



HAL
open science

Spatial disparities in accessibility to recreational amenities: the case of Pokémon GO

Florent Sari

► **To cite this version:**

Florent Sari. Spatial disparities in accessibility to recreational amenities: the case of Pokémon GO. *Annals of Regional Science*, 2020, 64 (2), pp.389-419. 10.1007/s00168-019-00961-8 . hal-03709678

HAL Id: hal-03709678

<https://nantes-universite.hal.science/hal-03709678v1>

Submitted on 6 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SPATIAL DISPARITIES IN ACCESSIBILITY TO RECREATIONAL AMENITIES: THE CASE OF POKÉMON GO

Florent Sari

University of Nantes, Lemna and Tepp-CNRS

florent.sari@univ-nantes.fr

Chemin de la censive du Tertre, 44322 Nantes Cedex 3

Abstract

In this study, we analyze spatial disparities in access to particular recreational amenities. We focus on the famous application for smart phones: *Pokémon GO*. The aim of this location-based augmented reality application is to capture and collect virtual creatures called *Pokémon*, using a cell phone. Collecting Pokemon is made possible by using certain items only available in particular places, called *Pokéstops*, that the player has to visit in order to obtain. Pokestops are also very useful to players because the chances of finding Pokémon within its vicinity are higher as opposed to elsewhere. Because these Pokéstops are not randomly located, some spatial disparities have been observed by players. Our objective is to explain these spatial disparities in French cities. We focus on the French department of *Loire-Atlantique* and Nantes metropolis that permit use, to compare situations and determinants in urban and rural neighborhoods. Results of different spatial econometric models reveal strong disparities between neighborhoods located in the Nantes metropolis, and others. It appears that determinants of these disparities may vary according to whether an individual is located in a rural or urban area. Location of Pokéstops does not seem to be correlated with population characteristics, but rather with local amenities such as historical monuments, shops, or green spaces.

KEYWORDS: *Geographic analysis; Leisure amenities; Location-based augmented reality application; Spatial disparities.*

JEL CODE: *H41; R19.*

Introduction

Pokemon GO is a location-based augmented reality application for smart phones released in early July 2016¹. Created by *The Pokémon Company* and *Niantic*, the game has enjoyed immediate success. It reached 75 million downloads (Android and iOS included) in only 19 days. One month later, the total number of downloads was more than 500 million. It was the most downloaded iOS application worldwide in 2016. During the Game Developers Conference held in San Francisco in March 2017, Niantic's founder and chief technology officer revealed the game had been downloaded over 650 million times. In May 2018, the total number of downloads had reached 800 million, and the total number of monthly users was 147 million worldwide (the highest peak since the game's launch)².

The concept of this free-to-play game is to capture and collect virtual creatures called *Pokémon* with a cell phone. The player also has the ability to battle their own creatures with others in arenas, with the objective of becoming the arena master. The creatures may be placed anywhere in public locations such as streets, parks, buildings etc. The player may observe their location on a local map before walking to certain places to capture the creature. Because creatures appear temporarily at a given location, they can be captured only during a limited time. It is important to note that capture by one player does not prevent another from capturing it as well. For this reason, creatures present the characteristics of a public good (*i.e.* non-rival and non-exclusive good).

Catching Pokémon is one primary way players can progress in the game. Collecting the existing 488 creatures is one objective for players. In addition, when players collect creatures, they earn experience points (called "XP") that permit them to make their creatures more powerful. The collection of Pokemon is made possible by using items, such as *Poké Balls* or incense (which catches or attracts them). These items are available in particular places called *Pokéstops* that players have to visit in order to obtain them³. These places are very useful to players because they permit them to receive items essential to capturing Pokémons, potions for healing them after battles in arenas, and even XP. Pokéstops also permit players to increase their chances of finding Pokémons, because the apparition frequency is higher in such places. As a result, a strong Pokéstop density in a territory is seen as an important advantage for players.

For all these reasons, a Pokéstop might be considered as a recreational amenity for individuals (for those interested in the game) in the same way as a green space, a playground, or a sporting facility. Various empirical works have already shown access to public amenities such as parks, recreational facilities, and social and cultural services may be beneficial to a resident's well-being and life satisfaction (Gidlöf-Gunnarsson and Öhrström, 2007; Shields *et al.*, 2009; Ambrey and Fleming, 2014; Carrus *et al.*, 2015 etc.). As a result, areas which include diversity and concentration of amenities are more likely to see important population growth, and attract highly-educated individuals (Hunter *et al.*, 2005; Carlino and Saiz, 2008; Frenkel *et al.*, 2013). However,

¹ Augmented reality consists in the integration of digital information with the user's environment in real time. It differs from virtual reality that creates an artificial environment.

² Source: Superdata, May 2018.

³ Pokéstops may be visited many times. However, the player has to wait 5 minutes each time.

the geography of these various amenities reveals some disparities between and among territories (Neckerman *et al.*, 2009; Dahman *et al.*, 2010) that are often linked to socio-economic characteristics of inhabitants (Moore *et al.*, 2008; Billaudeau *et al.*, 2011; Duncan *et al.*, 2013 etc.).

Even if Pokéstops represent a small subset of virtual recreational amenities among a variety of others, we want to assess their geography and determine whether correlations exist between their location and socio-economic characteristics of the neighborhoods where they are found. Indeed, in the U.S., some newspaper articles and research papers have shown that there are more Pokéstops in white-majority neighborhoods than in other neighborhoods, and it has been strongly argued that the application has increased socio-spatial segregation⁴. In the French situation, no empirical research has been released. Only a *Le Monde* newspaper has shown that Pokéstops are more abundant in urban rather than rural areas, and that their location does not seem to be linked to income criteria⁵. If so, it is important to provide explanations of Pokéstops geography in French cities because if the phenomenon is associated with spatial inequalities, a better spatial distribution could have a positive effect on an individuals well-being. In addition, recent studies have shown Pokémon GO to have positive effects on the physical well-being of players (Althoff *et al.*, 2016; Wong, 2017) and a small yet positive effect on local businesses associated with it (Pamuru *et al.*, 2018). Finally, as this particular recreational amenity might have consequences on life satisfaction, public health or even economic activity, it seems necessary to highlight the underlying mechanisms involved in order to determine the positive policies or regulations.

The objective of this study is twofold. *Firstly*, we studied the spatial distribution of Pokéstops in neighborhoods in the whole French department of Loire-Atlantique and more specifically in the Nantes metropolis. Focusing on this large territory allows us to compare situations in large and small cities, but also in urban and rural areas. We used Explanatory Spatial Data Analysis (ESDA) to reveal spatial patterns of this particular amenity and the presence of local clusters, or spatial dependency phenomena. *Secondly*, we used spatial econometrics models to explain spatial disparities in the accessibility of Pokéstops. These models allow us to assess the role of various local characteristics of the observed geography, but also to consider problems of spatial autocorrelation and/or spatial heterogeneity.

Results of our models reveal strong disparities between neighborhoods located in the Nantes metropolis and others. It appears that determinants of these observed disparities may vary according to whether an individual is located in a rural or urban area. Location of Pokéstops does not seem to be correlated with population characteristics, but rather with local amenities such as historical monuments, shops, or green spaces.

The remainder of the manuscript is structured as follows: After presenting the potential links between access to amenities and well-being, the first section focuses on research analyzing the role of a neighborhoods' socio-economic composition. We also review recent publications that

⁴ Huffaker C. "There are fewer Pokemon Go locations in black neighborhoods, but why?", *Miami Herald*, July 29, 2016; Akthar A. "Is Pokémon Go racist? How the app may be redlining communities of color", *USA Today*, August 9 2016.

⁵ Leloup D. and Grandin J. "Pokémon Go: les multiples facteurs des inégalités géographiques", *Le Monde*, August 8, 2016.

study the impact of Pokemon GO on its users and on local business. The second section describes the data and variables retained for the analysis. The third section presents results of the ESDA and descriptive statistics for the department of Loire-Atlantique. The fourth section describes the estimation strategy, and the final section the results of our different models.

1 Literature review

1.1 Access to amenities and well-being

Literature relating to urban economics and urban planning has studied the consequences of accessibility to natural and/or recreational amenities on well-being, in several contexts. For instance, Gidlöf-Gunnarsson and Öhrström (2007) examine whether perceived accessibility to nearby green areas affects various aspects of well-being for inhabitants exposed to high road-traffic noise. The results show that vicinity to green areas is important for individuals' well-being and daily behavior. Shields *et al.* (2009) also investigates the association between neighborhood characteristics and life satisfaction. They find the most important determinant is the extent of neighborly social interaction, while more tangible perceptions of the neighborhood seem to be unimportant. It turns out this is precisely one of the characteristics of Pokemon GO developed by *Niantic*. Indeed, it impacts social interactions by encouraging collaboration or even competition between players. In the same way, Ambrey *et al.* (2012) examines the influence of public greenspace access on the life satisfaction of residents of Australia's capital cities. The authors find a positive relationship between the presence of public greenspace in a residents' local area and their life satisfaction.

Other studies have focused on the links between access to recreational amenities and population growth. For example, Carlino and Saiz (2008) explore how leisure consumption opportunities affected metropolitan statistical population growth areas during the 1990s. Their findings suggest that population growth was about 2 % higher in an area with twice as many leisure options compared to another area. More recently, Frenkel *et al.* (2013) investigated the residential location choice of skilled-workers in metropolitan areas, and found that cultural and educational land-use are important factors. Through these findings, we observe that vicinity to some specific amenities is a source of attraction for people, and therefore is a factor of population growth. For this reason, urban planners must pay attention to spatial distribution of valued amenities when they consider their policies.

In our study, a virtual recreational amenity such as a Pokéstop represents only one particular amenity, among others. Its accessibility is valued by inhabitants, and that could impact their life satisfaction and well-being. In fact, various studies consider that well-being is multi-dimensional, and access to some services or amenities might be one of the various existing dimensions. This is proposed by Witten *et al.* (2003) whereby an index based on locational access to six domains, among which include recreational amenities, allows us to identify relationships between opportunity structures in the local environment, and residents' well-being⁶. In the case of Tovar

⁶ Others domains are: public transport and communication, shopping and banking facilities, educational services, health services, and social and cultural services).

and Bourdeau-Lepage (2013) a multidimensional well-being indicator was developed based on Sen's capability approach (Sen, 1985; 1999 etc.). Indeed, their capability indicator gathered different forms of well-being: well-being as freedom, well-being as a choice or agency freedom, and well-being as effective realizations. The last one takes into account accessibility to services and consumption, but could also consider accessibility to recreational amenities.

Finally, because Pokéstops may represent recreational amenities that determine the life satisfaction and well-being of inhabitants, it appears interesting to analyze their spatial distribution and determinants. In particular, we investigate if relationships exist with neighborhood socio-economic characteristics.

1.2 Neighborhood conditions and access to recreational amenities

Numerous studies have analyzed the disparities in spatial accessibility to recreational amenities, and their relationships with neighborhood compositional characteristics. Nicholls (2001) examined the spatial disparities in the distribution and accessibility levels of parks and green spaces in Bryan (Texas). Strong local disparities were observed: only 40 % of inhabitants have good access to these spaces, and ethnic minorities and low-income households seem to be relatively disadvantaged. Moore *et al.* (2008) analyzed the availability of recreational facilities and parks in North Carolina, New York and Maryland. The authors found that neighborhoods characterized by high percentages of ethnic minorities, or low-income levels, were less likely to have access to recreational facilities. More recently, Wen *et al.* (2013) examined the links between spatial access to parks or green spaces, and percentages of ethnic minorities and low-income inhabitants in urban and rural areas. The results reveal a negative relationship between poverty rates and percentage of green spaces in urban areas, but a positive relationship in rural areas. They also reveal a negative relationship between the percentage of ethnic minorities and presence of green spaces.

A second group of studies focused on accessibility to urban services, such as playgrounds or sport facilities. For example, Talen and Anselin (1998) assessed to what extent the distribution of urban public services is equitable. They were interested in the spatial distribution of playgrounds in Tulsa (Oklahoma) compared to the location of children and minority populations. The results were inconclusive because they found that the choices of measuring access affected conclusions about the existence of spatial inequities. Smoyer *et al.* (2004) compared the spatial accessibility to playgrounds in Edmonton (Canada) to demographic and social needs in certain neighborhoods. It appears playground location is positively correlated with social need, but weakly and negatively correlated with demographic need. Concerning spatial disparities in access to sport facilities, Estabrooks *et al.* (2008) studied whether accessibility differed according to neighborhood socio-economic status in a small U.S. city. They found that neighborhoods with low and medium socio-economic status had fewer physical activity resources available. Panter *et al.* (2008) examined correlations between households' income, access to sports facilities, and physical activity in Norwich (England), and observed that low income households are least well served by affordable facilities. In the French case, Billaudeau *et al.* (2011) analyzed the spatial distribution of sport facilities in the Paris region. They revealed that the number of immediately accessible tennis courts or private sporting spaces increases in-line with an area's income, while the number of accessible

public playgrounds or athletic facilities tends to decrease. Finally, Duncan *et al.* (2013) assessed the association between minority neighborhood ethnic composition, neighborhood poverty, and recreational open spaces in Boston, and found Black neighborhoods were less likely to have public recreational spaces.

In summary, most of these studies confirm some relationship between access to amenities for inhabitants, and a neighborhoods' socio-economic characteristics. However, effects may vary according to the cities studied, amenities concerned (parks, health facilities, recreational resources etc.) or even characteristics examined. Our study is somewhat different, because we are focused on a virtual recreational amenity whose characteristics do not vary (free access for everyone, and remains the same everywhere).

1.3 Effects of Pokemon GO on users and local business

The immediate success of the application and the large number of daily users has raised various questions and debates concerning Pokemon Go's observed and expected effects. Initial positive effects on physical activity were put forward. Because walking is encouraged by the game, it implies an increase of players' physical activity, and more generally, improvement of public health. This aspect has been observed in the recent studies of Althoff *et al.* (2016), Howe *et al.* (2016) and Wong (2017). Using different methodologies, the authors found the application increased the physical activity of players, but mainly in the short term. Whilst the effects were strong during the first weeks of playing, these diminished with time. However, they found that players who used to be sedentary benefited the most from the game.

Some positive effects on social relationships and local businesses could also be observed. Because players have to go outside in public spaces to find creatures, they can meet each other. Various groups or special events were created on social networks (Facebook, Twitter etc.) to encourage information exchange between players, and to organize meetings in parks, for example. In addition, players must choose between three different competing teams (yellow, blue and red). This choice encourages collaboration between players of the same team, to challenge the others. A recent study of Graells-Garrido *et al.* (2017) proposes a natural experiment (using data from mobile phone networks) to assess the effects of the game on floating population patterns in Santiago (Chile). The authors found there were more people on the street at commuting times, and that they were also playing in areas near their homes during the week-end. Moreover, Pamuru *et al.* (2018) investigated the economic implications of Pokemon GO on local activity in the U.S., and in particular, restaurants. They found that restaurants located near Pokéstops observed a higher number of customers compared to restaurants that do not have Pokéstops nearby.

Concerning negative effects, the risks of accidents for drivers or pedestrians were commonly put forward. It is difficult to measure the relevance of these risks, because real statistics on this topic are not available. Nevertheless, the French association "40 millions d'automobilistes" (40 million drivers) has estimated the number of individuals playing the game while driving is 3.4 million

(in August 2016)⁷. *Niantic* has considered measures to prevent this problem, namely by preventing the application from functioning normally if the player's detected speed is too high. However, according to the study of Ono *et al.* (2017), mobilizing regression discontinuity methods, and data from the Japanese Institute for Traffic Accident Research and Data Analysis, shows that the effect of Pokémon GO on fatal traffic injuries is negligible.

Because Pokémon GO is a location-based augmented reality application, the question relating to places that are advantaged or not has also quickly emerged. An advantaged place is an area where players can find large number of Pokéstops. It represents a serious advantage for players for at least two reasons. *Firstly*, it increases their chances of finding Pokémon, as Pokémon appear frequently around Pokéstops. *Secondly*, it is easier for a player to obtain items necessary to catch them. Indeed, a player can move between Pokéstops, instead of waiting for five minutes, before revisiting the same place. Therefore, the gaming experience is better for players living in areas with higher density Pokéstops, than for players who are not.

Some observers have highlighted spatial inequalities concerning Pokéstops' location, and the chances of finding Pokémon. In the U.S., newspaper articles have reported that there are more Pokémon GO locations in majority-white neighborhoods⁸. Even when accounting for population density, positive relationships between the percentage of white people present, and Pokéstop density, has been found. For Chicago and Detroit, Colley *et al.* (2017) found a significant, positive effect between the percentage of the population that is white, i.e. non-Hispanic, and Pokéstop density. In addition, their results reveal Pokéstop density has stronger links to race and ethnicity, than income. Juhász and Hochmair (2017) also analyze the distribution of Pokéstops in the United States. They want to identify which socio-economic variables and land use categories affect the density observed in urban areas of South Florida and Boston. Their results confirm that Black and Hispanic neighborhoods are disadvantaged and that Pokéstops are more frequently observed in commercial, recreational and touristic sites. In France, such a question has not been studied.

2 Data and variables

In this section, we present the data and the different variables retained in order to examine the links between Pokéstop location and neighborhood characteristics in the French context.

2.1 Pokéstop location

In this study, our main explanatory variable concerns Pokéstop location. To begin, we need geographical coordinates of these Pokéstops. Information about this has been collected from a

⁷ This behavior represented a serious concern for authorities who asked for caution on roads: the following messages "Don't Pokémon and drive" was readable on American highways or even "Celui qui conduit, c'est celui qui ne joue pas" (The one who drives is the one who does not play) on French highways.

⁸ Huffaker C. "Why are there fewer Pokemon locations in black neighborhoods?", *Belleville News Democrat*, July 14, 2016; Kooragayala S. and Srini T. "Pokémon GO is changing how cities use public space, but could it be more inclusive?", *Urban institute*, August 5, 2016.

previous game developed by *Niantic* in 2012, called *Ingress*⁹. In this game, players seek "portals" located in particular places, such as monuments, historical buildings, urban frescoes etc., and teams compete with each other to takeover these portals. The exact location of portals is obtained using a plugin called "IITC Ingress Portal CSV Export". It allows players to obtain the name, image, and coordinates of all *Ingress* portals within a viewport.

Because the two games are created by *Niantic*, Pokéstops location is based on the portals location. As already outlined, the location and choice of Portals is not random. It is often guided by the presence of historical monuments, but also by players' suggestions. This point will be discussed in the next subsection because it represents some problems that must be considered in our econometric approach. Finally, the geographical coordinates of these portals are used to determine the exact location of approximately 3,850 Pokéstops in the Loire-Atlantique department. This department is an interesting case study, because it allows us to highlight the respective determinants of Pokéstop locations for places in urban and rural areas. As the number of Pokéstops in a neighborhood may vary according to its area, we calculate the density in every neighborhood. It is defined as the ratio between the number of Pokéstops and the size of the area in square kilometers. It permits us to work with a continuous variable, taking into account some inequalities due to the size of the geographical unit observed.

2.2 Explanatory variables

We use the 2010 population census conducted by the French national institute of statistical and economic information (INSEE) in order to obtain information on the whole department at the neighborhood level. The population census represents an exhaustive dataset providing statistics on households, housing conditions, and residents, as well as a number of their socio-economic characteristics (gender, age, qualification, employment status, occupational status etc.). We retain the Iris code as a proxy of the French neighborhood. The Iris may represent a municipality, or a subdivision of a municipality, if the latter has more than 10,000 inhabitants. It represents homogeneous spatial subdivisions in terms of housing and population with an average of 2,344 inhabitants. We have information on the 514 Iris codes that make up the department. Varied information is extracted from census data.

We retain the variables relating to socio-economic characteristics of inhabitants for two reasons. Firstly, population characteristics are necessary for our analysis because it is likely to be an important determinant of Pokéstop density. We have argued that players could suggest to *Niantic* the location of *Ingress* portals, which then determines Pokéstop location. Indeed, it appears that locations for *Ingress* were mostly crowdsourced (so *Niantic* did not deliberately distribute markers unevenly). For example, in the U.S., the company started with the Historical Marker Database, a crowdsourced collection of historical markers from around the country, that they bought the rights to in 2011¹⁰. When the firm took the database and used the sites in *Ingress*, they also allowed players to submit new locations, so the crowdsourcing process continued

⁹ Huffaker C. "There are fewer Pokemon Go locations in black neighborhoods, but why?", *Miami Herald*, July 29, 2016.

¹⁰ O'Brien M. "'Pokemon Go' players stumble on hidden history", *Associated Press News*, July 25, 2016.

(Submission was shut down by the end of 2015). Secondly, we have to consider the fact that people at the origin of crowdsourcing may have some particular characteristics that must be taken into account in our analysis. Moreover, the large majority of studies assessing disparities in accessibility to recreational amenities look at the relationships with neighborhood socio-economic characteristics. This empirical strategy allows us to reveal and identify socio-spatial segregation in the distribution of these amenities, that could impact an individuals' well-being.

For these two main reasons, we retain a selection of variables providing information about local demography (percentage of young people – *i.e.* individuals between 15 and 24 years old), qualification level (percentage of individuals with at least a second year university level), occupational status (percentage of executives) and employment status (unemployment rate). Some of these are likely to correspond to characteristics of regular players (young or unemployed, for example).

We retain variables to characterize services and activities in the different neighborhoods of the sample. We include the percentage of social housing that is used as a *proxy* for deprived neighborhoods. This housing type are usually moderate rental properties and are largely reserved for moderate and lower-income tenants and families. Subsequently, their concentration in municipalities or neighborhoods may reveal the presence of fragile populations. Retaining this information is another way to establish the links between neighborhood conditions and access to recreational amenities. In addition, we retain the number of shops in each neighborhood. As previously argued, players immediately noticed that the gaming experience was different between urban and rural areas¹¹. Because this distinction between areas might reflect differences in terms of local economic activity, we retain the number of shops to characterize neighborhoods. For this variable, we restrict ourselves to the most common shops: mini-markets, supermarkets, grocers, bakeries, butcher shops, fishmongers, bookstores, clothing stores or shoe shops. This information is extracted from the French "Permanent Facilities Database" for 2015. By checking these different characteristics, we want to verify if Pokéstops are more prevalent in dynamic neighborhoods.

Given location of *Ingress* portals (and thereafter of Pokéstops) is linked to historical markers, statues, museums, old churches, skyscrapers, street art or even city parks¹², we also utilize data from the Regional Directorate for Cultural Affairs (DRAC), of the Pays de la Loire region, in order to obtain the location of monuments in the department of Loire-Atlantique. This dataset provides us with geographical coordinates for various elements of French cultural heritage in 2016. It includes museums, historical monuments, labeled buildings, gardens and sites, protected areas, public libraries, cinemas, music schools and conservatories. From this diverse information, we create a variable that represents the total number of monuments in each neighborhood of our sample. However, we do not make any distinction between the different types of monuments. By doing this, we want to assess the links between the location of these monuments, and the location of Pokéstops.

Finally, one can believe that games' developers have payed attention to the quality of mobile network coverage and then may have deliberately not located Pokéstops where a mobile network

¹¹ Hargarten J. "Why some places in Minnesota are better for playing Pokémon GO than others", *StarTribune*, August 3, 2016; Leloup D. and Grandin J. "Pokémon Go: les multiples facteurs des inégalités géographiques", *Le Monde*, August 8, 2016.

¹² O'Brien M. "Pokémon Go' players stumble on hidden history", *Associated Press News*, July 25, 2016.

is not available or inefficient. In such case, Pokéstop's location should be directly related to a networks coverage. To check this hypothesis, we mobilize data from the Regulatory Authority for Electronic Communications and Posts (ARCEP) providing information on the size of the 4G network deployed on the French territory in 2017. However, results of various econometric models do not reveal any significant relationship between mobile network coverage and Pokéstops' density¹³. This surprising finding might be due to the important number of relay antennas in Loire-Atlantique and then to an efficient network coverage¹⁴. For these reasons, this variable was not retained for our empirical analyzes.

Table 1: Description of the variables

Variable	Definition	Source
Pokéstops density	Ratio between the number of Pokéstops and the area in km ²	Ingress
<i>Socio-economic characteristics</i>		
% of young people	% of the total population between 15 and 24 years old	INSEE, 2010
% of qualified people	% of individuals with at least a second year university level	INSEE, 2010
% of blue-collars	% of blue-collars in the total population	INSEE, 2010
% of inactive	% of inactive in the total population	INSEE, 2010
<i>Neighborhood characteristics</i>		
% of public housing	% of public housing on the housing stock	INSEE, 2010
Number of shops	Total number of mini-markets, supermarkets, grocers, bakeries, butchers, fishmongers, bookstores, clothing stores and shoe shops	PFD 2015
Number of monuments	Total number of museums, historical monuments, labeled buildings, gardens and sites, protected areas, public libraries, cinemas, music schools and conservatories.	DRAC 2016

Notes: All the variables are defined at the neighborhood (IRIS) level.

Table 1 shows an overview of the different variables retained (and their sources) for the analysis. Mobilizing spatial econometric models will allow us to determine the role and influence of each of these variables. On one hand, we will be able to evaluate the importance of socio-spatial segregation in spatial distribution of this amenity, and verify to what extent it is linked to historical markers. On the other hand, our data does not permit us to properly measure the importance of crowdsourcing on Pokéstop locations. Indeed, one can assume players might be motivated by the idea of proposing Pokéstops close to their residential location, in which case correlations should be observed between some socio-economic characteristics, and Pokéstop location. However, population characteristics could also be correlated with the presence of some specific services and activities that could be themselves correlated with a Pokéstop location. Consequently, the objective of this analysis is to only highlight the most relevant factors of accessibility to this recreational amenity.

¹³ Results for these estimations are available from the author upon request.

¹⁴ A survey realized by the ARCEP in 2018 reveals that 98 % of the total population in Loire-Atlantique has access to good mobile coverage.

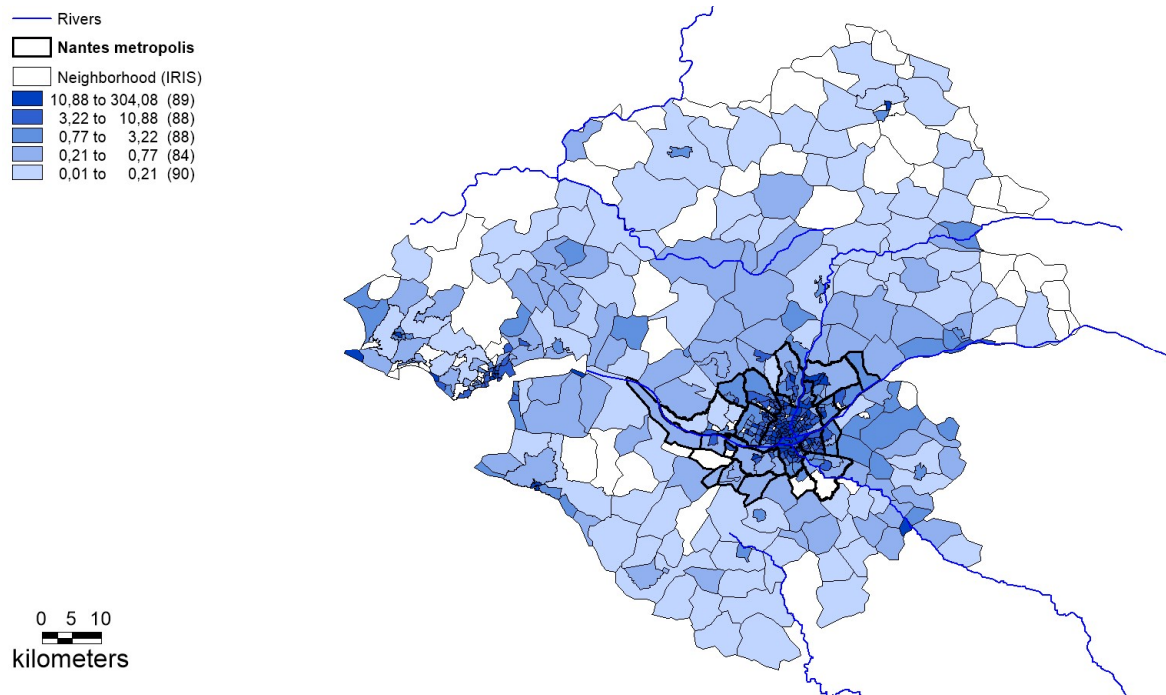
3 Descriptive statistics

3.1 Spatial distribution of Pokéstops in Loire-Atlantique

Pokéstops geography

Nantes is a city in the Loire-Atlantique department in the west of France, located on the banks of the Loire, and at a distance of 50 kilometers from the Atlantic Ocean. Nantes is the sixth largest city in France with approximately 300,000 inhabitants in the central city, and almost 610,000 inhabitants in the greater metropolis. The Nantes metropolis incorporates 24 municipalities in a total area of 523 square kilometers. The Loire-Atlantique department includes 212 municipalities, and represents almost 1,350,000 inhabitants.

Figure1: Pokéstops density in Loire-Atlantique



Source: Ingress Intel Map, 2017.

Figure 1 highlights disparities in terms of Pokéstop density in the Loire-Atlantique and Nantes metropolis. As a starting point, we observe strong disparities between the metropolis and the rest of the department. Most neighborhoods with a density greater than 10 Pokéstops per square kilometer are located in Nantes, or in its vicinity. On the contrary, Pokéstop density appears to be below 1 per square kilometer for a large number of municipalities outside the metropolis (in rural areas). In addition, the density seems to decrease as the distance from the center of Nantes increases; ie the most distant municipalities are least well represented. Some municipalities located on the Atlantic coast are distinguished by a high Pokéstop density. These include famous vacation spots such as Le Croisic and La Baule, and Saint-Nazaire, which is the the second

biggest economic center of the department. In rare cases, some municipalities do not have any Pokéstops, a fact that greatly limits the gaming experience of players. Finally, we cannot exclude the fact that the low density observed in large neighborhoods is justified by the size of the area: a large area tends to decrease the ratio between the number of Pokéstops and the measured surface.

Spatial autocorrelation of Pokéstops density

The previous figure also shows phenomena of spatial autocorrelation in the distribution of Pokéstops on the whole department. More precisely, we observe that neighboring areas often show similar values for our Pokéstops density index. Anselin (2001) defines this spatial effect as the coincidence of value similarity with locational similarity. A positive spatial autocorrelation reveals similar values coincide in neighboring areas, whereas a negative spatial autocorrelation reveals the presence of very different values in neighboring areas.

In order to confirm the presence of dependence between observations, we compute Moran's I autocorrelation coefficient (Moran, 1948, 1950) for Pokéstops density. The Moran's I may be interpreted as the ratio of the covariance between contiguous observations and total variance of the sample (Cliff and Ord, 1973, 1981). It is defined as follows:

$$I_M = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

Where n is the number of observations, x is the value of the variable at neighborhood i and \bar{x} the mean value of the variable in question and w_{ij} is an element of a spatial weight matrix (W) that accounts for the proximity of the neighborhoods i and j ¹⁵.

Table 2: Spatial autocorrelation of Pokéstops density

Spatial weight matrix	Loire-Atlantique		Nantes metropolis	
	Moran's I	P-value	Moran's I	P-value
Contiguity, order 1	0,5971	0,001	0,5162	0,001
Contiguity, order 2	0,4302	0,001	0,3453	0,001
4 nearest neighbors	0,6007	0,001	0,5106	0,001
5 nearest neighbors	0,6657	0,001	0,5781	0,001
6 nearest neighbors	0,6239	0,001	0,536	0,001
7 nearest neighbors	0,6109	0,001	0,5211	0,001
Minimum distance (8/4 km radius)	0,117	0,002	0,0938	0,003
Inverse squared distance (max 8/4 km)	0,117	0,002	0,0938	0,003

Notes: The distance of 8 kilometers is retained for the whole department, while the distance of 4 kilometers is only retained for the Nantes metropolis.

Source: Ingress Intel Map, 2017.

¹⁵ When $I > E[I] = -(n - 1)^{-1}$ (respectively $I < E[I]$), the values taken by our index, for two neighboring areas, are not randomly distributed but close (respectively distant). Then, neighborhoods geographically close are also statistically close (respectively distant) and we conclude that there is a positive spatial autocorrelation (respectively negative). When I is close to $E[I]$, we conclude that there is no link between statistical proximity and geographical proximity of units.

In fact, the computation of Moran's I index is sensitive to the definition of the spatial weights matrix W retained. The following table (Table 2) presents values of this index for different spatial weight matrices for the whole Loire-Atlantique department, and for the Nantes metropolis exclusively. Values for Moran's I demonstrate the existence of a positive and strong spatial autocorrelation. Two neighborhoods geographically close are also two neighborhoods statistically close. Unsurprisingly, we also observe that the dependence between areas generally decreases when the order of contiguity, or the rank for the nearest neighbor, is higher. It confirms that spatial dependency is mainly observed among immediate neighboring areas. This is confirmed for the Loire-Atlantique department as well as the Nantes metropolis.

If a global spatial autocorrelation is highlighted within our sample, it could be interesting to mobilize the local Indicators of Spatial Association (LISA) to identify some local clusters (Anselin, 1995). By using LISA cluster maps, we are able to detect the presence of local spatial clusters influencing the values of the global spatial autocorrelation (detected through Moran's I index)¹⁶. Figures 2a and 2b show local spatial autocorrelation for the whole department, and for the Nantes metropolis. In this case, we retain a spatial weight matrix W based on the k -nearest neighbors (here: $k = 5$). The following section (Section 4) demonstrates the superiority of this definition for our sample over other specifications. Technically, Euclidian distances are calculated between all the centroids of the neighborhoods in our sample. If all distances from each area i to all areas j are ranked, each element w_{ij} of W are defined as follows:

$$w_{ij} = \begin{cases} 1 & \text{if area } j \text{ belongs to the set of area } i\text{'s } k \text{ nearest neighbors} \\ 0 & \text{otherwise} \end{cases}$$

Both figures reveal significant local spatial correlations. In particular, a small cluster of neighboring areas with high levels of Pokéstop density is identified for the whole department, and when we focus on the Nantes metropolis exclusively (areas identified as High-High on the maps). Logically, this corresponds to the central business district of Nantes. Large parts of the department are characterized by low Pokéstop density neighborhoods (areas identified as Low-Low on the maps). This is the case for a high number of municipalities in both the North and South of Nantes. In addition, we observe the same phenomenon for neighborhoods of the metropolitan area that are located far away from the city center, on the borders of the Nantes metropolis. Some islands are highlighted by the spatial clusters map. For example, in the West of the Loire-Atlantique, we can find a neighborhood characterized by high Pokéstop density surrounded by neighborhoods with low densities (areas identified as High-Low on the maps). This corresponds with a very famous tourist attraction known as Le Croisic. Concerning the metropolis itself, the two places with a similar configuration are the city centers of Saint-Herblain and Sautron, two important cities within the territory. Finally, the two figures revealed by the LISA analysis highlight two different geographies between the Nantes metropolis, and the rest of the department. Namely, Nantes has

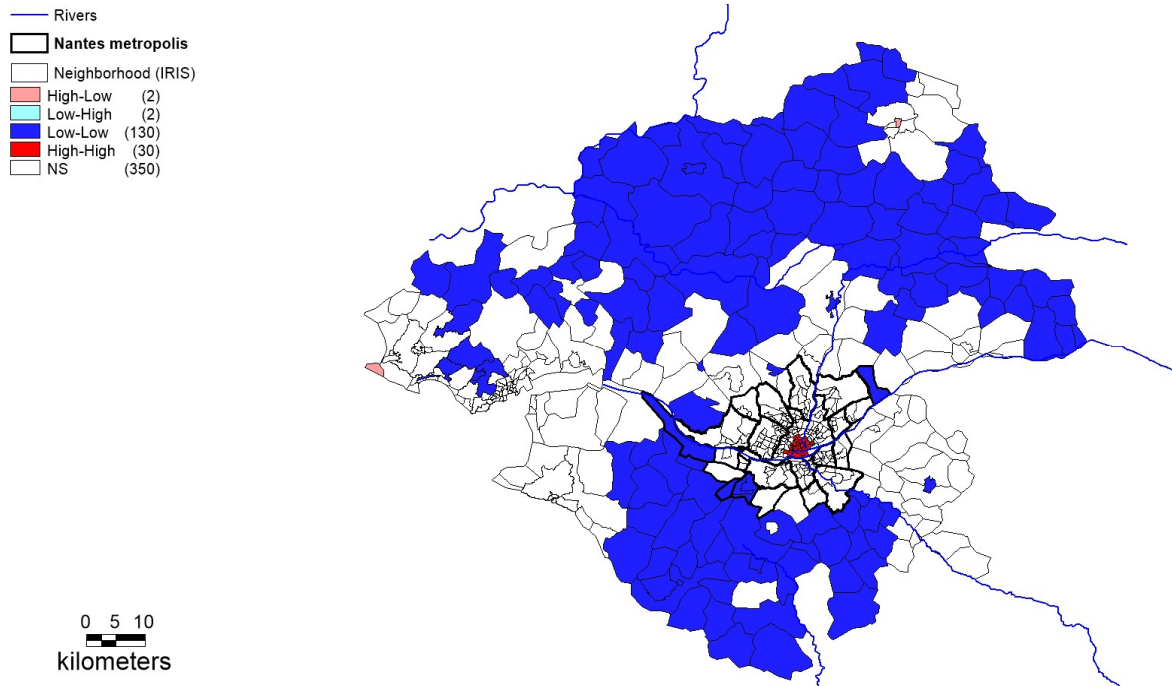
¹⁶ LISA represents a decomposition of the Moran's I index and is obtained as follows:

$$I_i = \frac{\sum_{j=1}^N w_{ij} (x_i - \bar{y})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

high densities of Pokéstops in its center with low densities surrounding it, while other areas of the department only show low densities.

Figure2: LISA clustering maps of Pokéstops density

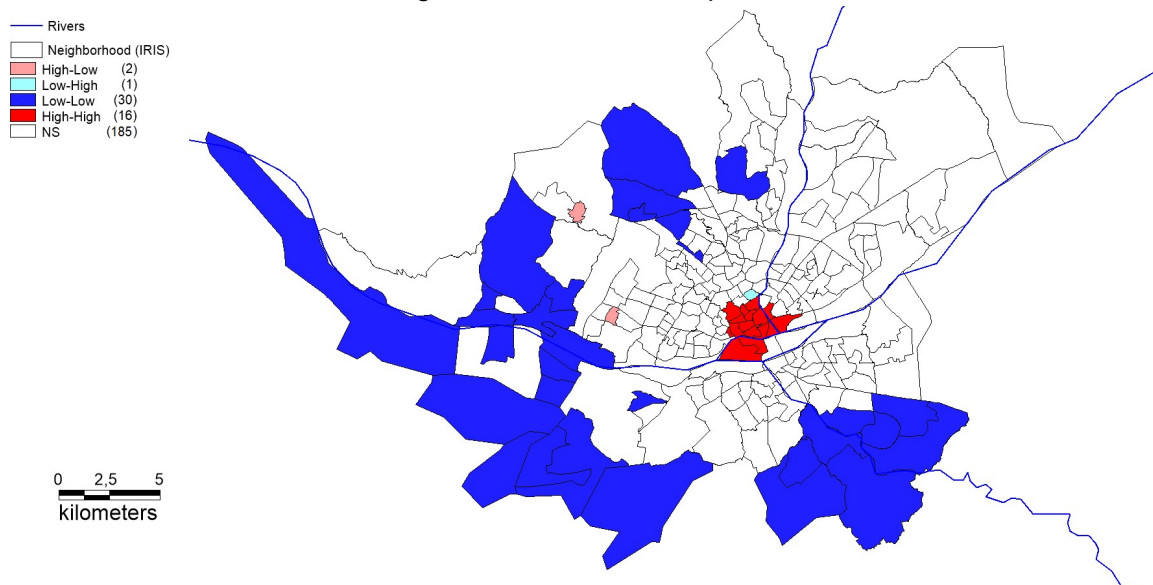
Figure 2.a: Loire-Atlantique department



Notes: LISA statistics are determined with a spatial weight matrix based on the 5 nearest neighbors.

Source: Ingress Intel Map, 2017.

Figure 2.b: Nantes metropolis



Notes: LISA statistics are determined using a spatial weight matrix based on the 5 nearest neighbors.

Source: Ingress Intel Map, 2017.

Previous analysis has shown strong disparities concerning access to this particular recreational amenity. They also shed light on strong and significant global and/or local spatial autocorrelation phenomena. As a result, in the next section, our objective is to provide explanations for this particular geography, by taking into account a likely problem of spatial autocorrelation in the residuals of our econometric model.

3.2 Sample characteristics

Table 3 presents some characteristics of our sample. We make the distinction between the whole department and the Nantes metropolis. Firstly, we observe that the average Pokéstop density is approximately 8 per square kilometer, with the average number of Pokéstops in a neighborhood also close to 8. In both cases, standard errors are very high and reflect an important variability between neighborhoods. Secondly, in accordance with the previous map, we find that inhabitants of Nantes enjoy a large number of Pokéstops: the density per square kilometers is close to 30 in Nantes, while it is below 6 outside the city.

When we look at population characteristics, we find that inhabitants of Nantes are younger, more educated, and of a higher social status than inhabitants located in other parts of the Loire-Atlantique. The percentages of 15-24 year olds, executives, and individuals with a high school diploma (or greater) are respectively 27%, 16.2% and 39.1%, against 19.1%, 11.3% and 26.3% for the whole department. Some of these characteristics could explain why Pokémon GO is more developed in some neighborhoods. Indeed, the Pokéstop location is partly determined by players of the previous game developed by *Niantic*. As a result, location could depend on the presence, or not, of players with particular characteristics in these neighborhoods. We consider the age of inhabitants could be an important determinant, as some reports have highlighted that gamer profiles can be quite young. For example, a TNS Sofres study on French video game consumption practices, released in 2013 for the National Film and Moving Image Centre (CNC), revealed that under-25 year olds represent 38.1% of players. The same study also showed that occupational status might be important: the higher occupational statuses are represented by a high concentration of players (67.5%), which is slightly more than lower occupational statuses (62.4%).

Concerning neighborhood characteristics, we observe some disparities between the metropolis and the other municipalities of our sample. Unsurprisingly, the number of shops (mini-markets, bakeries, butcher shops, bookstores etc.) is more important in the center of the metropolis. The number of monuments does not vary greatly across the whole sample. On average, each neighborhood has a museum, historical monument, labeled building, garden sites, and even a public library. This is not the case for neighborhoods not located in Nantes, but in the other parts of the metropolis. The percentage of social housing is the most important in the Nantes metropolis. It is partly due to some French laws requiring municipalities to have 20% social housing in municipalities with more than 3,500 inhabitants (SRU law). As a result, a large number of municipalities in the department are not concerned by these laws.

Table 3: Sample characteristics

Variables	Loire-Atlantique		Nantes		Rest of metro.		Outside metro.	
	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
<i>Pokéstops</i>								
Pokéstops density	8,23	24,517	29,76	49,037	6,17	9,04	1,78	5,75
Number of Pokéstops	7,80	11,685	13,18	19,514	6,88	7,313	6,38	9,027
<i>Population characteristics (in %)</i>								
15-24 years old	19,1%	0,076	27,0%	0,098	19,8%	0,078	16,0%	0,036
Executives	11,3%	0,072	16,2%	0,073	15,1%	0,071	7,7%	0,049
Unemployment rate	9,9%	0,055	14,3%	0,06	8,8%	0,06	8,9%	0,04
High school diploma and more	26,3%	0,123	39,1%	0,151	29,9%	0,101	20,2%	0,069
<i>Neighborhood characteristics</i>								
Social housing (in %)	10,6%	0,166	20,3%	0,261	11,6%	0,139	6,7%	0,113
Number of shops	4,58	16,876	10,66	32,146	3,58	7,399	2,96	11,2
Number of monuments	1,01	3,661	1,67	6,672	0,24	0,809	1,16	2,922
Population density	2 247,08	3 376,59	6 838,38	4 229,35	2 176,99	2 388,270	690,82	1 505,57
Observations	514		97		137		280	

Notes: “Rest of metropolis” corresponds to the neighborhoods of the whole metropolis without Nantes. “Outside metropolis” corresponds to the remaining part of the department outside the metropolis.

Sources: Population census 2010 (INSEE), DRAC (2016) and Ingress (2017).

4 Estimation strategy

4.1 Empirical model and specification strategy

We follow a conventional approach in the applied spatial econometric literature, and begin our specification with a non-spatial ordinary least squares (OLS) regression model as a benchmark (Elhorst, 2010). Thereafter, we test the possibility of the extension of the baseline model to include spatial interaction effects. Our model is defined as follows:

$$Y_n = \beta_0 + \beta_1 P_n + \beta_2 N_n + \varepsilon_n \quad (1)$$

$$\varepsilon \sim N(0, \sigma^2)$$

where Y_n is the $n \times 1$ vector of Pokéstops densities (with $n = 1, 2, \dots, 514$), P_n is a $n \times k$ matrix of variables relating to population characteristics, N_n is a $n \times k$ matrix of variables relating to neighborhood characteristics defined previously (in Table 1) and ε_n is a $n \times 1$ vector of error terms. β represent vectors of parameters to be estimated. Results for this baseline model are reported in Table 5.

To the extent we work on spatial data, we may have to retain an appropriate spatial econometric model that takes into account a likely dependence between neighboring areas, *i.e.* the problem of spatial autocorrelation (Le Gallo, 2002; LeSage and Pace, 2009). In that case, using the OLS method may conduct less efficient or biased estimators, because hypotheses necessary for its implementation are no longer verified (*i.e.* the assumption of independent error terms). This spatial dependence can take two forms. One form of dependence arises when the outcome of an area is affected by outcomes of neighboring areas. In order to account for this, a spatially lagged dependent variable is included among the explanatory variables. The classical

linear regression model in equation (1) is then adapted to specify a spatial autoregressive model (SAR model). When the spatially lagged variable is ignored in the model specification, while it is present in the data, then the OLS estimators will be biased and not convergent. Another form of dependence arises from a spatially correlated disturbance term among the observations. In that case, the spatial dependence is considered as a statistical nuisance, which can be explained by problems of bad specifications (omitted variables, bad geographical scale etc.). Unless the equation (1) is well specified, the estimated parameters are inefficient and potentially biased. In order to consider this problem, the specification should incorporate a spatial autoregressive process in the error term (SEM). Whatever the model specified, the parameters of the equation are generally estimated using the Maximum Likelihood method.

In practice, some tests permit us to decide between the two forms of spatial dependence: the Lagrange multiplier (LM) tests (lag and error) and their robust versions. These tests are developed for identifying a spatial autocorrelation in the dependent variable or in the error terms (Anselin and Florax, 1995). We present the values of the standard LM tests and their robust versions for various specifications of the spatial weight matrix in Table 4. These tests are conducted on the residuals of the estimated OLS model. They allow acceptance or rejection of the null hypothesis (absence of spatial autocorrelation), but also permit the choice between SAR model and SEM (if the null hypothesis is rejected). Whatever the spatial weight matrix retained, we find that LM tests reject the null hypothesis of absence of spatial autocorrelation in the error term of the model. Then, the tests provide information on which model is well suited to consider this problem. In all specifications, the standard and robust tests on spatial lag dependence are the most significant, except for the last two that are not significant at the 10% level.

Table 4: Results for Lagrange multiplier tests

Spatial weight matrix	LM Tests		Robust LM tests	
	Spatial lag	Spatial error	Spatial lag	Spatial error
1st order contiguity	188,9***	86,48***	103,89***	1,47
2nd order contiguity	233,48***	117,65***	119,15***	3,33*
4 nearest neighbors	183,23***	97,96***	85,69***	0,41
5 nearest neighbors	232,72***	109,16***	123,89***	0,33
6 nearest neighbors	230,33***	93,51***	136,84***	0,02
7 nearest neighbors	268,07***	119,68***	148,87***	0,49
Minimum distance (8 km radius)	3,79*	2,8*	1,12	0,14
Inverse squared distance (max 8 km)	110,05***	61,3***	48,87***	0,12

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

Following the study of LeSage and Pace (2009), Elhorst (2010) argues that some additional tests can be conducted to determine the appropriate specification. Namely, if one of the LM tests (or even both of them) is significant, it is preferable to estimate a spatial Durbin model (SDM), since the SAR model and SEM are particular forms of the latter. Indeed, the SDM simultaneously considers the presence of the spatially lagged dependent variable, and the spatially lagged independent variables. In our case, SDM can be specified as follows:

$$Y_n = \beta_0 + \rho WY_n + \beta_1 P_n + \beta_2 N_n + \theta_1 W P_n + \theta_2 W N_n + \varepsilon_n \quad (2)$$

$$\varepsilon \sim N(0, \sigma^2)$$

where W is a spatial weight matrix constructed to define relationships between areas, ρ is the spatial autoregressive parameter, θ and β represent vectors of parameters to be estimated and ε_n is a $n \times 1$ vector of error terms.

Because the SDM contains the SAR model (more appropriate than the SEM according to LM tests), it is necessary to discern these two specifications. We use the likelihood ratio in order to know which estimation strategy should be adopted. The SAR model corresponds to the specific case where spatial autocorrelation is only due to the influence of areas on the neighbors, which is the case when $\theta = 0$ ¹⁷. The likelihood ratio allows us to test this hypothesis. If the hypothesis holds, the SDM could be simplified to the SAR model. When the SEM is applied with the SDM, the hypothesis to be tested is: $\theta = -\rho\beta$. If this hypothesis holds, the SEM is equivalent to a restricted SDM, and should be preferred¹⁸.

In addition, before running a SDM, we need to make a choice concerning which spatial weight matrix to retain for the analysis. Elhorst (2010) shows that different selection procedures may be used to choose the best specification: procedures based on the Bayesian posterior model probability (LeSage and Pace, 2009), or based on a “goodness-of-fit” criterion such as J -type statistics (Anselin, 1986; Kelejian, 2008), and log-likelihood function values (Stakhovych and Bijmolt, 2009). In this study, we use this approach based on the log-likelihood function value for two main reasons. Firstly, it represents the most widely used criterion, and secondly, it produces robust estimations when spatial autocorrelation is strong (Stakhovych and Bijmolt, 2009).

We then follow a strategy proposed by Chasco *et al.* (2017) where we estimate the SDM using the different spatial weight matrices selected for Lagrange multiplier tests. Results in Table 5 show the spatial weight matrix with the highest log-likelihood function value is the “5 nearest neighbors”. In addition, Table 5 also shows that hypothesis $\theta = 0$ is always rejected. This means that the SAR model is not appropriate, and that SDM should be preferred. Next, we estimate the SDM presented in equation (2) with a “5 nearest neighbors” spatial weight matrix. A basic first-order contiguity matrix could also have been retained, because values for the log likelihood function is very similar, and because Stakhovych and Bijmolt (2009) have demonstrated spatial models which estimate such a matrix often perform better on average than others. Results of our estimated model with this matrix specification are shown in Appendix A. However, LeSage and Pace (2009) have highlighted that, in most situations, the specification of the weight matrix does not substantially change the estimated results of the model.

¹⁷ The SAR model can then be specified as follows: $Y_n = \beta_0 + \rho WY_n + \beta_1 P_n + \beta_2 N_n + \varepsilon_n$.

¹⁸ The SEM could be specified as follows: $Y_n = \beta X + u_n$ with $u_n = \rho W u + \varepsilon_n$, where X is the vector of independent variables. If the matrix is nonsingular, the model can be rewritten as follows: $Y_n = \beta X + (I - \rho W)^{-1} \varepsilon_n$. Finally, after multiplication of the equation terms by $(I - \rho W)$, we obtain a SDM specification: $Y_n = \rho W Y + \beta X - \rho \beta W X + \varepsilon_n$.

Table 5: Spatial weight matrices and model comparisons

Spatial weight matrix	Log likelihood function value	LR Test : $\theta = 0$
1st order contiguity	74,507	51,686***
2nd order contiguity	36,98	19,798**
4 nearest neighbors	73,828	43,31***
5 nearest neighbors	74,549	25,698***
6 nearest neighbors	50,198	39,477***
7 nearest neighbors	47,719	44,214***
Minimum distance (8 km radius)	0,018	16,338**
Inverse squared distance (max 8 km)	51,518	58,315***

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

4.2 Direct, indirect and total effects

In such a model, the effects of the explanatory variables do not coincide with the estimated coefficients, due to the presence of the spatially lagged dependent variable WY . A change of the main explanatory variable in a given area not only affects the explained variable in the area, but also the values of the explained variable in the neighboring area, which may further influence some areas far away. For these reasons, the coefficients of the SDM do not represent the marginal effects, and do not have the same interpretation as in a simple linear model. Including a spatial lag of the dependent variable implies the outcomes are determined simultaneously.

We support the distinction made by LeSage and Pace (2009) between direct, total, and indirect effects. Appendix B explains how these different effects are obtained with SDM. Namely, the *direct effect* gives the average effect of a change in a given explanatory variable occurring in a spatial area on the explained variable of the same area. The *indirect effect* corresponds to the change in an area's explained variable caused by a change in other areas' explanatory variable. Alternatively, a change in an area associated with any given explanatory variable may affect the area itself, but also other areas indirectly. The *total effect* represents the sum of these two effects. It gives the average change in the explained variable of a given area, attributable to simultaneously changing the explanatory variable for all areas (in the area itself but also the others).

Finally, as the presence of a spatially lagged dependent variable in our models induces feedback loops and spillover effects, it is preferable to refer to direct effects, rather than coefficients estimated by the models.

4.3 Spatial heterogeneity and spatial regime models

The LISA analysis revealed a particular pattern and important differences concerning accessibility to this recreational amenity between the Nantes metropolis and the rest of the department. As a result, it could be useful to implement a specific analysis for each area in order to understand what is at work, and what may differ between them. One can assume that determinants of Pokéstop locations may vary according to these areas. We could consider vicinity to monuments, historical places or shops may be more important in neighborhoods located in rural areas, while characteristics of the population could have a greater impact in neighborhoods located in urban areas (as is the case for most recreational and leisure amenities, according to

literature on urban economics and urban planning). Next, we consider whether there is a discrete spatial heterogeneity taking the form of different spatial regimes. This aspect could be considered as grouped heteroscedasticity, and/or a structural instability between the different regimes (Le Gallo, 2004). This heteroscedasticity is shown by the variability of variances in errors terms according to location and could come from omitted variables. We estimate the following model which permits us to incorporate different spatial regimes in equation (2) previously defined and a grouped heteroscedasticity:

$$\begin{bmatrix} Y_{1n} \\ Y_{2n} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \begin{bmatrix} WY_{1n} & 0 \\ 0 & WY_{2n} \end{bmatrix} + \begin{bmatrix} \beta_{11} \\ \beta_{21} \end{bmatrix} \begin{bmatrix} P_{1n} & 0 \\ 0 & P_{2n} \end{bmatrix} + \begin{bmatrix} \beta_{12} \\ \beta_{22} \end{bmatrix} \begin{bmatrix} N_{1n} & 0 \\ 0 & N_{2n} \end{bmatrix} + \begin{bmatrix} \theta_{11} \\ \theta_{21} \end{bmatrix} \begin{bmatrix} WP_{1n} & 0 \\ 0 & WP_{2n} \end{bmatrix} \\ + \begin{bmatrix} \theta_{12} \\ \theta_{22} \end{bmatrix} \begin{bmatrix} WN_{1n} & 0 \\ 0 & WN_{2n} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

where parameters ρ , β and θ and variables Y , P and N are the same as those defined previously. Parameters of the SDM are estimated by the Maximum-Likelihood method. The different regimes are defined as follows: The first regime is represented by all the neighborhoods in the Nantes metropolis, while the second regime is represented by all other neighborhoods of the department. The gathering of neighborhoods in different regimes allows us to test the hypothesis of the consistency of the parameters between the different regimes through statistical tests of structural instability (Chow asymptotic test).

5 Results

5.1 The role of population and neighborhood characteristics on Pokéstop density

We first estimate the baseline non-spatial model corresponding to equation (1) in Table 6. This model permits us to check if the spatial autocorrelation problem is effectively observed in residuals. If this is the case, the estimated parameters through the OLS method could be biased or inefficient. Result for Moran's I in the first column of the table shows that spatial autocorrelation is still present in our data (the value of 10.86 is significant at the 1 % level) and that the appropriate spatial econometrics model should be used. Next, we used the approach proposed by Elhorst (2010) in order to determine acceptable specification. However, this approach is based on the assumption the neighborhood matrix is known, the independent variables are exogenous, and the normality of the model residuals. In our case, the choice of the matrix, *a priori* unknown, has been made through a procedure based on a "goodness-of-fit" criterion. In our opinion, there is no reason to believe that some variables are endogenous. We also check the normality assumption of residuals through a histogram (see Appendix C).

Table 6: Estimation results for access to Pokéstops

Variables	OLS		SAR (direct effects)		SDM (direct effects)	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
Constant	-27,854***	2,166	-12,635***	1,783	-18,487***	2,887
<i>Population characteristics</i>						
% of 15-24 year olds	38,871***	10,772	13,826	9,070	6,951	8,850
% of high school diplomas (or greater)	80,153***	11,861	28,676***	9,798	17,616	11,47
% of executives	-38,069**	19,047	-9,949	15,643	-12,734	16,658
Unemployment rate	39,913**	16,978	12,796	13,706	16,956	13,671
<i>Neighborhood characteristics</i>						
% of social housing	15,581***	5,417	12,709***	4,576	9,187*	4,784
Number of monuments	2,042***	0,214	1,266***	0,180	1,299***	0,178
Number of shops	0,517***	0,045	0,489***	0,038	0,486***	0,035
ρ			0,564***	0,032	0,476***	0,051
<i>Spatial lag of independent variables (WX) (original estimators)</i>						
% of 15-24 year olds					38,125**	17,329
% of high school diplomas (or greater)					27,651	18,15
% of executives					-20,738	26,86
Unemployment rate					-4,302	26,027
% of social housing					-0,927	8,333
Number of monuments					1,221***	0,378
Number of shops					-0,178**	0,069
R ²	0,619					
Moran's I (error)	10,86***					
LM test : spatial lag	232,72***					
LM test : spatial error	109,16***					
RLM test : spatial lag	123,89***					
RLM test : spatial error	0,32					
LR Test : $\theta = 0$ (SAR vs SDM)					25,698***	
Log Likelihood			-2 015,12		-2 002,27	
AIC			4 050,2		4 038,5	
Observations	514		514		514	

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

Sources: Population census 2010 (INSEE), DRAC (2016) and Ingress (2017).

Estimated results highlight that neighborhood socio-economic characteristics appear to determine access to this particular recreational amenity. First of all, we observe positive effects of the percentage of 15-24 year olds, or the percentage of high-qualified individuals, in the neighborhood. A strong and positive effect is also observed for the local unemployment rate. Because some of these characteristics are often associated with individuals playing video games, one can believe the presence of some video game players has a significant impact on Pokéstop locations. It might also mean players of the previous game, *Ingress*, are likely to have an effect on their location, because they could propose portal creation close to their home (for example). In addition, it appears difficult to argue that access to this amenity is only for favored neighborhoods, since we observe a negative correlation with the percentage of executives in the total population of the neighborhood. Various studies explaining disparities in terms of access to amenities find the distribution is not always equitable: the most favored population generally have better access to equipment and facilities (Moore *et al.*, 2008; Wen *et al.*, 2013 etc.). In our case, such

mechanisms could be confirmed with the percentage of high school diplomas (or higher) possessed, but contradicted with the percentage of executives. For these reasons, it is difficult to argue that Pokémon GO increases socio-spatial disparities in French cities.

Concerning neighborhood characteristics, some of them appear to be important determinants. A positive relationship is observed between the percentage of social housing in the neighborhood and Pokéstop density. This effect is partly unexpected because this percentage could be interpreted as a poverty index. Indeed, this housing is generally reserved for the lowest socio-economic categories. Once again, we find that socio-spatial disparities are not necessarily reinforced by the game. The importance of monuments and shops in the neighborhood reveals a weak, but positive, correlation with accessibility to this recreational amenity. It is in accordance with the fact that the Pokéstops locations are initially linked to historical markers in cities. As a result, location choice can be considered exogenous because it seems to be determined by some historical facts, on which the current population has little or no impact. The positive effect of the number of shops shows the most dynamic neighborhoods, in terms of economic activities, benefit from important access to this amenity. This finding suggests some important disadvantages between rural and urban areas, as activities are less developed in the former than in the latter. It also confirms observations made by players and also highlighted in Colley *et al.* (2017) or Juhász and Hochmair (2017) for the U.S.

Results of Lagrange Multiplier tests applied to this model (Table 4), and Likelihood Ratio tests (Table 5), indicate SDM should be used. Estimated coefficients for this specification are shown in column 3 of Table 6. Direct effects are directly reported in order to compare estimated coefficients by OLS. Direct effects for the SAR model are also added to the table in order to check the robustness of our results. With SDM specification, we observe that none of the population characteristics are significant. If we do not take into account spatial dependency for our dependent variable, it appears that we overestimate the importance of the local population's characteristics. Concerning the neighborhood characteristics, we still observe positive and significant effects on the number of shops and the number of monuments on Pokéstop density. A positive effect for the percentage of social housing remains, but the effect is weaker and less significant than before.

Looking at the estimated coefficients for the spatially lagged independent variables could provide information on the influence of neighboring areas' characteristics. However, Elhorst (2014) has shown the estimated indirect effects of the independent explanatory variables should be used, rather than the coefficient estimates of these spatially lagged independent variables, when the objective is to highlight some spillover effects. For this reason, we make the distinction between *direct*, *indirect* and *total effects* for each independent variable, and summarized results are shown Table 7.

Table 7: Direct, indirect and total effects for SDM

Variables	Direct effect	Indirect effect	Total effect
<i>Population characteristics</i>			
% of 15-24 year olds	6,951	70,927***	77,877***
% of high school diplomas (or greater)	17,616	64,356***	81,972***
% of executives	-12,734	-47,476	-58,740
Unemployment rate	16,956	3,326	20,283
<i>Neighborhood characteristics</i>			
% of social housing	9,187*	6,578	15,765
Number of monuments	1,299***	3,189***	4,489***
Number of shops	0,486***	0,093	0,579***

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

Sources: Population census 2010 (INSEE), DRAC (2016) and Ingress (2017).

When we consider indirect effects, we find a positive and significant effect (at 1 % level) for the percentage of young people and percentage of highly qualified individuals. If we assume that most players using mobile phones are young people, this finding could confirm that Pokéstop location (or *Ingress*' portals location) is partly determined by the suggestions of players. Indeed, as previously argued, the crowdsourcing process was important in the development of these games. Surprisingly, while the percentage of nearby monuments impacts positively on Pokéstop density in a given area, no significant impact is observed for the number of shops present. This negative impact is partly unexpected. It means accessibility to this recreational amenity is strong in places where there is an immediate access to other facilities (as in cities' centers), while the effect becomes non-significant when they are located in neighboring areas. However, it is likely that this finding is only valid for neighborhoods in rural areas. In this particular case, distances between neighborhoods are more important, and spatial units, defined as neighbors according to our definition (5 nearest neighbors), can be located far away from each other.

5.2 Determinants of Pokéstop density: urban versus rural areas

Table 8 presents results of the SDM with spatial regimes. When we focus on the Nantes metropolis, we observe that only one population characteristic is still relevant: the percentage of high school diplomas or higher degrees attained by the population in the area. It may show some form of advantage for neighborhoods with a favored population. This result is more in accordance with the existing literature on accessibility to recreational amenities, and determinants of their spatial disparities (Estabrooks *et al.*, 2008; Panter *et al.*, 2008; Billaudeau *et al.*, 2011; Wen *et al.*, 2013 etc.). However, when we look at the neighborhood characteristics, we still find a positive effect on the percentage of social housing. This effect counterbalances the previous one and mitigates the likely disparities between socio-economic groups. The number of monuments and shops still play an important role, as positive relationships are observed with accessibility to Pokéstops.

Results for the rest of the department do not show any significant effect on population characteristics. The most educated neighborhoods are no longer advantaged in terms of access to this amenity. We can consider that socio-spatial disparities are not reinforced by the game in rural areas. The only characteristic that produces an effect on density is the concentration of shops

in the area. It seems a positive correlation exists between the economic activity of a neighborhood, and its Pokéstop density. The level of urbanization is therefore an important factor.

Table 8: Estimation results for urban and rural areas

Variables	SDM (direct effects) – Spatial regimes model			
	Nantes metropolis		Rest of dept.	
	Coeffs.	Std. Error	Coeff.	Std. Error
Constant	-20,886**	9,896	-20,429**	8,686
<i>Population characteristics</i>				
% of 15-24 year olds	-6,759	8,55	1,210	18,948
% of high school diplomas (or greater)	25,456**	12,454	10,793	22,179
% of executives	-25,865	20,106	-7,323	26,866
Unemployment rate	14,743	15,059	34,225	26,185
<i>Neighborhood characteristics</i>				
% of social housing	8,322*	4,988	-2,118	9,255
Number of monuments	1,572***	0,244	0,171	0,247
Number of shops	0,620***	0,045	0,108*	0,056
ρ			0,320***	
<i>Spatial lag of independent variables (WX)</i> <i>(original estimators)</i>				
% of 15-24 year olds	16,024	18,740	53,920	48,338
% of high school diplomas (or greater)	36,521*	20,064	28,096	32,979
% of executives	-28,109	32,674	-25,127	46,430
Unemployment rate	13,071	30,487	16,332	43,641
% of social housing	-0,562	10,582	-18,141	17,015
Number of monuments	1,322**	0,642	-0,193	0,509
Number of shops	0,036	0,125	-0,042	0,094
Log Likelihood			-1 953,75	
AIC			3 973,5	
Observations		234		280

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

Sources: Population census 2010 (INSEE), DRAC (2016) and Ingress (2017).

When we compare indirect effects on the Nantes metropolis and the rest of the department, we do not find any characteristics influencing accessibility to this recreational amenity, and on the whole our model fails to identify determinants of Pokéstop density in rural areas. Concerning the Nantes metropolis, we again find that the percentage of individuals with a high school diploma or greater, and the number of monuments in the neighboring areas, have a positive and significant impact on accessibility to this recreational amenity. Estimated impacts are even greater than those estimated for direct effects. Moreover, the focus on the metropolis reveals negative indirect effects for the variable “percentage of executives” (at 5 % level). Once again, this finding is quite surprising when we consider the literature on urban economics or urban planning, which study characteristics of neighborhoods with poor accessibility to sport facilities, green parks or recreational amenities. It confirms the fact that, in the French case, the game does not increase socio-spatial disparities. We also observe a positive effect on the proximity to shops, which means amenities are often concentrated in urban areas.

Table 9: Direct, indirect and total effects for spatial regime model

Variables	Nantes metropolis			Rest of dept.		
	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
<i>Population characteristics</i>						
% of 15-24 year olds	-6,759	-21,541	28,300	1,210	-47,331	-46,121
% of high school diplomas (or greater)	25,456**	66,753***	92,209***	10,793	22,724	33,517
% of executives	-25,865	-112,099**	-137,964***	-2,481	-26,007	-28,489
Unemployment rate	14,743	-1,169	13,574	34,225	0,497	34,721
<i>Neighborhood characteristics</i>						
% of social housing	8,322*	-9,786	-1,465	-2,118	9,207	7,089
Number of monuments	1,572***	3,791***	5,364***	0,171	-0,255	-0,084
Number of shops	0,620***	0,226	0,846***	0,108*	0,019	0,127

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

Sources: Population census 2010 (INSEE), DRAC (2016) and Ingress (2017).

Finally, various results have shown that neighborhood characteristics are the main determinants of accessibility to Pokéstops. In particular, we found positive effects for the number of shops, and the number of historical markers. Among the population characteristics, only the percentage of highly qualified people seems to produce a significant and positive effect. It is important to note that it is also the stronger effect observed. In addition, most of these findings are observed for the Nantes metropolis. When we restrict our analysis to the rest of the Loire-Atlantique department, we only observe a role for the number of shops in the neighborhood and to a lesser degree a role of the urbanization level.

Conclusion

In this study, we investigated the factors and variables that explain accessibility to particular recreational amenities (Pokéstops) in a French context. Our model examined the Nantes metropolis and the Loire-Atlantique department in order to compare situations between urban and rural areas. Firstly, we studied the spatial distribution of Pokéstops in the whole region. Secondly, we assessed whether correlations exist between their location and a neighborhoods' characteristics. It is important to determine if relationships exist with socio-economic characteristics, because Pokémon GO has been "accused" of increasing socio-spatial disparities.

To answer these questions, we utilized Spatial Durbin and spatial regime models that also permit us to study the differences between urban and rural areas. Our results do not show obvious relationships between characteristics of inhabitants and Pokéstop density. We do not observe that some groups are particularly advantaged or disadvantaged. For example, we found simultaneous positive relationships between the percentage of high school diplomas attained and Pokéstop density, and between the percentage of social housing and this density. This is quite surprising because these characteristics may socially reflect an advantaged population on one hand, and a disadvantaged population on the other. In contrast, we found that neighborhoods with a high number of amenities (such as shops, historical markers etc.) also have a high Pokéstop density. Pokéstop density in neighborhoods and municipalities outside the metropolis are dependent on the number of shops. This means that rural areas with low economic activity are somewhat disadvantaged.

Even though Pokémon GO is only a game for smart phones and Pokéstops are a very specific recreational amenity, questions raised by this study are not without interest. Indeed, spatial segregation is an important phenomenon in cities, which may have consequences on criminality, education, the labor market, people's well-being etc. It would be harmful if it had consequences on some optional activities as mobile games. It turns out that this is precisely the case in the United States, where non-white neighborhoods cannot take full advantage of the game. In the French case, disparities are not linked to the socio-economic characteristics of inhabitants. However, the game strengthens inequalities between inhabitants of urban and rural areas. Inhabitants of large cities already enjoy good access to recreational activities and leisure amenities, and such a game is likely to increase their satisfaction and well-being. In addition, as this game can have a positive impact on local businesses, rural areas might also be penalized in this regard.

As a result, in order to not widen the gap between rural and urban areas, developers of this game, and the next of its kind, must pay attention to the gaming experience of inhabitants in every territory. Since the end of 2018, *Niantic* has rolled out a program where players can submit potential Pokéstops in their area for review. The firm has listed the kinds of places that are likely to be approved as Pokéstops: A location with historical or educational value, an interesting piece of art or unique architecture (Statues, paintings etc.), public parks, public libraries, major transit stations hubs etc¹⁹. The implementation of this program could therefore help to increase the ability for rural players to play effectively. On the whole, all questions and consequences that have arisen from this game may re-emerge because *Niantic* is already working on their next game which could surpass *Pokemon Go* in popularity, given it is based on another well-known franchise: *Harry Potter*. The launch is scheduled for 2019.

References

- Althoff T, White R, Horvitz E (2016) Influence of Pokémon GO on physical activity: Study and implications. *Journal of Medical Internet Research* 18(12).
- Ambrey C, Fleming C (2014) Public Greenspace and Life Satisfaction in Urban Australia, *Urban Studies* 51(6): 1290-1321.
- Anselin L (1986) Some further notes on spatial models and regional science. *Journal of Regional Science* 26(4): 799-802.
- Anselin L (1995) Local indicators of spatial association – LISA. *Geographical Analysis*, 27(2): 93-115.
- Anselin L (2001) Spatial econometrics, in: Baltagi B (ed.), *Companion to theoretical econometrics*. Oxford: Blackwell Scientific Publications. Pp: 310–330.
- Anselin L, Florax R (1995) Small sample properties of test for spatial dependence in regression models: Some further results. In: Anselin L, Florax R (eds.), *New direction in spatial econometrics*. Berlin: New York.
- Billaudeau N, Oppert J-M, Simon C, Charreire H, Casey R, Salze P, Badariotti D, Banos A, Weber C, Chaix B (2011) Investigating disparities in spatial accessibility to and characteristics of sport

¹⁹ Tassi P. "Rural Players Rejoice, 'Pokémon GO' Will Soon Let You Nominate Your Own PokéStops", *Forbes*, September 12, 2018.

facilities: Direction, strength, and spatial scale of associations with area income. *Health & Place*, 17(1): 114-121.

Carlino G, Saiz A (2008) *City Beautiful*. IZA Discussion Paper n°3778.

Carrus G, Scopelliti M, Laforteza R, Colangelo G, Ferrini F, Salbitano F, Agrimi M, Portoghesi L, Semenzato P, Giovanni S (2015) Go greener, feel better? The positive effects of biodiversity on the well-being of individuals visiting urban and peri-urban green areas. *Landscape and Urban Planning* 134: 221–228

Chasco C, Lacalle-Calderón M, Alfonso-Gil J (2017) Spatial contagion of civil liberty: some evidence from a spatial econometrics analysis. *Paradigma Económico* 9(2): 5-32.

Cliff A, Ord J (1981) *Spatial processes: models and applications*. Taylor & Francis.

Colley A, Thebault-Spieker J, Lin A, Degraen D, Fischman B, Häkkinen J (2017) The geography of Pokémon go: Beneficial and problematic effects on places and movement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2017*.

Dahman N, Wolch J, Joassart-Marcelli p, Reynolds K, Jerrett M (2010) The active city? Disparities in provision of urban public recreation resources. *Health & Place* 16: 431-445.

Duncan D, Kawachi I, White K, Williams D (2013) The geography of recreational open space: influence of neighborhood racial composition and neighborhood poverty. *Journal of Urban Health* 90(4): 618-631.

Elhorst J (2010) Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis* 5(1):9–28.

Elhorst J (2014). *Spatial Econometrics: From Cross-sectional Data to Spatial Panels*. Berlin: Springer.

Estabrooks P, Lee R, Gyurcsik N (2003). Resources for physical activity participation: Does availability and accessibility differ by neighborhood socioeconomic status? *Annals of Behavioral Medicine*, 25(2):100–104.

Frenkel A, Bendit E, Kaplan S (2013) Residential location choice of knowledge-workers: The role of amenities, workplace and lifestyle. *Cities* 35: 33-41.

Gidlöf-Gunnarsson A, Öhrström E (2007) Noise and well-being in urban residential environments: The potential role of perceived availability to nearby green areas. *Landscape and Urban Planning* 83: 115–126.

Graelles-Garrido, Ferrer L, Bravo L (2017) The effect of Pokémon Go on the pulse of the city: a natural experiment. *EPJ Data Science* 6.

Howe K, Suharlim C, Ueda P, Howe D, Kawachi I, Rimm E. Gotta catch'em all! pokémon go and physical activity among young adults: difference in differences study. *BMJ* 355.

Hunter L, Boardman J, Saint Onge J (2005) The Association Between Natural Amenities, Rural Population Growth, and Long-Term Residents' Economic Well-Being. *Rural Sociology* 70(4): 452-469.

Juhász L, Hochmair H (2017) Where to catch 'em all? A geographic analysis of Pokémon Go locations. *Geo-Spatial Information Science* 20(3): 241-251.

Kelejian H (2008) A spatial J-test for model specification against a single or a set of non-nested alternatives. *Letters in Spatial and Resource Sciences* 1 (1): 3-11.

Le Gallo J (2002) *Econométrie spatiale: l'autocorrélation spatiale dans les modèles de régression linéaire*. *Economie et Prévision* 155(4):139–158.

Le Gallo J (2004) Hétérogénéité spatiale, principes et méthodes. *Economie et Prévision* 162(1):151–172.

LeSage J, Pace K (2009) *Introduction to Spatial Econometrics*. Boca Raton: Taylor and Francis.

- Moore L, Diez Roux A, Evenson K, McGinn A, Brines S (2008) Availability of recreational resources in minority and low socioeconomic status areas. *American Journal of Preventive Medicine* 34(1): 16-22.
- Moran P (1948) The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society* 10: 243-251.
- Moran P (1950) A Test for the Serial Independence of Residuals. *Biometrika* 37: 178-181
- Neckerman K, Lovasi G, Purciel-Hill M, Rundle A (2009) Disparities in Urban Neighborhood Conditions: Evidence from GIS Measures and Field Observation in New York City. *Journal of Public Health Policy* 30: 264-285;
- Nicholls S (2001) Measuring the accessibility and equity of public parks: a case study using GIS. *Managing Leisure* 6: 201–2019.
- Ono S, Ono Y, Michihata N, Sasabuchi Y, Yasunaga H (2017) Effect of Pokémon GO on incidence of fatal traffic injuries: a population-based quasi-experimental study using the national traffic collisions database in Japan. *Injury Prevention* 24(6).
- Pamuru V, Khern-am-nuai W, Kannan K (2018) The Impact of an Augmented Reality Game on Local Businesses: A Study of Pokemon Go on Restaurants. Working paper.
- Panter J, Jones A, Hillsdon M (2008) Equity of access to physical activity facilities in an English city. *Preventive medicine* 46: 303–307.
- Sen A (1985) *Commodities and Capabilities*. Amsterdam: North-Holland.
- Sen A (1999) *Development as Freedom*. Oxford: Oxford University Press.
- Shields M, Weathley Price S, Wooden M (2009) Life Satisfaction and the Economic and Social Characteristics of Neighbourhoods. *Journal of Population Economics* 22(2): 421-443.
- Smoyer-Tomic K, Hewko J, Hodgson J (2004) Spatial accessibility and equity of playgrounds in edmonton, canada. *The Canadian Geographer* 3: 287–302.
- Stakhovych S, Bijmolt, T (2009) Specification of spatial models: a simulation study on weights matrices. *Papers in Regional Science* 88: 389-408.
- Talen E, Anselin L (1998) Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds. *Environment and Planning A* 30: 595–613.
- Tovar E, Bourdeau-Lepage L (2013) Well-being Disparities within the Paris Region: A Capabilist Spatialised Outlook. *Urban Studies* 50(8): 1575-1591.
- Wen M, Zhang X, Harris C, Holt J, Croft J (2013) Spatial disparities in the distribution of parks and green spaces in the USA. *Annals of Behavioral Medicine*, 45(1): 18–27.
- Witten K, Exeter D, Field A (2003) The Quality of Urban Environments: Mapping Variation in Access to Community Resources. *Urban Studies* 40(1): 161-177.
- Wong F (2017) Influence of Pokémon GO on physical activity levels of university players: a cross sectional study. *International Journal of Health Geographics* 16(8).

Appendices

Appendix A – Estimated results with a first-order contiguity matrix

Variables	SAR <i>(direct effects)</i>		SDM <i>(direct effects)</i>	
	Coeff.	Std. Error	Coeff.	Std. Error
Constant	-13,149***	1,858	-19,549***	2,855
<i>Population characteristics</i>				
% of 15-24 year olds	14,962	9,230	2,484	8,760
% of high school diplomas (or greater)	34,038***	10,535	21,480***	11,169
% of executives	-16,197	16,241	-17,249	17,093
Unemployment rate	15,453	14,353	13,027	15,046
<i>Neighborhood characteristics</i>				
% of social housing	14,227***	4,545	10,439***	4,826
Number of monuments	1,315***	0,189	1,316***	0,177
Number of shops	0,501***	0,04	0,503***	0,037
ρ	0,574***	0,036	0,49**	0,05
<i>Spatial lag of independent variables (WX)</i> <i>(original estimators)</i>				
% of 15-24 year olds			41,813***	16,282
% of high diploma or more			21,848	18,744
% of executives			-5,048	28,207
Unemployment rate			11,741	29,012
% of social housing			2,875	9,478
Number of monuments			2,064***	0,404
Number of shops			-0,333***	0,075
Log Likelihood	-2 033,3		-2 007,456	
AIC	4 086,6		4 048,9	
Observations	514		514	

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

Sources: Population census 2010 (INSEE), DRAC (2016) and Ingress (2017).

Appendix B – Total, direct and indirect effects in SDM

The Spatial Durbin model (SDM) is based on spatial lagged dependent variable WY and spatial lagged independent variables WX . Then, it is specified as follows:

$$Y = \rho WY + \beta X + \vartheta WX + \varepsilon$$

In fact, if the model includes WY , parameters estimated for the independent variables cannot be interpreted as is the case in OLS models or SEM, for example. Because there are spatial interactions, a change in variable X for a given area directly impacts its outcome Y , and indirectly, outcomes of all other areas. In order to show how *direct* and *indirect effects* are obtained, we rewrite SDM as follows:

$$Y = (I - \rho W)^{-1} \beta X + (I - \rho W)^{-1} W \theta X + (I - \rho W)^{-1} \varepsilon$$

$$Y = \sum_{r=1}^k (I - \rho W)^{-1} \beta_r X_r + (I - \rho W)^{-1} W \theta_r X_r + (I - \rho W)^{-1} \varepsilon$$

$$Y = S_r(W) X_r + (I - \rho W)^{-1} \varepsilon$$

where $S_r(W) = (I - \rho W)^{-1} (I_n \beta_r + W \theta_r)$ and S_r are square matrices whose size is equal to the number of observations.

We observe that the effect of a change in variable X_r for an area i is not β_r anymore but $S_r(W)_{ii}$, that is to say the diagonal value of rank i of the matrix. The diagonal terms of the matrix S_r are then the *direct effects* for each area of a change in variable X_r in the same area. Finally, the average direct effect corresponds to the diagonal terms' average of matrix S_r (*average direct effect* = $\frac{1}{n} \text{trace}(S_r)$). Its interpretation is close to that of the coefficients β of a model estimated by OLS. The other terms of the matrix (the off-diagonal elements) are *indirect effects* and represent impacts of change in variable X_r for an area on another area. It is obtained by making the difference between *average total effect* and *average direct effect*, where *average direct effect* is the average for all the terms of matrix S_r (*average total effect* = $\frac{1}{n} \sum_i [\sum_k S_r(W)_{ik}]$).

It is important to note that *average direct effect* can be interpreted in two ways. Firstly, it can be considered as average of n effects of a one-unit change in variable X_r in an area i on all the areas ($\sum_k S_r(W)_{ki}$). Secondly, it can also be interpreted as the average of the n effects on an area i of a modification of a one-unit change in variable X_r in all areas ($\sum_k S_r(W)_{ik}$).

Appendix C – Model residuals from equation (1)

