



A longitudinal assessment of racial and ethnic inequities in food environment exposure and retail market concentration

Qianxia Jiang^{1,*} , Debarchana Ghosh², Sandro Steinbach³ and Kristen Cooksey Stowers⁴

¹Center for Children's Healthy Lifestyles and Nutrition, Children's Mercy Kansas City, Kansas City, MO, USA:

²Department of Geography, University of Connecticut, Storrs, CT, USA: ³Department of Agricultural and Resource Economics, University of Connecticut, Storrs, CT, USA: ⁴Department of Allied Health Sciences, University of Connecticut, Storrs, CT, USA

Submitted 6 August 2022: Final revision received 10 May 2023: Accepted 5 June 2023: First published online 16 June 2023

Abstract

Objective: This paper assesses trends in food environment and market concentration and racial and ethnic inequities in food environment exposure and food retail market concentration at the US census tract level from 2000 to 2019.

Design: Establishment-level data from the National Establishment Time Series were used to measure food environment exposure and food retail market concentration. We linked that dataset to race, ethnicity and social vulnerability information from the American Community Survey and the Agency for Toxic Substances and Disease Registry. A geospatial hot-spot analysis was conducted to identify relatively low and high healthy food access clusters based on the modified Retail Food Environment Index (mRFEI). The associations were assessed using two-way fixed effects regression models.

Setting: Census tracts spanning all US states.

Participants: 69 904 US census tracts.

Results: The geospatial analysis revealed clear patterns of areas with high and low mRFEI values. Our empirical findings point to disparities in food environment exposure and market concentration by race. The analysis shows that Asian Americans are likelier to live in neighbourhoods with a low food environment exposure and low retail market concentration. These adverse effects are more pronounced in metro areas. The robustness analysis for the social vulnerability index confirms these results.

Conclusion: US food policies must address disparities in neighbourhood food environments and foster a healthy, profitable, equitable and sustainable food system. Our findings may inform equity-oriented neighbourhood, land use and food systems planning. Identifying priority areas for investment and policy interventions is essential for equity-oriented neighbourhood planning.

Keywords

Racial and ethnic inequities
Food environment exposure
Food retail market concentration
Longitudinal analysis

Neighbourhood disparities in access to food outlets are a significant concern for the US food system⁽¹⁾. Differences in food access can affect dietary intake and increase the risk of several diet-related negative health outcomes⁽²⁾. Racial and ethnic minority populations mostly living in neighbourhoods with limited access to grocery stores and other services such as healthcare and social support experience disproportionately poor health outcomes (e.g. healthy diet and physical and mental health) than their White counterparts due to systemic socio-economic inequities⁽²⁾.

Neighbourhoods with higher concentrations of racial and ethnic minority populations likely have food deserts and food swamps^(3,4). Among the several definitions of food deserts, the United States Department of Agriculture's⁽⁵⁾ description of a food desert is the most commonly used, where US census tracts are identified as food deserts if they satisfy the following two conditions of (1) 'low-income communities', based on having a poverty rate of 20% or greater or a median family income at or below 80% of the area median family income and (2) 'low-access communities', based on the determination that at

*Corresponding author: Email qjiang@cmh.edu

© The Author(s), 2023. Published by Cambridge University Press on behalf of The Nutrition Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.





least 500 persons and at least 33 % of the census tract's population live more than *one mile* from a supermarket or a large grocery store. Food swamp is defined as an area where there is an overabundance of convenience store and fast-food restaurants, swamping the surrounding areas with limited access to healthy food options⁽⁶⁾. Previous studies have shown that food access is associated with the racial composition of a neighbourhood⁽⁷⁾, with more supermarkets and grocery stores in predominantly White areas and fewer in mostly Black areas⁽⁸⁾. Cooksey *et al.*, in a national study of US adults, showed that residents of food swamps report poorer dietary habits and that Black Americans are likely to live in a food swamp⁽⁴⁾. Previous research also links neighbourhoods with food swamps⁽⁶⁾ to residential segregation⁽⁷⁾, disparities in individuals' dietary behaviours and diet-related health outcomes^(3,4). However, a neighbourhood's food environment is seldom healthy or unhealthy. Instead, it is on a continuum of *healthiness* where we have food deserts and food swamps on the one end and food oases (*areas with an abundance of supermarkets and a variety of grocery stores*) on the other⁽⁹⁾. A recent study measured food exposure based on perceived measures of a food desert and food swamp⁽⁴⁾.

The market structure of US food retailing is changing rapidly, both in rural and urban areas⁽¹⁰⁾. This trend has significantly impacted retail competition and, more specifically, consumer choices for healthy food; as a result of such structural changes in the retail market over the past decades, how and where people purchase food has changed⁽¹⁰⁾. For example, several empirical studies show that concentration in the food retail market may influence access to grocery stores and supermarkets and pricing^(11,12). For instance, Ma *et al.* show that supermarkets do not raise prices in local markets in response to firm market shares.⁽¹¹⁾ Instead, smaller food retailers charge substantially higher prices on average than supermarkets. Another study has shown associations between evolving food retail landscape with inequities in access to healthy food, leading to health disparities related to poor diet and diet-related health outcomes⁽¹³⁾. The studies, however, either perceive food access primarily from economic outcomes⁽¹⁴⁾ or for the general population, missing the critical associations of racial and ethnic inequities in food access. However, few studies have explored the racial and ethnic inequities in combining food environment exposure and food market concentration at the neighbourhood level over time. Such insights are crucial for food system planners to identify racial and ethnic minority neighbourhoods disproportionately burdened by inequitable access to healthy food and often have unhealthy food retail environments⁽¹⁵⁾.

In this study, we conduct a longitudinal investigation of racial and ethnic inequities in food environment exposure across the USA. Using establishment-level food retail data from the National Establishment Time Series (NETS) dataset from Walls & Associates⁽¹⁶⁾ for 2000–2019, we

assess food environment exposure and market concentration at the census tract level. Although the NETS dataset was used in large-scale studies⁽¹⁴⁾ focused on food markets and measuring food swamps before, to the best of our knowledge, this is the first longitudinal study that will reveal the associations between race and ethnicity and food environment exposure and retail food market concentration. In addition, the longitudinal dataset will enable controlling for unobserved characteristics at the census tract level using two-way fixed effects regression models.

Methodology

We used US census tracts (n 69 904) as the neighbourhood unit of analysis. Typically encompassing 2500–8000 people, a census tract is smaller than a city but larger than a block group or census block.⁽¹⁷⁾ The census tract is equivalent to a neighbourhood established by the Bureau of Census for analysing stable population characteristics and is typically integrated with extensive national surveys at the individual level⁽¹⁸⁾.

Food retail market concentration measurement

The food retail market concentration was constructed from NETS, a longitudinal dataset compiled from Dun & Bradstreet archival establishment information. NETS contains longitudinal data from 1990 on various dynamics of the US economy, such as establishment job creation and destruction, sales growth performance, survivability of business start-ups, mobility patterns and changes in primary markets⁽¹⁶⁾. To align with the years that publicly available US census data on racial and ethnic characteristics are available, we used NETS data from 2000 to 2019. Each retail establishment can be tracked over time and classified based on the North American Industry Classification System (NAICS). NAICS is a classification of business establishments by the type of economic activity. Each food retail establishment's location information (geographic coordinates) was extracted from NETS, geocoded and then aggregated at the census tract level using the Python GeoPandas library. Geocoding is the process of converting addresses into geographic coordinates used to place markers on a map. We used the Herfindahl–Hirschman index (HHI) to measure the food retail market concentration⁽¹⁹⁾. The index is calculated by squaring the market share of each firm competing in a census tract and then summing the resulting numbers. A market with an HHI of less than 1500 is considered competitive, an HHI of 1500–2500 is moderately concentrated and an HHI of 2500 or greater is highly concentrated. The Department of Justice considers highly concentrated markets, according to the HHI, as markets where companies have considerable market power⁽²⁰⁾. Such market power leads to anticompetitive behaviour and often causes the Federal Trade

Commission to intervene in merger and acquisition cases following the Federal Horizontal Merger Guidelines⁽²⁰⁾.

Measuring food environment exposure

Following previous research^(4,6,21), we categorised food retailers into seven categories using the following NAICS codes: supermarkets/grocery stores (445 110), fruit and vegetables markets (445 230), supercentres (452 311), convenience stores (445 120), dollar stores (452 319), full-service restaurants (722 511) and limited-service restaurants (722 513). A detailed description of retail formats is provided in Appendix 1. The food retail data from NETS need to differentiate between smaller and independently owned grocery stores. Previous studies have shown that many smaller stores are considered corner stores⁽²²⁾. Therefore, we identified these stores as convenience stores if they had a NAICS code designation for a grocery store and fewer than five employees⁽²²⁾. Dollar stores are categorised as general merchandise stores in NAICS code 452319. To incorporate dollar stores exclusively, we followed the previous research⁽²³⁾ and restricted our sample to establishments with 'dollar' in the establishment name.

Consistent with previous studies on food swamps, each food retail outlet mentioned above was categorised into unhealthy, intermediate and healthy⁽²⁴⁾. The unhealthy category included convenience stores and limited-service restaurants. Full-service restaurants were in the intermediate food retail outlet category drawing from prior studies that found no association between full-service restaurant access and the risk of obesity⁽²⁵⁾. Supermarkets/grocery stores, fruit and vegetable markets and supercentres are in the healthy retail food outlet category. Various associations between these food outlets and dietary and health outcomes are documented in the literature^(25,26). We computed the modified Retail Food Environment Index (mRFEI) at the census tract level using these food retailers. The mRFEI is a score measuring the relative healthiness of a food environment, a continuum where we have limited access to healthy food (food desert) or higher availability of unhealthy food (food swamp) on the one end and higher availability and easy access to healthy food (food oasis) on the other end. The mRFEI is calculated as follows⁽⁶⁾:

$$mRFEI = 100 \times \text{count of } \frac{\text{supermarkets} + \text{fruit and vegetable markets} + \text{supercenters}}{\text{all food retail outlets}} \quad (1)$$

Explanatory and control variables

We used racial (i.e. White, Black and Asian) and ethnicity (i.e. Hispanic) proportions at the census tract level as the primary explanatory variables. These variables for 2000 and 2010 were obtained from the Decennial Census and 2019 from the 2015–2019 5-year estimate American Community Survey sample data. Several other factors

may influence food environment exposure and market concentration. In line with previous research, we included the following control variables at the tract level: the overall population size (the log of the population)⁽⁷⁾; educational attainment (the percentage of people having a Bachelor's degree or higher)⁽¹⁶⁾ and the poverty rate⁽²⁷⁾. The above control variables were also obtained from the Decennial Census (2000 and 2010) and the American Community (2015–2019). The metropolitan and non-metropolitan classification of the census tract is based on the 2013 United States Department of Agriculture Rural-Urban Continuum Codes shown in Appendix 2⁽²⁸⁾.

Moreover, we used the social vulnerability index (SVI) as a robustness check. The SVI indicates the potential negative effects of US census tract areas⁽²⁹⁾, which includes an overall score and four sub-theme scores (i.e. socio-economic status, household composition and disability, minority status and language, and housing type and transportation). The overall SVI census tract rating was calculated based on four sub-themes. We used the overall census tract rating and four summary sub-themes from 2000 to 2019. Appendix 3 provides further detail, and Appendix 4 shows the descriptive statistics.

Analytical approach

Statistical analysis

We used two-way fixed effects regression models to investigate racial and ethnic inequities in the food environment⁽³⁰⁾. The baseline model is specified as follows:

$$y_{it} = \exp(\alpha_i + \alpha_t + \beta X_{it} + \gamma Z'_{it}) \varepsilon_{it}, \quad (2)$$

where we denote the census tract with i and the year with t . The outcomes of interest (y_{it}) are the mRFEI and HHI. Because the data are left-skewed, we use an exponential regression model to identify the parameters of interest. We control for unobserved time-invariant factors with census-tract fixed effects (α_i) and account for common shocks over time with time fixed effects (α_t). We balanced the dataset at the census tract level and kept census tracts with zero outcomes available from 2000 to 2019. The variable of interest X_{it} represents the racial and ethnic composition (%) within a given census tract and year. The statistical model identifies the association between the outcome and treatment variables through variation within census tracts and over time. We denote the set of control variables with Z'_{it} , and the multiplicate error term with ε_{it} . After dropping observations that were either singletons or separated by a fixed effect, 39 965 (mRFEI model) and 47 486 (HHI model) census tracts were included in the regression analysis.

We follow the standard practice of dealing with abundant zeros and rely on the Poisson pseudo maximum likelihood estimator to identify the parameters of interest in Equation 1^(31,32). Silva and Tenreiro showed that Poisson

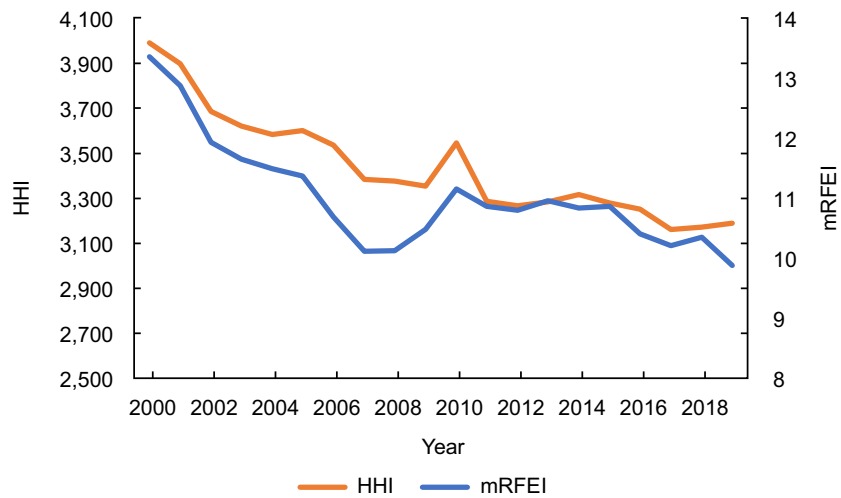


Fig. 1 Trends in food environment exposure and market concentration 2000–2019

pseudo maximum likelihood is robust to different patterns of heteroskedasticity and measurement error⁽³³⁾. The estimator also allows us to address the large share of zeros in the dataset consistently. We account for the high-dimensional fixed effects using a modified version of the iteratively re-weighted least-squares algorithm robust to statistical separation and convergence issues⁽³⁴⁾. Lastly, we suspect that the standard errors are correlated within census tracts, prompting us to cluster them at this level⁽³⁵⁾.

Geospatial analysis

Geospatial analyses of food environments, including supermarkets, convenience stores, restaurants or fast-food centres, and their impact on dietary health have been commonly used over the past 25 years⁽³⁶⁾. As spatial data and geographic information systems (GIS) software became widely available to researchers and practitioners across the social sciences, public health and medicine, mapping and geospatial analysis have been the predominant methods. One review published in 2012 found that 53% of published research on the food environment used geospatial analysis⁽³⁷⁾. Another systematic review from 2017 found that 49.6% of articles focused on food access included geospatial analysis before 2007; from 2007 through 2015, that percentage increased to 65.3%⁽³⁸⁾.

Following this trend, we used cluster analysis in a GIS framework to identify statistically significant clusters of census tracts with low and high mRFEI values. Low mRFEI indicates limited access to healthy food options, while a high mRFEI suggests greater access to healthy food choice.

First, we prepared the GIS files by merging the boundary files of the census tracts⁽³⁹⁾ with mRFEI score and the racial (i.e. White, Black and Asian) and ethnicity (i.e. Hispanic) proportions.

Second, we conducted a hot-spot analysis using ESRI’s ArcGIS Pro 2.8 GIS software⁽⁴⁰⁾. We implemented this analysis for 2000, 2010 and 2019. Based on a set of weighted features (here, mRFEI scores), the hot-spot

analysis function identifies statistically significant hot spots and cold spots using the Getis-Ord G_i^* statistic⁽⁴¹⁾. Hot spots are regions with higher mRFEI scores indicating better access to healthy foods or food oases. Cold spots are regions with lower mRFEI scores indicating higher relative access to unhealthy food retailers⁽⁴²⁾, showing likely situations of food deserts or food swamps. The hot-spot analysis tool creates an output with a z-score, *P*-value and confidence levels for each census tract. The z-scores and *P*-values are measures of statistical significance that tell us whether to reject the null hypothesis of no clustering pattern tract by tract. In effect, they indicate whether the observed geographical clustering of high or low values is more pronounced than expected in a random distribution of those same values. A high z-score and small *P*-value for a census tract indicate a spatial clustering of high mRFEI values or food oasis. A low negative z-score and small *P*-value indicate a spatial clustering of low mRFEI values, indicating food swamps or deserts. We use three confidence intervals for the identified clusters for statistical significance: 99%, 95% and 90%. We want to briefly note that hot-spot analysis of HHI value (i.e. identifying areas with high and low concentrations of retail markets) did not provide meaningful results, perhaps due to large numbers of census tracts with null or zero HHI values.

Results

Temporal trends in food environment exposure and retail market concentration

Figure 1 shows trends in food environment exposure and market concentration from 2000 to 2019. On average, we find that food environment exposure decreased slightly (mRFEI decreased to 26% in 2019 compared to 2000), while market concentration also reduced during the study period (HHI decreased 20.1% in 2019 compared to 2000).

Table 1 Food environment and racial and ethnic proportions

	Food environment exposure		Market concentration	
	β	sig.	β	Sig.
White proportion	0.00128	0.006	-0.00066	0.003
Black proportion	-0.00038	0.586	-0.00046	0.129
Asian proportion	-0.00281	0.008	-0.00531	< 0.001
Hispanic proportion*	0.00065	0.265	-0.00156	< 0.001
Poverty percentage	0.00003	0.700	-0.00012	0.002
Bachelor or higher percentage	0.00032	0.001	0.00007	0.083
Population	-0.11 987	< 0.001	-0.17 633	< 0.001

Food environment exposure: Number of observations = 798 500; residual df = 39 965. R squared = 0.46.

Market concentration: Number of observations = 948 571; residual df = 47 486. R squared = 0.74.

*The impact of the Hispanic proportion was assessed in a separate model with the same outcome and control variables.

Table 2 Food environment and racial and ethnic proportions for metro v. non-metro areas

	Food environment exposure				Market concentration			
	Metro		Non-metro		Metro		Non-metro	
	β	Sig.	β	Sig.	β	Sig.	β	Sig.
White proportion	0.00158	0.002	-0.00075	0.504	-0.00099	< 0.001	0.00172	0.003
Black proportion	0.00019	0.803	-0.00353	0.040	-0.00038	0.228	0.00095	0.256
Asian proportion	-0.00195	0.083	-0.00281	0.478	-0.00421	< 0.001	-0.00176	0.319
Hispanic proportion*	0.00136	0.027	-0.00250	0.145	-0.00011	0.713	-0.00129	0.133
Poverty percentage	-0.00002	0.828	0.0026	0.095	-0.00004	0.307	-0.00021	0.014
Bachelor or higher percentage	0.00033	0.002	0.00016	0.499	0.00002	0.600	0.00006	0.621
Population	-0.09667	< 0.001	-0.21 248	0.001	-0.15 471	< 0.001	-0.18 526	< 0.001

Food environment exposure (metro): Number of observations = 671 745; residual df = 33 627. R squared = 0.46.

Food environment exposure (non-metro): Number of observations = 126 755; residual df = 6338. R squared = 0.46.

Market concentration (metro): Number of observations = 804 657; residual df = 40 291. R squared = 0.75.

Market concentration (non-metro): Number of observations = 143 113; residual df = 7194. R squared = 0.75.

*The Hispanic proportion was assessed in a separate model with the same outcome and control variables.

Statistical analysis results

Table 1 summarises the baseline results for the associations between the racial and ethnic proportions and food environment exposure measure, that is, mRFEI. We found that the proportion of White people at the census tract level is positively associated with the mRFEI. In contrast, the proportion of Asian people was negatively associated with the mRFEI, suggesting that census tracts with more White people experience healthier food environments. The opposite relationship holds for the proportion of Asian people. The ratio of White, Asian and Hispanic at the census tract level was negatively associated with HHI, suggesting that census tracts with more White, Asian and Hispanic people were associated with more competitive markets.

Table 2 assesses the associations for the baseline model separately for metro and non-metro areas. We found that the White and Hispanic proportions were positively associated with mRFEI for metro census tracts. In contrast, the Black ratio was negatively associated with mRFEI for non-metro tracts. This result aligns with the geospatial finding that the size or the geographic coverage of the cold

spots of mRFEI scores indicating limited access to healthy foods increased over the years from the selected metros to its suburbs. In addition, we found that White and Asian proportions were negatively associated with HHI in metro areas for market concentration. However, only the White ratio was negatively associated with HHI among non-metro census tract areas.

Table 3 investigates the robustness of our baseline results using SVI scores as the explanatory variables. These estimates indicate that minority status and language sub-theme scores were negatively associated with mRFEI, suggesting that the greater the diversity of minority status and language, the healthier the food environment. Furthermore, the sub-theme scores for socio-economic status were negatively associated with HHI, suggesting that the greater the vulnerability of socio-economic status and minority status and language, the higher the market concentration in each census tract. At the same time, we found that the household composition and disability score were positively associated with HHI, and the relationships between the total SVI score and mRFEI and HHI were not statistically significant.

Table 3 Food environment and the social vulnerability index

	Food environment exposure		Market concentration	
	β	Sig.	β	Sig.
Model 1				
SVI total score	-0.012	0.690	-0.018	0.172
Model 2				
Socio-economic status	0.009	0.740	-0.039	0.001
Household composition and disability	-0.017	0.333	0.096	< 0.001
Minority status and language	-0.042	0.034	-0.112	< 0.001
Housing type and transportation	0.013	0.513	-0.016	0.070

Food environment exposure (model 1): Number of observations = 871 871; residual df = 52 866. R squared = 0.12.
 Food environment exposure (model 2): Number of observations = 871 871; residual df = 52 866. R squared = 0.12.
 Market concentration (model 1): Number of observations = 1 060 563; residual df = 65 264. R squared = 0.09.
 Market concentration (model 2): Number of observations = 1 060 563; residual df = 65 264. R squared = 0.09.

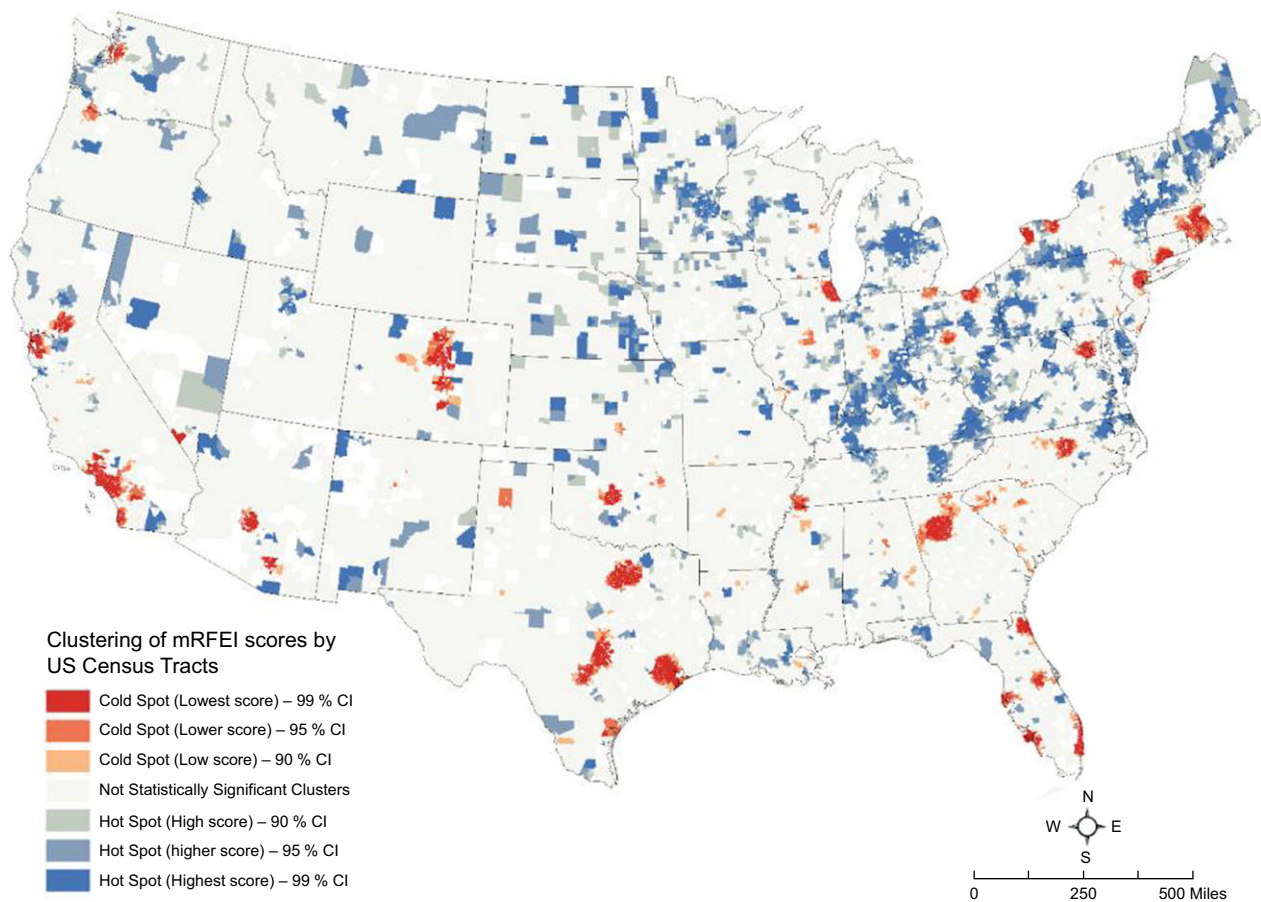


Fig. 2 Hotspot analysis of modified Retail Food Environment Index mRFEI scores in 2000 (tracts in the lower 48 US states)

Geospatial analysis results

Figures 2, 3 and 4 show the results of the geospatial analysis that identifies statistically significant hot and cold spots of mRFEI scores in 2000, 2010 and 2019, respectively. Hot spots, that is, clusters of census tracts (in shades of blue) with higher mRFEI scores, indicate neighbourhoods with higher availability and easy access to healthy food retailers. Hot spots can be related to food oases. Conversely, cold spots, that is, clusters of census tracts (in shades of red) with lower mRFEI scores, indicate

neighbourhoods with lower availability (or higher availability of unhealthy foods) and limited access to healthy food retailers. Cold spots can be related to food deserts or food swamps.

Several spatial patterns of mRFEI scores were observed during the study period. First, the number of statistically significant clusters, including hot and cold spots, increased from 2000 to 2019, with the number of hot spots (higher mRFEI scores; easy access to healthy food) growing more than cold spots. Second, in some cold spots, that is, areas

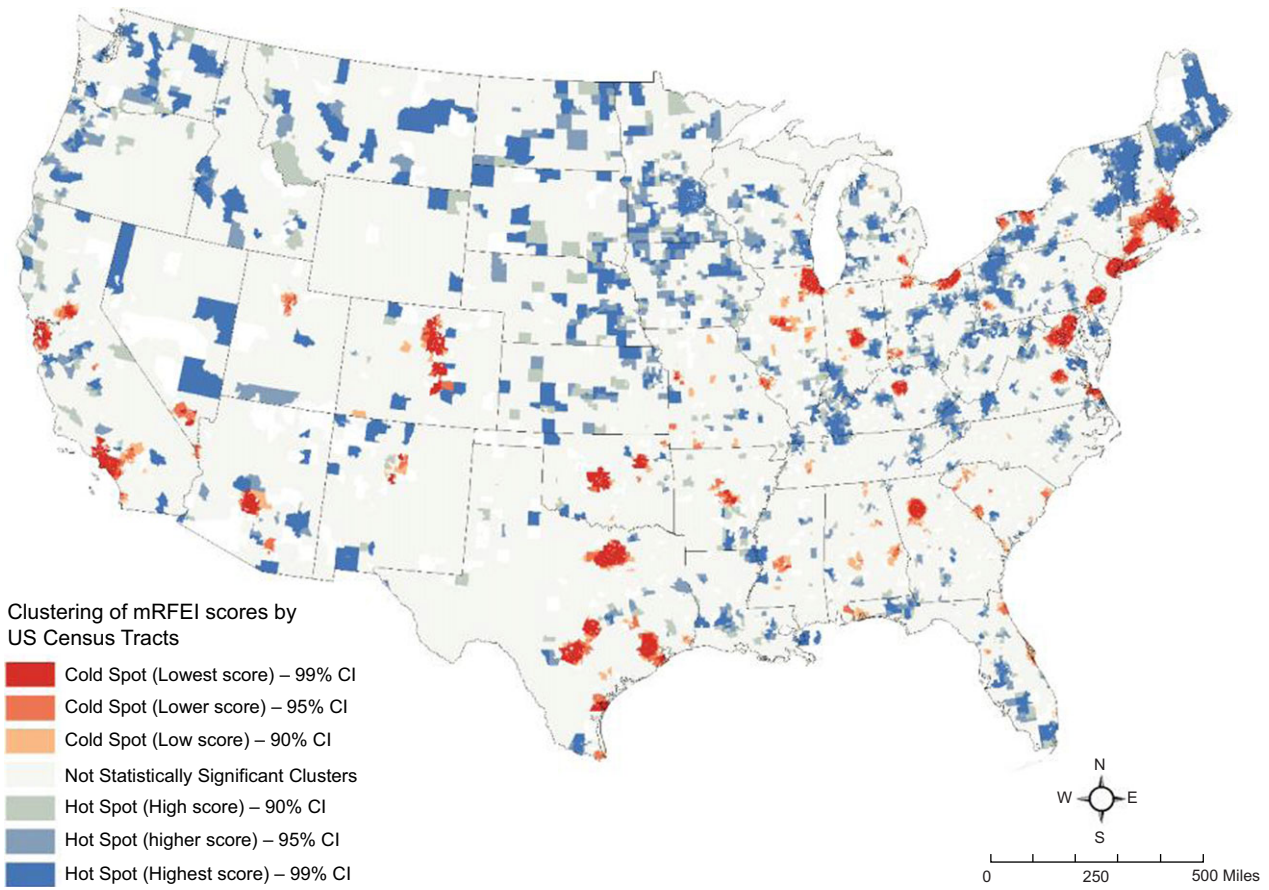


Fig. 3 Hotspot analysis of modified Retail Food Environment Index (mRFEI) scores in 2010 (tracts in the lower 48 US states)

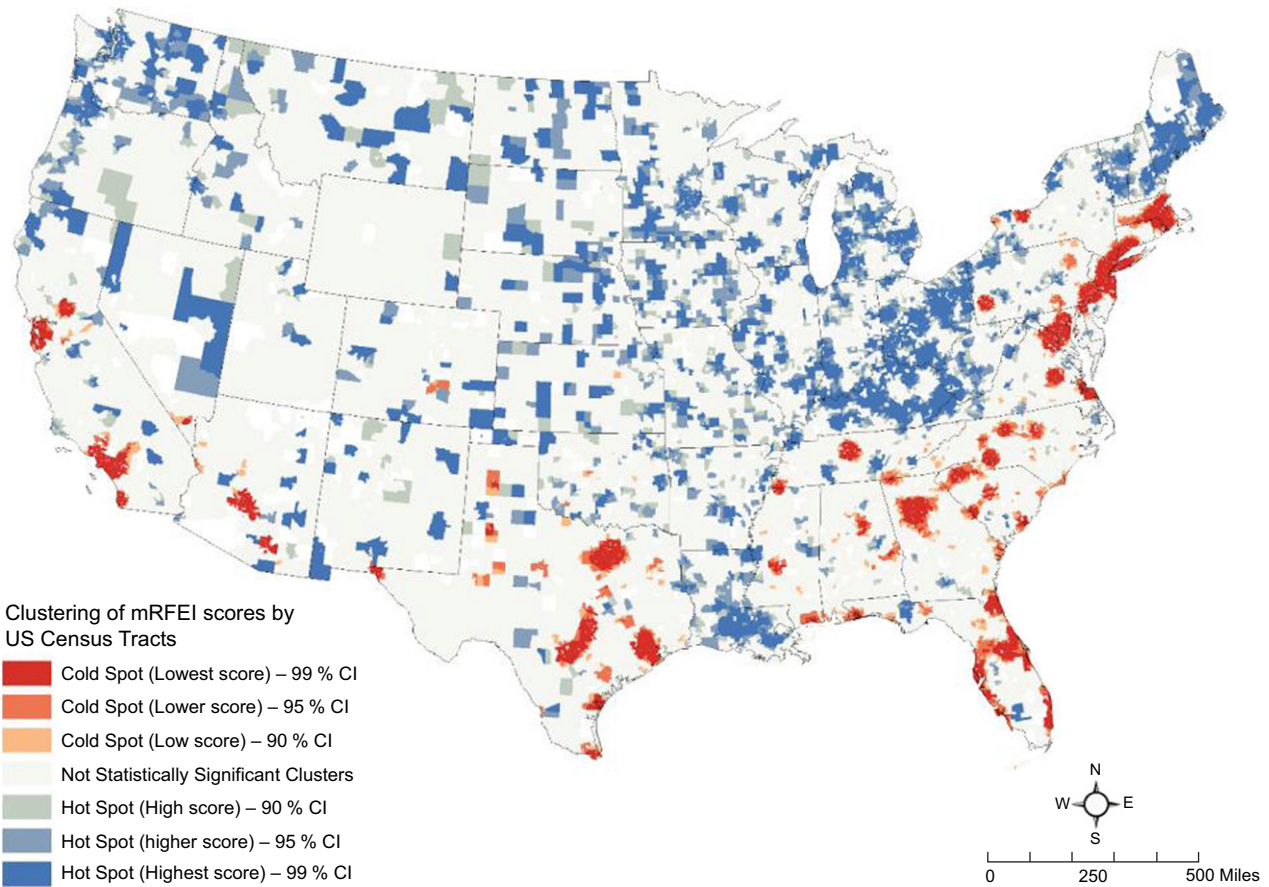


Fig. 4 Hotspot analysis of modified Retail Food Environment Index (mRFEI) scores in 2019 (tracts in 48 US states)



with limited access to healthy food (a possible situation of food deserts) or higher availability of unhealthy *v.* healthy foods (a possible case of food swamps), accessibility to healthy food improved over time, and they changed to hot spots. Some examples include census tracts in cities in Illinois (e.g. Chicago), Ohio (e.g. Cleveland, Columbus) and Colorado (e.g. Denver) (see Figs 2 and 4). Third, even though no new cold spot emerged over time, the area of the existing cold spots increased spatially (i.e. the spot began wider or longer), indicating more inequity (or segregation) in the relative access to unhealthy compared to healthy foods. Such clusters are in Texas, Florida, the southern states, including Alabama, Georgia and South Carolina, and the northeastern states, including New Jersey, New York, Connecticut and Massachusetts (Fig. 4). Fourth, the spatiotemporal pattern of the hot and cold spots of mRFEI scores or food environment exposure in the western states did not change significantly.

Conclusion

This paper provides the first systematic assessment of racial and ethnic inequities in food environment exposure and market concentration at the census tract level for the USA over the past two decades. According to the descriptive results, food swamps expanded from 2000 to 2019 in areas where food retail outlets with limited healthy food (and more unhealthy food) choices grew substantially while independent full-service grocery or supermarket stores fell behind, creating a situation of supermarket redlining⁽²⁾. The term supermarket redlining describes the disinclination of chain supermarkets to locate or relocate existing stores from impoverished inner-city neighbourhoods to affluent suburbs. As with more familiar forms of banking and residential redlining, the driving force behind supermarket redlining is also an abstraction based on perceived 'urban obstacles', including racial segregation⁽⁴³⁾. Supermarkets also tend to drive smaller local grocery stores out of business when they move in. When they relocate or close, residents face difficulties in accessing healthy and affordable food, thus widening the grocery gap, increasing food insecurity and unhealthiness of the food environment. In a case study of the city of Hartford⁽²⁾, with a higher proportion of Black American and low-income communities, mapped critical areas in the inner-city region where if a nearby supermarket closes or relocates to a suburb with little mitigation efforts to fill the grocery gap, a large number of minority, poor and disadvantaged residents will experience difficulties to access healthy food. This is of concern as earlier studies showed that unhealthy food environment such as food swamp strongly predicts geographic disparities in adult obesity prevalence⁽⁶⁾.

We found strong evidence for racial and ethnic inequities in food environment exposure at the census tract level. Our regression model with the SVI also indicated that the greater diversity of minority status and language

was associated with an unhealthier food environment. We also found that these racial and ethnic inequities in food environments differed among the metro and non-metro areas. Some of our findings align with previous studies on structural racism in the food system^(2,44). These studies found that racial and ethnic minorities are likelier to live near unhealthy food retail outlets than White. For example, Black and Hispanic people tended to have greater access to fast-food establishments than their white counterparts⁽⁴⁴⁾. Previous research also pointed to racial and ethnic disparities in the likelihood of residents living in a food swamp or desert⁽⁴⁾. Additional research examining structural racism and residential segregation as a fundamental cause⁽⁴⁵⁾ of health inequities is needed, particularly as it relates to the clustering of health restricting resources (e.g. unhealthy food outlets). Additional research on these topics is warranted to better inform federal nutrition policies, neighbourhood revitalisation efforts, land use zoning and housing policies to combat racial residential segregation.

Except for minority status and language score, we found no significant correlations between the social vulnerability sub-theme scores (i.e. socio-economic status, household composition and disability, housing type and transportation). The previous literature suggested that vacant homes may influence the composition of food outlets in urban neighbourhoods⁽⁴⁶⁾. Specifically, a study found that an increase in the vacancy home rate was negatively associated with the food swamp index from 2001 to 2012, in Baltimore, among non-African American neighbourhoods. Furthermore, research also showed that neighbourhood food access varies significantly with a neighbourhood's socio-economic composition⁽⁴⁷⁾. However, we did not detect evidence for these associations at the national level. Therefore, a more comprehensive understanding of the specific roles of housing type and socio-economic composition in shaping the food environment may be helpful for policymakers, city planners and public health practitioners to promote healthy food access among neighbourhoods.

The current study does have a few limitations. First, the secondary data source we used to measure the local food environment may need to accurately explore the associations we found for some study areas because we cannot observe the product range offered by the stores. Realistically, food environment field audits cannot measure food establishments in a large region or during a longer historical period because such efforts are generally cost-prohibitive⁽⁴⁸⁾. Previous studies suggest that InfoUSA and government food registries have a higher level of agreement than reported by other secondary data sources⁽⁴⁹⁾. Future work might improve consistency in data gathering, geocoding, editing and analysing secondary data sources. Previous evidence also showed that data providers might apply classification schemes inconsistently⁽⁵⁰⁾. Lastly, our study could have undercounted the number of dollar stores

as there is no specific NAICS code for this retail outlet type. Further research may involve within-store measures of previous menu audit studies to classify different formats of retail food outlets⁽¹⁸⁾.

Our study findings, nevertheless, may be used to identify racial and ethnic minority neighbourhoods burdened by unhealthiness of food environments with food deserts and food swamps and to inform equity-oriented food systems and neighbourhood planning. In addition, initiatives aiming to simultaneously incentivise additional healthy food retailers (e.g. Healthy Food Financing Initiative, which provides subsidies for supermarket operators that locate in food deserts) and disincentivise unhealthy food retail environments (e.g. zoning restrictions, stocking standards for SNAP-authorized corner stores) could play a crucial role in ensuring equitable access to healthy food, particularly in communities of colour. Our findings suggest that US food policies and economic incentives must foster a healthy, fair and sustainable food system to address disparities in neighbourhood healthy food environments. Finally, future research may explore how citizenship and immigrant status affects widening food environment disparities between Hispanic populations and other people of colour.

Acknowledgements

Acknowledgements: The authors may acknowledge individuals or organisations that provided advice and/or support (non-financial). **Financial support:** This work was supported by a Research Excellence Award provided by the University of Connecticut. **Conflict of interest:** The authors declare that they have no conflicts of interest. **Authorship:** Q.J. led all phases of the study, including the statistical analysis and manuscript preparation. D.G. conducted geospatial analysis and reviewed and edited the manuscript. S.S. and K.C.S. provided input on data analyses, presentation and reviewed and edited the manuscript. **Ethics of human subject participation:** Manuscripts describing research did not involve human participants.

References

1. Singleton CR (2022) Improving urban food systems requires emphasizing nutrition equity in interventions and policy action. *Am J Clin Nutr* **115**, 981–982.
2. Zhang M, Ghosh D & Debarchana G (2016) Spatial supermarket redlining and neighborhood vulnerability: a case study of Hartford, Connecticut. *Trans Gis* **20**, 79–100.
3. Hager ER, Cockerham A, O'Reilly N *et al.* (2017) Food swamps and food deserts in Baltimore City, MD, USA: associations with dietary behaviours among urban adolescent girls. *Public Health Nutr* **20**, 2598–2607.

4. Stowers KC, Jiang Q, Atoloye A *et al.* (2020) Racial differences in perceived food swamp and food desert exposure and disparities in self-reported dietary habits. *Int J Environ Res Public Health* **17**, 1–14.
5. USDA (2023) Food Access Research Atlas; available at <https://www.ers.usda.gov/data-products/food-access-research-atlas/> (accessed May 2023).
6. Cooksey-Stowers K, Schwartz MB, Brownell KD *et al.* (2017) Food swamps predict obesity rates better than food deserts in the United States. *Int J Environ Res Public Health* **14**, 1366.
7. Havewala F (2021) The dynamics between the food environment and residential segregation: an analysis of metropolitan areas. *Food Policy* **103**, 102015.
8. Morland K & Filomena S (2007) Disparities in the availability of fruits and vegetables between racially segregated urban neighbourhoods. *Public Health Nutr* **10**, 1481–1489.
9. Food Desert or Oasis (2019) Triangle Area NC Food Market Analysis; available at <https://storymaps.arcgis.com/stories/74612f38cfe54b02808674a2e9e87c29> (accessed April 2023).
10. Cho C & Volpe R (2017) Independent Grocery Stores in the Changing Landscape of the U.S. Food Retail Industry. USDA Economic Research Service; available at <https://www.ers.usda.gov/webdocs/publications/85783/err-240.pdf?v=5083> (accessed April 2023).
11. Ma M, Saitone TL, Volpe RJ *et al.* (2019) Market concentration, market shares, and retail food prices: evidence from the U.S. women, infants, and children program. *Appl Econ Perspect Policy* **41**, 542–562.
12. Zeballos E, Dong X & Islamaj E (2023) A Disaggregated View of Market Concentration in the Food Retail Industry. Amber Waves 2023; available at <https://www.ers.usda.gov/webdocs/publications/105558/err-314.pdf?v=8156.2> (accessed April 2023).
13. Caillavet F, Kyureghian G, Nayga RM *et al.* (2015) Does healthy food access matter in a French urban setting? *Am J Agric Econ* **97**, 1400–1416.
14. Chapple K (2014) The highest and best use? Urban industrial land and job creation. *Econ Dev Q* **28**, 300–313.
15. Zenk SN, Schulz AJ, Israel BA *et al.* (2005) Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in metropolitan Detroit. *Am J Public Health* **95**, 660–667.
16. Dun and Bradstreet (2019) National Establishment Time-Series Database; available at <https://www.dnb.com/> (accessed April 2023).
17. U.S. Census Bureau (2021) Geographic Terms and Definitions; available at <https://www.census.gov/programs-surveys/popest/about/glossary/geo-terms.html> (accessed April 2023).
18. Martin KS, Ghosh D, Page M *et al.* (2014) What role do local grocery stores play in urban food environments? A case study of Hartford-Connecticut. *PLoS One* **9**, e94033.
19. Rhoades SA (1993) The Herfindahl-Hirschman index. *Fed Reserv Bull* **79**, 188.
20. The U.S. Department of Justice (2018) Herfindahl-Hirschman Index; available at <https://www.justice.gov/atr/herfindahl-hirschman-index> (accessed April 2023).
21. Centers for Disease Control and Prevention (2013) *Census Tract Level State Maps of the Modified Retail Food Environment Index (mRFEI)*. Atlanta, GA: Centers for Disease Control and Prevention.
22. Wong MS, Peyton JM, Shields TM *et al.* (2017) Comparing the accuracy of food outlet datasets in an urban environment. *Geospat Health* **12**, 546.
23. Stevens A, Cho C, Çakır M *et al.* (2021) The Food Retail Landscape Across Rural America; available at <https://www.ers.usda.gov/webdocs/publications/101356/eib-223.pdf?v=2593.3> (accessed April 2023).
24. Mui Y, Gittelsohn J & Jones-Smith JC (2017) Longitudinal associations between change in neighborhood social



- disorder and change in food swamps in an urban setting. *J Urban Heal* **94**, 75–86.
25. Jia P, Yang H, Cao X *et al.* (2021) Association between access to full-service restaurants and childhood obesity. *Obes Rev* **22**, e13076.
 26. Mazidi M & Speakman JR (2017) Higher densities of fast-food and full-service restaurants are not associated with obesity prevalence. *Am J Clin Nutr* **106**, 603–613.
 27. Handbury J, Rahkovsky IM & Schnell M (2015) *What Drives Nutritional Disparities? Retail Access and Food Purchases Across the Socioeconomic Spectrum*. Cambridge, MA: National Bureau of Economic Research.
 28. US Department of Agriculture Economic Research Service (2013) USDA ERS – Rural-Urban Continuum Codes. US Department of Agriculture and Economic Research Service. <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx> (accessed June 2022).
 29. GRASP (2022) CDC/ATSDR SVI Data and Documentation Download | Place and Health | ATSDR. https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html (accessed May 2022).
 30. Wooldridge JM (2021) *Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators*. Rochester, NY: Social Science Research Network.
 31. Gourieroux C, Monfort A & Trognon A (1984) Pseudo maximum likelihood methods: applications to Poisson models. *Econometrica* **52**, 701.
 32. Gong G & Samaniego FJ (1981) Pseudo maximum likelihood estimation: theory and applications. *Ann Stat* **9**, 861–869.
 33. Santos Silva JMC & Tenreiro S (2006) The log of gravity. *Rev Econ Stat* **88**, 641–658.
 34. Correia S, Guimarães P & Zylkin T (2019) Verifying the existence of maximum likelihood estimates for generalized linear models; available at <https://arxiv.org/abs/1903.01633> (accessed August 2022).
 35. Colin Cameron A & Miller DL (2015) A practitioner's guide to cluster-robust inference. *J Hum Resour* **50**, 317–372.
 36. Shannon J, Reese AM, Ghosh D *et al.* (2021) More than mapping: improving methods for studying the geographies of food access. *Am J Public Health* **118**, 1418–1422.
 37. Caspi CE, Sorensen G, Subramanian SV *et al.* (2012) The local food environment and diet: a systematic review. *Heal Place* **18**, 1172–1187.
 38. Lassen AD, Fagt S, Lennernäs M *et al.* (2018) The impact of worksite interventions promoting healthier food and/or physical activity habits among employees working 'around the clock' hours: a systematic review. *Food Nutr Res* Published online: 02 August 2018. doi: 10.29219/fnr.v62.1115.
 39. U.S. Census Bureau (2019) TIGER/Line Shapefiles. <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2019.html> (accessed May 2022).
 40. ESRI (2020) 2D, 3D & 4D GIS Mapping Software | ArcGIS Pro. ArcGIS Pro (Version 2.5.1). <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview> (accessed May 2022).
 41. Ord JK & Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. *Geogr Anal* **27**, 286–306.
 42. Jeffery RW, Baxter J, McGuire M *et al.* (2006) Are fast food restaurants an environmental risk factor for obesity? *Int J Behav Nutr Phys Act* **3**, 2.
 43. Eisenhauer E (2001) In poor health: supermarket redlining and urban nutrition. *Geojournal* **53**, 125–133.
 44. Dunn RA, Sharkey JR & Horel S (2012) The effect of fast-food availability on fast-food consumption and obesity among rural residents: an analysis by race/ethnicity. *Econ Hum Biol* **10**, 1–13.
 45. Phelan JC & Link BG (2015) Is racism a fundamental cause of inequalities in health? *Annu Rev Sociol* **41**, 311–330. <https://doi.org/10.1146/annurev-soc-073014-112305>
 46. Mui Y, Jones-Smith JC, Thornton RLJ *et al.* (2017) Relationships between vacant homes and food swamps: a longitudinal study of an urban food environment. *Int J Environ Res Public Health* **14**, 1426.
 47. Moore LV & Diez Roux AV (2006) Associations of neighborhood characteristics with the location and type of food stores. *Am J Public Health* **96**, 325–331.
 48. Auchincloss AH, Moore KAB, Moore LV *et al.* (2012) Improving retrospective characterization of the food environment for a large region in the United States during a historic time period. *Health Place* **18**, 1341–1347.
 49. Fleischhacker SE, Evenson KR, Sharkey J *et al.* (2013) Validity of secondary retail food outlet data: a systematic review. *Am J Prev Med* **45**, 462–473.
 50. Forsyth A, Lytle L & Van Riper D (2010) Finding food: issues and challenges in using Geographic Information Systems (GIS) to measure food access. *J Transp Land Use* **3**, 43.

Appendix 1: Definitions of Retail Format

Format	NAICS code	Definition
Supermarkets/ grocery stores	445110	The industry comprises establishments generally known as supermarkets and grocery stores primarily engaged in retailing a general line of food, such as canned and frozen foods; fresh fruits and vegetables; and fresh and prepared meats, fish and poultry
Fruit and vegeta- bles markets	445230	This industry comprises establishments primarily engaged in retailing fresh fruits and vegetables
Supercentres and warehouse clubs	452311	This industry comprises establishments known as warehouse clubs, super-stores or supercentres primarily engaged in retailing a general line of groceries in combination with general lines of new merchandise, such as apparel, furniture and appliances
Convenience stores	445120 and those with employees four or less in NAICS code 445110	This industry comprises establishments known as convenience stores or food marts primarily engaged in retailing a limited line of goods that generally includes milk, bread, soda and snacks
Dollar stores	452319	Stores with 'dollar' in the name with NACIS code 452319
Full-service restaurants	722511	This industry comprises establishments primarily engaged in providing food services to patrons who order and are served while seated (i.e. waiter/ waitress service) and pay after eating
Limited-service restaurants	722513	This industry comprises establishments primarily engaged in providing food services where patrons generally order or select items and pay before eating. Food and drink may be consumed on premises, taken out or delivered to the customer's location

Sourced from the North American Industry Classification System code definitions.

Appendix 2: Categories for metro and non-metro counties

Code	Description
Metro counties	
1	Counties in metro areas of 1 million population or more
2	Counties in metro areas of 250 000–1 million population
3	Counties in metro areas of fewer than 25 000 population
Non-metro counties	
4	Urban population of 20 000 or more, adjacent to a metro area
5	Urban population of 20 000 or more, not adjacent to a metro area
6	Urban population of 2500–19 999, adjacent to a metro area
7	Urban population of 2500–19 999, adjacent to a metro area
8	Completely rural or less than 2500 urban population, adjacent to a metro area
9	Completely rural or less than 2500 urban population, not adjacent to a metro area

Appendix 3: Components of the social vulnerability index⁽²⁹⁾

Overall vulnerability	Sub themes	Components
Social vulnerability	Socio-economic status	Below poverty Unemployed Income No high school diploma
	Household composition and disability	Aged 65 or older Aged 17 or younger Civilian with a disability Single-parent households
	Minority status and language	Minority Speaks English 'less than well'
	Housing type and transportation	Multi-unit structures Mobile homes Crowding No vehicle Group quarters

**Appendix 4: Summary statistics**

Variable	Mean	SD	Min	Max
Population	4338.89	1880.40	1	59 947
White proportion	72.90	26.14	0	100
Black proportion	14.12	23.15	0	100
Asian proportion	4.30	8.59	0	96
Hispanic proportion	14.76	21.21	0	100
Poverty percentage	15.74	22.24	0	100
Bachelor or higher percentage	17.54	22.78	0	100
HHI	3438.79	2262.78	87.96	10 000
mRFEI	11.02	14.37	0	100
SVI total score	0.511	0.284	0	1
Socio-economic status	0.505	0.283	0	1
Household composition and disability	0.507	0.278	0	1
Minority status and language	0.504	0.280	0	1
Housing type and transportation	0.517	0.276	0	1