ESTUDIOS BCRA

Working Paper 2015 | 66

Spatial analysis of the financial system in urban areas: The case of Buenos Aires city

Andrés Denes / Gastón Repetto
Central Bank of Argentina

October, 2015

ie | BCRA
INVESTIGACIONES ECONÓMICAS

Banco Central de la República Argentina
ie | Investigaciones Económicas

October, 2015
ISSN 1850-3977
Electronic Edition

Reconquista 266, C1003ABF
C.A. de Buenos Aires, Argentina
Phone: (5411) 4348-3582
Fax: (5411) 4348-3794
Email: investig@bcra.gov.ar
Web Page: www.bcra.gov.ar

Spatial analysis of the financial system in urban areas: The case of Buenos Aires city*

Denes, Andrés Federico
BCRA

Repetto, Gastón Luis
BCRA

May 2015

Abstract

Broadening the geographical coverage of the financial system is one of the concerns of the Central Bank of Argentina. In this regard, efforts have been made to achieve a diagnosis of this issue at a national level. These led to concrete policy measures. Nevertheless, the spatial distribution of financial systems in large urban areas has been less explored. This paper aims at providing a comprehensive study of the geographical distribution of the supply and demand for financial services in the city of Buenos Aires. Our results are useful as drivers for economic and financial policy making at a urban level.

JEL Classification: C31,G21, G28, R12
Keywords: Spatial Analysis, Financial System, Buenos Aires city

*The opinions expressed in this paper are those of the authors and do not necessarily reflect the position of the Central Bank of Argentina (BCRA). We want to thank all who helped in the working process of this paper. First of all, we are grateful to all the organizations and areas that provided us with the raw material for this study, namely: the Federal Administration of Public Revenues (AFIP), the Department of Monetary Statistics –especially Ricardo Martinez- of the BCRA, the General Department of Statistics and Censuses of the city of Buenos Aires (BA), the Department of National Social Statistics and Census –primarily to Rubén Nigita and Guillermo Krieger- of the National Institute of Statistics and Census (INDEC), and the Authorization Department – in particular Marcelo Retorta, Viviana Novales and Romina Platania- and the Department of Information Management –mainly Hernán Rodriguez- of the Superintendency of Financial Institutions (SEFyC). Processing million of registers would not have been possible without the help of the staff of the Database Department –specially Carlos Bodini and Natalia Berrueta– of the BCRA. Finally, we thank our co-workers in the Economic Research Department of the BCRA for their continuous support and valuable contributions.
Emails: andres.denes@bcra.gob.ar and grepetto@bcra.gob.ar

1 Introduction

Argentina's Central Bank Chart was reformed in 2012 by the law 26.739. This reform explicitly reintroduces, among other issues, the goal for an inclusive and equitative financial system. Therefore, two articles can be understood as baselines of this goal. Article 3 extends the mandate of the Central Bank. Not only is monetary stability an objective, but also financial stability and employment and economic development under social equity. On the other hand, Article 14 reinforces the need to promote universal access to financial services by expanding the geographical coverage of the banking system, reaching areas with lower economic potential and lower population density.

In this regard, the Central Bank has not only made efforts to carry out an analysis of the geographical distribution of the financial system and economic activity in Argentina's territory[1], but has also implemented specific measures aimed at promoting greater geographic reach of the services provided by the financial institutions (FI) under its supervision. In particular, regulation Com. A 5355 refers to the opening of branches. This regulation mandates a redefinition of Argentine territory zoning scheme at a city level[2]. Built on this regulation, a renewed set of criteria was established to promote branches and ATMs in cities with less developed financial infrastructure[3].

The new zoning scheme and the regulations based on it are important steps towards horizontal equity (equal treatment for equal status). However, there are certain aspects that are subject to improvement. For example, not all cities have similar characteristics all through their territory. Using the whole city as the unit of analysis may not be the most appropriate way to proceed. Large cities, because of their size and heterogeneity, may require a special treatment.

In this sense, this paper aims to conduct a comprehensive study of the geographical distribution of the supply and demand for financial services in the city of Buenos Aires (BA), in order to draw conclusions that might be used for economic and financial policy making at an intra-city level.

The paper is structured as follows: in the next section we describe the data and explain the process of transformation and standardization of it. In the third section, we analyze how some relevant variables behave through out the territory of BA using geographic pattern analysis and mapping techniques. In the fourth chapter, we conduct an econometric analysis to explore the determinants of the localization strategy of FI's. In the fifth section we try to find optimal locations for opening branches considering different criteria by performing a suitability analysis. Finally we conclude and draw some challenges for the future.

[1] See "Mapa económico y financiero de Argentina: Un sistema geo-referenciado de indicadores de demanda, oferta y mercado de servicios financieros a nivel de localidad", Blanco, Denes, Repetto, Documento de trabajo N° 2012/59 BCRA.

[2] Cities of the 2001 National Census

[3] See "Políticas de Estímulo a la Bancarización en base al Nuevo Esquema de Zonificación del Sistema Financiero", Apartado 3 del Boletín de Estabilidad Financiera (Segundo semestre 2012).

2 Data and unit of analysis

Putting together a complete set of consistent and homogeneous financial and socio-economic variables for all cities in Argentina is a complex task. The main difficulty lies in finding variables at a national level that can be imputed to each of the Argentine cities. The imputation process is tough since not all national agencies use the same lists of cities and within the databases the quality of the field that assigns the name of the city is usually low (name of the city written in different ways or even different names referring to a single city). Finally, to perform spatial analysis techniques (simple map reports, spatial autocorrelation analysis or complex spatial econometric analysis) we have to associate each city to the pair of geographic coordinates of its geographic center.

The work described above is the basis for defining a new zoning scheme of Argentina and for the public policies that were derived from it. However, if we want to go further and analyze what happens within each city (and we will try to do so for BA), processing data becomes even more complex. The simple association of the number of firms, individuals, banks, ATMs, etc. to a pair of geographic coordinates (ie the center of a city) is not enough: each record must be associated with the geographical coordinates of its location.

To do this we need a software capable of performing geocoding (that is assigning geographic coordinates to addresses) with two key inputs: 1) a vectorial street map of the city and 2) the standardized list of addresses to which we ant to assign their geographic coordinates.

Several of the so called Geographic Information System (GIS) software have geocoding functionality and we already have a vectorial street map of BA. Once again, the biggest difficulty lies in the standardization of addresses of the databases of the different public agencies such us the BCRA, Superintendency of Financial Institutions (SEFyC) and the Federal Administration of Public Revenues (AFIP). If the standardization process of the city variable is complex, the task of standardizing the address is even harder. Besides having to standardize two attributes (street name and street number) instead of one (name of the city), the quality of the field that assigns the address is usually worse than the field that assigns the city (more errors in the street name and more names for the same street). As in the case of city standarization, address standardization is crucial for efficient geocoding databases and the subsequent spatial analysis. Using geocoding functionality we managed to locate on BA territory all branches, ATMs and other banking agencies, all firms and all individuals who's salary is paid through a bank account or has a loan from a FI (from now on we will call them simply individuals).

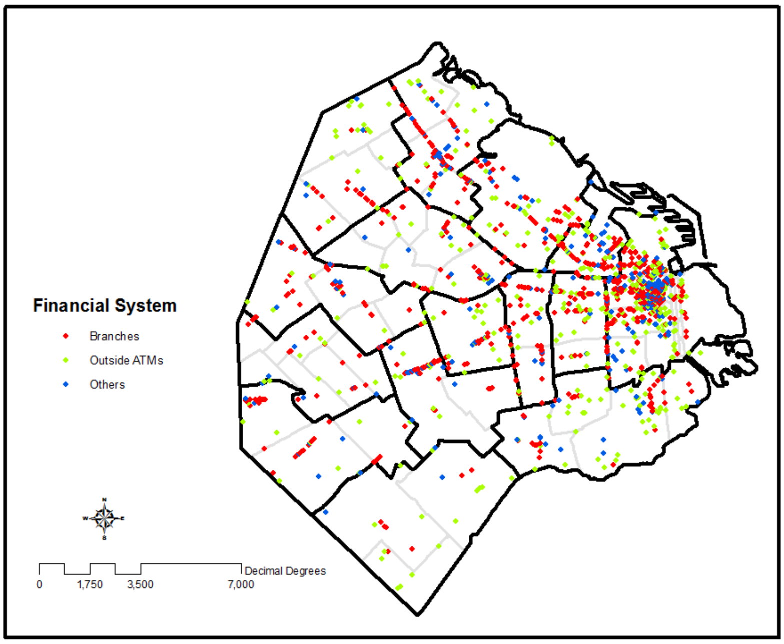
Regarding the socio-economic information of population and households, we have the data of the National Census of Population, Households and Housing (CNPHV) done by the National Institute of Statistics and Census (INDEC) in 2010. Because of the Statistical Secrecy Act (Law 17,622) the information of the CNPHV can not be linked to a specific address. Instead it is presented in different units of analysis: census radius, census fraction, department and province. The smallest unit of analysis is the radius followed by the fraction and finally the department (which coincides with what is known as communes in BA). For these units of analysis we have all the information provided by the CNPHV in addition to its surface and perimeter.

We also have useful geo-referenced information for BA: political division (communes and districts) and basic infrastructure (hospitals and their areas of influence, police stations and sections, school districts, green spaces, subway, premetro and rail stations and the street map).

So, we have five georeferenced datasets to perform our spatial analysis in the following sections:

1. Financial Infrastructure (Branches, agencies and ATM's, deposits and credits and number of accounts and loans). Point Shape file[4] (Map 1)
2. Firms and their main characteristics (sector, legal form, date of startup, loans). Point Shape file (Map 2)
3. Individuals. Point Shape File (Map 3)
4. Radius and fractions with their main characteristics (socio-economic data from CNPHV, area and perimeter). Polygon Shape file (Map 4)
5. Political division and infrastructure of BA. Point, Polygon and Polylines Shape files (Map 5)

Map 1. Financial System Infrastructure

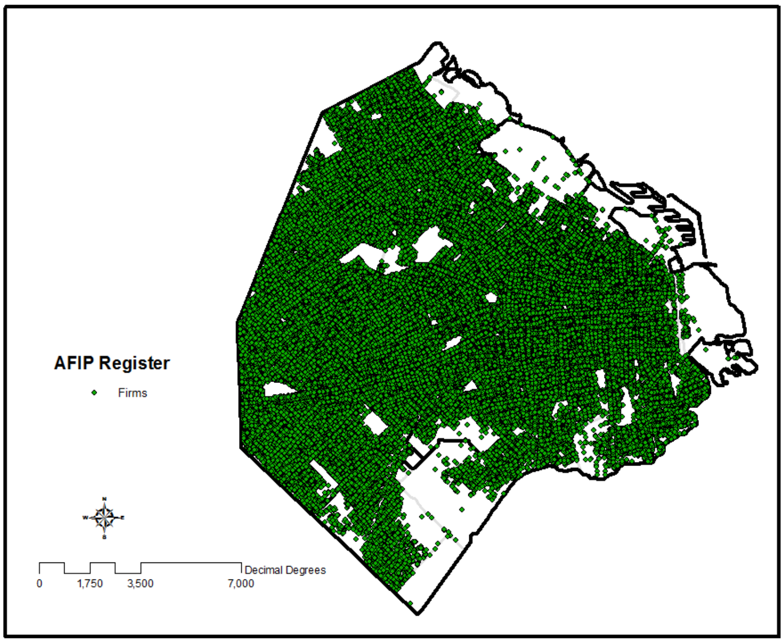


Map 1 shows all branches (including head offices), ATMs and other agencies (branches within companies, promotion points, etc.) within BA territory. We can see a great concentration of points in the area known as Downtown and also all along main avenues.

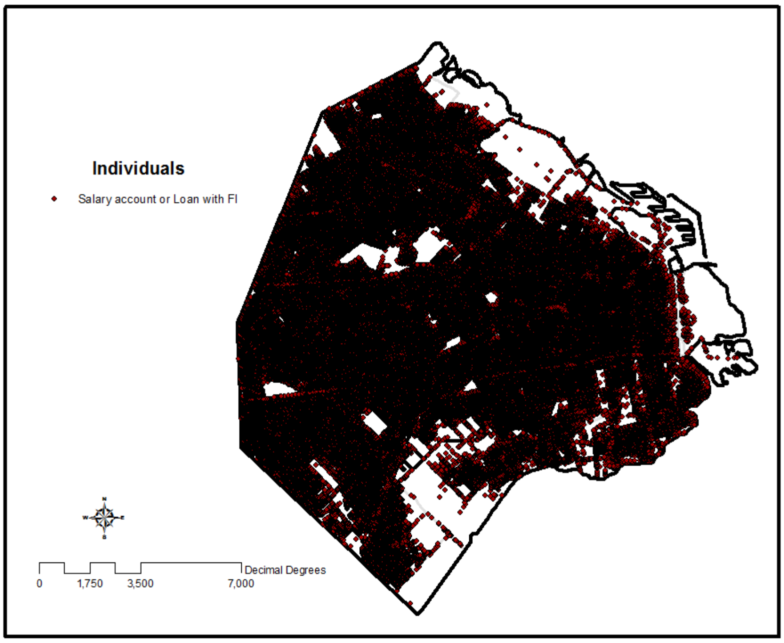
Map 2 and Map 3 show the spatial distribution of firms and individuals in BA. The points are located all across BA except for some white zones that are mostly green areas. Due to the scale of the map it is not possible to accurately observe if there are different densities of points in different parts of BA. At the same time if several firms or individuals are located in the same building we only observed one point (since the points representing each company or individual are superimposed one above the other).

[4]Shape files (.shp extension) are the basic inputs of GIS. These files may represent points (addresses, points of interest, etc.), polygons (states, districts, census fractions, etc.) or polylines. (streets, rivers, etc.).

Map 2. Firms

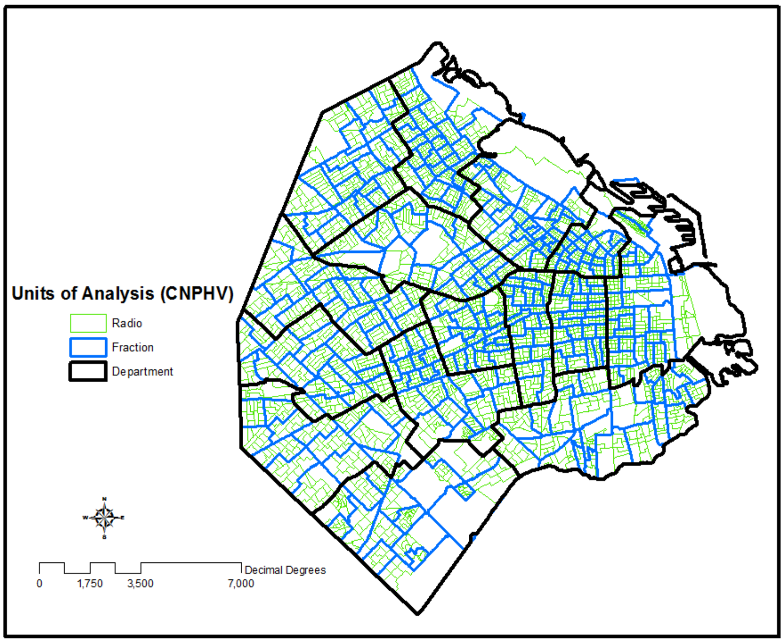


Map 3. Individuals



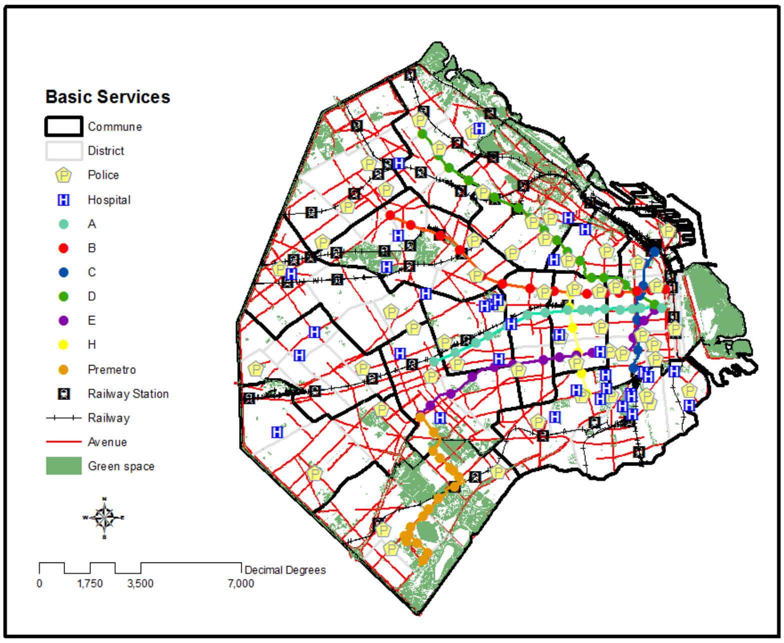
Map 4 polygons are the units of analysis for which we have information from the CNPHV. Black polygons are the 15 departments of the BA. In blue we have the 351 fractions and in green the 3554 radius in which BA was divided for the CNPHV.

Map 4. Unit of Analysis of CNPHV



Finally, in Map 5, we have the information of the political division and basic infrastructure. BA's 15 communes and 48 districts are represented in the map. With respect to public services, we mapped hospitals, police stations, main avenues and subway, premetro and train stations.

Map 5. Basic Infrastructure of BA



Once we have standardized and geocoded our data, the problem that arises is to choose the proper unit of analysis for spatial analysis. This problem has at least two dimensions.

First of all this choice is subject to the characteristics of the data that we have available. For example, we might want to work with the highest level of accuracy of all our variables, that is, with the geocoded addresses of all records. However the CNPHV data, due to statistical confidentiality, is presented in more aggregated units of analysis (radius, fraction, department and province, that are polygons). Therefore, if we want to work with an uniform unit for the five datasets, we have to transform the information of Financial Structure, Firms, Individuals and Basic Infrastructure from points to polygons. Using tools available in the GIS, it is possible to add, count or average points (firms, branches, public transport stations, etc) which are located within each department, section or radius based on its characteristics or calculate distances from the polygons centroids[5] to some reference point. We can, for example, determine how many head offices, branches, agencies or ATMs are located within each radius or in the case of firms, we can add the number of firms by their sector or legal form. We can also determine how far the nearest bank branch is from the polygon's centroid.

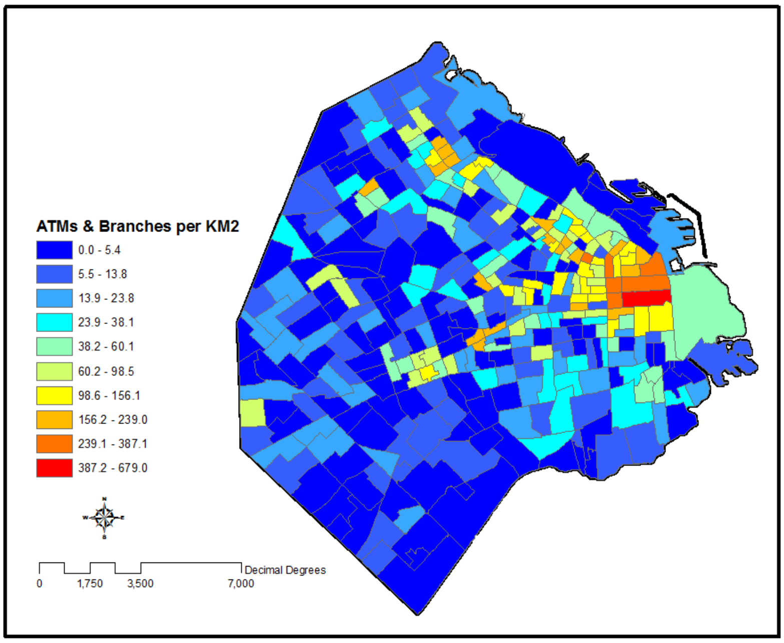
The second dimension of the problem is that the most precise unit of analysis is not always the right one for policy decisions. Given the characteristics of our data, the most accurate analysis unit for our spatial analysis of the demand and supply of financial services in BA is the radius. However, the 3554 BA's radius represent a very atomized universe to formulate economic policy. By contrast, within the 15 departments we find the same heterogeneities that led us to investigate the spatial distribution of the financial system in large urban areas. So, we chose the 351 fractions as our units of analysis.

The area of each fraction can differ significantly. It is likely that, for example the number of branches or firms, are positively correlated with the size of the fractions surface. That is, the larger the area of the fraction, the higher the value of the variables. Using variables measured in terms of squared kilometers is the way to guarantee comparability. On the other hand, as the size of the fractions is subject to their green spaces (parks, squares, boulevards, etc.), we subtract those areas from the total area. So, to express the variables in terms of square kilometers we have divided them by the total area of the fraction without taking in account the green spaces.

In maps 6, 7 and 8 we show some relevant variables at a fraction level. Map 6 shows the financial infrastructure (branches, ATM´s and other agencies) per square kilometer. We can see a very high concentration (red colored fractions) in the downtown area (San Nicolas District) and then some fractions with high concentration (orange colored fractions) scattered in different zones of BA (Districts of Caballito, Floresta, Balvanera, Villa Crespo, Recoleta and Belgrano).

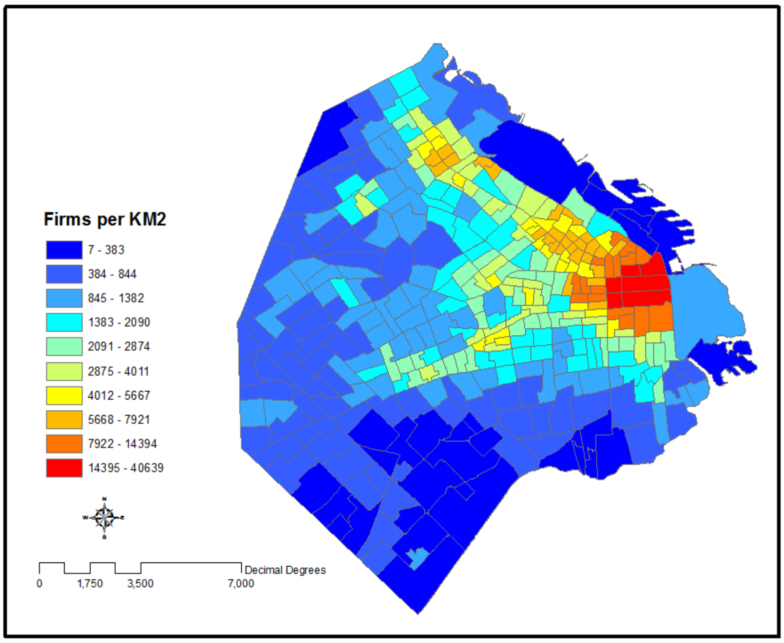
[5] Center of symmetry of a polygon.

Map 6. Financial Infrastructure per Km^2



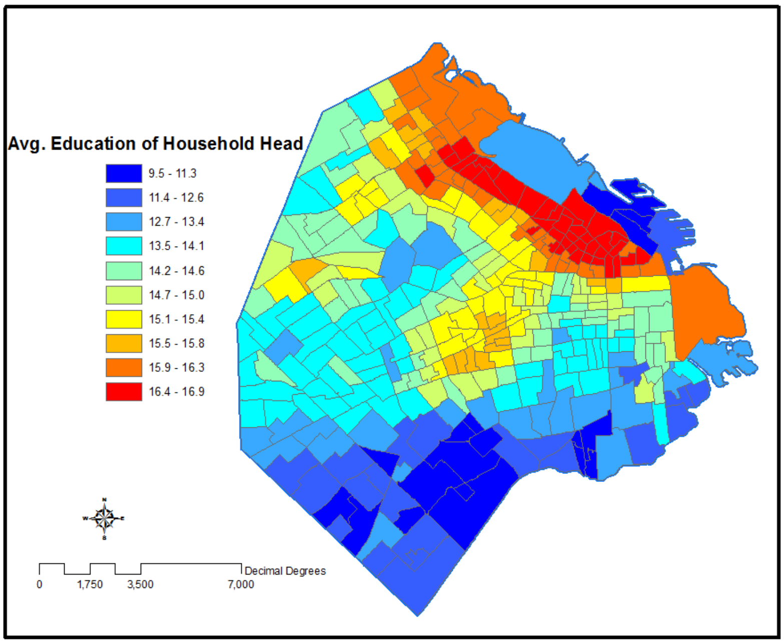
The spatial distribution of firms at the fraction level (Map 7), allows a better understanding of the variable than in Map 2. In that map we only could see that there were firms all across BA, except in green spaces. In Map 7 we can see that this is so (in all fractions we have at least 7 firms per square kilometer), but we also find, as in the case of financial infrastructure, a large concentration in downtown and its surroundings and very low levels throughout the southern part of BA. If we compare Map 6 and Map 7, we can see that in the latter there are fewer "jumps of colors", which would indicate that the distribution of firms is smoother than the one of branches, ATMs and agencies of FIs.

Map 7. Firms per Km^2



Finally, we map a variable that plays a major role when it comes to estimate the socioeconomic level of the household: the education level of the household head. In Map 8 we show the average education of household heads. In this case we also have a fairly smooth distribution (no great leaps of colors) with a significant concentration of high levels of education in northern BA (Districts of Recoleta, Palermo, Colegiales, Belgrano and Nuñez) compared to the high concentration of lower educational levels in the south (Districts of La Boca, Nueva Pompeya, Villa Soldati, Villa Riachuelo and Mataderos).

Map 8. Years of education of the household head



As we have seen, an overview of the spatial distribution of relevant variables can be reached by simply mapping them. However, there are tools that provide a more precise output when it comes to characterize spatial distributions. These tools are the ones we use in the next section to define the geographic patterns of our variables of interest.

3 Geographic Patterns

One of the main issues addressed by some spatial analysis techniques is the existence of geographical patterns. That is, if a variable is randomly distributed within a given territory or if instead the location of that variable has a specific type (dispersed or agglomerated) of spatial pattern.

In this section we try to establish whether the spatial distribution of demand (firms and general population) and supply (branches, ATMs and other agencies) of financial services have a specific pattern in the territory of the BA and if so which this is.

On the demand side we analyze the following variables: 1) population per square kilometer, 2) percentage of economically active household heads, 3) average years of education of the household head and 4) firms per square kilometer. On the supply side the selected variables are: 1) branches per square kilometer, 2) ATMs per square kilometer and 3) average number of ATMs per branch. To determine the existence of a geographic pattern of each of these variables we perform two global tests (Moran's I and Getis-Ord General G) and finally we map the results of a Local Indicator of Spatial Association (LISA).

The global spatial autocorrelation index, known as Moran's I, is widely used when it comes to detect geographic patterns. This test allows you to determine whether a variable is positively or negatively autocorrelated, or if it is randomly distributed throughout the territory.

The index is computed as follows:

$$I = \frac{N \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(x_i - \bar{x}\right) \left(x_j - \bar{x}\right)}{\left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \right) \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}$$

where

N	number of observations (fractions)
$\bar{x}$	variable mean
x_i	variable value in a given fraction.
x_j	variable value in the rest of the fractions
w_{ij}	element i,j of a matrix of spatial weights

The meaning of a positive and significant index is that the general behavior of the variable is similar in near locations. For example, fractions with a high levels education are surrounded by fractions with high levels of education and fractions with a reduced educational level are surrounded by similar fractions. Thus, a positive autocorrelation is associated with the existence of clusters. By contrast, the negative spatial autocorrelation indicates that fractions with a high level of education are surrounded by fractions with low levels of education and vice versa. A negative autocorrelation is associated with high geographical dispersion of a variable. Finally, if the value of the index is not significant, we can not reject the null hypothesis and must concluded that the variable can be generated by a random spatial process.

The spatial weights matrix used to calculate the index is a symmetric matrix of 351x351 elements. All elements of the main diagonal and the elements of fractions that are separated by a distance of more than 1425 meters[6], are zeros. These elements are not considered neighbors so its value is null. On the other hand the value of the elements corresponding to neighboring fractions (those within 1425 meters away) is the inverse distance between them. The value of each element becomes smaller as the distance between the two fractions increases.

Table 1 contains the statistical value of Moran's index with the respective Z-score and P-value for all variables of interest. Both demand and supply variables have a significant and positive spatial autocorrelation index. That is, all variables are geographically distributed as clusters. Fractions with high values are surrounded by similar fractions and fractions with low values have low value neighbors.

Table 1. Spatial autocorrelation Index

Variable	Moran's I	Z-score	P-value
Average years of education of the household head	0.685823	33.248676	0.000000
Population per km^2	0.682445	33.230643	0.000000
% of economically active household heads	0.181219	8.931639	0.000000
Firms per km^2	0.704776	35.743656	0.000000
Branches per km^2	0.399196	19.813541	0.000000
ATMs per km^2	0.420913	20.942429	0.000000
ATMs per branch	0.101593	5.029247	0.000000

[6]Minimum distance at which every fraction has at least one neighboring fraction.

The second global test is the Getis - Ord's General G. This test determines if high values clusters predominate or if low values clusters prevail in a geographical pattern. General Statistical G has a different interpretation to Moran's I and is calculated as follows:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_i x_j}{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j}$$

where
x_i variable value in a given fraction.
x_j variable value in the rest of the fractions
w_{ij} element i,j of a matrix of spatial weights

If the value of the index exceeds the expected G[7] the variable presents high value clusters more often than would be expected from random geographic distribution. On the other hand, if its value is less than the expected G, low value clusters occur more frequently.

Given that we have corroborate, through Moran's I global index, a positive spatial autocorrelation of all variables, the General Statistical G is a good instrument to figure out if the low value clusters predominate over the high value clusters, or vice versa.

Table 2. High/Low Clustering

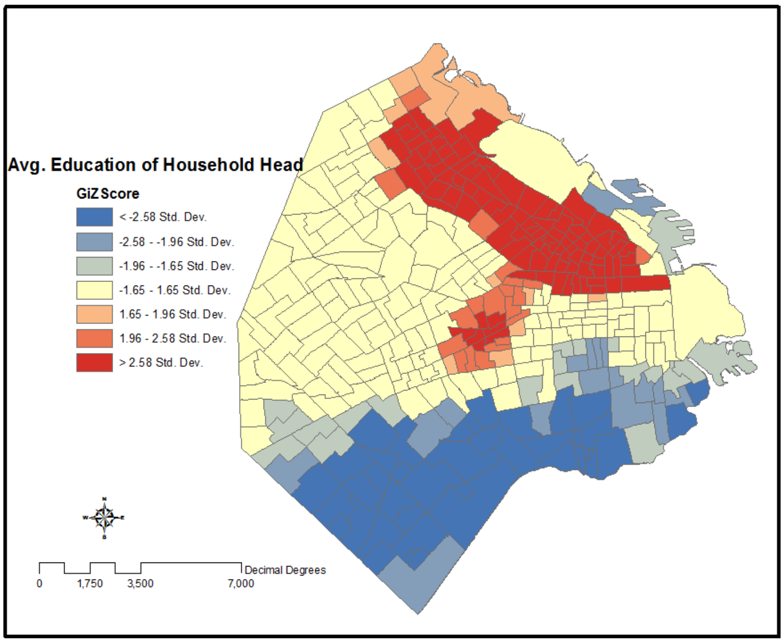
Variable	General G	Z-score	P-value
Average years of education of the household head	0.000054	12.160011	0.000000
Population per km^2	0.000088	18.739646	0.000000
% of economically active household heads	0.000049	1.114274	0.265162
Firms per km^2	0.00157	21.568258	0.000000
Branches per km^2	0.000142	14.826048	0.000000
ATMs per km^2	0.000150	16.016540	0.000000
ATMs per branch	0.000066	5.410371	0.000000

Table 2 presents General G values and their significance levels for all variables. We can see that all of them are significant and bigger than the expected value of G (in our case the value is 0.000049) except for the percentage of economically active household heads. This means that high value clusters prevail for most of our variables. In the case of the percentage of economically active household heads, General G is not statistically significant so we can not say what kind of cluster is dominant.

After determining the existence of a global geographical pattern characterized by high spatial autocorrelation and dominance of high values clusters for all variables (except for the percentage of economically active household heads) we might want to identify the location and extent of these patterns. To perform this kind of analysis we use LISA. For each unit of analysis we compute a local indicator. The "hot spot analysis", which uses the local version of the General Statistical G, allows to map clusters of high values (hot spots) and low values (cold spots). We present the Hot-Spot maps resulting of this technique. Fractions in red are related with high value clusters, blue fractions with low value clusters and yellow fractions indicate that the local indicator is not statistically significant. The intensities of the colors vary by Z-score corresponding to the fraction indicator. The greater the value of Z-score and therefore its statistical significance, the more intense the color is.

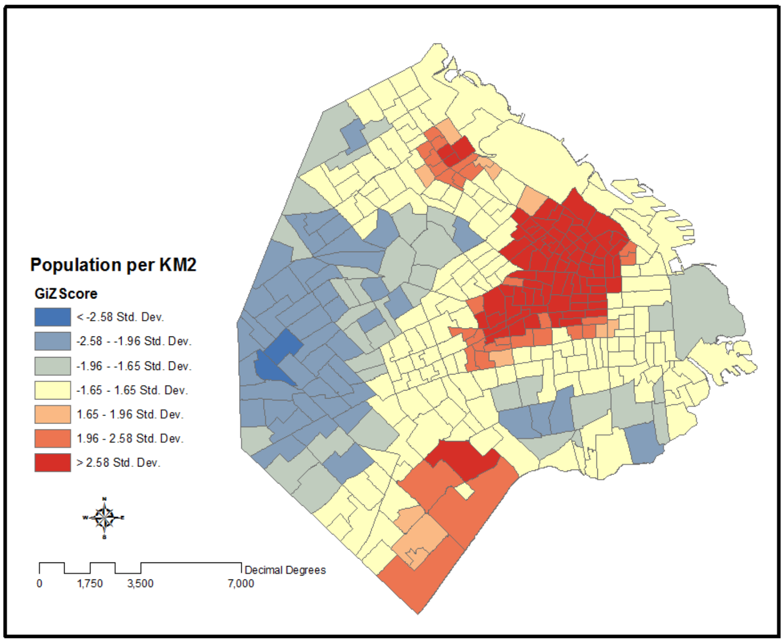
[7] Value that the indicator would take for a random spatial distribution.

Map 9. Hot Spot Analysis: Household head education



Map 9 is eloquent. We can see that hot spots (clusters of fractions with high levels of education) are located in the northeast area of BA and in the neighborhood of Caballito while cold spots concentrate in the south part of BA.

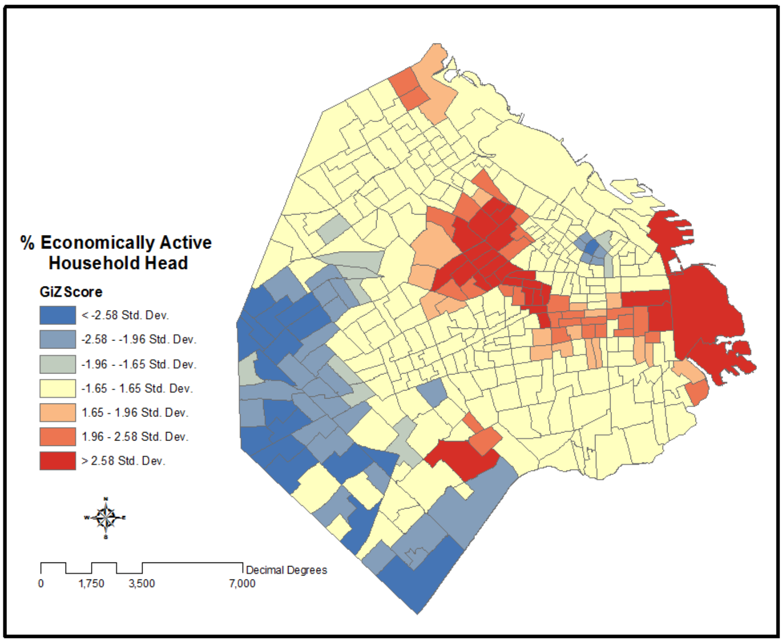
Map 10. Hot Spot Analysis: Population per km^2



Population density Hot Spots are represented on the map 10. We can see them in three areas: in the northern area (the center of Belgrano district), in the east-center (Districts of Recoleta, Balvanera, Almagro and Caballito) and in the south (Districts of Villa Riachuelo, Villa Lugano and

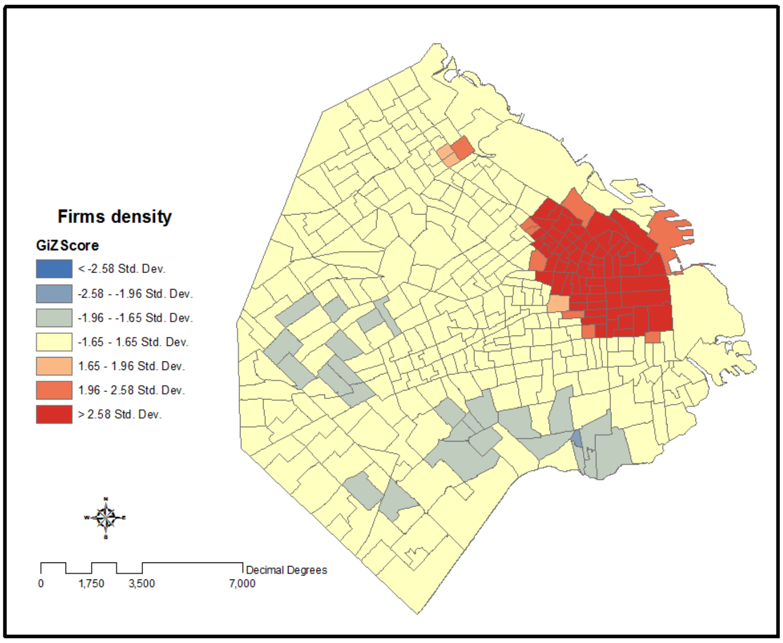
Villa Soldati). The location of the cold spots is mostly concentrated in the western area near to General Paz Avenue.

Map 11. Hot Spot Analysis: % of economically active household heads



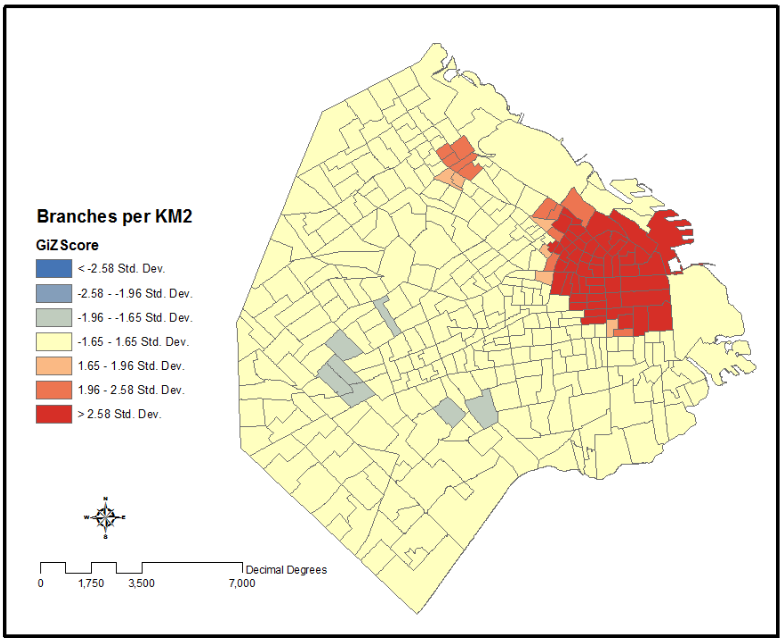
Although we could not define, through the Global General G indicator, which kind of clusters prevail in the case of the percentage of economically active household heads, the Hot Spot analysis allows us to map those clusters. In Map 11 we can see that the percentage of economically active household heads has a high concentration of hot spots in Downtown and Puerto Madero districts and in the north-center of BA (Districts of Chacarita, Villa Crespo and part of Palermo and Almagro). Cold spots are mostly in the south-west area (Districts of Villa Riachuelo, Villa Lugano, Mataderos, Liniers, Versalles, Villa Real, Montecastro and Villa Devoto), nevertheless a small cluster in the district of Recoleta is also observed.

Map 12. Hot Spot Analysis: Firms per km^2



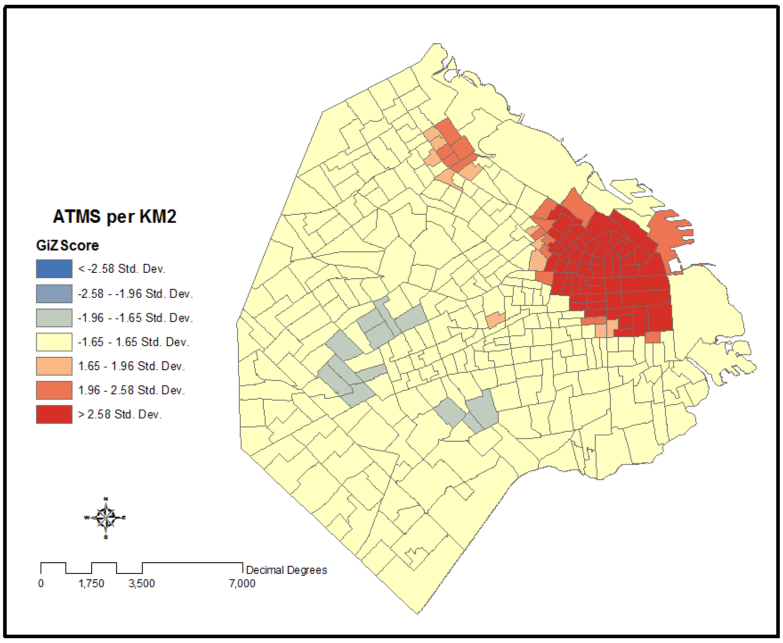
Firm density (shown on Map 12) has a clear concentration of hot spots in the Downtown area of BA (districts of San Nicolás and Montserrat) and its surroundings (Districts of Balvanera, Recoleta and Retiro). A second area where hot spots concentrate is Belgrano district, but it is not so significant. Cold spots are scattered all through the south-west area of BA but their significance is low.

Map 13. Hot Spot Analysis: Branches per km^2



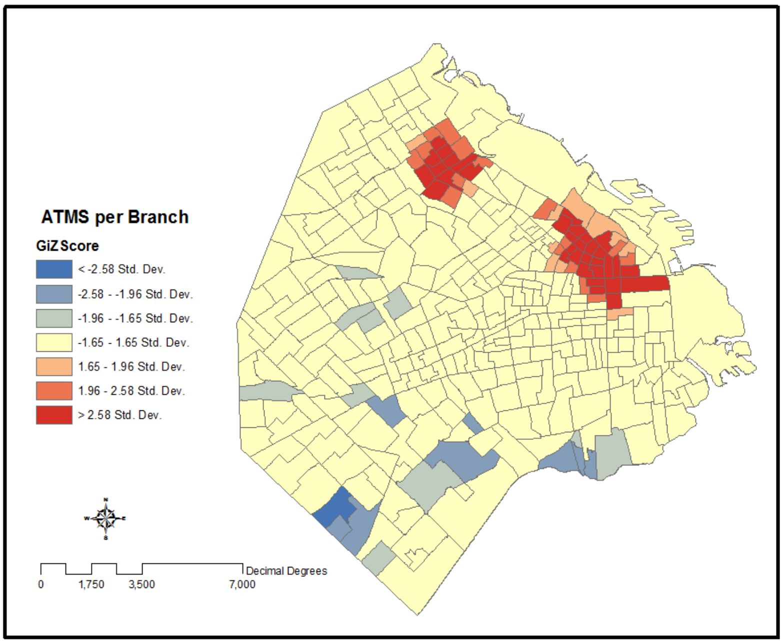
Map 13 (hot spots of branches density) is very similar to the one of firms. Downtown and its surrounding and the district of Belgrano concentrate all hot spots. Cold spots are even fewer than in the case of firm density.

Map 14. Hot Spot Analysis:ATMs per km^2



In Map 14 we can see that the density of ATMs has an almost identical distribution of hot spots as the branches density. One reason for this is that within the branches we have an important proportion of all ATMs, but ATMs outside branches have a similar geographic pattern.

Map 15. Hot Spot Analysis:ATMs per branch



When we analyze the size of branches (taking in account how many ATMs per branch we have in every fraction), we can see on Map 15 that hot spots cluster in the northeast area of BA (districts of of Belgrano, Recoleta and San Nicolas). Cold spots (branches with few ATMs) cluster instead mostly in the southwest of the BA.

By analyzing the spatial pattern of variables related to demand and supply of financial services we can have an intuition of what may be the determinants of the localization strategy of FIs. In the next section we do a detailed econometric analysis of this issue. In addition to the variables we already have analyzed, we include variables that refer to accessibility (public transport infrastructure) in our econometric approach.

4 Localization strategy of FIs determinants

As mentioned above, in this section we try to identify, using different types of econometric regressions, the determinants of the localization strategy of financial services supply. That is, we try to explain, using demand and infrastructure variables (independent variables), which are the characteristics of the fractions that influence the location of financial services supply (dependent variable). We choose the total number of ATMs expressed in terms of square kilometers as the dependent variable. We prefer this variable over the number of branches because it includes financial services supply outside of branches and also it introduces implicitly the issue of the size of branches. This variable will be transformed as required by the econometric model used[8]. The variables that characterize the fractions (independent variables) are: 1) population per square kilometer, 2) average education of the household head, 3) percentage of economically active household heads, 4) firns per square kilometer, 5) distance between the centroid and a subway, rail or premetro station, 6) meters of avenues per square kilometer[9]. In Table 3 we present a descriptive analysis of these variables.

Table 3. Descriptive Analysis

Variable	Mean	Max.	Min.	Std. Dev.
ATMs per km^2	33.01	494.18	0.00	54.96
Binary variable: 0- Witohut ATM 1- With ATM	0.80	1.00	0.00	0.40
Ln ATMs per km^2	2.24	6.20	-0.38	1.78
Population per km^2	23,928.20	104,168.80	4.79	14,378.56
Average education of the household head	14.49	16.93	9.50	1.53
% of economically active household heads	77.73	100.00	66.67	3.56
Firms per km^2	2,876.39	40,638.51	7.24	3,952.32
Distance to nearest station	681.56	2,829.56	63.88	465.31
Meters of avenue per km^2	4,336.30	13,367.67	0.00	2,819.32

We consider two central issues when we analyze the determinants of the location of financial services supply. First, whether there are ATMs in the fraction or not and second the number of ATMs per square kilometer that are located in the fraction. We analyze with a Probit and Logit model, how characteristics of the fractions that have at least one ATM are different from the fractions that have none. We use a binary variable equal to 1 in fractions that at least have one ATM and to 0 in those that have none. Then we try, through an ordinary least squares (OLS) and a Tobit model, to establish the determinants of the density level of ATMs within a fraction. With the Heckman selection model[10] we address both issues within a single theoretical framework. Finally, considering the high spatial autocorrelation of the dependent and independent variables, we test our OLS model to find out if there is any spatial dependence evidence.

Table 4 shows the coefficients, the level of significance and standard errors of the Probit and Logit regressions. The results are similar in terms of signs of the coefficients and their significance

[8] The variable will be transformed into a dichotomous variable to run the Probit and Logit models and a logarithmic transformation will apply for OLS, Tobit and Heckman models SLX.

[9] The total area of the fractions was used in this case, because green spaces are often crossed by avenues.

[10] See, Heckman, J.(1979), "Sample selection bias as a specification error",.Econometrica 47, pp.153-61.

in both models. The density of firms and the amount of meters of avenue per square kilometer have positive and significant coefficient (at a significance level of 1% and 5% respectively). In contrast, the population density has a negative and significant coefficient (at 1% significance level). The rest of the variables are not statistically significant. These results indicate that in fractions with low population density, high firm density and crossed by avenues it is more likely to find at least one ATM. These characteristics are a clear reference to fractions where economic activity is located. It is not a surprise that financial services supply is placed in those locations. The results that the average of education of the household head, which is a good indicator of the socio-eonomic level of the fraction, is not significant is at least unusual.

Table 4. *Dichotomous Analysis*

Variable	Probit	Logit
Population per km^2	-0.0000614***	-0.0001069***
	[0.0000137]	[0.0000244]
Average education of the household head	-0.1321274	-0.2342075
	[0.0874565]	[0.1537344]
% of economically active household heads	-0.0271081	-0.0452084
	[0.0241858]	[0.0407584]
Firms per km^2	0.0006291***	0.0011272***
	[0.0001439]	[0.0002675]
Distance to nearest station	-0.0000604	-0.0000681
	[0.0002065]	[0.0003583]
Meters of avenue per km^2	0.0000833**	0.0001483**
	[0.0000351]	[0.0000622]
Constant	4.765925**	8.056308**
	[2.328761]	[3.969181]
Pseudo R2	0.1681	0.1675

Standard errors in brackets
* significant at 10% ** significant at 5% *** significant at 1%

After finding out what are the determinants of the existence of financial services supply, the question comes to what determines the amount of ATMs. For this purpose we perform first an OLS regression and then a Tobit regression (considering that 69 out of 351 fractions have no ATMs). We use the logarithm of the density of ATMs as dependent variable and maintain the independent variables of previous regressions. The results of both models are shown in Table 5.

Table 5. Quantitative Analysis

Variable	OLS	Tobit
Population per km^2	0.00000284	-0..00000301
	[0.00000596]	[0.0000075]
Average education of the household head	0.2602547***	0.3263145***
	[0.0595115]	[0.0750458]
% of economically active household heads	-0.0428257**	-0.0571066**
	[0.021608]	[0.0267032]
Firms per km^2	0.0001653***	0.0001745***
	[0.0000224]	[0.0000269]
Distance to nearest station	-0.0003366*	-0.0003885*
	[0.0001911]	[0.0002332]
Meters of avenue per km^2	0.0001049***	0.0001213***
	[0.000028]	[0.0000341]
Constant	1.026646	1.083663
	[1.977936]	[2.417294]
Adj. R2/Pseudo R2	0.3669	0.0981

Standard errors in brackets
* significant at 10% ** significant at 5% *** significant at 1%

In these two models we find some similarities and some differences from the earlier dichotomous analyses. The firm density and the amount of avenue meters per square kilometer have a positive and very significant coefficient as in the Probit and Logit models. In contrast, population density variable ceases to be significant and the average education of the household heads and the percentage of economically active heads become significant. The education variable, which was mentioned as a good approximation to the socio-economic level of the fraction, has a positive and very significant coefficient in both the OLS and Tobit regressions. This result confirms the intuition of a wider range of financial services in areas of higher socioeconomic status. The percentage of economically active household heads has a negative coefficient, and this is hard to understand since a positive relation between activity status of the head and the provision of financial services could be expected. The coefficient of the distance to a train, subway or premetro station is negative and significant at 10% in both models. Being near to a station seems to be a relevant fact when we try to explain ATM,s density.

Table 6. *Heckman selection Model*

Variable	Selection	Model
Population per km^2	-0.0000627***	0.0000292***
	[0.0000136]	[0.00000584]
Average education of the household head	-0.1733539**	0.1426515***
	[0.0879597]	[0.0550223]
% of economically active household heads	-0.0218405	-0.0058029
	[0.0223622]	[0.0184879]
Firms per km^2	0.0006651***	0.000111***
	[0.0001448]	[0.0000171]
Distance to nearest station	-0.0000203	-0.0002482*
	[0.000195]	[0.0001447]
Meters of avenue per km^2	0.0000583	0.0000393
	[0.0000416]	[0.0000245]
Constant	4.970396**	0.3831505
	[2.236207]	[1.51775]
Rho		-0.6919694**
		[0.6919694]

Test LR (Rho=0): chi2(1)=4.89 Prob>chi2=0.0270
Standard errors in brackets
* significant at 10% ** significant at 5% *** significant at 1%

With Heckman's Model we address both dichotomous and quantitative analysis in the same theoretical structure. Heckman's model estimates the probability of having (or not) financial services supply with a Probit model. That probability is transformed and included as explanatory variable in an OLS regression. So, the independence of both models is tested. The Heckman model results are shown Table 6. Its important to point out that the coefficient related to probability of the Probit model (rho) is significant. This means that you can reject the null hypothesis of independence of both models. When we analyze the signs and significance levels of the rest of the coefficients, we observe that the average education, firm density and distance to a station preserve their sign and significance level of previous models. Both the percentage of economically active heads and the amount of avenue meters per square kilometer are no longer significant. However, the most important change is that the population density changes from being not significant to have a positive and significant at 1%. That is, once the selection bias is corrected, in addition to average education, firm density and distance to the nearest station; population density is also important to explain the location of the financial services supply.

Finally, using robust Lagrange Multipliers we test our OLS regression for spatial dependency. The results of the test indicate that we should use a Spatial Lag Model (SLM) Nevertheless, before doing so we introduce spatial lags in the explanatory variables and run the test again. This time the results indicate that there is no evidence for spatial dependency. The coefficients and standard errors of the OLS regression using spatially lagged explanatory variables (SLX[11]) are presented in Table 7.

[11]Spatial Lag of X Model.

Table 7. *Spatial Regression*

Variable	SLX
Population per km^2	-0.0000000474
	[0.00000708]
Average education of the household head	0.2510771**
	[0.1002899]
% of economically active household heads	-0.584707**
	[0.0257518]
Firms per km^2	0.0001352***
	[0.0000298]
Distance to nearest station	-0.0004316*
	[0.0002415]
Meters of avenue per km^2	0.0001235***
	[0.000003]
Spatial Lag - Population per km^2	0.0000592
	[0.0000605]
Spatial Lag - Average education of the household head	-0.595715
	[0.5748766]
Spatial Lag- % of economically active household heads	0.6034429
	[0.3875393]
Spatial Lag - Firms per km^2	0.0003901*
	[0.0002022]
Spatial Lag - Distance to nearest station	0.0028449
	[0.0034075]
Spatial Lag - Meters of avenue per km^2	-0.0009467**
	[0.0004264]
Constant	-36.09105
	[33.39507]
Adj. R2	0.3820

Standard errors in brackets
* significant at 10% ** significant at 5% *** significant at 1%

The coefficients of the explanatory variables without the spatial effect have similar characteristics to the ones of the OLS and Tobit model, where the average education of the household, firms density and the amount of avenue meters per square kilometer have a positive a significative coefficient. The coefficients of the percentage of economically active household heads and the distance to the nearest station are both negative and significant. Looking at the spatially lagged variables we find two significative coefficients. We have a positive relation between the density of firms in the neighboring fractions and the density of ATMs. In contrast we have a negative relationship between the amount of avenue meters per square kilometers in neighboring fractions and the density of ATMs.

5 Suitability Analysis

After analyzing the geographical pattern of demand and supply of financial services in BA and establishing the determinants of the localization strategy of IFs, we might wonder what are the best locations to enhance the development of financial infrastructure, according to the new mandate of the BCRA. Within the different spatial analysis techniques "Suitability Analysis" seeks to answer these kinds of questions. From the standpoint of the BCRA the question might be: Which is the most suitable location to promote the opening of new branches in order to extend the geographical coverage of the financial system, and ultimately move toward universal access to financial services?

To carry out this kind of analysis we choose the variables with which we will try to answer that question and standardize them to have them all to the same scale. Then we build a synthetic index that will be used to characterize the different units of analysis. The index variables may have the same weights or weights may change according to the importance we want to attribute to each variable.

We standardize our variables as follows:

$$X_i\_std = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$$

So, the range of our standardize variables will be 0 to 1.

An equiproportional synthetic index (ESI) in which each variable has the same weight could be the following:

$$ESI = \sum_{i=1}^{n} \frac{X_i\_std}{n}$$

On the contrary, if we want to choose different weights for each variable, the weighted synthetic index (WSI) might be the following:

$$ISP = \sum_{i=1}^{n} p_i * X_i\_std$$

$$\textit{where } 0 < p_i < 1 \quad \textit{and}$$

$$\sum_{i=1}^{n} p_i = 1$$

So, if we try to find locations where the supply of financial services and the bankarization levels are low and the potential demand is high, we could use the following variables: 1) weighted number of branches[12], 2) number of salary accounts per people older than 18 years, 3) percentage of firms with credit, 4) population density and 5) firm density. The first variable refers to the supply of financial services, the second and third to the bankarization level and the fourth and the fifth to the potential demand of the location.

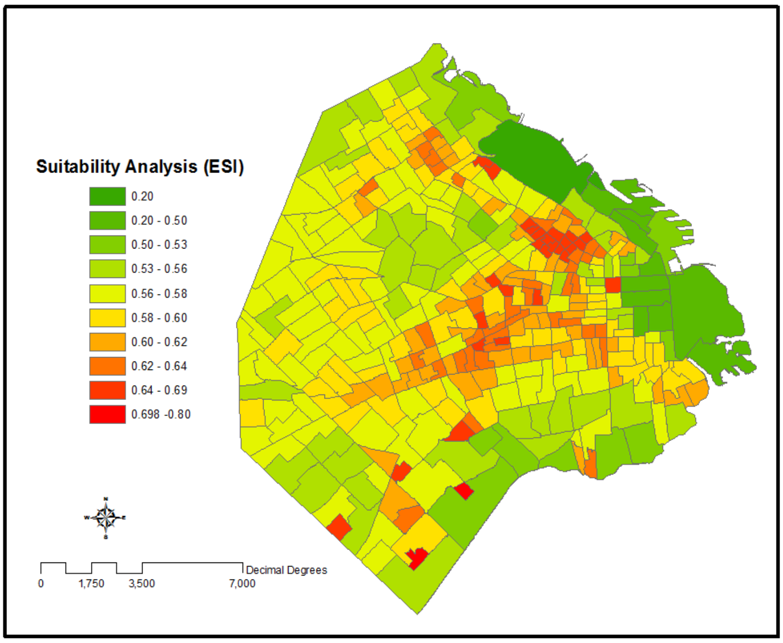
ESI for this specification would be:

$$ESI = \frac{(1 - bra\_wei_{std}) + (1 - sac\_o18_{std}) + (1 - pct\_loan_{std}) + de\_pop_{std} + de\_firm_{std}}{5}$$

As we mentioned above, by standardizing our variables all values are within the range 0 to 1. However, it is possible that in some cases we must make one last transformation to be consistent with the objective of the indicator. For example, if we want a ESI with high values to identify locations where supply and bankarization are low and the potential demand is high, then we need that fractions where we have few branches and levels of salary accounts and firms with credit are low, standardized variables need to have high values. That is why we have this terms $(1 - bra\_wei_{std})$, $(1 - sac\_o18_{std})$ and $(1 - pct\_loan_{std})$ in the ESI formula.

[12]We compute all branches that are within 1642 meters of the fractions centroids weighted by the inverse distance.

Map 16. Suitability Analysis (ESI)



Map 16 shows our ESI. Red colored fractions are those with a higher ESI and therefore where the BCRA should encourage opening of new branches. Actions should be focused broadly on the center area of BA (Districts of Caballito, Flores and Almagro), in the northeast area (Districts of Recoleta, Palermo and Belgrano) and in some specific fractions in the southwest area of CABA. Fractions corresponding to downtown (Districts of San Nicolás and Montserrat) and Puerto Madero are colored in an intense green and are therefore location where no effort should be made.

If we want to prioritize locations where the financial infrastructure is reduced, we will have to assign greater weight to the variable of number of branches. A WSI may then be as follows:

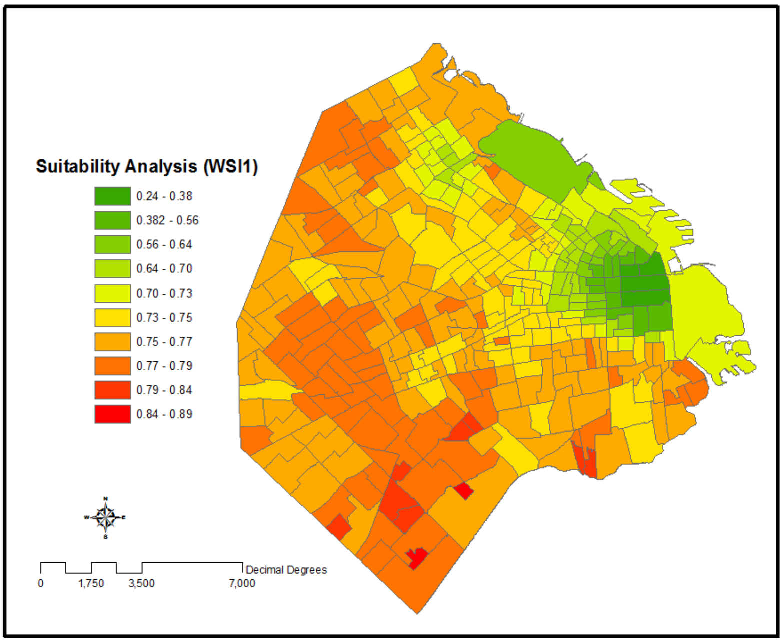
$$WSI1 = \frac{6}{10} * (1 - bra\_wei_{std}) + \frac{1}{10} * (1 - sac\_o18_{std}) + \frac{1}{10} * (1 - pct\_loan_{std}) + \frac{1}{10} * de\_pop_{std} + \frac{1}{10} * de\_firm_{std}$$

Conversely, if the idea is to promote credit to firms, the greatest weight should be assigned to the percentage of firms with credit, and our WSI may be the following:

$$WSI2 = \frac{1}{10} * (1 - bra\_wei_{std}) + \frac{1}{10} * (1 - sac\_o18_{std}) + \frac{6}{10} * (1 - pct\_loan_{std}) + \frac{1}{10} * de\_pop_{std} + \frac{1}{10} * de\_firm_{std}$$

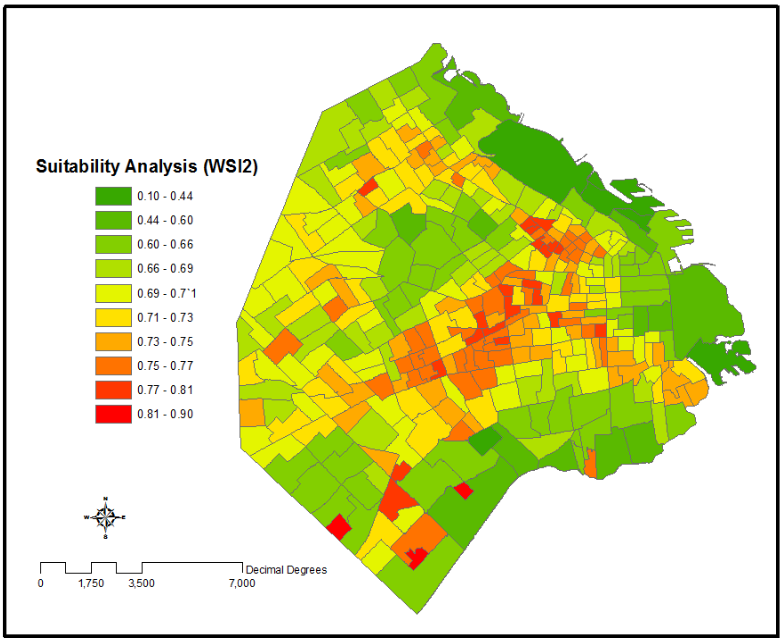
Maps 17 and 18 represent the geographical distribution of WSI1 and WSI2 respectively. We can see the changes in colors as the WSI specification changes.

Map 17. Suitability Analysis (WSI1)



In the case of WSI1 in which a greater weight is assigned to the financial infrastructure, we can be see that Downtown and his surrounding areas and the district of Belgrano are the locations where no incentives are required (green colored fractions). Red colored fractions, which are those where the opening of new branches should be encouraged, are mainly in the southern part of BA and near to the limit of BA established by the General Paz Avenue

Map 18. Suitability Analysis (WSI2)



Finally, we note that the geographic distribution of WSI2 is quite similar to the one of ISE. That is, the firms located in fractions of the center and northeast area and some disperse of the south area of BA are more restricted in terms of their access to credit.

These indicators should not literally be understood as the objectives of the Central Bank in terms of financial system expansion. They rather are examples to highlight the potential and versatility of this type of analysis.

6 Conclusions

The object of study and interest of the BCRA extends well beyond of BA. This document tries to emphasize different spatial analysis techniques that can be helpful to thinking about and conducting public policy throughout the Argentine territory. We chose BA because it is the urban area for which we have the largest amount of data and thus provides a richer context for spatial analysis. We mapped elements of both the supply and demand of financial services and defined the geographic patterns that characterize them. Using different econometric models we determine which characteristics of infrastructure and demand influence the location of the supply of financial services. Finally we presented a versatile tool which can help in the process of thinking and implementing public policies at a local level.

Results obtained in the different sections are not only consistent with each other, but also are consistent with the presumption that we, as residents of the BA, have about the geographical distribution of the demand and supply of financial services. This last fact may lower the impact of the document, because it presents no surprises to the conscious observer of the surrounding reality. But the concurrence between the analyzed data and the perceived daily context is what gives the final indication of robustness needed to perform this type of analysis in urban areas that are not familiar and to objectively contrast policy management.

Some challenges both in terms of research and economic and financial policy implementation, arise from this paper. First, based on the geo-referenced information processed for BA further additional studies must be done. The analyses of the geographical distribution of a range of financial instruments or different types of FIs are some examples. Second, expanding the geographical scope of the study is vital to have an overall assessment of the situation in all major urban areas of Argentina. This is the biggest challenge we face. The data that we already have for the other major urban centers has a lower quality than that of BA. Therefore it will require lots of processing efforts to standardize all the datasets. Third, the use of geo-referenced information within urban areas is essential if we want to take another step towards horizontal equity when we take economic and financial policy decisions at national level.

References

[1] Anselin, L. (2005). *Exploring Spatial Data with GeoDa: A Workbook. Center for Spatially Integrated Social Science.*

[2] BCRA. (2012) "Políticas de Estímulo a la Bancarización en base al Nuevo Esquema de Zonificación del Sistema Financiero", Apartado 3 del Boletín de Estabilidad Financiera (Segunda semestre 2012).

[3] Blanco, E., Denes, A. y Repetto, G.(2012). "Mapa económico y financiero de Agentina: un sistema geo-referenciado de indicadores de demanda y oferta de servicios financieros a nivel de localidad". Documento de trabajo n°$2012/59$. Banco Central de la República Argentina.

[4] Cameron, A. C. and Triverdi, P. K. (2009).*Microeconometrics Using Stata*. Stata Press.

[5] Drukker, D. M., Prucha, I. R. and Raciborski, R. (2013). "Maximum-likelihood and generalized spatial two-stage least-squaers estimators for a spatial-autoregressive model with spatial-autoregressive disturbances". *The Stata Journal,* 13, n°2, pp.221-241

[6] Drukker, D. M., Peng, H.,Prucha, I. R. and Raciborski, R. (2013). "Creating and managing spatial-weighting matrices using the spmat command" *The Stata Journal,* 13, n°2, pp.242-286

[7] Garrocho-Rangel, C. y Campos-Alanis, J. (2010). "Organización espacial del sistema bancario dentro de la ciudad: estartegia territorial, accesibilidad y factores de localización" *Economía, Sociedad y Territorio*, vol. X, n°33, pp 413-453

[8] Heckman, J. (1979) "Sample selection bias as a specification error",.*Econometrica 47*, pp.153-61.

[9] INDEC (2013). "Base de datos REDATAM. Definiciones de la base de datos". Censo Nacional de Población, Hogares y Viviendas 2010.

[10] LeSage, J. and Pace, R.K. (2009) *Introduction to Spatial Econometrics*. CRC Press

[11] Yrigoyen, C. C. (2003). *Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales*. Comunidad de Madrid. Consejería de Economía e Innovación Tecnológica.