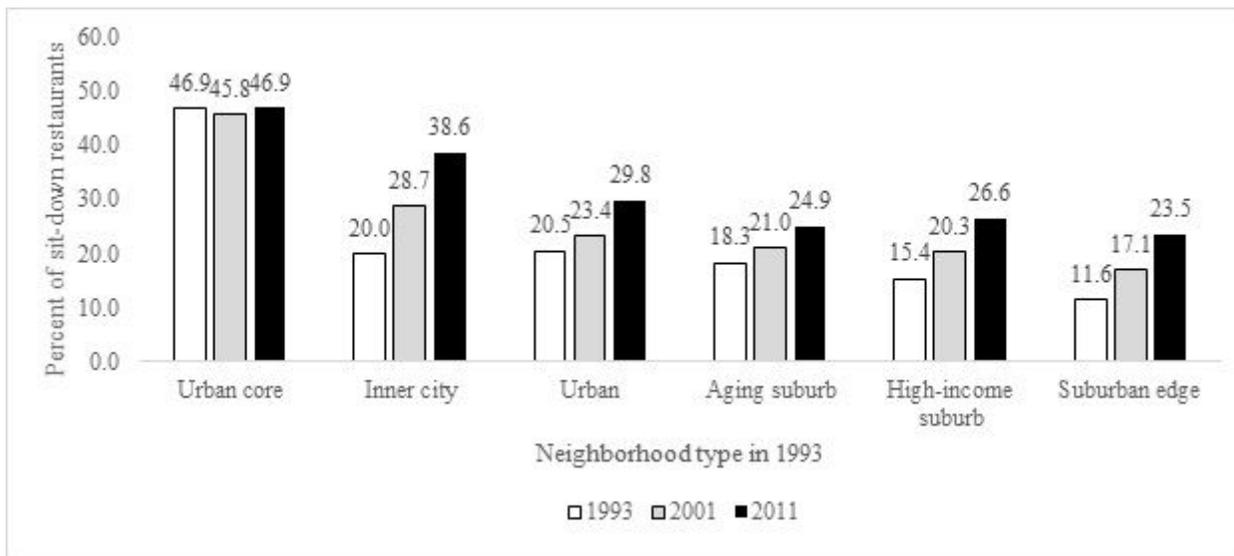


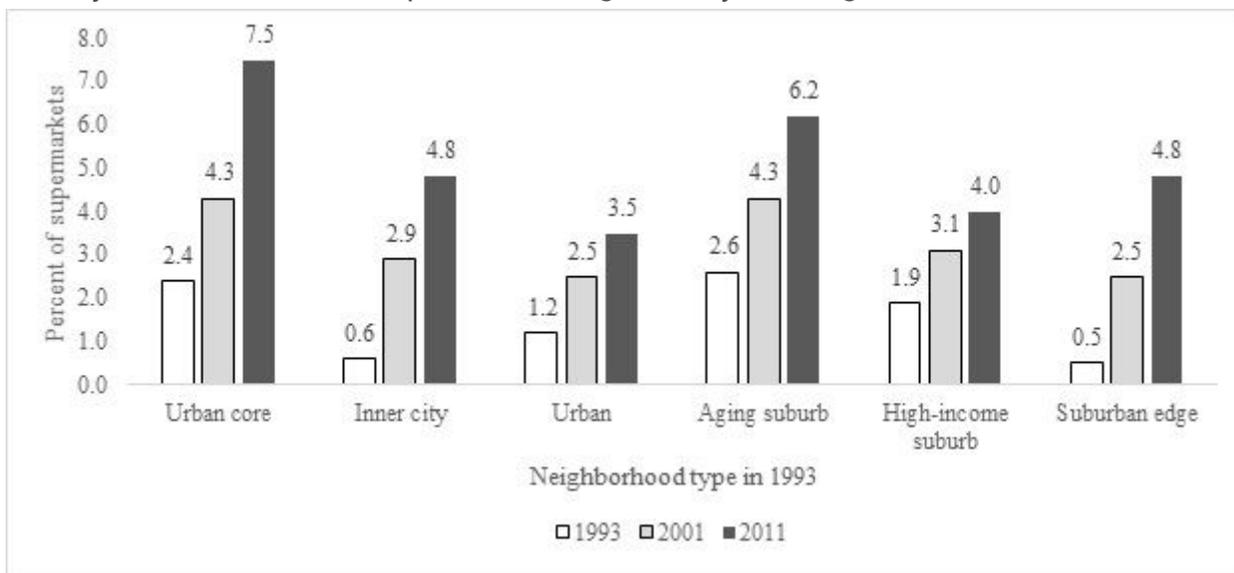
Figure 1

Neighborhood types in 1993 in the Twin Cities Region of Minnesota



**Figure 2**

Estimated mean a of percent of sit-down restaurants by six types of neighborhoods b a Multivariable mixed effects regression modeling percent of sit-down restaurants relative to total sit-down restaurants and fast food restaurants in each neighborhood as a function of neighborhood type in 1993, time elapsed since 1993, interaction between neighborhood type in 1993 and time elapsed, changes in residential population density, median household income, percent of white and percent of single-family housing since 1993, total sit-down restaurants and fast food restaurants and a random intercept for each neighborhood. b Derived from cluster analysis of block-group level data in 1993: percent of age under 14, age aged 15-29, 30-44, 45-64, and aged above 65, percent of education of college and above, percent of white, percent of black, median household income, residential population density, employment population density, mix of land use and percent of single-family housing.



**Figure 3**

Estimated mean a of percent of supermarkets relative to total supermarkets, grocery stores and convenience stores by six types of neighborhoods b: Twin Cities Region, 1993-2011. a Multivariable mixed effects regression modeling percent of supermarkets relative to total supermarkets, grocery stores and convenience stores in each neighborhood as a function of neighborhood type in 1993, time elapsed since 1993, interaction between neighborhood type in 1993 and time elapsed, changes in residential population density, median household income, percent of white and percent of single-family housing since 1993, total supermarkets, grocery stores and convenience stores and a random intercept for each neighborhood. b Derived from cluster analysis of block-group level data in 1993: percent of age under 14, aged 15-29, 30-44, 45-64, and aged above 65, percent of education of college or above, percent of white, percent of black, median household income, residential population density, employment population density, mix of land use and percent of single-family housing.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [PengRodriguezHirschGordonLarsenAdditionalfile.pdf](#)