Obesity in Urban Food Markets: Evidence from Geo-referenced Micro Data

Susan E. Chen

Raymond J.G.M. Florax

And

Samantha D. Snyder

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009
AAEA & ACCI Joint Annual Meeting, Wisconsin, July 26-29, 2009

Obesity in Urban Food Markets: Evidence from Georeferenced Micro Data*

Susan E. Chen[†1], Raymond J.G.M. Florax[1,2], and Samantha D. Snyder[1]

[1] Purdue University, West Lafayette, Indiana, USA
[2] VU University Amsterdam, Amsterdam, The Netherlands

March 31, 2009

Abstract. This paper provides quantitative estimates of the effect of proximity to fast food restaurants and grocery stores on obesity in urban food markets. Our empirical model combined georeferenced micro data on access to fast food restaurants and grocery stores with data about salient personal characteristics, individual behaviors, and neighborhood characteristics. We defined a "local food environment" for every individual utilizing 0.5-mile buffers around a person's home address. Local food landscapes are potentially endogenous due to spatial sorting of the population and food outlets, and the body mass index (BMI) values for individuals living close to each other are likely to be spatially correlated because of observed and unobserved individual and neighborhood effects. The potential biases associated with endogeneity and spatial correlation were handled using spatial econometric estimation techniques. Our policy simulations for Indianapolis, Indiana, focused on the importance of reducing the density of fast food restaurants or increasing access to grocery stores. We accounted for spatial heterogeneity in both the policy instruments and individual neighborhoods, and consistently found small but statistically significant effects for the hypothesized relationships between individual BMI values and the densities of fast food restaurants and grocery stores.

JEL codes: C31, D12, I12, I18
Keywords: obesity, fast food, grocery store, spatial econometrics, micro data

* The authors thank Joseph Gibson of the Health and Hospital Corporation of Marion County for sharing the data. They gratefully acknowledge financial support from the University of Michigan Poverty Center through a grant from the United States Department of Agriculture (USDA). The authors also thank Inas Rashad, William Masters, James Binkley, Jos van Ommeren, participants of the USDA conference "Understanding The Economic Characteristics of Food Access", and workshop participants at Purdue University and the University of Maryland, Baltimore, for their helpful suggestions and comments.

† Contact Author: Department of Agricultural Economics, Purdue University, 403 W. State Street, West Lafayette, IN 47907–2056, USA. Telephone: 765–494–7545. Fax: 765–494–9176. E-mail: sechen@purdue.edu.

1 Introduction

Effective public health interventions depend critically on identifying and understanding the health-related behaviors that cause the obesity epidemic. For example, although individual-level factors such as biological (genetic) and socioeconomic conditions have been shown to be associated with obesity, there is a growing literature on the role the food environment plays in the prevalence of obesity in communities. The main questions addressed in this literature are whether close proximity to fast food restaurants makes people obese, whether lack of access to retail grocers contributes to obesity, or whether a combination of these factors is at work. The dominant research finding thus far is that a lack of access to grocery retailers is positively associated with increased obesity rates, whereas the evidence for access to fast food is less clear (Morland et al., 2006; Cummins and Macintyre, 2006). Some studies have documented a positive correlation between obesity and access to fast food (Maddock, 2004; Chou et al., 2004), but others have not (Burdette and Whitaker, 2004; Jeffery et al., 2006).

Although the abovementioned studies have persuasively argued that there is a correlation between some subset of access to fast food and grocery stores, the findings are not causal because they do not account for the importance of choice. People select where they want to live based on some subset of neighborhood characteristics and individual preferences. Although simple regression analyses and bivariate correlations reflect the current state of the literature, these approaches do not account for selection bias and the potential of endogeneity of the food landscape variables. The food landscape variables may be endogenous because where a person chooses to live may be driven by underlying preferences, which in turn may also be correlated with factors responsible for obesity. Often, these factors are unobserved. Ignoring this endogeneity introduces bias in analyses of the effects of the food landscape on obesity.

Other shortcomings of previous work in this area include the aggregate nature of the areal unit under consideration, and the failure to use appropriate spatial econometric techniques. Confidentiality restrictions often prohibit the release of geographical identifiers in publicly available health surveys and economic surveys.[1] As a result, many of the studies that examine the relationship between the food landscape and body mass index (BMI) have had to use arbitrarily designated regions (e.g., census tracts or counties), which are largely based on administrative compatibility (for example: Maddock, 2004; Chou et al., 2004; Morland et al., 2002, 2006; Moore and Diez Roux, 2006). The problem with using arbitrarily designated neighborhoods and large areal units is that these samples may lead to biased results because of ecological inference fallacy and the assumption that people do not shop outside of their census tract. These "spillover" behaviors can only be adequately accounted for by considering spatial dependence across administrative units.

In this study, we estimated a reduced-form model for the determinants of BMI. We were able to overcome some of the limitations of previous studies by using two unique data sources that included geographical identifiers for individuals and all retail food establishments, along with demographic, economic, and health data for a group of citizens living in Indianapolis, Indiana, in 2005. Unlike in previous analyses, these unique datasets allowed us to create local food

[1] For example, two key national surveys that could have been used to address our study question are the Behavioral Risk Factor Surveillance Survey and the Census of Retail Trade. These surveys, however, only release information at the census tract level.

landscapes for each individual. We were also able to include variables for other neighborhood characteristics such as crime, measured at the level of the individual. Arguably, our analysis controls for neighborhood characteristics that have been ignored in previous studies, and we were able to isolate the effects of fast food restaurant and chain grocers on individual health.

The second contribution to the literature that our research provides stems from the consideration that individuals select where they want to live based on neighborhood amenities. (The food landscape is only one subset of these amenities.) To control for individual neighborhood selection, we used an instrumental variables approach based on city zoning regulations. The commercial zoning instrument that we used affects where fast food restaurants can be located, but we assert that it is uncorrelated with other unobserved determinants of BMI.

As a final contribution to the literature, we argue that dependence across individuals should be accounted for. Given the inherently spatial nature of the dataset, it is likely that observations are not independent across space due to unobserved social network ties among individuals or shared unobserved neighborhood characteristics across individuals living in proximate neighborhoods. We therefore used instrumental variables and generalized method of moments techniques recently developed in spatial econometrics to account for spatial dependence and heteroskedasticity (Kelejian and Prucha, 2007). This approach provides an even stronger test for the effect of access to fast food restaurants and chain grocers on the BMI of individuals.

Recent policy action to influence the food environment, particularly among populations deemed most “at risk” for obesity (e.g., minority and low-income groups), aims to restrict the number of fast food restaurants. In South Los Angeles, “health zoning” has recently been proposed; the ordinance would put in place a 2-year moratorium on new fast food restaurants. The goal of another proposed law in California is to increase the availability of nutritious foods, particularly in underserved areas (Abdollah, 2007). Drawing on past policies that focused on limiting liquor store licenses in response to alcohol abuse problems, and not withstanding that these policies have faced lawsuits challenging their constitutionality, municipalities and other local governing bodies are considering similar laws focused on where and how fast food restaurants operate (Mair and Teret, 2005). Our research focused on informing the development and implementation of policies such as these by quantifying the estimated effects on BMI of access to fast food restaurants and grocery stores, while explicitly accounting for the spatial variability of the effects.

2 Behavioral Model

Economists have suggested that technological change is one of the primary causes of the increasing prevalence of obesity because it has increased the real value of time over the last 30 years (Cutler et al., 2003; Philipson and Posner, 2003). The rise in the real value of time increases the demand for food away from home and increases the consumption of prepared foods and highly processed high-caloric foods at home. If individuals are more time-constrained today than in the past, travel time to food retailers is an implicit price of the cost of food. Formally, if health, H, is a function of food choices, F, and a numeraire good for other goods that affect weight (e.g., a gym membership), C, an individual chooses a consumption bundle that satisfies:

$$\max_{F,C} U(H, C), \tag{1}$$

subject to:

$$\begin{aligned} pF + C &= I, & (2) \\ H &= f(F, C), & (3) \end{aligned}$$

where p is the price of food, and I is total income for an individual. Assuming an interior solution, optimization results in the following reduced-form demand equations for food, the composite good, and health:

$$\begin{aligned} F^* &= g(I, p), & (4) \\ C^* &= h(I, p), & (5) \\ H^* &= j(I, p), & (6) \end{aligned}$$

where C^* and F^* are the optimal consumption goods that produce health, H^*. The vector of food prices, p, is hypothesized to be a function of both the price of food and the travel time to food retailers.

Travel time is obviously affected by the location choices of food suppliers, and there is increasing evidence of clear geographical patterns in where grocery and food retailers choose to locate. Morland et al. (2002), Moore and Diez Roux (2006), and Zenk et al. (2005) found that food retailer patterns closely follow the residential distribution of income, minority populations, and other neighborhood characteristics. Grocery stores tend to be located in affluent White neighborhoods, whereas fast food restaurants are disproportionately located in lower-income neighborhoods (Block et al., 2004). One reason suggested for this growing disparity in access to food retailers is the consolidation of retail grocery stores into large grocery chains over the last 30 years. Chung and Myers (1999) incorporated food prices in their research on the grocery retailer environment. They found that price discrepancies are starkest between chain and non-chain retailers, with the latter charging significantly higher prices. In accordance with previous studies, they found that larger chain grocery stores were less likely to be located in inner-city, lower-income neighborhoods.

Research on obesity prevalence has concurrently provided evidence of spatial clustering of obese individuals. Both at the state level and at more disaggregated geographical levels such as the county or census tract, researchers have found that people with similar BMI cluster spatially (Mobley et al., 2004; Eid et al., 2008). This clustering phenomenon supports the hypothesis that contextual factors, whether social (e.g., crime, peer effects) or physical (e.g., the food landscape), are affecting the health of the individuals that live within these neighborhoods. It has, therefore, been suggested that limited access to food retailers that offer affordable and healthy food options has led to an increase in the prevalence of obesity, particularly in urban neighborhoods with a low mean income, predominantly minority residents, or both (Cummins and Macintyre, 2006).

The growing literature on the relationship between the food landscape and the modern obesity epidemic focuses largely on the consumption of food away from home, and particularly on food obtained from fast food restaurants, and not on retail grocers. Chou et al. (2004), using individual-level data from the Behavioral Risk Factor Surveillance Survey, statewide counts of restaurants, and data on restaurant expenditures, found that residents of states with a higher number of restaurants tended to have higher BMI values. In addition, lower prices at restaurants (and lower food prices in general) were correlated with higher BMI. Although this analysis provided valuable insights into the factors that contribute to obesity and excessive weight, limitations arose from the less-specific state-wide proxies used for food consumption behavior.

Other notable studies in this literature have used less-aggregated data to study the association between access to retailers and BMI (Morland et al., 2002, 2006; Rose and Richards, 2004; Jeffery et al., 2006). Morland and colleagues used data at the level of census tracts to examine the relationships between food access, consumption, and the proportions of obesity and excessive weight. Morland et al. (2002) defined an individual's local food environment in terms of the number and type of food retailers within the census tract where the person resides. They found that for Blacks, fruit and vegetable consumption increased by 32% for each additional supermarket located in their census tract. Additional work by Morland et al. (2006) divided food retailers into three categories: supermarkets, grocery stores, and convenience stores. They found a lower prevalence of obesity and excessive weight to be associated with the presence of supermarkets, whereas higher prevalence rates were associated with the types of stores characterized by less-healthy dietary options.

There are two main limitations in this literature. The majority of studies such as Morland's used the census tract to define an individual's market for food. However, census tracts vary in size, and their boundaries often do not reflect any substantively significant food-related delineation. Moreover, using samples from arbitrarily designated neighborhoods and large areal units may create spurious relationships as a result of the ecological fallacy and boundary issues. The ecological fallacy argument stems from the fact that we are inferring characteristics of smaller areal units or individuals from aggregate data available at the census tract level. For example, the number of grocery stores, median income, race, or any other aggregate statistic at the census tract level may be biased and may not accurately reflect what is truly happening in the smaller geographical neighborhoods within the census tract. In addition, individuals do not confine their retail activities to the census tract where they live. It may, in fact, be more convenient to shop at a grocery store in another census tract for someone who lives close to the border of that tract. The second shortcoming is that these studies do not account for spatial correlation, which may bias the effect of grocery store access on healthy eating and BMI.

Some studies have used disaggregated, individual-level data to study the relationships between access to various retailers, dietary choices, and health outcomes. Jeffery et al. (2006) studied a set of survey respondents in Minnesota and linked the frequency of consumption of food from fast food restaurants to health outcomes. They found that BMI tended to increase with increasing frequency of fast food meals. However, there was no significant relationship between proximity to fast food restaurants and either the consumption of fast food meals or higher BMI. Rose and Richards (2004) used data from the federally funded Food Stamp Program to assess the impact of retail access on fruit and vegetable consumption. This study differed from the preceding ones in that the authors had information on the actual fruits and vegetables consumed by Program participants as well as information on the retail outlets where they purchased their food. They used the distance and travel time to a store to quantify the ease of access, and found that distance matters when choosing to consume healthy food at home. Although both of these studies were novel in their use of individual data to study this problem, neither accounted for the unobserved environmental and social factors that affect people's eating choices.

The relationship between food consumption and health outcomes is complex. Previous research has shown that food consumption decisions are affected by food availability. More recent research on the relationship between obesity and social networks has been mixed. It has been suggested that obesity can be spread through social networks. Using data from the Framingham Heart Study over a period of 30 years, Christakis and Fowler (2007) found that individuals were far more likely to be obese if their friends and family were also obese. They maintained that friends, colleagues, and family affected a person's perceptions of weight and their eating habits. In other words, prevailing norms in a person's social network about how much to eat, exercise, and what constitutes an appropriate weight affect our decisions on food choices, physical activity, and body image (Blanchflower et al., 2009). Cohen-Cole and Fletcher (2008), using a national sample of adolescents from the Add Health data, found no evidence that obesity spreads through social networks. Although our study does not try to explicitly address or identify the social network aspect of the obesity epidemic, our spatial econometric modeling approach allowed us to model the potential effect of social networks, at least to the extent that network effects can be proxied by the salient characteristics and the behavior of neighbors.

The newest generation of studies (Dunn, 2008; Anderson and Matsa, 2009; Currie et al., 2009) has used more rigorous statistical methods to examine the relationship between obesity and the food landscape. Dunn (2008) and Anderson and Matsa (2009) employed instrumental variable techniques to account for the endogeneity of store locations. If stores are located near obese people, the use of a naïve least-squares estimator introduces bias in the effect of proximity of fast food restaurants on BMI. Both authors therefore utilized a feature of the built environment, specifically the proximity of highway interstate exits, as an instrumental variable for restaurants. Although these studies were methodologically similar, they used data at differing levels of geography for different populations. The Dunn (2008) research was based on county-level data and a sample of densely populated counties. Anderson and Matsa (2009) used individual-level data for a sample of rural counties and actual distances from the centroid of an individual's town of residence to an interstate highway. The results from these studies were, however, inconclusive. Anderson and Matsa (2009) found no effect of the availability of restaurant food on obesity, whereas Dunn (2008) found that an increase in the mean number of fast food restaurants in a county led to increases in BMI.

The study by Currie et al. (2009) is perhaps the most similar to ours in terms of the disaggregation of the data and the level of detail used to characterize the food landscape. They used individual data from natality records linked to actual distances to fast food restaurants to study the effects of fast food availability on the BMI of pregnant women. They introduced both individual-level and zip-code-level fixed effects to control for unobserved heterogeneity that may be correlated with proximity to fast food restaurants. Currie et al. (2009) found that the density of fast food restaurants within a 0.5-mile radius of a person's residence affected the amount of weight gain in pregnant women.

The present paper is laid out as follows. First, we outline the research methods and the data that we used in our study. Subsequently, we present empirical estimates of the effect of access to fast food and chain grocers on BMI using models that account for both the sorting of people into neighborhoods and the spatial spillover effects across people. Using the results from our models, we then simulate the marginal effects for two policy experiments that are of interest to policymakers. The first experiment examines the effect of setting a density limit on the number of fast food restaurants in high-density fast food areas. The second policy experiment increases the number of chain grocers in neighborhoods that are particularly vulnerable to the obesity epidemic.

3 Econometric Model

We developed a reduced-form production function for health in which health is measured by BMI. To operationalize this model, we assumed that the BMI of individual i living in community j (or more generally, at location j, identified by coordinates x and y) is a function of food prices:

$$h_{ij} = p'_{ij}\gamma_1 + \varepsilon_{ij}, \tag{7}$$

where h_{ij} is the health outcome of interest (measured as BMI), p'_{ij} is a vector of individual, location-specific counts for the number of grocers and fast food restaurants within a 0.5-mile radius, and ε_{ij} is an error term. In the absence of price competition, and assuming product homogeneity, p'_{ij} can be seen as a price vector because it is inversely related to the generalized transportation costs incurred to obtain food.

Note that equation (7) differs from the majority of studies that have investigated the effect of the food environment on health because it considers both fast food restaurants and grocery stores in the same model. The previous literature focused on either fast food restaurants or grocery stores without considering that any change in BMI could result from cumulative exposure to both fast food restaurants and grocery stores. Equation (7) is flexible because it allows both these channels to affect BMI.

If individuals select where they want to live based on neighborhood amenities—the food landscape being only one subset of these amenities—the use of equation (7) may yield biased estimation results for γ_1. One way to solve this problem is to control for heterogeneity across individuals as well as across the neighborhoods where they live:

$$h_{ij} = p'_{ij}\gamma_1 + x'_{ij}\beta + n'_{ij}\gamma_2 + \varepsilon_{ij}, \tag{8}$$

where x'_{ij} is a series of individual demographic and behavioral characteristics, including income, and n'_{ij} is a set of neighborhood characteristics for a specific individual. Including controls for both individual and neighborhood heterogeneity can solve the endogeneity problem if the selection is based on observable factors. If selection, as is likely the case, is based on both observable and unobservable neighborhood and individual characteristics, then the estimator of γ_1 in equation (8) may still be biased.

We argue that, in particular, the number of fast food restaurants within an individual's local food landscape of a 0.5-mile radius around the individual's residence is likely to be endogenous. When selecting a neighborhood, individuals may make location decisions based on proximity to fast food restaurants if they value the combination of services and convenience that fast food restaurants provide. The 0.5-mile radius describes the area within which an individual is willing to travel to purchase food quickly and conveniently, presumably under a time constraint. It is much less likely that planned grocery store trips are subject to the same convenience-driven impulses as fast food consumption, and we therefore treat the number of grocery stores as exogenous.

A valid instrumental variable will only affect BMI through its effect on fast food restaurant locations, will not itself be affected by BMI, and will be highly correlated with fast food. On this basis, we assert that the amount of land that is zoned non-residential within a 0.5-mile radius of a respondent's residence is a valid instrumental variable for fast food. Utilizing zoning maps for Indianapolis, we constructed this measure by calculating the proportion of non-residentially zoned property within a 0.5-mile radius of where a person lives. This created individualized zones for each person in our sample. The variable proportion of the area zoned non-residential within a 0.5-mile radius of where a person lives is highly correlated with the density of fast food restaurants within that distance.

Arguably, the individual observations are correlated across space. This may be caused by, for instance, the existence of social networks that are partly formed on the basis of spatial proximity, or by shared local environmental characteristics among individuals living in proximate neighborhoods. Extending (8) to allow for both types of (spatial) correlation, and including all individuals in the sample, we obtain (in matrix notation):

$$h = \lambda Wh + P\gamma_1 + X\beta + N\gamma_2 + \varepsilon, \varepsilon = \rho W\varepsilon + \mu, \tag{9}$$

where P, X and N are matrices containing individual, location-specific counts for the number of grocers and fast food restaurants within a 0.5-mile radius, individual demographic and behavioral characteristics, and neighborhood characteristics, respectively. The $(n \times n)$ matrix W defines who is a neighbor of whom by means of values of either 0 (not neighbors) or 1 (neighbors), and μ is an error term that is assumed to be independently distributed. The spatial weights matrix is typically standardized so that the sum of each row equals 1, which implies that the spatially lagged dependent variable, Wh, contains the average h value of the neighbors. Erroneously omitting spatial correlation creates both bias and inefficiency in the estimated parameters (Anselin, 2006).

The spatially explicit version presented in equation (9) has several interesting features with important implications for the policy recommendations that can be derived from our model. The model in (9) contains a spatially lagged dependent variable. In order to see how this affects the interpretation of our model, it helps to rewrite (9) as:

$$h = (1 - \lambda W)^{-1}[P\gamma_1 + X\beta + N\gamma_2 + (I - \rho W)^{-1}\varepsilon], \tag{10}$$

where $(1 - \lambda W)^{-1}$ is a spatial multiplier. The spatial multiplier can be written as an infinite power series, $I + \lambda W + \lambda^2 W^2 + ...$, where W contains the neighbors of an individual, W^2 the neighbors of the neighbors, and so forth (Anselin, 2003). Effectively, our model implies that a person's BMI is not only determined by his or her own characteristics in terms of P, X, and N, but also by a person's location in terms of the average values of P, X, and N of the neighbors, the neighbors of the neighbors, and so on. For instance, if vigorous physical activity has a negative effect on BMI, an environment in which your neighbors are also active strengthens this effect. The spillover and feedback effects may be due to social network effects (e.g., imitation behavior, peer effects) or to shared neighborhood characteristics (e.g., availability of a park), and they will follow a smooth distance decay pattern across space. Equation (10) is very general because it not only accommodates spatial spillover and feedback effects in terms of BMI, but also simultaneously allows for the effects of spatially correlated omitted variables to be part of the error term.

In terms of estimation, the specification in equation (10) is not entirely straightforward, because the spatially lagged dependent variable is obviously endogenous, and the spatially correlated errors create a non-spherical error variance-covariance structure. We follow the estimation theory for spatial ARAR models developed recently by Kelejian and Prucha (2007). They propose a combination of instrumental variables and generalized method of moments techniques. As outlined in Arraiz et al. (2008), the estimation procedure comprises a series of steps. In the first step, we use the spatial two-stage least-squares or instrumental variables estimator to estimate equation (9), ignoring the spatially correlated errors.[2] Subsequently, we used the estimated residuals of the first step in a spatial general moments estimator to obtain an estimate of ρ. With that estimate in hand, we applied a Cochrane-Orcutt transformation and re-estimated the model with a spatial two-stage least-squares estimator. The asymptotic variance-covariance matrix for this estimator was derived by Arraiz et al. (2008) under the assumption that the errors are heteroskedastic.

Finally, the specification in equation (10) shows that the marginal effects of policy-induced changes in the explanatory variables are not simply equal to the associated estimated coefficients. Effectively, the marginal effects depend on the spatial location of the people that are directly affected by the policy-induced change, and by the resulting spillover and feedback effects among individuals. For example, the effect on BMI of improving the food landscape through "health zoning" no longer solely depends on the strictness of the zoning policy measure alone, but also on the location where the zoning is implemented. Pace and LeSage (2007) provide details about the computation of marginal effects and statistical inference procedures.

4 Data and Research Design

We gathered the data for this study from a variety of sources. The individual health data were obtained from the Adult Obesity Needs Assessment telephone survey conducted by the Marion County Health Department from February through June 2005. These self-reported data include age, sex, education, income, labor force participation, and physical activity level, along with weight and height information. Unlike most health surveys, these data also included the location of the respondent's home.[3]

[2] We have followed standard practice and instrumented the spatially lagged dependent variable with the first- and second-order spatial lags of the explanatory variables. In addition, we replaced the fast food variable by the instrumental variable for the proportion of the area zoned as non-residential within a 0.5-mile radius.

[3] The data are geo-masked and provide the self-reported intersection closest to home rather than the exact coordinates. In processing this intersection data, we geocoded the data using projected UTM NAD 1983 16N coordinates with units in meters as the coordinate system, and perturbed each x and y coordinate randomly to produce a new set of coordinates within a circle with a 100-m radius from the original point. This avoided the

The food landscape data were obtained from the Marion County Health Department's health safety inspection records in 2005. These records included the name and location of all food retailers in Marion County. Using this data, it was possible to construct a representation of the "food landscape" for the county by geocoding the addresses both automatically and manually using ArcMap 9.1. Retail food stores were classified into four categories (large chain grocer, small grocery store, convenience store, and specialty store), of which we used the large chain grocer data. Restaurants were classified into two categories: fast food and "sit down". Fast food restaurants were defined based on the North American Industry Classification System (NAICS) definitions of Limited-Service Eating Places by the U.S. Census Bureau. Data on neighborhood characteristics were based on geocoded crime data (Indianapolis Metropolitan Police Department, 2007) and zoning regulations (Indiana University, 2008).

The unique level of geographic detail in this data allowed us to create individual-specific landscapes for food, criminal activity, and zoning. Using the x and y coordinates for each respondent and the locational coordinates of food retailers, crimes committed, and zoning information, we calculated distance and density measures for each respondent within a 0.5-mile buffer of where they lived. We chose this buffer because empirical data suggests that people in the U.S. do not travel distances further than this for shopping (Agrawal and Schimek, 2007). Other studies have also used a 0.5-mile buffer as the food market radius in urban areas (Rose and Richards, 2004). In addition, we used the geographical scale of Indianapolis and sensitivity analysis to test different measures of market diameter and determine the optimal market radius. Based on that analysis, a 0.5-mile buffer seemed most appropriate for this urban environment.

Figure 1 shows an example of the resulting spatial data. This localized view outlines explicitly how we defined individualized markets for food for each person. We have chosen two people arbitrarily and drawn circles of a 0.5-mile radius around each one. In both cases, there are no chain grocers within a 0.5-mile radius of where the individuals live. In the second case, there is one fast food outlet within that distance. We repeated this exercise using the data on crime to obtain a count of the number of crimes committed within a 0.5-mile radius of each person. Finally, we used GIS zoning maps to estimate the proportion of the land zoned as non-residential within a 0.5-mile buffer of each person's residence.

The sample was restricted to adults between the ages of 21 and 75 years. Implausibly high and low BMI values were deleted from the original sample.[4] The sample consisted of 3550 individuals. The descriptive statistics are reported in Table 1. The sample was predominantly White and female. Approximately 58% of those interviewed were women, and about 30% classified themselves as non-White. The average age was 47 years. Approximately 65% had pursued some post-secondary education. Almost 21% lived in a household that earned an annual income of less than 200% of the Federal Poverty Level (FPL), as defined by the 2003 standards of the U.S. Department of Health and Human Services (Office of the Federal Register, 2003). We used a cutoff of 200% of the FPL income based on research conducted by the National

---

problem of respondents with identical coordinates, which would create operational difficulties in most software environments because of unidentifiable neighbor assignments. We verified that there were no apparent biases in the spatial sampling design. The sampling probability of individuals by census tract was generally between 0 and 1%.

[4] Here, we defined an implausible BMI as a value greater than 70 or less than 16. In total, we removed 156 individuals from the sample on this basis.

Academy of Sciences, and later augmented by the National Center for Children in Poverty, which suggested that a household actually requires about twice the FPL to meet its basic needs (Cauthen and Fass, 2008).

The behavioral variables used in this analysis were weekly physical activity, physical activity on the job, and smoking habits. On average, respondents engaged in vigorous physical activity 3 days per week. Here, vigorous physical activity is defined as an activity of at least 10 minutes that requires harder than normal effort (e.g., heavy lifting, aerobics, or fast cycling). Just over 41% of the respondents reported that their job keeps them physically active and 26% of the respondents currently smoke.

In terms of the "food landscape", most respondents did not have a large chain grocer within their immediate neighborhood. On average, there were at least two fast food restaurants within a 0.5-mile radius of where a respondent lived.

The neighborhood variable that we considered to affect health outcomes was the number of serious crimes committed within a 0.5-mile radius of where a person lives. Here, a serious crime was defined as an attempted or accomplished rape, homicide, robbery of a residence, or assault on a person. The mean number of serious crimes committed within a 0.5-mile radius of a respondent's residence was 41 per year. Since people will choose a neighborhood based on amenities such as safety, we included this variable in our regression to control for residential choice.

In order to assist with the visualization of individual data points and to operationalize the spatial econometric model, and specifically the spatial weights matrix, we transformed the individual point data shown in Figure 1 into Thiessen polygons. This transformation assigns every point in space to the nearest point for which an actual observation is available, and creates artificial areal units that enable neighbors to be determined on the basis of contiguity. For the spatial regressions we used first- and second-order queen contiguity to define who is a neighbor of whom. Effectively, this implies that individuals are considered neighbors if their Thiessen polygons share either a common border or a vertex. Including second-order neighbors (i.e., neighbors of the neighbors) meant that on average, each individual had approximately 20 neighbors.[5]

5 Empirical Results

We started our analysis by estimating BMI using equations (7) and (8). In these formulations, BMI is assumed to be a function of individual demographic, behavioral, and neighborhood characteristics. The results from ordinary least-squares estimation of these models with White-adjusted standard errors, are provided in Table 2, columns (1)–(3). Column (1) reports the results of a naïve model that only accounts for the food environment. These results suggest that each additional fast food restaurant within the specified radius increases BMI by 0.06 points, whereas each additional large grocery store decreases BMI by 0.30 points. Only the

[5] Alternatively we could have used a distance metric. However, for everybody to have at least one neighbor, the cut-off distance would have to be relatively high because sampling is rather sparse in the rural outskirts of Marion County. This would result in people in the city center having a disproportionately large number of neighbors.

effect of chain grocers was significantly different from zero at the 5% level. When individual heterogeneity is controlled for, as outlined in equation (8), the magnitude of the associations between fast food and large grocers changes, but the results are no longer significant for both fast food restaurants and grocery stores, as shown in column (2) of the table. The demographic variables included in this model have the predicted sign and significance reported in the literature (Mokdad et al., 2003). For example, BMI is negatively correlated with education and income, but positively correlated with being non-White. There was a quadratic relationship between BMI and age, and increased physical activity and smoking both decreased BMI. In column (3), we provide estimates based on equation (8), in which we attempted to control for sorting using an observable neighborhood variable, the amount of crime within a 0.5-mile radius of a person's residence. We conjectured that the choice of where to live, measured in our analysis by the presence of fast food restaurants, would be strongly correlated with the level of crime. Although crime may not "belong" in the BMI equation, including this neighborhood characteristic as a proxy can potentially absorb the bias caused by the endogeneity of fast food restaurants. However, adding this control did not result in any change in the magnitude or significance of the food landscape variables. The number of serious crimes committed or attempted within a 0.5-mile of where an individual lives also had no statistically significant effect.

In addition to the selection effect, it is likely that observations are not independent across space because of either observable or unobservable social network ties or shared observable or unobservable neighborhood characteristics across individuals living in proximate neighborhoods. To explore this issue, we first investigated a simple *ad hoc* specification in which we estimated a random-effects model that assumes the composite error term can be decomposed into a neighborhood effect measured at the census tract level and an additional random disturbance so that $\varepsilon_{ij} = \omega_{j \in J} + \mu_{ij}$, where J represents the census tracts. The random effects model, estimated with standard errors allowing for heteroskedasticity and clustering at the census tract level, is presented in column (4) of Table 2. These results are not notably different from the previous ordinary least-squares estimates. We also note that the random effects specification, which is popular in applied research (e.g., Morland et al., 2002, 2006), does not really account for spatial correlation based on a distance decay pattern across individuals. Arguably, the errors are correlated within census tracts, but they are uncorrelated between census tracts. The correlation structure is therefore due to the incorporation of spatial heterogeneity by means of random neighborhood effects rather than spatial dependence among observations due to proximity. To explore this issue further we abandoned the *ad hoc* specifications and the assumption that markets are defined at the census tract level, and we undertook a more systematic analysis of the spatial structure of the data.

We used Lagrange Multiplier tests for spatial dependence to determine whether the suggested ARAR specification in equation (9) is potentially adequate. The full results of the spatial diagnostics are presented in conjunction with the ordinary least-squares results for the extended specification in column (3). Comparison of the Lagrange Multiplier statistics and significance levels was conducted as suggested by Anselin et al. (1996). The spatial diagnostic tests suggest an ARAR spatial process model, in addition to the heteroskedasticity tests that suggest the model specification should allow for a general form of heteroskedasticity.

As discussed earlier, there is reason to believe that the food environment variable is still endogenous in equation (8). Empirical evidence suggests that people usually drive or take public transit to grocers even if they live in low-income neighborhoods (Clifton, 2004). Consequently, although access and distance may enter into the price of food, the actual time cost of going an extra mile for grocery food is lower. On the other hand, when people prefer to eat quickly, they will go to the closest restaurant they can find. As such, distance is very important when choosing among restaurants. Since we use an instrumental variables estimator to account for the endogeneity of the spatially lagged dependent variable, we can easily extend the instrumentation to the number of fast food restaurants. We therefore instrumented the number of fast food restaurants using individualized residential zoning patterns.[6] There is a strong positive correlation between non-residential zoning within a 0.5-mile radius of where a person lives and the number of fast food restaurants in that same radius.

Estimation of the spatial ARAR model was conducted using code that we wrote to implement the general moments and instrumental variables techniques described above using the R software, version 2.8.0.[7] The results of the spatial ARAR model, as outlined in equation (9), are reported in column (5) of Table 2. There are at least three interesting findings from this specification. First, the effects of proximity to fast food restaurants and to chain grocers are both significant. As explained earlier, the estimated coefficients cannot be interpreted as partial effects since they are affected by the spatial multiplier, as outlined in equation (10). In order to estimate partial effects, we must select particular geographical areas where a health policy will be implemented and simulate the effects of the policy change on the individuals. This exercise will be conducted in the next section. The sign of the estimated coefficients is, however, still relevant, and shows that increasing access to fast food increases BMI, whereas increasing access to chain grocers decreases BMI.

The second interesting finding is that the value of ρ is significantly different from zero, suggesting that it is important to account for unobservable factors that potentially work through social network ties or neighborhood effects. The negative sign of ρ notwithstanding, the spillover and feedback effects of the unobserved spatially correlated effects are actually positive, because ultimately the sign of these effects is determined by the interaction of the two spatial multipliers distinguished in equation (9), and hence by ρ and λ simultaneously.

Finally, the statistically significant and positive value of λ suggests that the BMI of an individual's neighbors indirectly affects their own BMI. The reduced form in equation (9) shows that this implies that an individual's BMI is also determined by the exogenous characteristics of the individual's neighbors. The positive sign of λ suggests that the explanatory variables of the neighbors will affect an individual's BMI in the same direction as the parameter estimate β. For instance, the estimated impact of being non-White is to increase BMI. The spatial pattern implicit in the reduced form reveals that the effect is increased to the extent that the neighbors are also non-White. Similarly, the effect of vigorous physical activity on BMI is beneficial (i.e., a decrease). This effect is reinforced, either through network effects (e.g., peer pressure) or

[6] We also note that the geo-masking of individual home addresses (see footnote 3) creates an error-in-variables problem that most likely has the severest impact on the fast food access variable. As a result, the ordinary least-squares and random effects estimators suffer from attenuation bias, which implies that the coefficients are biased towards zero. This provides another motivation for the instrumental variables approach.

[7] R is available at http://www.R-project.org (R Development Core Team, 2008). The R code used to implement this estimator is available from the authors upon request.

shared neighborhood characteristics (e.g., availability of a park), if the individual lives among neighbors who perceive exercising as the norm.

6 Policy Experiments

Municipal ordinances have traditionally been used to limit the number of liquor stores or pornographic outlets within cities. More recently, however, cities have looked at zoning regulations as a means to improve health outcomes, especially in disadvantaged neighborhoods. Some cities have targeted the "bad" aspects of restaurants by instituting mandatory posting of nutritional information at fast food restaurants, moratoriums on the openings of new fast food restaurants, or outright bans on the use of unhealthy trans fats in food preparation. Other municipalities have focused on the "good" by providing increased access to healthy foods in neighborhoods that formerly lacked such access.

The previously estimated parameter values can now be used to simulate two policy experiments. For the "bad" scenario, we simulated the effect of restricting the density of fast food in areas that are currently overserved. In this experiment, we considered an area to be overserved if it had more than six fast food restaurants per km^2. In the "good" experiment, we investigated the implications of providing better access to healthy food by identifying areas that have more than 40% of the population below the Federal Poverty Level and more than 40% with less than a high school diploma.

In both the "good" and the "bad" scenarios, we used data gathered as part of the SEDAC project developed at Columbia University.[8] The SEDAC project provides data on Marion county at the grid cell level (1 km $\times$ 1 km rasters) with sociodemographic and economic characteristics attached to each cell. We used the geographical information on food establishments to attach the density of grocery stores and fast food restaurants to each of these grid cells, and subsequently used the characteristics of the grid cells to identify our Policy Implementation Areas for the "good" and "bad" policies. We used the indirect approach of first selecting Policy Implementation Areas on the basis of an area-related criterion and subsequently assessing the impact of the policy measure on individuals living within and beyond the Policy Implementation Area to mimic the actual policy design process. Zoning ordinances are primarily targeted at areas, and only indirectly at people.

There were 563 fast food restaurants in our dataset. In the first policy scenario, we randomly removed one fast food restaurant from high-density fast food areas (i.e., those with six or more fast food restaurants). In total, this scenario decreased the number of fast food restaurants by 15. When we recalculated the number of fast food restaurants within a 0.5-mile radius of each person, we found that 178 people were directly affected by the policy change (i.e., the number of fast food restaurants in their local food environment decreased).

The marginal effects of restricting fast food density are presented in Figure 2 and summarized in Table 3. Since the marginal effects of the policy change are different for each individual, depending on their geographic location, we reported the average direct, indirect, and total effects of restricting access to fast food restaurants. The direct effect reported in column (1) of Table 3 is the average partial effect on the individuals who are directly affected by the policy

[8] Shapefiles for the SEDAC project are available from http://sedac.ciesin.columbia.edu/usgrid/.

experiment. The mean direct effect of reducing their local food environment by one fast food outlet amounts to a –0.22 point change in their BMI. Due to the working of the spatial multiplier, all individuals in the sample are ultimately affected through spillover effects. The average indirect effect is fairly small, but statistically significant. On average, the total impact of the policy scenario that reduces access to fast food restaurants decreases the population's BMI by 0.04 points.[9]

The spatial distribution of the average total effects is displayed in Figure 2 by means of Thiessen polygons drawn around each of the individual data points. Hence, each area represents an individual from our sample at their residence's location. It is clear from the map that the change in policy has a different effect on each person based on how far they live from the neighborhood where the policy is introduced. Moreover, in our model, the effect of a new policy that attempts to change the food environment in one location will have ripple effects across space and will thereby affect the BMI of individuals living in neighboring locations. As a result, even individuals who do not experience a policy change in the number of fast food restaurants or grocery stores in their immediate local food environment (i.e., the 0.5-mile buffer) are affected by the spatial distribution of these restaurants through spillover effects. The spatial heterogeneity of the partial effects and the policy diffusion process that occurs is evident from the shading patterns in Figure 2. It is also clear from examining this map that the actual location of the Policy Implementation Area is important because the marginal impacts change according to where the policy is implemented.

In the second policy scenario, we increased access to healthy foods by locating chain grocery stores at the geographical centroid of disadvantaged neighborhoods. This policy experiment increased the number of chain grocery stores from 94 to 107, and directly increased access to these stores for 74 people in our sample. The average marginal effects for increasing access to healthy foods are displayed in Table 4. All marginal effects were significant at the 5% confidence level. The average direct effect for those individuals who live in the Policy Implementation Area equaled –0.58. The average indirect effects were –0.04, indicating that people who were located in neighborhoods where the policy change was not implemented would also benefit from more chain grocery stores in proximate neighborhoods. The average total effect of increasing access to healthy food was –0.05, which is comparable in magnitude to the average total impact of the fast food reduction scenario. However, Figure 3 clearly shows that the spatial distribution of the impacts clearly differs between the two policies.

7 Conclusion

This study differs from previous research because of the spatially explicit nature of our analysis. Our results show how the magnitude of the marginal effects depends on the exact geographical location of the individuals for which marginal (policy) effects are determined. We also show how spatial econometric methodology can be used as a tool to inform local policymakers who want to understand how specific neighborhood-based policies can affect the health of the local

[9] The average direct and indirect impacts add up to the average total impact, but in the main text, we have averaged the direct impact over the individuals for whom a non-zero change in their local food environment occurs instead of all individuals in the sample. See also the footnote to Table 3.

community, and to what extent these policies will have spatial spillovers that affect neighboring communities.

Our results suggest that past attempts to explore the relationship between the food landscape and obesity have been hindered by issues related to sample selection and endogeneity. In addition, the spatial nature of the data should not be ignored, because erroneously failing to account for spatial dependence creates bias and inefficiency in the estimation process, possibly leading to erroneous conclusions regarding the magnitude and the statistical significance of the impact of the food environment on obesity. The incorporation of spatial heterogeneity also appears to be a crucial factor when it comes to designing spatial policy scenarios to combat the obesity epidemic.

When sample selection effects and spatial dependence are accounted for in our estimation, the marginal effects of changing access to either fast food restaurants or chain grocers is small but statistically significant. Translating our results from BMI to a more widely understood measure such as weight, the average total effect of improving grocery store access in targeted disadvantaged neighborhoods would be a decrease of less than 0.25 lb, calculated using a population of 5'6" women who weigh 160 lb. Restricting access to fast food has an average total effect of the same magnitude. For individuals who live in the area where an additional grocery store is placed, however, the effects would be much larger, and would translate into a 3-lb decrease in weight. For fast food restaurants, the impact would be a 1-lb decrease for the people who live in the areas where fast food is restricted.

Our finding are in contrast with those of Anderson and Matsa (2009), who found no evidence of a causal link between restaurants and obesity, and much smaller than those of Currie et al. (2009), who found that the presence of a fast food restaurant within a 0.5-mile radius of the residence of a sample of pregnant women resulted in a 2.5% increase in the probability of gaining more than 40 lb.

A cornerstone of urban renewal is the development of the kind of infrastructure that creates and maintains prosperous neighborhoods. Often, this includes incentives to attract large chain groceries or to increase the healthy food offerings of existing small grocery and convenience stores. The analysis conducted in the present study demonstrates a spatially explicit way to investigate the impact of policy-induced changes to either the "good" or the "bad" aspects of the food landscape. Future work in this area could examine the role that food prices have on shopping behavior, as this would contribute to our understanding of what is offered in existing food retail restaurants in disadvantaged areas and to what extent policy incentives can be used to induce retailers to improve their healthy food offerings.

References

Abdollah, T. (2007). A strict order for fast food. *Los Angeles Times, September 10.*

Agrawal, A. and P. Schimek (2007). Extent and correlates of walking in the USA. *Transportation Research Part D 12*(8), 548–563.

Anderson, M. and D. Matsa (2009). Are restaurants really supersizing America? *Available at the SSRN eLibrary Paper No. 1079584.*

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review 26*(2), 153–166.

Anselin, L. (2006). Spatial econometrics in T.C. Mills and K. Patterson (Eds.). *Palgrave Handbook of Econometrics: Volume 1, Econometric Theory*, Basingstoke: Palgrave Macmillan, 901–969.

Anselin, L., A. Bera, R. Florax, and M. Yoon (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics 26*(1), 77–104.

Arraiz, I., D. Drukker, H. Kelejian, and I. Prucha (2008). A spatial Cliff-Ord-type model with heteroskedastic innovations: Small and large sample results. *Available at the SSRN eLibrary Paper No. 1313030.*

Blanchflower, D., A. Oswald, and B. Van Landeghem (2009). Imitative obesity and relative utility. *Available as NBER Working Paper, No. W14377.*

Block, J., R. Scribner, and K. DeSalvo (2004). Fast food, race/ethnicity, and income: A geographic analysis. *American Journal of Preventive Medicine 27*(3), 211–217.

Burdette, H. and R. Whitaker (2004). Neighborhood playgrounds, fast food restaurants, and crime: Relationships to overweight in low-income preschool children. *Preventive Medicine 38*(1), 57–63.

Cauthen, N. and S. Fass (2008). *Measuring poverty in the United States, Fact sheet.* Columbia University, Mailman School of Public Health: National Center for Children in Poverty.

Chou, S., M. Grossman, and H. Saffer (2004). An economic analysis of adult obesity: Results from the behavioral risk factor surveillance system. *Journal of Health Economics 23*(3), 565–587.

Christakis, N. A. and J. H. Fowler (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine 357*(4), 370–379.

Chung, C. and S. Myers (1999). Do the poor pay more for food? An analysis of grocery store availability and food price disparities. *Journal of Consumer Affairs 33*(2), 276–296.

Clifton, K. (2004). Mobility strategies and food shopping for low-income families: A case study. *Journal of Planning Education and Research 23*(4), 402.

Cohen-Cole, E. and J. Fletcher (2008). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics 27*(5), 1382–1387.

Cummins, S. and S. Macintyre (2006). Food environments and obesity—Neighbourhood or nation? *International Journal of Epidemiology 35*(1), 100–104.

Currie, J., S. DellaVigna, E. Moretti, and V. Pathania (2009). The effect of fast food restaurants on obesity. NBER Working Paper No. 14721.

Cutler, D., E. Glaeser, and J. Shapiro (2003). Why have Americans become more obese? *Journal of Economic Perspectives 17*(3), 93–118.

Dunn, R. A. (2008). Obesity and the availability of fast-food: An instrumental variables approach. *Available at the SSRN eLibrary Paper No. 989363*.

Eid, J., H. Overman, D. Puga, and M. Turner (2008). Fat city: Questioning the relationship between urban sprawl and obesity. *Journal of Urban Economics 63*, 385–404.

Indiana University (2008). *2004 Marion County Vector Data*. Geographic Information System Resource: Indiana Spatial Data Portal.

Indianapolis Metropolitan Police Department (2007). *Uniform Crime Reporting*. Indianapolis Metropolitan Police Department, updated February 20, 2008 (accessed online May 5, 2008).

Jeffery, R., J. Baxter, M. McGuire, and J. Linde (2006). Are fast food restaurants an environmental risk factor for obesity. *International Journal of Behavioral Nutrition and Physical Activity 3*(2).

Kelejian, H. and I. Prucha (2007). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *University of Maryland, forthcoming in Journal of Econometrics*.

Maddock, J. (2004). The relationship between obesity and the prevalence of fast food restaurants: State-level analysis. *American Journal of Health Promotion 19*(2), 137–143.

Mair, Julie Samia, M. W. P. and S. P. Teret (2005). The use of zoning to restrict fast food outlets: A potential strategy to combat obesity. *The Center for Law and the Public's Health at Johns Hopkins and Georgetown Universities*.

Mobley, L., E. Finkelstein, O. Khavjou, and J. Will (2004). Spatial analysis of Body Mass Index and smoking behavior among WISEWOMAN participants. *Journal of Women's Health 13*(5), 519–528.

Mokdad, A., E. Ford, B. Bowman, W. Dietz, F. Vinicor, V. Bales, and J. Marks (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA 289*(1), 76–79.

Moore, L. and A. Diez Roux (2006). Associations of neighborhood characteristics with the location and type of food stores. *American Journal of Public Health 96*(2), 325–331.

Morland, K., A. Diez Roux, and S. Wing (2006). Supermarkets, other food stores, and obesity: The atherosclerosis risk in communities study. *American Journal of Preventive Medicine 30*(4), 333–339.

Morland, K., S. Wing, and A. Diez Roux (2002). The contextual effect of the local food environment on residents' diets: The atherosclerosis risk in communities study. *American Journal of Public Health 92*(11), 1761–1768.

Morland, K., S. Wing, A. Diez Roux, and C. Poole (2002). Neighborhood characteristics associated with the location of food stores and food service places. *American Journal of Preventive Medicine 22*(1), 23–29.

Office of the Federal Register (2003, February 7). *Federal Register of the United States*, Volume 68, Chapter Federal Poverty Level Standards, pp. 6456–6458. National Archives and Records Administration.

Pace, R. and J. LeSage (2007). Interpreting spatial econometric models. Working Paper, Texas State University-San Marcos.

Philipson, T. and R. Posner (2003). The long-run growth in obesity as a function of technological change. *Perspectives in Biology and Medicine 46*(3), S87–S107.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rose, D. and R. Richards (2004). Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Health Nutrition 7*(8), 1081–1088.

Zenk, S., A. Schulz, B. Israel, S. James, S. Bao, and M. Wilson (2005). Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in metropolitan Detroit. *American Journal of Public Health 95*(4), 660–667.

Table 1: Summary statistics[a]

Variable	Mean	Std. dev.
Body Mass Index	27.675	6.099
Obese (proportion)	0.272	0.445
Overweight (proportion)	0.355	0.479
Healthy weight (proportion)	0.362	0.481
No. fast food restaurants (0.5-mile radius)	2.033	2.860
No. large grocery stores (0.5-mile radius)	0.354	0.722
Non-White (proportion)	0.303	0.460
Female (proportion)	0.583	0.493
Age (years)	47.011	14.231
Less than 200% of the FPL (proportion)[b]	0.207	0.405
More than high school education (proportion)	0.645	0.479
Vigorous physical activity (days per week)	2.682	2.348
Physically demanding job (proportion)	0.413	0.492
Smoker (proportion)	0.259	0.438
No. serious crimes per year (0.5-mile radius)	41.406	43.860
Proportion zoned as non-residential (0.5-mile radius)	0.322	0.214

[a] Sample size is 3550 observations.

[b] The Federal Poverty Level (FPL) for a family of four in 2003 was $18400 (Office of the Federal Register, 2003).

Table 2: Estimation results[a]

Variable	(1)	(2)	(3)	(4)	(5)
No. fast food restaurants (0.5-mile radius)	0.055	0.056	0.052	0.054	0.201**
	(0.046)	(0.044)	(0.044)	(0.045)	(0.100)
No. large grocery stores (0.5-mile radius)	−0.303*	−0.167	−0.159	−0.156	−0.481*
	(0.181)	(0.175)	(0.176)	(0.182)	(0.283)
Non-White		1.035**	0.972**	0.944**	0.506**
		(0.224)	(0.229)	(0.244)	(0.218)
Female		−0.357*	−0.343*	−0.347*	−0.336*
		(0.204)	(0.204)	(0.194)	(0.186)
Age		0.403**	0.404**	0.405**	0.393**
		(0.048)	(0.048)	(0.047)	(0.046)
Age squared		−0.004**	−0.004**	−0.004**	−0.004**
		(0.000)	(0.000)	(0.0005)	(0.0005)
Less than 200% of the FPL		1.207**	1.158**	1.136**	1.059**
		(0.265)	(0.268)	(0.304)	(0.294)
More than high school		−0.875**	−0.854**	−0.819**	−0.641**
		(0.224)	(0.225)	(0.238)	(0.233)
Vigorous physical activity per week		−0.319**	−0.320**	−0.318**	−0.302**
		(0.044)	(0.044)	(0.044)	(0.043)
Physically demanding job		−0.611**	−0.603**	−0.614**	−0.576**
		(0.213)	(0.214)	(0.209)	(0.207)
Smoker		−1.321**	−1.339**	−1.349**	−1.368**
		(0.233)	(0.233)	(0.228)	(0.219)
No. serious crimes (0.5-mile radius)			0.003	0.003	−0.001
			(0.002)	(0.003)	(0.002)
Constant	27.669**	19.907**	19.777**	19.730**	−0.821
	(0.126)	(1.129)	(1.133)	(1.099)	(2.942)
λ (Wh)					0.753**
					(0.097)
ρ ($W\varepsilon$)					−0.651**
					(0.131)
R^2	0.0002	0.0635	0.0637		
Breusch-Pagan test			248.218**	[0.000]	
Lagrange Multiplier spatial lag			5.269**	[0.022]	
Robust Lagrange Multiplier spatial lag			5.156**	[0.023]	
Lagrange Multiplier spatial error			3.365*	[0.067]	
Robust Lagrange Multiplier spatial error			3.252*	[0.071]	
Lagrange Multiplier spatial ARAR			8.521**	[0.014]	

[a] Columns (1)–(3), ordinary least-squares estimator with White-adjusted standard errors; column (4), random-effects estimator with clustered standard errors; and column (5), spatial ARAR estimator. Standard errors are in parentheses. The probability of falsely rejecting the null hypothesis is labeled with ** and * for $p < 0.05$ and $0.05 < p < 0.10$, respectively. The Breusch-Pagan test has random coefficients as the alternative hypothesis. The spatial diagnostic tests are based on queen first- and second-order contiguity weights matrices, with p values in square brackets.

Table 3: Simulated marginal effects on BMI of restricting access to fast food

	Average direct effect[a]	Average indirect effect	Average total effect
Mean	−0.011	−0.034	−0.045
Standard error[b]	0.000	0.001	0.001
z value	−44.409	−27.302	−31.875

[a] Direct effect averaged over the entire sample. The local food environment of 178 out of a total of 3550 individuals was directly affected. The average direct effect for those individuals was –0.22, based on rescaling with 3550/178. This is the figure used in the main text. The indirect and total effects are averaged over the entire sample, because all individuals are impacted through spillover effects.

[b] The standard errors are calculated based on repeated evaluation of the marginal effects utilizing 100 random draws from a multivariate normal distribution centered on $\hat{\lambda}$ and $\hat{\gamma}$ from the ARAR model and the associated part of the variance-covariance matrix. Values greater than 0.95 for the simulated λ are discarded to avoid singularity.

Table 4: Simulated marginal effects on BMI of increasing access to grocery stores

	Average direct effect[a]	Average indirect effect	Average total effect
Mean	−0.012	−0.042	−0.054
Standard error[b]	0.001	0.004	0.005
z value	−17.080	−9.961	−11.483

[a] Direct effect averaged over the entire sample. The local food environment of 74 out of a total of 3550 individuals was directly affected. The average direct effect for those individuals was –0.58, based on rescaling with 3550/74. This is the figure used in the main text. The indirect and total effects are averaged over the entire sample, because all individuals are impacted through spillover effects.

[b] The standard errors are calculated based on repeated evaluation of the marginal effects utilizing 100 random draws from a multivariate normal distribution centered on $\hat{\lambda}$ and $\hat{\gamma}$ from the ARAR model and the associated part of the variance-covariance matrix. Values greater than 0.95 for the simulated λ are discarded to avoid singularity.

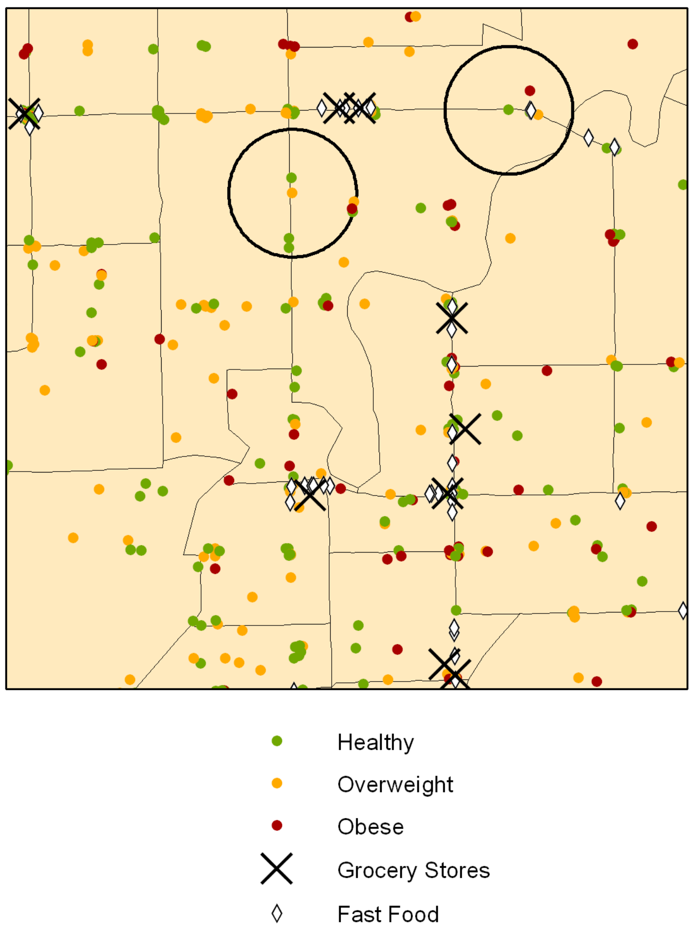

Figure 1: Illustration of an example of the local food environment within a 0.5-mile radius. Circles indicate this radius for two sample individuals.

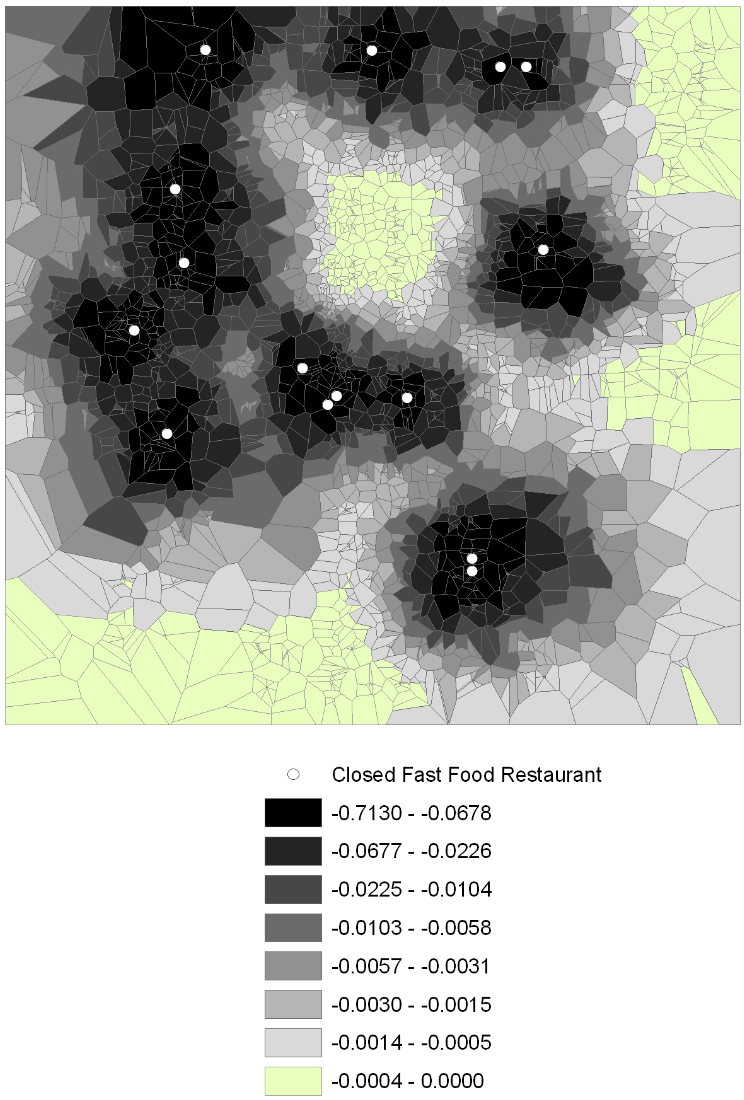

Figure 2: Total marginal effects on BMI of restricting access to fast food.

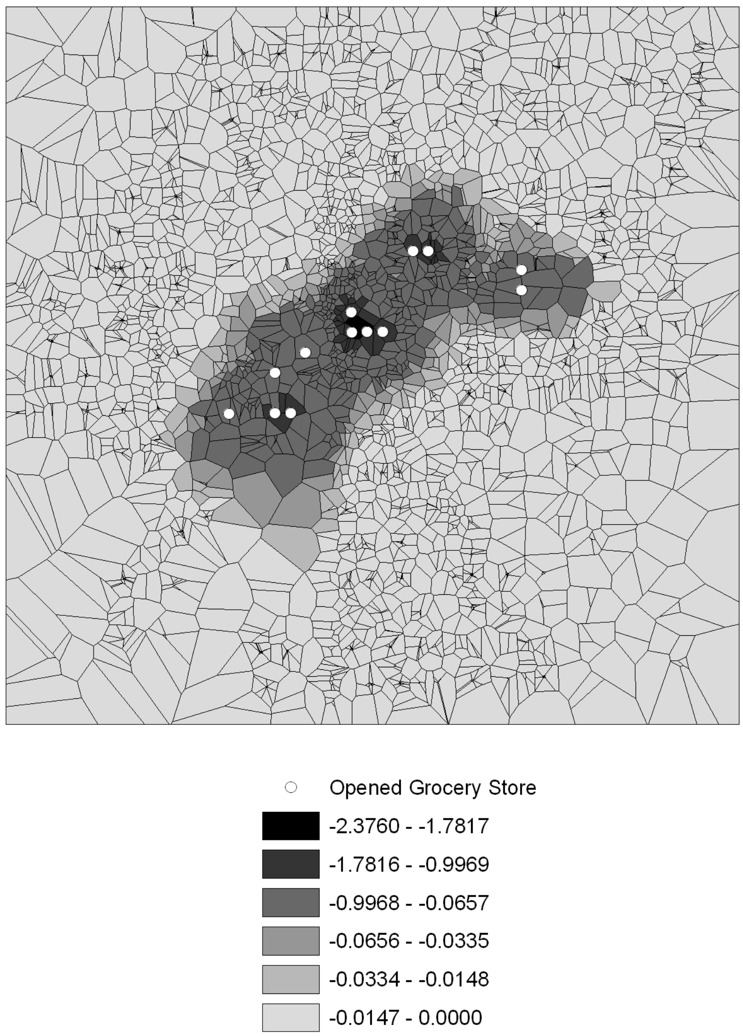

Figure 3: Total marginal effects on BMI of increasing access to grocery stores.