Spatial spillovers, living environment and obesity in France: Evidence from a spatial econometric framework

Céline Bonnet, Cécile Détang-Dessendre, Valérie Orozco, Elodie Rouviere

Spatial spillovers, living environment and obesity in France: Evidence from a spatial econometric framework

Bonnet Céline[*], Détang-Dessendre Cécile[†], Orozco Valérie[‡], Rouvière Elodie[§]

March, 2022

[*]Toulouse School of Economics, INRAE, Université Toulouse Capitole, 1, Esplanade de L'Université 31080 Toulouse, France; celine.bonnet@tse-fr.eu

[†]CESAER, Agrosup Dijon, INRAE, Université Bourgogne-Franche-Comté, 26, bd Docteur Petitjean, BP 87999, 21079 Dijon Cedex France; cecile.detang-dessendre@inrae.fr

[‡]Toulouse School of Economics, INRAE, Université Toulouse Capitole,1, Esplanade de L'Université 31080 Toulouse, France; valerie.orozco@tse-fr.eu

[§]**Corresponding Author**: SADAPT, AgroParisTech, INRAE, Université Paris Saclay, 16, rue Claude Bernard 75231 Paris cedex 05, France; elodie.rouviere@agroparistech.fr, Tel. +331 44 08 72 38. Fax +331 44 08 16 57

Spatial spillovers, living environment and obesity in France: Evidence from a spatial econometric framework

March, 2022

Abstract

In 2019, obesity affected 17% of French adults. In this article, we use a unique data set that combines individual-level health and consumption data with living environment data (food, sports and health amenities). We develop a spatial econometric framework to address French health disparities in obesity prevalence across space. We find that regulations on fast food restaurant locations could be a policy instrument to counter the prevalence of obesity. We also establish the existence of spatial spillovers of sports and medical amenities on obesity. This new evidence points to the need to consider a wider context than just the immediate local environment in the fight against the obesity pandemic.

Keywords: health inequalities, spatial patterns, living environment

1 Introduction

The prevalence of obesity has increased substantially around the world and has had considerable negative economic and health impacts. In France, the prevalence of obesity increased from 5% of the entire population in 1981 to 15% in 2012 (Roche, 2012). This figure has been roughly stable over the past ten years. In 2019, obesity affected 17% of adults.

Many studies have established a strong association between nutrition and certain chronic diseases, such as obesity. The etiology of obesity is multifactorial and involves a combination of individual and external factors. Individual factors include an energy imbalance as well as genetic, epigenetic and metabolic profiles. External factors refer to the living environment. In other words, the factors that cause obesity are related to the individual per se (the host), the choices he or she makes (the vectors: eating, exercising, drinking) and the set of external factors to which he or she is exposed, that is, the environment (Rosengren and Lissner, 2008). Consequently, choices made by individuals or firms are constrained by this environment. Economists study those constrained choices and their social, spatial, and organizational interactions, which might have consequences for the obesity prevalence. The consequences of the environment then need to be addressed by public policies.

In this article, we are interested in the association between obesity and the living environment in relation to the relevant spatial factors. We define the living environment as the place where people live and might have access to food, health and sports amenities.

The originality of our article is threefold. First, we provide an original database that merges individual-level data from a representative survey of French households on the health status of adult people, their individual socioeconomic and demographic characteristics, and their food consumption patterns, with data from an original database on food, health and sports amenities. Second, we analyze the association between obesity and the living environment in France using a spatial econometric framework. We establish the existence of spatial spillovers that might affect health outcomes and inequalities. These spatial spillovers suggest that a geographically broader environmental context might exert an influence on obesity risk in addition to the effects of the more immediate local environment. Third, we contribute to the developing literature that approaches the obesity epidemic as the result of systemic effects between the living environment and individual health.

Environmental factors have been noted as a policy target in reducing obesity by research for more than 20 years. In a large number of disciplines (epidemiology/population health/health geography/economics/etc.), scholars have investigated the associations between urban sprawl and density, food environment, access to physical amenities, and obesity. North American economists have analyzed the association between urban sprawl and obesity. They show that once individual effects are controlled for, the shape of the city and the extent of its urban sprawl might explain the rapid expansion in overweight rates: individuals living in low-

density suburbs are more prone to obesity than are individuals who live in more densely populated suburbs (Ewing et al., 2003; Frank et al., 2004). In other words, the greater the distance one lives from downtown, the higher the probability of being obese. However, other economics research counters this result. Plantinga and Bernell (2007) examine the influence of urban sprawl on obesity and show that individuals who moved to denser locations lost weight. However, such locations are unlikely to be selected by individuals with a high body mass index (BMI). Similarly, Eid et al. (2008) find no evidence that urban sprawl causes obesity. Caspi et al. (2012) propose a comprehensive review of the food environment-diet relationship. They focus on characterization of the food environment, questioning food accessibility across studies. They establish that, despite the large amount of literature, there is moderate evidence on how local food environments influence diet. They argue that metrics and overall definition matter in providing a strong body of evidence of such a link. Currie et al. (2010) question the influence of the exact location of fast food restaurants on children and pregnant women, controlling for endogenous fast food locations. They find a significant effect of proximity to fast food restaurants on the risk of obesity but a different magnitude of the effect for children and adults. Courtemanche and Carden (2011) examine the effects of the entry of new supercenters in the market on BMI and show that the proliferation of Walmart has had an effect on obesity. Allcott et al. (2019) assess nutritional inequalities in the United States using supermarket entry into the market. They reject the hypothesis that supermarkets contribute meaningfully to nutritional inequality.

Other analyses show evidence of the relationship between access to a heterogeneous physical activity environment and health outcomes (see, for example, James et al. (2015), Lachowycz and Jones (2011), Mackenbach et al. (2014), Soga et al. (2017)). To our knowledge, Van Holle et al. (2012) is the first review on European results on the relationship between attributes of the physical environment (e.g., access to shops/services/work, residential density) and physical activity as a health outcome. Led across 27 European countries, they question the applicability of US results in a European context. Ellaway et al. (2006) show that a high level of greenery and a low level of graffiti increase the propensity to be active and thereby decrease the obesity risk.

Feng et al. (2010) provide a systematic review of the epidemiological analyses of the link between health outcomes and neighborhood built environments (physical activity, land use, food environment) and show inconsistent results. Two strong limits have been highlighted to sustain these inconsistencies. First, these studies never consider the multidimensional features of neighborhoods. Second, most of the time, these studies focus on one specific geographic location, specifically the US or UK, overwhelming the potential spatial interactions that might exist.

A new strand of the literature has thus emerged to consider the possibility of heterogeneity of the living environment across space (cities and suburbs) and the interactions between environmental features. This intuition has been sustained through empirical facts. In France, De Saint Pol (2007) establishes that people

living in rural areas have a higher BMI than people living in urban areas. Recently, (Ncd Risk Factor Collaboration, 2019) and Popkin (2019) show that BMI within the rural population has sharply increased and is sustaining the global adult obesity epidemic in developed and low- and middle-income regions. In the US, Befort et al. (2012) observe a higher obesity rate among urban adults than among rural adults. Chen et al. (2019) show i) the existence of an obesity cluster across the US and ii) the variation in geographical disparities in the association between the community food environment and obesity across the US. They also claim that BMI clusters do not follow political boundaries. Mason et al. (2020) adopt a global view of the relationship between health and the living environment and analyze how neighborhood characteristics act together to influence BMI. Their results indicate that space matters. They sustain that the existence of spatial patterns might influence the propensity to be obese. Mason et al. (2021) propose clues on how to address the neighborhood heterogeneity effect on health outcomes across space in the UK. They establish that a wider context than proximity would help to better understand the relationship between the living environment and health. They also claim that administrative boundaries might not be the correct space division. In this article, we analyze how the living environment (e.g., access to food, sports, and health amenities), and its spatial interactions, could affect the obesity rate. We consider a wide definition of neighborhood (food, sports and health) amenities and tackle potential spatial effects through a spatial econometric approach. Moreover, we use a spatial division that takes into account the living area definition and not an administrative boundary. To the our knowledge, this association between the living environment and health status, using spatial econometrics, has been unexplored in France.

2 Materials and Method

In this cross-sectional analysis, we combine individual health and consumption data with living environment data in a single unique data set. All data sources are detailed in Section 5.1 of the Supplementary Material. The health and consumption data provide individual-level information on the weight and food consumption of a representative sample of French households in 2008. The living environment data include variables on food, sports and health amenities. We also provide control variables to account for the income, education level, age and employment status of the French population (in the same year).

2.1 Data

2.1.1 Health and Consumption Data

We use home-scan data from the Kantar Worldpanel company, which provides information on household purchases during 2008 in France for 11,422 households that are representative of the French population.

This database provides information on a detailed set of household characteristics, such as the composition of the household (number of adults and children) and the commune of residence. At the individual level, we have information on age, gender, weight and height, enabling us to compute the BMI for each individual. Concerning household purchases, we observe the categories of food items purchased, their price, their quantities, and their characteristics (brands, flavor, etc.), which represent at-home consumption. We also use the *NutriXConso* database, which links the Kantar home-scan data with the Food Composition Data (de Mouzon and Orozco, 2011). This database provides the nutritional value (micronutrients, macronutrients, and calories) contained in each purchased food item listed in the Kantar database.

2.1.2 Living Environment

We describe the living environment in one unique setting in terms of food, health and sports amenities. To describe the food and health amenities, we use the French Census of Permanent Facilities provided by the French National Institute of Statistics and Economic Studies (INSEE). Data are available at the commune level, which is the smallest administrative subdivision in France. We extract the number of supermarkets (stores larger than 400 m^2) and the number of general practitioners in the 36,186 existing communes in mainland France on January 1st, 2008. We also know the GPS coordinates of the main French fast food restaurant chains (McDonald's, Quick, KFC) and compute the number of such restaurants at the commune level. Considering sports amenities, we use the French Census of Sports Amenities (e.g., number of soccer fields, tennis courts, swimming pools, skate parks, playgrounds, outdoor recreational areas (e.g., ski resorts) and natural sites). We also use the annual census of sports licenses provided by the French ministry in charge of sports. In France, a sport license allows participation in activities organized by the sports federation, in particular competitions, and it provides insurance coverage for damages caused during these activities. Sports licenses measure a consumption variable for official sports practice in clubs.
For communes where some (or all) amenities are missing, we consider the road distance to the closest commune equipped with the considered amenity using the *Odomatrix* software. *Odomatrix* has been developed to help researchers analyze the accessibility of local services (Hilal et al., 2018) and provides, in a systematic and comprehensive manner, the center-to-center driving distance in minutes between communes.

2.1.3 Controls

We extract commune-level data from the 2008 National Population Census conducted by INSEE: the total population, land area, labor force, levels of education, and share of people who are more than 65 years old. Household net income data come from the French tax authority (Direction Générale des Impôts) and are formatted by INSEE.

2.2 Aggregation Method

2.2.1 Spatial Division

The Kantar Worldpanel data are representative of purchases of the French population and cover 7,940 of the 36,186 French communes. In 2008, among the 36,186 communes, we observe substantial heterogeneity from the smallest commune with 1 inhabitant (Rochefourchat) to the commune of Paris with 2.2 million inhabitants. This wide variability and the Kantar coverage of communes explain why we need to gather communes to create groups of comparable communes according to the type of living areas with sufficient Kantar adults.

To analyze the association between the obesity rate and the living environment, we define our unit of observation from a spatial division based on the 2010 INSEE French Urban Area Zoning. The 2010 Zoning is the governmental and representative definition of the types of living areas of French people. It provides the distribution of communes in urban areas, which describes the influence of cities on the entire French territory, distinguishing urban centers, their crowns and nonurban areas.

Urban centers contain at least 1,500 jobs. Their areas of influence are based on the commutes of the population in neighboring municipalities (Floch and Levy, 2011). First, we select urban areas with more than 80,000 inhabitants, and we distinguish between city centers and crowns (99 city centers and 86 crowns). Crowns comprise communes around the city center, and at least 40% of their active population works in the city center. These urban areas contain 41,891,096 people (68% of the French population). Communes that are not part of an urban area are categorized as nonurban areas, and we connect them to the closest city (center-to-center distance in minutes) to build our third category of areas with weak urban connections using *Odomatrix*.

From the 312 initial divisions, we exclude several divisions. First, 26 divisions do not have sufficient (fewer than 10) adults in the Kantar panel. Second, in 4 divisions, some food items composing the food basket price are not purchased. Third, one division has no 'neighborhood' subdivision, which is important for our spatial analysis. These 31 divisions represent 1.4% of the French population. Our final sample therefore contains 281 divisions or units of observation that are mapped in **Figure 1**.

PLEASE INSERT FIGURE 1 HERE

The color gradient indicates the type of division. The darkest colored divisions represent the 99 city centers (with an average density close to 830 inhabitants/km^2). The lightest colored divisions represent the 96 nonurban areas (with an average density equal to 67 inhabitants/km^2), which make up a very large proportion of the French territory. In between, we can observe the 86 crowns (with an average density equal to 84 inhabitants/km^2). White areas are the 31 excluded areas.

2.2.2 Aggregation of variables at the division level

We present how we aggregated raw data at the division level.

Caloric Intake

The average per capita daily intake of calories in a spatial division d for a category c, where c={whole diet, fruit and vegetables}, is based on the number of calories contained in the food products purchased by all the households in the division and can be written as follows:

$$Q_d^c = \frac{1}{366} \sum_{h \in d} \frac{O_h}{D_h} \sum_{t=1}^{13} \sum_{p \in P_t^h} Kcal_p W_{ht}^d \tag{1}$$

where P_t^h is the set of purchases of household h in period t, O_h is a per capita coefficient allowing for redressing for out-of-home consumption: $O_h = \frac{14}{14-M_h}$, where M_h is the per capita number of meals eaten outside the home in a week for household h, 14 represents the total number of meals per week, and D_h is a consumption unit of household h, which is the sum of individual weights in household h (1 for adults and 0.5 for children between 3 and 18 years old). We also tested the FAO consumption unit definition ((FAO/WHO/UNU, 2004), (Bermudez et al., 2012)). The results are presented in Model D of Table 4 of the Supplementary Material. W_{ht}^d is a normalized weighted variable that enables us to ensure the representativeness of each household in the sample and to redress for inactivity periods to obtain a yearly consumption using T_h, the total number of active periods (that is, with a nonzero weight W_{ht}):

$$W_{ht}^d = \frac{13}{T_h} \frac{W_{ht}}{\sum_{h \in d} W_{ht}}, \tag{2}$$

where W_{ht} is the four-week period weight for household h (the Kantar data have 13 periods of four weeks each) and $Kcal_p$ is the number of kilocalories contained in purchase p.

Food Basket Prices

The objective is to define food basket prices capture the price heterogeneity among divisions. The basket is composed of fixed quantities of several staple (brand-specific) food products such that the only variation is in the observed level of prices across divisions. The food basket prices can be written as follows:

$$FP^d = \sum_{f \in B} \overline{pr_d^f} \dot{q}_f \tag{3}$$

where $\overline{pr_d^f}$ is defined as the ratio of weighted (by W_{ht}^d) expenditures over the weighted (by W_{ht}^d) quantities of all purchases of food product f in division d, $\dot{q}_f$ is the per capita purchased quantity of product f in France, that is, over all households in the sample, and B is the set of brands (national brand or private label) in the

following food product categories: yogurt, nonsparkling water, fresh desserts, milk, fruit juices and nectars, colas, butter, pastas, cookies, appetizers, chocolate bars, ham, fresh cream, and coffee.

Living Environment

We must aggregate the variables to the division level. Each division d is composed of n_d communes. First, we sum the number of sports amenities, the sports licenses for 71 sports federations, supermarkets larger than 400 m2 and fast food restaurants per division d and give them per 1,000 inhabitants. Second, we compute the rate of general practitioners, which represents the percentage of people within a division d who have at least one general practitioner in their own commune, the percentage of people over 65 years old in each division d, and the rate of people who have less than a high school diploma in each division d. Third, we compute a density variable that gives the number of individuals per square kilometer in division d, the employment rate variable in each division d, which is the share of individuals with a job among those who are of working age, and an income variable that represents the average income of the 'fiscal entities' in each division d. Finally, we compute the time to amenities, which represents the weighted (by population) average driving time (center-to-center) for individuals who live in communes where the amenity is missing. We compute this average time for supermarkets and fast food restaurants. Further details about the construction of the variables are provided in Section 5.3 of the Supplementary Material.

2.2.3 Descriptive Statistics

Table 1 provides descriptive statistics for the variables of interest and controls for the 281 spatial divisions.
PLEASE INSERT Table 1 HERE

Health and Consumption Information

From the individuals' BMI, we compute the average obesity rate for adults over 18 years old in each division as the total number of adults with a BMI greater than 30 over the total number of adults in each division. **Figure 2** maps the quantile distribution of the obesity rate across our 281 divisions. The darkest color represents divisions where more than 17% of the population is obese. The lightest color represents divisions where less than 6.6% of the population is obese. Among the 281 divisions, the mean percentage of obese people is 12.3%, with a standard deviation of 6%. These results are in line with the 2006 Obepi Roche survey (12.4%). The median is 12.2%, and 14 divisions have an obesity rate of more than 22%.

PLEASE INSERT Figure 2 HERE

Figure 2 suggests spatial disparities in obesity: obesity is less prevalent in the western region, Bretagne, and the southeast region than it is in the north and center west regions. Regarding food consumption, the per capita daily total caloric intake is 3,303 kcal on average, with a standard deviation of 575 kcal. The

minimum daily intake is 2,070 kcal, and the maximum daily consumption is 6,303 kcal. These results are in line with those of (Caillavet et al., 2019), who found a total caloric intake of 2,800 kcal in 2010, with 2,222 kcal coming from at-home consumption based on estimates of the caloric intake from meals outside the home equaling 20%. As shown in Figure 3 of the Supplementary Material, we observe spatial heterogeneity in food consumption with more calorie-rich diets in nonurban areas. This result is confirmed by a test of the equality of means (see Table 3 in the Supplementary Material).

The average individual daily caloric intake from the consumption of fresh fruits and vegetables (FV) is 252 kcal, with a standard deviation of 60 kcal. The minimum kcal from eating FV is 105 kcal, and the maximum is 602 kcal. Regarding the spatial disparities, Figure 5 in the Supplementary Material clearly shows a large difference between northern and southern France: individuals in southeastern France consume more calories from FV than do individuals in other regions. These individuals are closer to the French agricultural production areas. We also find some spatial disparities in the food basket price, as it ranges from €121.73 to €172.70 (Figure 7 in the Supplementary Material).

Living Environment

On average, there are approximately 0.2 supermarkets per 1,000 people. Some divisions have no supermarkets, while others have almost 0.4 supermarkets per 1,000 people. As shown in Figure 9 in the Supplementary Material, there is spatial heterogeneity in the supermarket ratio, but it is relatively small given the close values of the different quantiles. Concerning fast food restaurants, on average, there are 0.02 fast food outlets per 1,000 people, ranging from 0 to 0.085 across the divisions. As expected, amenities in city centers are, on average, more numerous than in nonurban divisions (Figure 13 in the Supplementary Material). Concerning the access costs, on average, individuals drive 7.2 minutes to reach the closest commune with a supermarket (with a 2.5 minute standard deviation) and approximately 17 minutes (with a 6.5 minute standard deviation) to access the closest commune with a fast food restaurant.

Regarding the spatial disparities in this driving time, we observe a clear pattern for access to supermarkets (Figure 11 in the Supplementary Material). The time needed to access a supermarket from communes without supermarkets is shorter in the northwest than it is in the other regions of France. As expected, the access time is shorter in urban areas than in nonurban areas.

There are, on average, 174 official sports licenses per 1,000 inhabitants, with a standard deviation of 29 and a range of 112 to 301. For health amenities, in our 281 divisions, 77.2% of the population had at least one doctor in their commune. This percentage ranges from 14.2% to 100%, and city centers have the highest rate of access (Figure 19 in the Supplementary Material).

Concerning sports amenities, there are, on average, 3.2 per 1,000 inhabitants, with a standard deviation of 1.6 and a range of 1 to 13. **Figure 3** presents the quantile distribution of sports facilities across the 281

divisions. The darkest color on the map indicates divisions where people have access to more than 4 sports amenities, and the lightest color indicates fewer than 2 sports facilities per 1,000 inhabitants. **Figure 3** shows that the north and southeastern divisions have fewer sports facilities than do other regions.

PLEASE INSERT Figure 3 and Figure 4 HERE

Controls

In 2008, on average, the employment rate in our areas is 64% (which is comparable to the OECD data) and ranges from 48% to 76%. The average annual net income for households is €21,424, and 64% of the population has less than a high school degree. The average population density is 34.1 people per square meter, and 17.6% of the population is over 65 years old (see Figures 21, 23, 25, and 27 in the Supplementary Material for more details).

3 Spatial Analysis

Regional habits may induce similarities as well as differences in consumption patterns. The locations of amenities may also induce differences in accessibility and generate spatial spillover effects. Thus, the whole living environment may impact lifestyle, consumption patterns and health status. Therefore, it is important to account for the spatial links between divisions.

Different mechanisms can explain the spatial patterns in behavior: regional habits and culture, regional specialization in production, geographical characteristics, spillover effects, etc. To address the spatial factors in the association between health status and the living environment, we thus must investigate two properties in our data and address both simultaneously: i) spatial autocorrelation, which represents the interdependence of observations across space (i.e., spatial dependence) and ii) spatial heterogeneity, which refers to variations in relationships because of the locations of the observations. In other words, we must consider both the systemic dimension of the living environment and its spatial effects on health status. Indeed, the spatial locations of food, health and sports amenities are not independent, and those spatial correlations could affect their estimated associations with the obesity rate.

3.1 Spatial Correlations

To investigate spatial autocorrelation, we adopt a three-step approach. First, we define the neighborhood structure of our observations through a spatial weight matrix. The weight coefficients in the matrix represent the spatial links between observations. We select a contiguity matrix, which is widely used in the spatial econometric literature. Contiguities are defined in terms of sets of neighbors for location i. The

weight matrix W_{ij} is a zero diagonal matrix. For the off-diagonal elements, W_{ij} = 1 if i and j are neighbors and 0 otherwise. We scale the off-diagonal elements to sum to unity in each row. We thus obtain a neighborhood matrix for our 281 mutually exclusive divisions by relying on the first-order contiguity approach. Our weight matrix has 1,148 nonzero links, and the average number of links is 4.1.

The definition of the spatial weight matrix is at stake in spatial analysis. We then also test two other specifications of the spatial weight matrices as a robustness check. First, we consider larger spatial interrelations with a second-order rank neighbors weighting matrix. Second, we use a matrix based on the first-order rank neighbors, where we add a neighborhood link between the nonurban areas and the city center division to which the nonurban area is directly connected. The results of the analysis are robust to these two specifications (see Table 4 in the Supplementary Material).

In the second step, we quantify the influence of the spatial correlation using a Moran diagram, which provides a clear visualization of the spatial structure of each variable. It is a two-way scatter plot with the spatially lagged variable on the y-axis, that is, the average of the values of the variable within the neighboring divisions, and the value of the same variable in the division on the x-axis. The slope of the linear fit to the scatter plot equals Moran's I. The Moran diagrams are provided in Section 5.2 of the Supplementary Material.

For the sake of clarity, we present only the process for sports amenities. Figures 3 and 4 imply that the sports facilities ratio variable is not randomly distributed across space and might exhibit some positive spatial autocorrelation. Indeed, there is a positive relationship between the sports facilities ratio in a division and the average ratio in its neighboring divisions. In other words, in the case of sports amenities, each division is similar to its neighbors. For other "Living Environment" variables, we observe from Figure 20 in the Supplementary Material no spatial autocorrelation for the General Practitioner Rate but a positive spatial pattern for the access costs to food amenities (Figures 12 and 16 in the Supplementary Material) and a negative spatial pattern for both food amenities (Figures 10 and 14 in the Supplementary Material).

In the third step, we assess the significance of this influence in all our variables by conducting a Moran test. The last column of Table 1 provides Moran's I statistic and whether it is significant at the 5% level. Fourteen variables exhibit significant spatial autocorrelation.

Different economic mechanisms can lead to this nonrandom spatial distribution. For instance, both the Supermarket Ratio and Fast Food Ratio reveal a negative and significant spatial autocorrelation. This phenomenon emphasizes that the dispersion of data might reflect the competition process: one division with a high value for a variable inhibits other divisions from having similarly high values. In other words, these food amenities are located in areas with no competitors. From an economic perspective, this correlation implies a nonhazardous location for food amenities.

History matters as much as comparative advantages (Goffette-Nagot and Schmitt, 1999) in explaining the

spatial organization of amenities. For example, regional food habits (more fish is eaten in the West, more fruit in the South, etc.) contribute to the spatial sorting in health status and food expenditure. Individual preferences for regional amenities—for example, a preference for sunny locations within the elderly population—can produce the same results. Last, the organization of markets and the links between demand and supply can explain either positive or negative spatial correlations (distance can be physical, social, or organizational). The foundations of urban agglomeration economies, based on sharing, matching and learning mechanisms (Duranton and Puga, 2004), explain the locations of activities and people: they become concentrated in order to share indivisible goods and facilities, to gain from variety and to improve the quality of labor market matches (counterbalanced by the risk of competition), as well as because of learning mechanisms. As a result, highly educated people are concentrated in urban areas, as are wealthy people. Concerning sales and distribution activities, proximity to the final consumers and the level of competition are key to choosing a location: not too close to competitors and not too far from demand. Land prices constitute another factor that can explain location choices. In addition, supermarkets are concentrated in crowns, which leads to negative autocorrelation in the Supermarket Ratio, which is the highest across divisions.

3.2 Spatial Econometrics

The association between health status and the living environment is related to geographical features and is a great public health concern. Table 1 shows the spatial autocorrelation in our variables of interest, which leads us to a spatial econometric framework. The analysis of the residuals from an estimated OLS model that does not take the spatial structure of our data into account confirms this intuition. Indeed, Table 5 in the Supplementary Material shows a significant Moran test on the residuals. We thus must take the spatial structure of our data into account to analyze spatial health inequalities.

In this section, we investigate the association between the obesity rate and the observable characteristics of the living environment in a division and in its neighboring divisions. This strategy allows us to investigate the geographical nature of health inequalities. According to LeSage (2014), economic intuition should guide model specification. Obesity is a noncommunicable disease: there is no reason why obesity in neighboring divisions should impact obesity in the local division. At fine geographical scales, various studies show the influence of shared social norms on the prevalence of obesity (Hruschka et al., 2011). Van Rongen et al. (2020) show that the impact of neighborhood fast food supply on consumption is indirect and mediated by neighborhood social norms regarding fast food consumption. However, our geographical scale is sufficiently broad that these mechanisms are negligible. Rather, we aim to test the impact of the environment in neighboring divisions (food, sport and health amenities available) on the health status of the population in a given division. In other words, we want to assess and account for potential spatial spillovers.

Halleck Vega and Elhorst (2015) recommend investigating the SLX model as a 'point of departure when

having any empirical evidence that the observations in the sample are spatially dependent.' However, we also test more general models that have SLX nested within them. We thus estimate a spatial Durbin error model (SDEM) that takes into account both types of local spillover: *WX* and *Wu*.

$$Y_i = X_i.\beta + W_{ij}X_j.\gamma + u_i \tag{4}$$

with $u_i = \lambda W_{ij}u_j + \varepsilon_i$ and $\varepsilon \sim N(\mu, \sigma^2)$

In Equation (4), Y_i represents the population health status in division i. X_i is a vector that contains the consumption variables (food and sports), the living environment variables, and the control variables for division i. W_{ij} is a contiguity matrix, and $W_{ij}X_j$ is its exogenous interaction with X_j and contains all information on neighboring divisions. We then estimate the parameters of (4) using maximum likelihood estimation. All estimations are performed in *Stata 17.0* using the *spregress* command and in *R 4.1.0* with the package *spatialreg* as a robustness exercise. We also perform multicollinearity checks on our explanatory variables (X and WX). These tests lead us to exclude two variables from our WX vector: (No High School Diploma Rate) and (Employment Rate). We also test two other specifications that are nested within the SDEM: the spatial lag of X (SLX) model, where $\lambda = 0$, that is, no spatial dependence in disturbances, and the spatial error model (SEM), where $\theta = 0$, that is, no spatial dependence in explanatory variables. We provide all these results in Section 5.5 of the Supplementary Material. Table 5 of the Supplementary Material provides the likelihood ratio tests used to select the model. We find that testing $H_0 : \theta = 0$ after estimation of the SLX model leads us to reject H_0 (p-value of 0.01) at a 5% level of significance, so θ is not null: we should choose an SLX specification rather than the OLS model. Testing $H_0 : \lambda = 0$ after estimation of the SDEM leads us to conclude that we cannot reject the null (implying that $\lambda = 0$) (p-values of 0.24), and we should choose an SLX model for an SDEM (i.e., there is no spatial residual component). SLX is then the model that best describes our data. Moreover, the lowest AIC is observed for the SLX estimation. Finally, our postestimation Moran test to check for autocorrelation in the residuals indicates that we cannot reject the null hypothesis of no spatial autocorrelation at the 5% level of significance (p-value of 0.13). In other words, we have purged the spatial autocorrelation and corrected the residuals. Last, comparing the OLS estimates and the total effects of the SLX model, which are the sum of the direct and indirect effects (see Table 6 in the Supplementary Material), we show that failing to consider the spatial disparities in the explanatory variables results in misestimated effects.

Our estimation results are presented in **Table 2**. Our main model is Model A, which shows the SLX estimation. Model B provides an analysis without the food and sports consumption variables that could capture a part of the effect of the living environment. Model C proposes the SLX model that introduces interactions between the fast food ratio and the three types of spatial division.

PLEASE INSERT Table 2 HERE

X Variables-Consumption

As expected, a significant and positive effect between daily total caloric intake and the obesity rate is observed: the more calories people eat within a spatial division, the higher the rate of obesity is. Likewise, there is a significant and negative association between caloric intake from FV and the obesity rate. This result supports the nutritional recommendations that suggest that FV are a very important part of a balanced diet. Holding total caloric intake constant, the more calories from FV in a region's diet, the lower the obesity rate is. Moreover, Model A in Table 2 shows a nonsignificant association between the food basket price and sports license ratio and the obesity rate. Looking at Model B, we can see that the results are robust with and without food consumption variables.

X Variables–Food, Health and Sports Amenities

There is no direct association between the obesity rate and the supermarket ratio. If supermarkets provide consumers with access to a large variety of food products at low prices and enable them to have a healthier diet, access to supermarkets is not a good policy instrument. However, our result does not mean that the prices and quality of food product regulations at the supermarket level could not be considered as a policy tool. This result is in line with that of Allcott et al. (2019), who show that the entry of supermarkets has no effect on the nutritional quality of food choices. However, we observe a positive and significant association between the fast food ratio and the obesity rate: the more fast food restaurants there are per inhabitant, the higher the obesity rate in the division. This positive correlation between the fast food ratio and health status is in line with previous individual-level results in the literature (Anderson et al., 2011; Crawford et al., 2008; Garcia et al., 2012). In model C in Table 2, we interact the fast food ratio with the three types of spatial divisions. The results show that the fast food ratio particularly affects the obesity rate in crowns. Concerning fast food restaurants, this result suggests that it is immediate proximity that impacts health indicators, ceteris paribus, which implies that public authorities should more carefully consider and perhaps regulate the number of fast food outlets on behalf of the public due to these health concerns. Moreover, the access times to supermarkets and fast food restaurants do not affect the health indicators. We observe no significant direct effects on the obesity rate from either the sports amenities or the health amenities variables.

WX Variables–Spatial effects

The estimated spatial effects (WX) are also presented in Table 2. Our spatial framework, as suggested by LeSage (2014), helps us to identify significant spatial spillover effects on the obesity rate. We observe significant spatial effects that are linked to two out of three amenities considered (sports and health) that we

use to describe the living environment. For instance, the sports amenities ratio and the general practitioner rate in the neighboring divisions have significant and negative effects on the obesity rate. We also observe that the price level in the neighboring divisions has a positive and significant effect on the obesity rate. The positive sign of the sports license variable in the neighboring divisions can be explained by the fact that, at a constant sports amenities level, the more sports licenses in the neighboring division j, the fewer opportunities to take a license for people in the analyzed division i, and the higher the rate of obesity. The existence of these local spillovers suggests that the spatial dispersion of food prices, health and sports amenities as sports consumption might affect local health indicators.

These results suggest that to address health inequalities, we should think about the spatial organization of living environments not only in terms of proximity but also on a larger scale because of these spatial spillovers. The heterogeneity in the magnitude and sign of these effects also suggests that the scale of intervention might differ according to the type of amenity.

4 Discussion

In this article, we analyzed the association between obesity and the living environment in France. To address this association, we first conducted an analysis to indicate that our original data exhibited spatial autocorrelation linked to the spatial organization of human activities.

To address the spatial autocorrelation in the data and to identify existing spatial spillovers, we conducted a spatial econometric analysis. We were then able to appraise the relevance of spatial disparities to health inequalities. Indeed, our spatial econometric framework, which controls for spatial spillover effects from the living environment, both purges the spatial autocorrelation and clarifies the influence of the locations of food, sports and health amenities on the obesity rate.

The strength of our article is that we provide a wide picture of the living environment by considering food, sports and health amenities, controlling for food and sports consumption within a unique framework. Concerning food consumption, our analysis shows that individual food consumption patterns are significant explanatory factors for obesity, with caloric intake having a positive effect and FV consumption having a negative effect. As expected, we established that the living environment of a division and its neighbors both influence the obesity rate. Concerning food amenities, we established that the presence of fast food restaurants could affect the prevalence of obesity, specifically in crowns. This is not the case for supermarkets. Public authorities should therefore continue directing consumers toward healthier diets, both by improving the nutritional quality of food purchases and by regulating the location of new fast food restaurants.

We have also demonstrated that the spatial organization of the French territory implies significant spatial spillover effects that could affect the obesity rate. Indeed, sports and health amenities influence the obe-

sity rate through these spillovers. While we do not reveal any causal effects, the significant correlations, their magnitudes and their directions suggest that this question is relevant from a public policy perspective. There is great concern about how public authorities organize their activities over their territories. France is currently struggling with the consequences of medical deserts in rural areas. Our results suggest that at least for obesity, the concept of proximity to health amenities should be thought of on a larger scale than simply within communes. In particular, the medical field could offer personalized monitoring and personalized health and nutritional messages.

Our analysis has some limitations that are mostly linked to data availability. First, we do not reveal causal effects; further work will be required to derive individual behaviors using both precise individual data sets and exogenous shocks on the living environment to determine their true causality links (regulation of the number of fast food, for instance). Second, people do not randomly choose where they live, and social and spatial networks should thus be considered. Third, the availability of individual data would also allow us to consider edge effects for people living near boundaries. These edge effects could play an important role in the spatial spillovers because living on the far side of a boundary will be much farther away than living near the boundary. Moreover, the Kantar data set suggested that we slightly underestimated the prevalence of obesity. This could result from the self-declaration of household members of height and weight measures or from the lack of household representativity regarding this health status that Kantar does not consider in their sample setting. Last, extending the data linking and the analysis on a larger and more recent period would be interesting.

References

Allcott, H., Diamond, R., Dubé, J.-P., Handbury, J., Rahkovsky, I. and Schnell, M. (2019), 'Food deserts and the causes of nutritional inequality', *The Quarterly Journal of Economics* **134**(4), 1793–1844.

Anderson, B., Lyon-Callo, S., Fussman, C., Imes, G. and Rafferty, A. P. (2011), 'Peer reviewed: Fast-food consumption and obesity among michigan adults', *Preventing chronic disease* **8**(4).

Befort, C. A., Nazir, N. and Perri, M. G. (2012), 'Prevalence of obesity among adults from rural and urban areas of the united states: findings from nhanes (2005-2008)', *The Journal of Rural Health* **28**(4), 392–397.

Bermudez, O. I., Lividini, K., Smitz, M.-F. and Fiedler, J. L. (2012), 'Estimating micronutrient intakes from household consumption and expenditures surveys (hces): an example from bangladesh', *Food and nutrition bulletin* **33**(3_suppl2), S208–S213.

Caillavet, F., Darmon, N., Létoile, F. and Nichèle, V. (2019), 'Four decades of household food purchases: Changes in inequalities of nutritional quality in france, 1971-2010', *Economie et Statistique / Economics and Statistics* (513), 69 – 89.

Caspi, C. E., Sorensen, G., Subramanian, S. and Kawachi, I. (2012), 'The local food environment and diet: a systematic review', *Health & place* **18**(5), 1172–1187.

Chen, M., Creger, T., Howard, V., Judd, S. E., Harrington, K. F. and Fontaine, K. R. (2019), 'Association of community food environment and obesity among us adults: a geographical information system analysis', *J Epidemiol Community Health* **73**(2), 148–155.

Courtemanche, C. and Carden, A. (2011), 'Supersizing supercenters? the impact of walmart supercenters on body mass index and obesity', *Journal of Urban Economics* **69**(2), 165–181.

Crawford, D. A., Timperio, A. F., Salmon, J. A., Baur, L., Giles-Corti, B., Roberts, R. J., Jackson, M. L., Andrianopoulos, N. and Ball, K. (2008), 'Neighbourhood fast food outlets and obesity in children and adults: the clan study', *International Journal of Pediatric Obesity* **3**(4), 249–256.

Currie, J., DellaVigna, S., Moretti, E. and Pathania, V. (2010), 'The effect of fast food restaurants on obesity and weight gain', *American Economic Journal: Economic Policy* **2**(3), 32–63.

de Mouzon, O. and Orozco, V. (2011), 'NutriXConso : Recherche et appariement de données d'achats et de données nutritionnelles', *Cahier des techniques de l'INRA* .

De Saint Pol, T. (2007), 'L'obésité en France: les écarts entre catégories sociales s'accroissent', *INSEE Première* .

Duranton, G. and Puga, D. (2004), Micro-foundations of urban agglomeration economies, *in* 'Handbook of regional and urban economics', Vol. 4, Elsevier, pp. 2063–2117.

Eid, J., Overman, H. G., Puga, D. and Turner, M. A. (2008), 'Fat city: Questioning the relationship between urban sprawl and obesity', *Journal of Urban Economics* **63**(2), 385–404.

Ellaway, A., Curtice, J., Morris, G., Robertson, C., Allardice, G. and Robertson, R. (2006), 'Perceptions of the local environment and their associations with health', *Epidemiology* **17**(6), S198.

Ewing, R., Schmid, T., Killingsworth, R., Zlot, A. and Raudenbush, S. (2003), 'Relationship between urban sprawl and physical activity, obesity, and morbidity', *American journal of health promotion: AJHP* **18**(1), 47–57.

FAO/WHO/UNU, J. (2004), 'Human energy requirements. report of a joint fao/who/unu expert consultation, rome, 17-24 october 2001'.

Feng, J., Glass, T. A., Curriero, F. C., Stewart, W. F. and Schwartz, B. S. (2010), 'The built environment and obesity: a systematic review of the epidemiologic evidence', *Health & place* **16**(2), 175–190.

Floch, J.-M. and Levy, D. (2011), 'Le nouveau zonage en aires urbaines de 2010: poursuite de la périurbanisation et croissance des grandes aires urbaines'.

Frank, L. D., Andresen, M. A. and Schmid, T. L. (2004), 'Obesity relationships with community design, physical activity, and time spent in cars - American Journal of Preventive Medicine', *American Journal of Preventive Medicine* pp. 87 – 96.

Garcia, G., Sunil, T. S. and Hinojosa, P. (2012), 'The fast food and obesity link: consumption patterns and severity of obesity', *Obesity surgery* **22**(5), 810–818.

Goffette-Nagot, F. and Schmitt, B. (1999), 'Agglomeration economies and spatial configurations in rural areas', *Environment and Planning A* **31**(7), 1239–1257.

Halleck Vega, S. and Elhorst, J. P. (2015), 'The slx model', *Journal of Regional Science* **55**(3), 339–363.

Hilal, M., Aline, F.-S. and Gojard, S. (2018), Odomatrix - mesurer des temps de trajet pour améliorer les conditions de vie, Technical report.

Hruschka, D. J., Brewis, A. A., Wutich, A. and Morin, B. (2011), 'Shared norms and their explanation for the social clustering of obesity', *American journal of public health* **101**(S1), S295–S300.

James, P., Banay, R. F., Hart, J. E. and Laden, F. (2015), 'A review of the health benefits of greenness', *Current epidemiology reports* **2**(2), 131–142.

Lachowycz, K. and Jones, A. P. (2011), 'Greenspace and obesity: a systematic review of the evidence', *Obesity reviews* **12**(5), e183–e189.

LeSage, J. P. (2014), 'What regional scientists need to know about spatial econometrics', *Available at SSRN 2420725* .

Mackenbach, J. D., Rutter, H., Compernolle, S., Glonti, K., Oppert, J.-M., Charreire, H., De Bourdeaudhuij, I., Brug, J., Nijpels, G. and Lakerveld, J. (2014), 'Obesogenic environments: a systematic review of the association between the physical environment and adult weight status, the spotlight project', *BMC public health* **14**(1), 1–15.

Mason, K. E., Pearce, N. and Cummins, S. (2020), 'Do neighbourhood characteristics act together to influence bmi? a cross-sectional study of urban parks and takeaway/fast-food stores as modifiers of the effect of physical activity facilities', *Social Science & Medicine* **261**, 113242.

Mason, K. E., Pearce, N. and Cummins, S. (2021), 'Geographical heterogeneity across england in associations between the neighbourhood built environment and body mass index', *Health & place* **71**, 102645.

Ncd Risk Factor Collaboration, a. (2019), 'Rising rural body-mass index is the main driver of the global obesity epidemic in adults', *Nature* **569**(7755), 260.

Plantinga, A. J. and Bernell, S. (2007), 'The association between urban sprawl and obesity: is it a two-way street?', *Journal of regional science* **47**(5), 857–879.

Popkin, B. M. (2019), 'Rural areas drive increases in global obesity', *Nature* **569**(7755), 200–201.

Roche, O. (2012), Obépi 2012 : Enquête épidémiologique nationale sur le surpoids et l'obésité, enquête 6ème édition.

Rosengren, A. and Lissner, L. (2008), 'The sociology of obesity', *Obesity and metabolism* **36**, 260–270.

Soga, M., Gaston, K. J. and Yamaura, Y. (2017), 'Gardening is beneficial for health: A meta-analysis', *Preventive medicine reports* **5**, 92–99.

Van Holle, V., Deforche, B., Van Cauwenberg, J., Goubert, L., Maes, L., Van de Weghe, N. and De Bourdeaudhuij, I. (2012), 'Relationship between the physical environment and different domains of physical activity in european adults: a systematic review', *BMC public health* **12**(1), 1–17.

Van Rongen, S., Poelman, M. P., Thornton, L., Abbott, G., Lu, M., Kamphuis, C. B., Verkooijen, K. and De Vet, E. (2020), 'Neighbourhood fast food exposure and consumption: the mediating role of neighbourhood social norms', *International journal of behavioral nutrition and physical activity* **17**(1), 1–9.

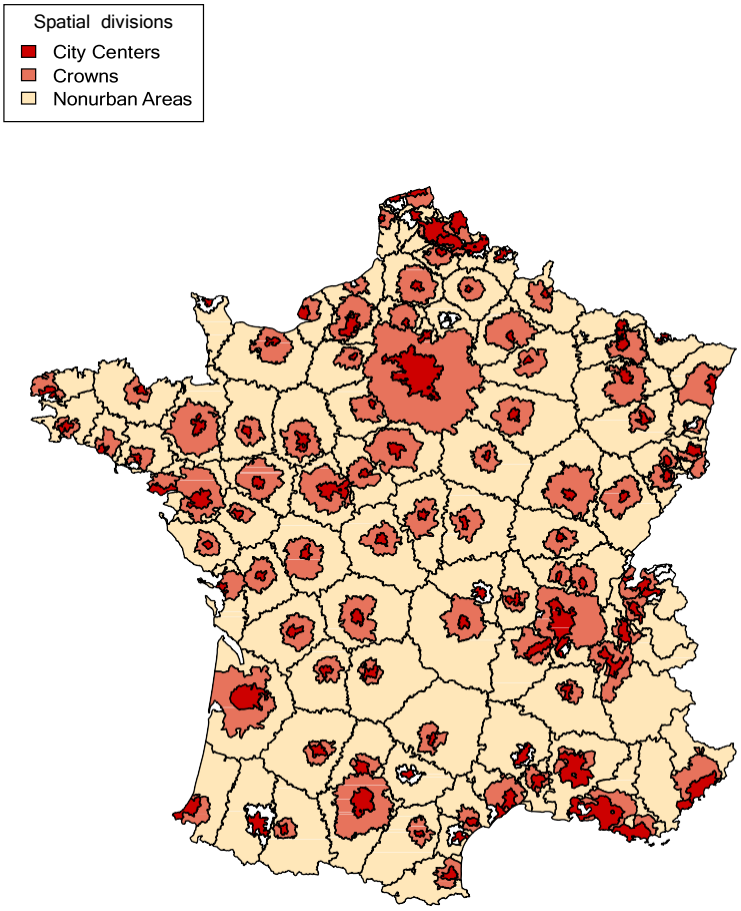

Figure 1: Map of our 281 spatial divisions.

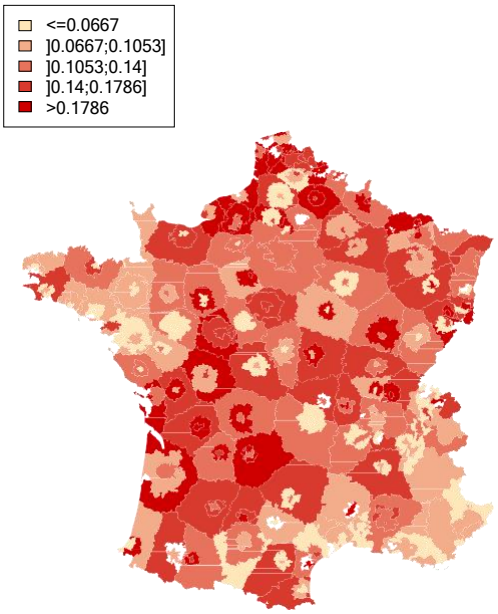

Figure 2: Map of the Quantile Distribution of the Obesity Rate variable.

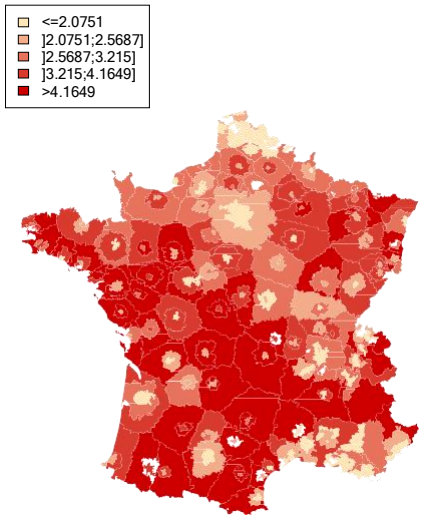

Figure 3: Spatial Distribution of the Sports Facilities Ratio variable.

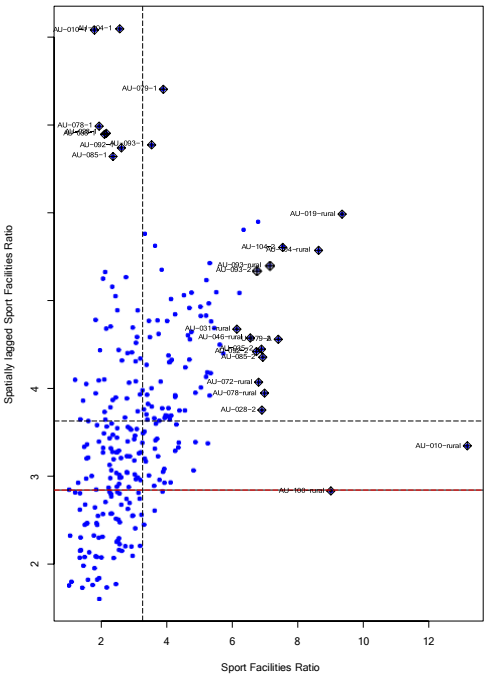

Figure 4: Moran Diagram for the Sport Facilities Ratio variable

Supplementary material for the paper:

Spatial spillovers, living environment and obesity in France: Evidence from a spatial econometric framework

5 Supplementary material

5.1 Data Sources

Table 1 gives the source, year, type of data (open or private) and the version used in our article. Table 2 provides information about the variables we used (name of the variable, detailed definition, units).

Table 1: Data sources

Source (Name and Producer)	***Year***	***Type of Data (Private/Open)***	***Description***	***Version/Identifier/Link***
Kantar Worldpanel from the Kantar France company	2008	Private INRAE Contract with the Kantar Worldpanel company for research use (limited terms of use)	A representative panel of French households that gives information about their food purchases (home-scan data)	Version 2 ID ODALIM: 11-worldpanel-2008-v2 https://odalim.inrae.fr/fr/dataset/11-worldpanel-2008-v2
The French Census of Permanent Facilities (BPE -INSEE)	2008	Open data	Offers the number of localized equipment and services each year, built from various administrative sources.	http://www.progedo-adisp.fr/enquetes/XML/lil.php?lil=lil-0452
Free GPS POI Files for McDonald's, KFC and Quick in France	2011	Open Data	Indicates the location of certain fast food restaurants (such as McDonald's, KFC and Quick)	http://www.gps-data-team.com/poi/france/restaurants/ ; Data Downloads: 22/07/2011 for McDonald's and KFC; 5/03/2011 for Quick
The French Census of Sporting Facilities (RES- French Sport Ministry)	2010	Open Data	Offers a national census of all sports equipment, spaces and practice sites	http://www.res.sports.gouv.fr/
The French Census of Sports Licenses (French Sport Ministry)	2010	Open Data	Offers a census of sports licenses by federation, by sex and by age group	https://www.sports.gouv.fr/organisation/publications/statistiques/Donnees-detaillees/Licences-et-groupements-sportifs-
Public Finances Directorate General, INSEE processing	2008	Open Data	The system for "localized tax revenues" provides local statistics on the tax revenues of households	https://www.insee.fr/fr/statistiques/1893295
National Population Census (INSEE)	2008	Open Data	Provides information on the population of France, its diversity and its evolution.	https://www.insee.fr/fr/statistiques/2409062
2010 Zoning into Urban Areas (INSEE)	2010	Open Data	Provides a spatial zoning established in reference to the known population in the 2007 census and the geography of the territory as of January 1, 2010.	https://www.insee.fr/fr/statistiques/1280970#documentation

Table 2: Information about variables

Sources	***Name***	***More details: * All our computations are made at the subdivision level (N=281)***	***Unit***
Kantar Worldpanel	Obesity Rate	% of people who are obese (BMI>30)	%
	BMI	Average body mass index	kg/(m^2)
	Caloric Intake (Tot)	Average individual total daily intake of calories	kcal
	Caloric Intake (FV)	Average individual daily intake of calories from fruits and vegetables	kcal
	Food Basket Price	Average expenditure on a defined basket of foods (based on leading national brands and private labels and the national average quantity purchased for certain staple goods: yogurt, non-sparkling water, fresh desserts, milk, fruit juices and nectars, colas, butter, pastas, cookies, appetizers, chocolate bars, ham, fresh cream, coffee)	euro
The French Census of Permanent Facilities (BPE - INSEE)	Supermarket Ratio	Number of supermarkets per 1,000 inhabitants	#/1000 inhabitants
	Time to Supermarkets	Average driving time to a supermarket (store with more than 400 m^2) for inhabitants of divisions without a supermarket	minutes
	General Practitioner Rate	% of the population with at least one general practitioner in their division	%
Free GPS POI Files for McDonald's, KFC and Quick in France	Fast Food Ratio	Number of McDonald's, KFC and Quick restaurants per 1,000 inhabitants	#/1000 inhabitants
Free GPS POI Files of McDonald's, KFC and Quick in France * *Odomatrix*	Time to Fast Food	Average driving time to a commune with a McDonald's, KFC or Quick restaurant for inhabitants of divisions without a fast food restaurant	minutes
The French Census of Sporting Facilities (RES-French Sport Ministry)	Sport Facilities Ratio	Number of sports facilities per 1,000 inhabitants	#/1000 inhabitants
The French Census of Sport Licenses (French Sport Ministry)	Sport Licenses Ratio	Number of sports licenses per 1,000 inhabitants	#/1000 inhabitants
Public Finances Directorate General, INSEE processing	Income	Average net adjusted disposable household income	euro
National Population Census (INSEE)	No High School Diploma Rate	% of individuals with a high school education (no graduation)	%
	Empl. Rate	% of the population (15-64) that is actively employed	%
	Density	Population divided by area	inhabitants/km^2
	Senior Citizen Rate	% of individuals over 65 years old	%
2010 Zoning into Urban Areas (INSEE)	City Centers	Cities with more than 80,000 inhabitants	#
	Crowns	Towns with more than 40% of their active population connected to a city center	#
	Nonurban Areas	Towns with less than 40% of their active population connected to a city	#

5.2 Additional spatial distributions and Moran diagrams

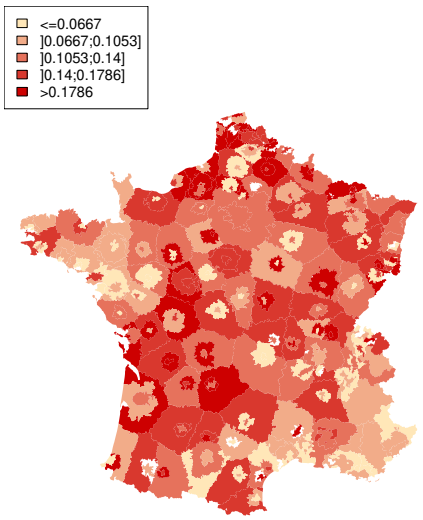

Figure 1: Spatial Distribution of the Obesity Rate variable.

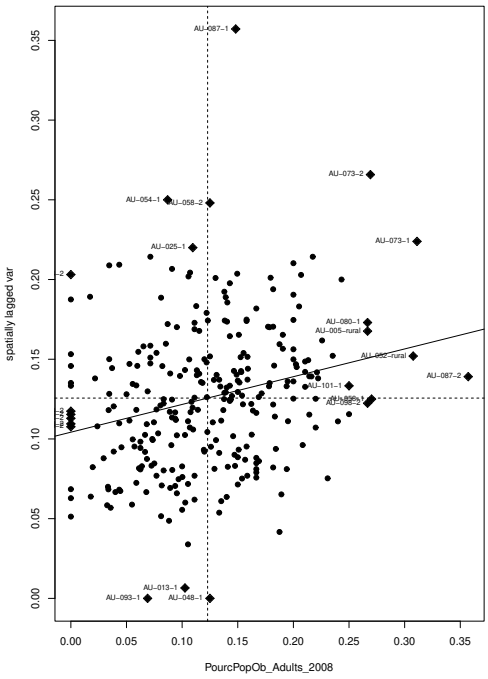

Figure 2: Moran Diagram for the Obesity Rate variable.

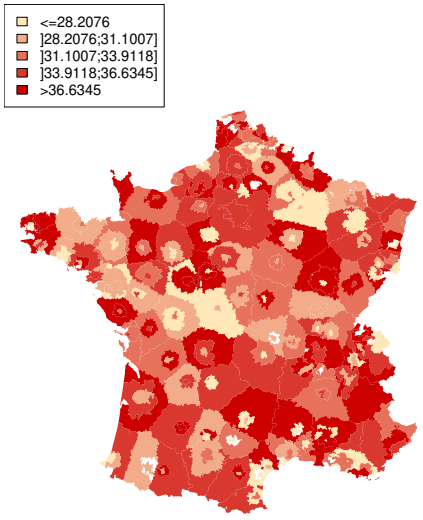

Figure 3: Spatial Distribution of Caloric Intake (Tot).

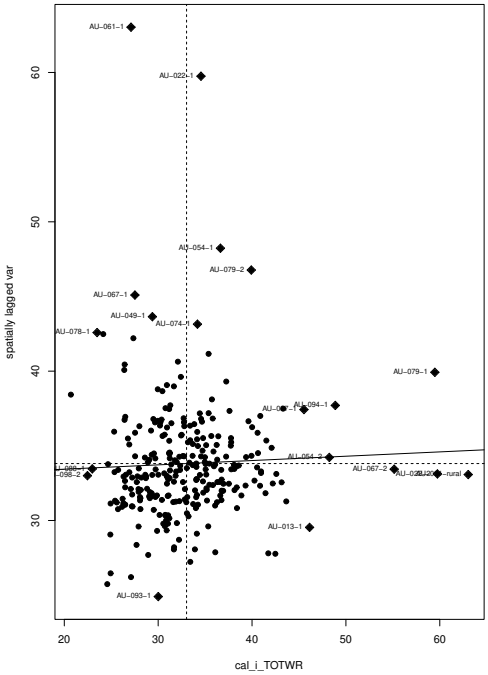

Figure 4: Moran Diagram for Caloric Intake (Tot).

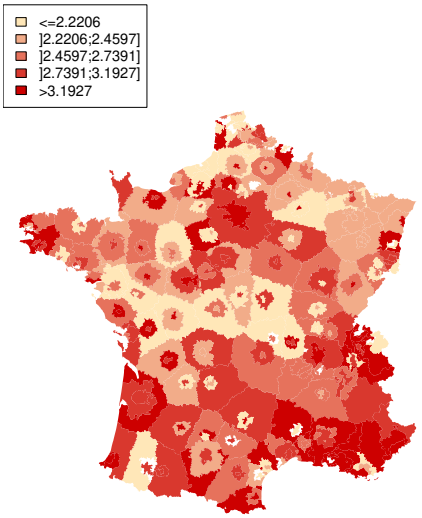

Figure 5: Spatial Distribution of Caloric Intake (FV).

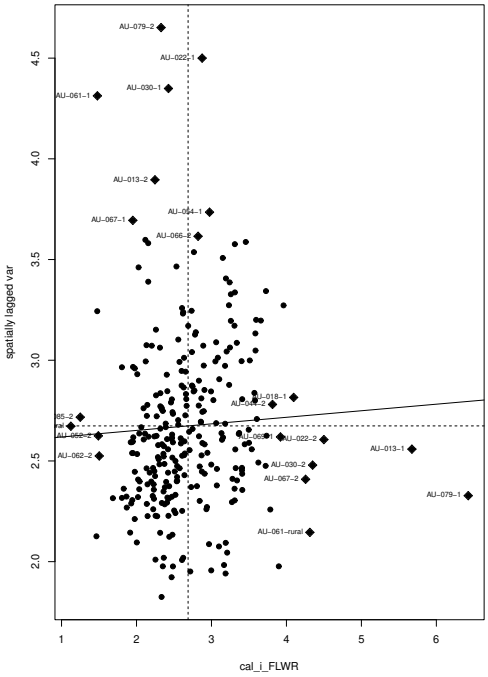

Figure 6: Moran Diagram for Caloric Intake (FV).

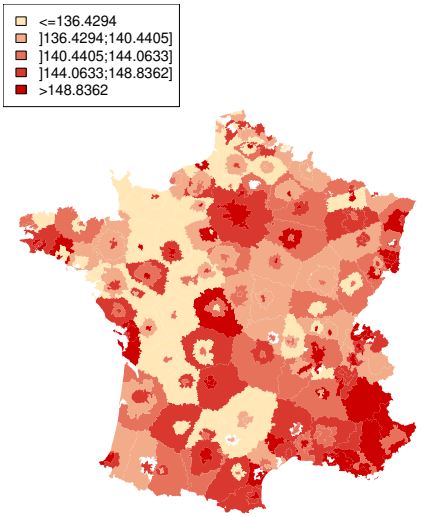

Figure 7: Spatial Distribution of Food Basket Prices.

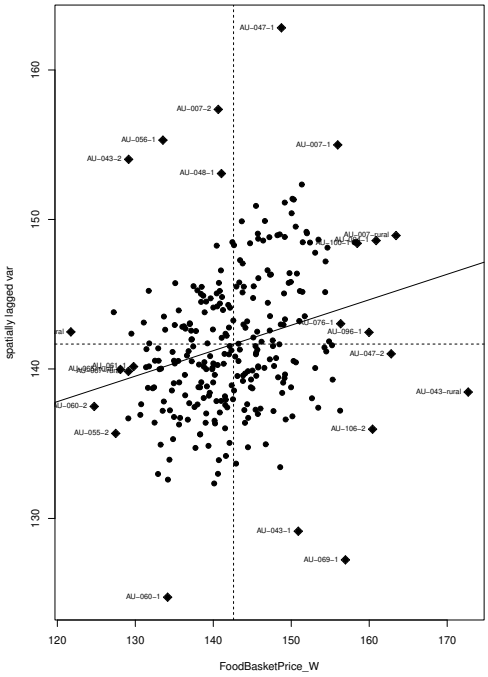

Figure 8: Moran Diagram for Food Basket Prices.

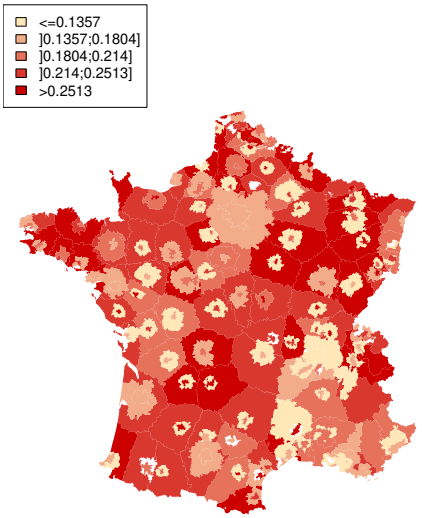

Figure 9: Spatial Distribution of the Supermarket Ratio.

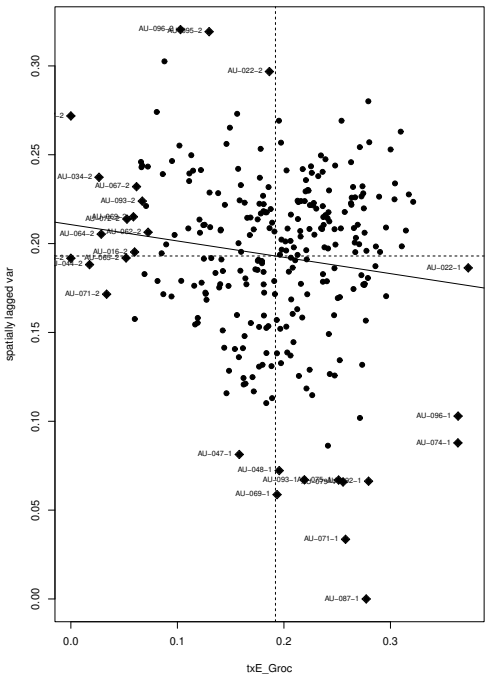

Figure 10: Moran Diagram for the Supermarket Ratio.

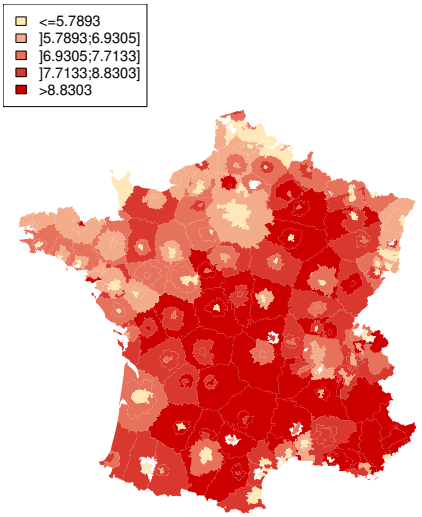

Figure 11: Spatial Distribution of the Travel Time to a Supermarket.

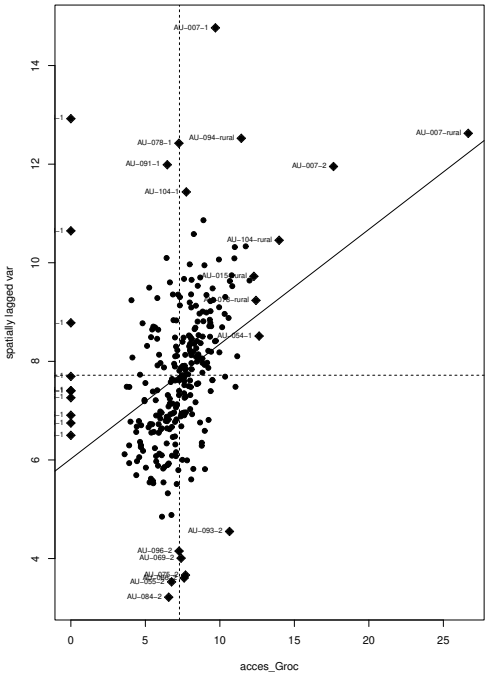

Figure 12: Moran Diagram for the Travel Time to a Supermarket.

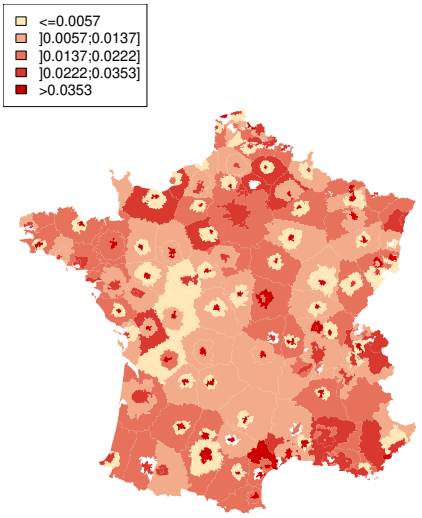

Figure 13: Spatial Distribution of the Fast Food Ratio.

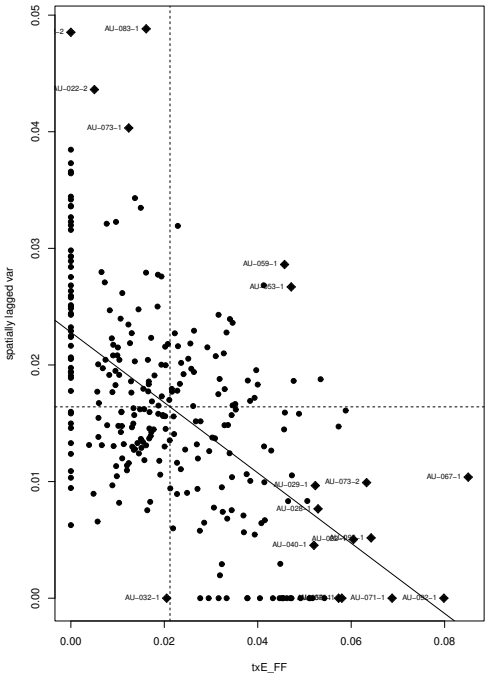

Figure 14: Moran Diagram for the Fast Food Ratio.

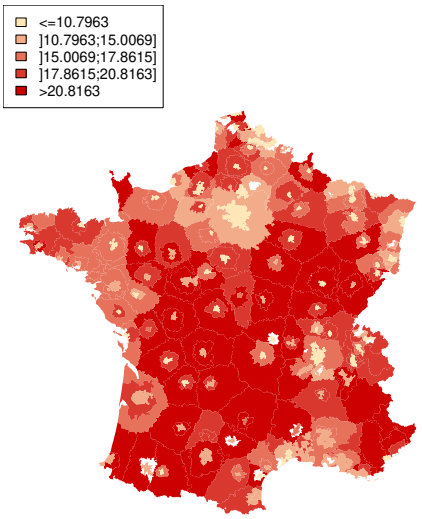

Figure 15: Spatial Distribution of the Travel Time to a Fast Food Restaurant.

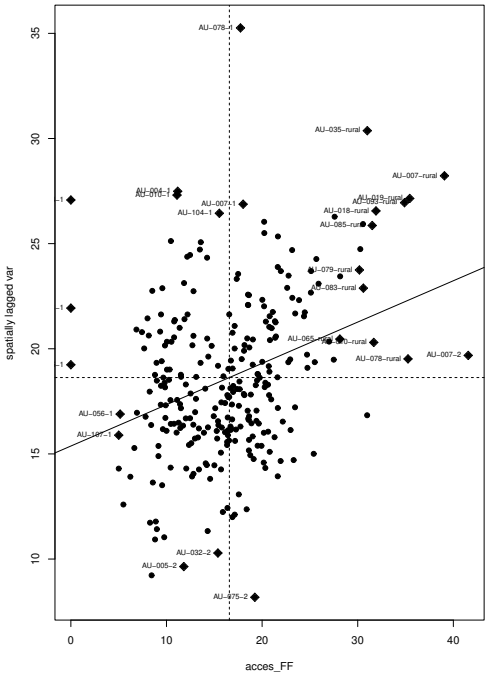

Figure 16: Moran Diagram for the Travel Time to a Fast Food Restaurant.

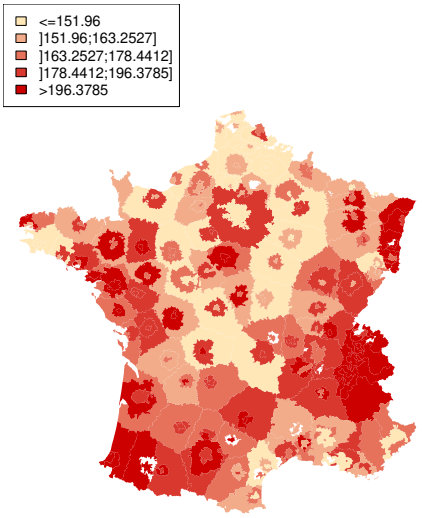

Figure 17: Spatial Distribution of the Sports Licenses Ratio.

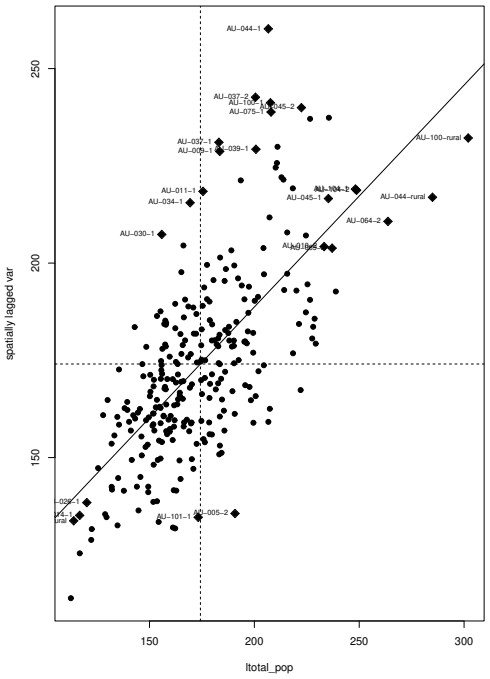

Figure 18: Moran Diagram for the Sports Licenses Ratio.

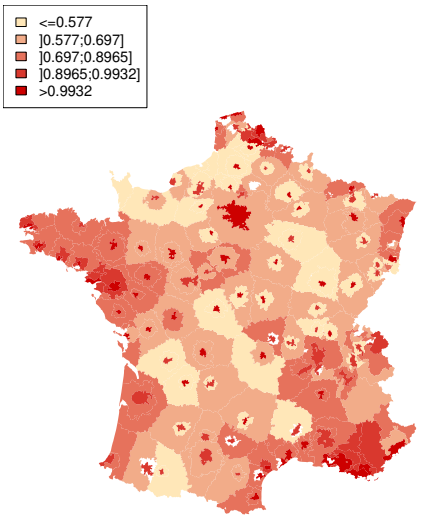

Figure 19: Spatial Distribution of the General Practitioner Rate.

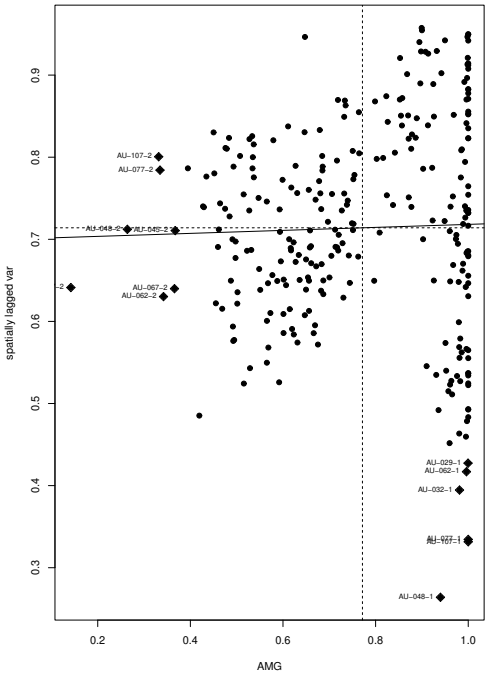

Figure 20: Moran Diagram for the General Practitioner Rate.

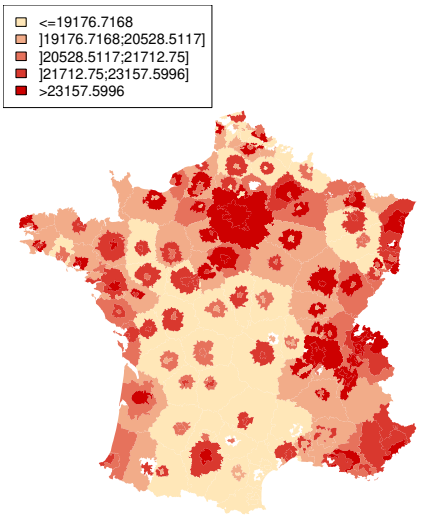

Figure 21: Spatial Distribution of Income.

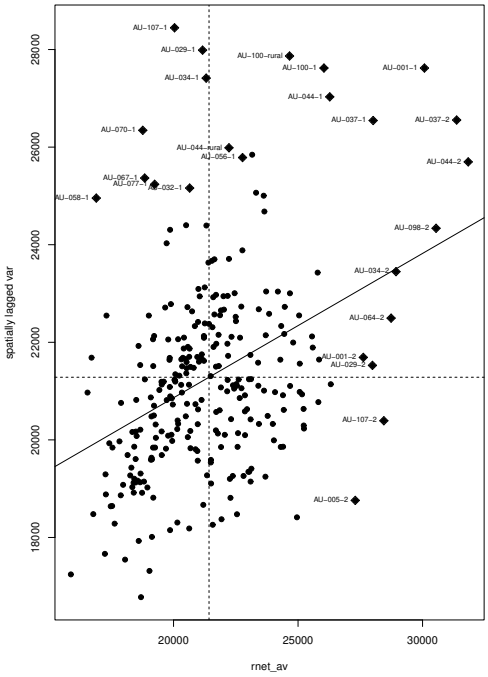

Figure 22: Moran Diagram for Income.

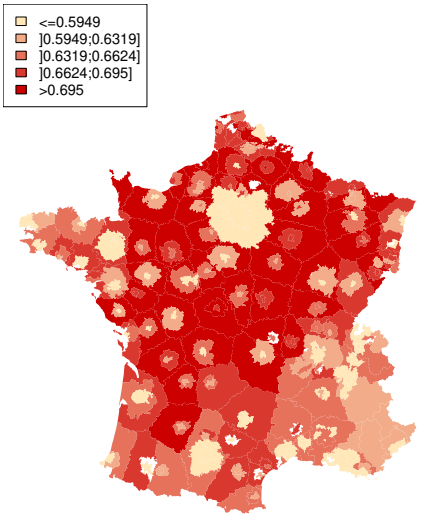

Figure 23: Spatial Distribution of the No High School Diploma Rate.

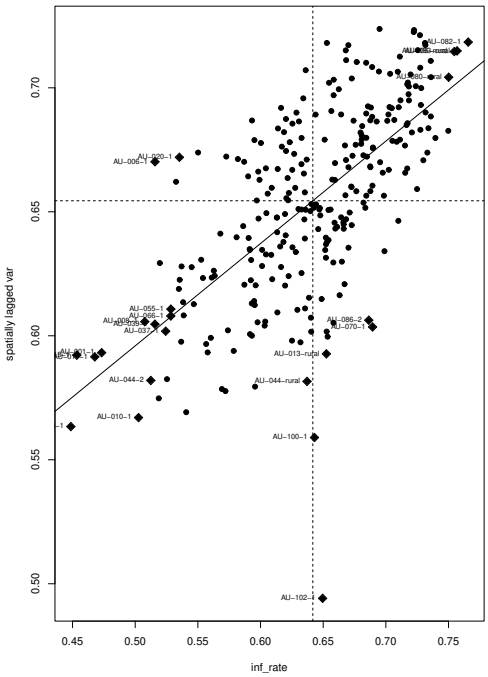

Figure 24: Moran Diagram for the No High School Diploma Rate.

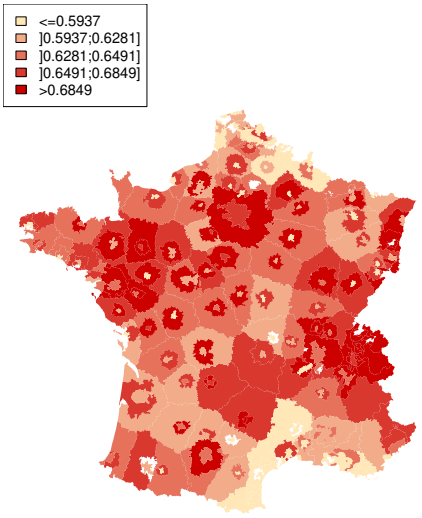

Figure 25: Spatial Distribution of the Employment Rate.

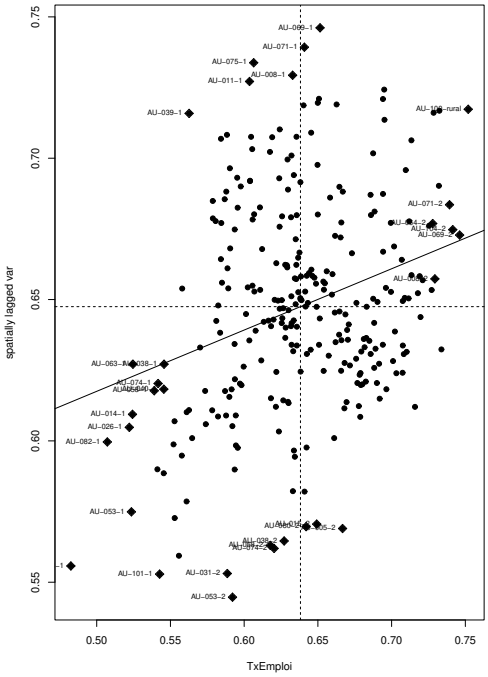

Figure 26: Moran Diagram for the Employment Rate.

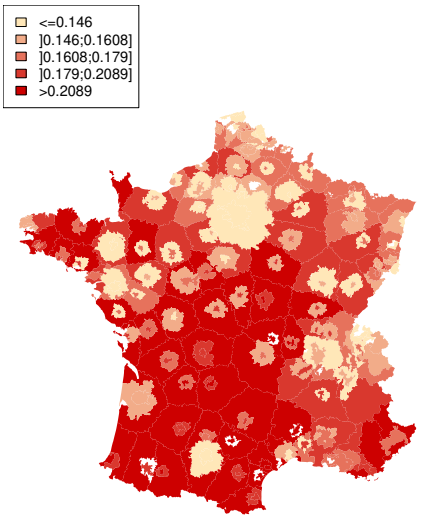

Figure 27: Spatial Distribution of the Senior Citizens Rate.

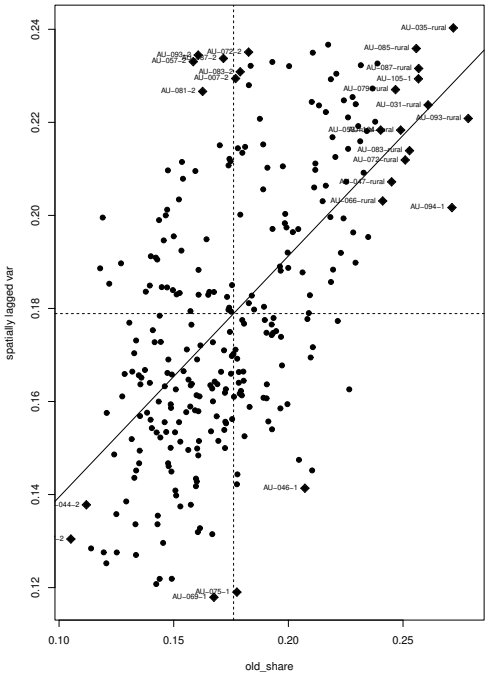

Figure 28: Moran Diagram for the Senior Citizens Rate.

5.3 Variable construction

5.3.1 Living Environment

Our raw data are at the commune level. We therefore must aggregate the variables to the division level, d, where d ranges from 1 to 281. Each division d is composed of i communes, and the number of communes in one division ranges from 1 to n_d. The population of division d is thus given by

$$Population_d = \sum_{1}^{n_d} Population_i \tag{5}$$

Ratios

Our availability ratio gives the number of 'amenities X' per 1,000 inhabitants in division d. X_d is either the number of supermarkets (SM_d), the number of fast-food restaurants (FF_d), the number of sports amenities (SA_d) or the number of sports licenses (SL_d) in division d. Then, the number of amenities at the division level (X_d) is given by:

$$X_d = \sum_{1}^{n_d} X_i \tag{6}$$

- FF_d, the **Number of fast food restaurants** comes from summing the number of fast food restaurants at the commune level.
- SM_d, the **number of supermarkets** gives the number of stores larger than 400 m^2 in division d. In the raw data, we have the number of stores larger than 2500 m^2 and the number of stores between 400 m^2 and 2500 m^2 at the commune level. We thus start by computing SM_i, where SM_i = (number of stores larger than 2500 m^2 + number of stores between 400 m^2 and 2500 m^2) in commune i.
- SA_d, **the number of sports amenities** gives the number of sports amenities in division d. In the raw data, we observe the number of sports amenities by type j in commune i (j can be a soccer field, playground, swimming pool, tennis court, etc.), and j ranges from 1 to 77. Therefore, $SA_i = \sum_{1}^{77} nb\ Sports\ Amenities_j$.
- SL_d, **the number of sports licenses** gives the number of sports licenses in division d. In the raw data, we observe this number for commune i and sport federation k, where k ranges from 1 to 71. Therefore, SL_i is given by $\sum_{1}^{71} Number\ of\ Licenses\ for\ Sport\ k$.

The amenities X ratio is thus given by:

$$X_d\ Ratio = \frac{X_d}{Population_d} * 1{,}000 \text{ inhabitants} \tag{7}$$

Rates

We have four rates to be defined:

- The **General practitioner rate** is the percentage of people within a division who have at least one general practitioner in their own commune (from the raw data), where $GP_d = \sum_{1}^{n_d} Population_i$ if the number of GPs >1 in commune i.

  Then,

$$GP_d\ Rate = \frac{GP_d}{Population_d} * 100 \tag{8}$$

- The **Senior Citizens Rate** is the percentage of people over 65 years old in each division d, where $65_d = \sum_{1}^{n_d} Population > 65_i$ (the number of individuals over 65 years old in commune i).

  Then,

$$65_d\ Rate = \frac{65_d}{Population_d} * 100 \tag{9}$$

- The **No High School Diploma Rate** is the percentage of people who have less than a high school diploma in each division. The raw data give the number of individuals who have less than a high school diploma in commune i, $noHSD\_i$. We thus compute $noHSD_d = \sum_{1}^{n_d} noHSD_i$

  Then, the **No High School Diploma Rate** is given by:

$$NHSD_d\ Rate = \frac{noHSD_d}{Population_d} * 100 \tag{10}$$

- The **Employment Rate** is the share of individuals with a job among those who are of working age. Our raw data are at the commune level. We thus compute ER_d at the division level as:

$$ER_d = \frac{\sum_{1}^{n_d} Population\ occupied_i}{\sum_{1}^{n_d} Working\ Age\ Population_i\ (15\text{-}64)} \tag{11}$$

Others

- **Density** gives the number of individuals per square kilometer in division d. In the raw data, we have the land area of each commune i in square kilometers, so we can compute $Area_d = \sum_{1}^{n_d} Area_i$, which gives $Density_d$ as

$$Density_d = \frac{Population_d}{Area_d} \tag{12}$$

- **Income** is the average income of the 'fiscal entities' in each division. In the raw data, we have the total taxable income and the total number of fiscal entities at the commune level. To aggregate the data to the division level, we compute I_d as a mean:

$$I_d = \frac{\sum_{1}^{n_d} Total\ Taxable\ Income_i}{\sum_{1}^{n_d} Number\ of\ Fiscal\ Entities_{\mathrm{i}}}$$

- **Time to Amenity Y**, in division d, represents the weighted average driving time for individuals who live in communes where amenity Y is missing (SM_i= 0 (no supermarkets) or FF_i= 0 (no fast food restaurants)). We first select those communes where Y is missing and calculate the division-level total number of individuals who do not have access to amenity Y in their commune as $Population_{dY=0} = \sum_{1}^{n_d} Population_{iY=0}$, where $Population_{iY=0}$ is the population in commune i, which does not have amenity Y. Next, we collect the driving time, Time to j, which is the driving time from commune $\mathrm{i}_{Y=0}$ to the nearest commune j with amenity Y using Odomatrix. We thus compute the following variable at the division level:

$$Time\ to\ Y_d = \sum_{1}^{nD} TimetoY_i * \frac{Population_i}{Population_{dY=0}} \tag{13}$$

5.4 Equality of Means

Table 3: Test of Equality of Means (ANOVA Test)

	City Centers	Crowns	Nonurban areas	pvalue	W pvalue	
Health and Consumption						
Obesity Rate	0.12	0.11	0.14	0.041	0.027	0.046
sd	0.07	0.08	0.05	.	.	.
BMI	24.84	24.92	25.20	0.023	0.014	0.026
sd	1.03	1.00	0.79	.	.	.
Caloric Intake (Tot, Kcal/100)	31.62	33.55	34.01	0.008	0.005	0.014
sd	5.38	6.55	5.10	.	.	.
Caloric Intake (F&V, Kcal/100)	2.83	2.59	2.63	0.022	0.040	0.022
sd	0.73	0.63	0.52	.	.	.
Food Basket Price	145.00	141.80	140.85	0.000	0.000	0.000
sd	7.08	7.30	7.35	.	.	.
Living Environment						
Food Amenities						
Supermarket Ratio	0.21	0.12	0.24	0.000	0.000	0.000
sd	0.05	0.05	0.04	.	.	.
Time to Supermarkets	5.80	7.78	8.38	0.000	0.000	0.000
sd	2.59	1.69	2.56	.	.	.
Convenience store Ratio	0.34	0.33	0.47	0.000	0.000	0.000
sd	0.13	0.15	0.24	.	.	.
Fast Food Ratio	0.04	0.01	0.02	0.000	0.000	0.000
sd	0.01	0.01	0.01	.	.	.
Time to Fast Food	10.48	18.52	21.12	0.000	0.000	0.000
sd	3.28	3.91	6.08	.	.	.
Sport Amenities						
Sports Amenities Ratio	2.19	3.41	4.25	0.000	0.000	0.000
sd	0.68	1.38	1.79	.	.	.
Sports Licenses Ratio	171.22	186.91	166.10	0.000	0.000	0.000
sd	26.39	28.44	28.99	.	.	.
Health Amenities						
General Practitioner Rate	0.98	0.62	0.69	0.000	0.000	0.000
sd	0.03	0.18	0.12	.	.	.
Controls						
Income	21261.73	23466.62	19764.40	0.000	0.000	0.000
sd	2355.44	2652.25	1525.61	.	.	.
No High School Diploma Rate	0.61	0.63	0.69	0.000	0.000	0.000
sd	0.06	0.04	0.04	.	.	.
Empl. Rate	0.60	0.68	0.64	0.000	0.000	0.000
sd	0.04	0.04	0.04	.	.	.
Senior Citizen Rate	0.17	0.15	0.20	0.000	0.000	0.000
sd	0.03	0.02	0.03	.	.	.
Nb of divisions	99.00	86.00	96.00	.	.	.

W test and Simanova test are added to test the case of unequal variances across divisions.
Indicated values, in the first three columns, are means and standard deviations (sd).

5.5 SLX estimation: Robustness checks

Table 4: Estimation results (SLX models : robustness checks)

	Model A		Model D		Model E		Model F	
X Variables								
Consumption								
Caloric Intake (Tot, Kcal/100)	0.002*	(0.001)	0.002**	(0.001)	0.001	(0.001)	0.002	(0.001)
Caloric Intake (F&V, Kcal/100)	-0.016*	(0.009)	-0.015**	(0.007)	-0.015	(0.009)	-0.016*	(0.009)
Food Basket Price	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)
Living Environment								
Food Amenities								
Supermarket Ratio	-0.112	(0.092)	-0.107	(0.092)	-0.129	(0.093)	-0.112	(0.092)
Time to Supermarkets	-0.002	(0.002)	-0.002	(0.002)	-0.002	(0.002)	-0.002	(0.002)
Fast Food Ratio	0.856**	(0.390)	0.853**	(0.389)	0.645*	(0.389)	0.778**	(0.391)
Time to Fast Food	0.001	(0.001)	0.001	(0.001)	0.001	(0.001)	0.001	(0.001)
Sport Amenities								
Sports Amenities Ratio	-0.005	(0.003)	-0.005	(0.003)	-0.005*	(0.003)	-0.004	(0.003)
Sports Licenses Ratio	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Health Amenities								
General Practitioner Rate	0.014	(0.038)	0.014	(0.038)	0.020	(0.038)	0.020	(0.039)
Controls								
Income	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)
No High School Diploma Rate	0.391***	(0.124)	0.393***	(0.124)	0.336**	(0.133)	0.376***	(0.126)
Empl. Rate	0.126	(0.166)	0.122	(0.166)	0.211	(0.167)	0.127	(0.173)
Senior Citizen Rate	-0.314	(0.227)	-0.334	(0.227)	0.078	(0.213)	-0.160	(0.226)
Density	0.000*	(0.000)	0.000*	(0.000)	0.000	(0.000)	0.000*	(0.000)
Crowns	-0.010	(0.023)	-0.011	(0.023)	-0.009	(0.024)	-0.001	(0.024)
Nonurban Areas	0.017	(0.018)	0.016	(0.018)	0.010	(0.019)	0.020	(0.018)
Constant	-0.395*	(0.213)	-0.387*	(0.210)	-0.381	(0.315)	-0.341	(0.228)
WX Variables								
Consumption								
Caloric Intake (Tot, Kcal/100)	0.002	(0.001)	0.001	(0.001)	0.006**	(0.003)	0.002	(0.002)
Caloric Intake (F&V, Kcal/100)	-0.020	(0.014)	-0.015	(0.012)	-0.045	(0.032)	-0.018	(0.017)
Food Basket Price	0.003***	(0.001)	0.003***	(0.001)	0.003*	(0.002)	0.003***	(0.001)
Living Environment								
Food Amenities								
Supermarket Ratio	0.138	(0.139)	0.126	(0.139)	0.174	(0.289)	0.147	(0.153)
Time to Supermarkets	0.002	(0.004)	0.001	(0.004)	0.005	(0.007)	0.001	(0.004)
Fast Food Ratio	-0.497	(0.669)	-0.490	(0.663)	-2.234	(1.438)	-0.400	(0.782)
Time to Fast Food	-0.002	(0.002)	-0.002	(0.002)	-0.007**	(0.003)	-0.001	(0.002)
Sport Amenities								
Sports Amenities Ratio	-0.010**	(0.004)	-0.010**	(0.004)	-0.020***	(0.007)	-0.011**	(0.005)
Sports Licenses Ratio	0.001**	(0.000)	0.001**	(0.000)	0.001***	(0.000)	0.001**	(0.000)
Health Amenities								
General Practitioner Rate	-0.159***	(0.043)	-0.157***	(0.043)	-0.236***	(0.065)	-0.174***	(0.048)
Controls								
Income	-0.000*	(0.000)	-0.000*	(0.000)	-0.000	(0.000)	-0.000*	(0.000)
Senior Citizen Rate	0.678**	(0.296)	0.684**	(0.294)	1.054**	(0.467)	0.450	(0.314)
Density	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Crowns	-0.011	(0.045)	-0.013	(0.045)	-0.129	(0.101)	-0.010	(0.051)
Nonurban Areas	-0.019	(0.042)	-0.018	(0.042)	-0.096	(0.089)	-0.039	(0.046)
/								
var(λ)	0.003***	(0.000)	0.003***	(0.000)	0.003***	(0.000)	0.003***	(0.000)
Observations	281		281		281		281	
Pseudo R^2	0.282		0.286		0.279		0.266	
AIC	-758.625		-760.195		-757.509		-752.338	

Standard errors in parentheses
Model A: SLX estimation (base model).
Model D: SLX estimation with the per capita consumption based on Adult Male Equivalent units - FAO.
Model E: SLX estimation where the neighborhood matrix used is a second order contiguity weighting matrix.
Model F: SLX estimation where the neighborhood matrix used is based on the contiguity matrix (first-order neighbors) where we add a neighborhood link between nonurban areas and the city center division to which the nonurban area is directly attached.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Estimation results (Estimation strategy)

	OLS		SDEM		SLX (Model A)		slx_gs2sls		SEM	
X Variables										
Consumption										
Caloric Intake (Tot, Kcal/100)	0.0018*	(0.001)	0.0017*	(0.001)	0.0017*	(0.001)	0.0017*	(0.001)	0.0016	(0.001)
Caloric Intake (F&V, Kcal/100)	-0.0171*	(0.009)	-0.0152*	(0.009)	-0.0155*	(0.009)	-0.0155	(0.010)	-0.0147	(0.009)
Food Basket Price	-0.0004	(0.001)	-0.0006	(0.001)	-0.0006	(0.001)	-0.0006	(0.001)	-0.0007	(0.001)
Living Environment										
Food Amenities										
Supermarket Ratio	-0.0339	(0.095)	-0.1090	(0.092)	-0.1118	(0.092)	-0.1118	(0.107)	-0.0580	(0.093)
Time to Supermarkets	-0.0008	(0.002)	-0.0024	(0.002)	-0.0023	(0.002)	-0.0023	(0.002)	-0.0012	(0.002)
Fast Food Ratio	0.7075*	(0.387)	0.8642**	(0.395)	0.8562**	(0.390)	0.8562*	(0.439)	0.7288**	(0.364)
Time to Fast Food	0.0005	(0.001)	0.0010	(0.001)	0.0010	(0.001)	0.0010	(0.001)	0.0007	(0.001)
Sport Amenities										
Sports Amenities Ratio	-0.0077**	(0.003)	-0.0049	(0.003)	-0.0046	(0.003)	-0.0046	(0.003)	-0.0071**	(0.003)
Sports Licenses Ratio	0.0002	(0.000)	0.0000	(0.000)	0.0000	(0.000)	0.0000	(0.000)	0.0001	(0.000)
Health Amenities										
General Practitioner Rate	-0.0519	(0.035)	0.0157	(0.037)	0.0145	(0.038)	0.0145	(0.044)	-0.0285	(0.038)
Controls										
Income	-0.0000	(0.000)	-0.0000	(0.000)	-0.0000	(0.000)	-0.0000	(0.000)	-0.0000	(0.000)
No High School Diploma Rate	0.3472***	(0.117)	0.4074***	(0.129)	0.3911***	(0.124)	0.3911***	(0.148)	0.3527***	(0.119)
Empl. Rate	0.1052	(0.159)	0.1360	(0.172)	0.1261	(0.166)	0.1261	(0.165)	0.1464	(0.163)
Senior Citizen Rate	-0.0305	(0.162)	-0.3555	(0.226)	-0.3144	(0.227)	-0.3144	(0.232)	-0.0901	(0.168)
Density	0.0000*	(0.000)	0.0000*	(0.000)	0.0000*	(0.000)	0.0000**	(0.000)	0.0000	(0.000)
Crowns	0.0003	(0.021)	-0.0109	(0.023)	-0.0102	(0.023)	-0.0102	(0.023)	0.0001	(0.020)
Nonurban Areas	0.0079	(0.018)	0.0162	(0.018)	0.0166	(0.018)	0.0166	(0.018)	0.0098	(0.018)
Constant	-0.0386	(0.161)	-0.3921*	(0.223)	-0.3947*	(0.213)	-0.3947*	(0.233)	-0.0249	(0.162)
WX Variables										
Consumption										
Caloric Intake (Tot, Kcal/100)			0.0015	(0.002)	0.0019	(0.001)	0.0019	(0.001)		
Caloric Intake (F&V, Kcal/100)			-0.0173	(0.015)	-0.0202	(0.014)	-0.0202	(0.015)		
Food Basket Price			0.0027***	(0.001)	0.0028***	(0.001)	0.0028***	(0.001)		
Living Environment										
Food Amenities										
Supermarket Ratio			0.1435	(0.142)	0.1379	(0.139)	0.1379	(0.166)		
Time to Supermarkets			0.0022	(0.004)	0.0016	(0.004)	0.0016	(0.004)		
Fast Food Ratio			-0.4516	(0.681)	-0.4970	(0.669)	-0.4970	(1.003)		
Time to Fast Food			-0.0017	(0.002)	-0.0018	(0.002)	-0.0018	(0.002)		
Sport Amenities										
Sports Amenities Ratio			-0.0094**	(0.004)	-0.0101**	(0.004)	-0.0101**	(0.004)		
Sports Licenses Ratio			0.0006**	(0.000)	0.0006**	(0.000)	0.0006**	(0.000)		
Health Amenities										
General Practitioner Rate			-0.1528***	(0.043)	-0.1586***	(0.043)	-0.1586***	(0.047)		
Controls										
Income			-0.0000*	(0.000)	-0.0000*	(0.000)	-0.0000*	(0.000)		
Senior Citizen Rate			0.6705**	(0.296)	0.6785**	(0.296)	0.6785*	(0.355)		
Density			0.0000	(0.000)	0.0000	(0.000)	0.0000	(0.000)		
Crowns			-0.0122	(0.046)	-0.0114	(0.045)	-0.0114	(0.043)		
Nonurban Areas			-0.0190	(0.042)	-0.0190	(0.042)	-0.0190	(0.042)		
λ			0.0915	(0.077)					0.1395*	(0.079)
/										
var(λ)			0.0031***	(0.000)	0.0031***	(0.000)			0.0034***	(0.000)
Obs.	281		281		281		281		281	
Adj.r2	0.14									
AIC	-757.38		-758.03		-758.62		.		-756.44	
Moran Test	0.059									
Test LM-Error	0.12									
Test LM-Lag	0.04									
Robust test LM-Error	0.16									
Robust test LM-Lag	0.05									
LR test SDEM/SLX					0.24					
LR test SDEM/SEM									0.01	
LR test SLX/OLS					0.00					

Standard errors in parentheses.
SLX model (Spatial Lag of X Model): $y = Xb + WX\theta + \varepsilon$
SEM (Spatial Error Model : $y = Xb + u$ with $u = \lambda Wu + \varepsilon$
SDEM (Spatial Durbin Error Model) : $y = Xb + WX\theta + u$ with $u = \lambda Wu + \varepsilon$
SAR, SDEM, SLX, SEM, SDM, SAC use maximum likelihood estimator.
slx_gs2sls uses generalized spatial two-stage least-squares estimator.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: OLS and SLX estimation results (with total, direct and indirect effects)

	OLS		SLX (Model A)							
			Total effect		Direct effects (X Variables)		Spatial effects (WX Variables)		/	
Consumption										
Caloric Intake (Tot, Kcal/100)	0.0018*	(0.001)	0.0036**	(0.002)	0.0017*	(0.001)	0.0019	(0.001)		
Caloric Intake (F&V, Kcal/100)	-0.0171*	(0.009)	-0.0358**	(0.018)	-0.0155*	(0.009)	-0.0202	(0.014)		
Food Basket Price	-0.0004	(0.001)	0.0022**	(0.001)	-0.0006	(0.001)	0.0028***	(0.001)		
Living Environment										
Food Amenities										
Supermarket Ratio	-0.0339	(0.095)	0.0261	(0.158)	-0.1118	(0.092)	0.1379	(0.139)		
Time to Supermarkets	-0.0008	(0.002)	-0.0008	(0.004)	-0.0023	(0.002)	0.0016	(0.004)		
Fast Food Ratio	0.7075*	(0.387)	0.3591	(0.872)	0.8562**	(0.390)	-0.4970	(0.669)		
Time to Fast Food	0.0005	(0.001)	-0.0008	(0.002)	0.0010	(0.001)	-0.0018	(0.002)		
Sport Amenities										
Sports Amenities Ratio	-0.0077**	(0.003)	-0.0148***	(0.004)	-0.0046	(0.003)	-0.0101**	(0.004)		
Sports Licenses Ratio	0.0002	(0.000)	0.0007***	(0.000)	0.0000	(0.000)	0.0006**	(0.000)		
Health Amenities										
General Practitioner Rate	-0.0519	(0.035)	-0.1441***	(0.043)	0.0145	(0.038)	-0.1586***	(0.043)		
Controls										
Income	-0.0000	(0.000)	-0.0000*	(0.000)	-0.0000	(0.000)	-0.0000*	(0.000)		
No High School Diploma Rate	0.3472***	(0.117)	0.3911***	(0.124)	0.3911***	(0.124)				
Empl. Rate	0.1052	(0.159)	0.1261	(0.166)	0.1261	(0.166)				
Senior Citizen Rate	-0.0305	(0.162)	0.3641	(0.254)	-0.3144	(0.227)	0.6785**	(0.296)		
Density	0.0000*	(0.000)	0.0001	(0.000)	0.0000*	(0.000)	0.0000	(0.000)		
Crowns	0.0003	(0.021)	-0.0216	(0.054)	-0.0102	(0.023)	-0.0114	(0.045)		
Nonurban Areas	0.0079	(0.018)	-0.0024	(0.045)	0.0166	(0.018)	-0.0190	(0.042)		
var(e.ObesityRate)									0.0031***	(0.000)
Constant	-0.0386	(0.161)			-0.3947*	(0.213)				
Observations	281				281					
Pseudo R^2					0.282					
AIC	-757.3846		.		-758.6245					

Standard errors in parentheses.
Standard errors are calculated by the delta method.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$