Towards an Explanation of the Location Pattern of a Banking System in the Intrametropolitan Space: a case study of Toluca, Mexico

Carlos Félix Garrocho Rangel

El Colegio Mexiquense, A.C.

carlosgarrochorangel@yahoo.com.mx

José Antonio Álvarez Lobato

El Colegio Mexiquense, A.C.

jalvar@cmq.edu.mx

April, 2010

Abstract

Previous studies on unbanked population, conducted in Mexico and other countries, have not taken into account key issues affecting the geographic organization of the banking systems in the intrametropolitan space. In this work we explore the influence of banking services supply and demand on the spatial organization of bank branches operating in the Toluca Metropolitan Area (TMA), the fifth largest in Mexico. Our results reveal the logic of spatial behavior of the banking system within the TMA, characterized by: i. a tendency towards strong spatial agglomeration; ii. prioritizing the spatial association (co-localization) between certain banking firms; and iii. orienting the localization of bank branches toward strategic segments of demand.

Introduction: *Unbanking and Geographic Organization of The Banking System in the Intrametropolitan Space.*

Financial intermediation via commercial banking is a fundamental service for efficient operation of an economy and for ensuring that the benefits from its growth spread to all segments of society. In Mexico, the inability of the financial system to reach the entire population and serve micro-enterprises and small businesses has led to the existence of a minority sector that enjoys the services of the financial system (*banked*) and a majority sector that can access these services only partially or not at all (*unbanked*) (Rojas, 2006; Ruiz, 2004; WB 2005). The situation is similar in many developing nations (Kumar et al, 2005).

The low level of utilization of the formal financial system in Mexico means that the unbanked sectors face excessive *transaction costs* in the payment of utility bills and processing of money orders and checks (including those from government programs aimed at combatting poverty). They also face very high interest charges from the use of alternative credit mechanisms (e.g. pawnbrokers, moneylenders) and little or no income from savings and investments (WB, 2005). Being unbanked, therefore, as an individual, a micro-enterprise or a small business, directly affects income and consequently economic level and competitive position.

The phenomenon of unbanking is due to a number of barriers to the utilization of banking services, which can be grouped into three major categories: i. *Cost Barriers*, ii. *Cultural Barriers* and iii.

Location Barriers (Connolly and Hajaj, 2001; Solo, 2008; Rojas, 2006). *Cost barriers* include, for example, service fees charged by banks or minimum deposit requirements imposed on new checking accounts. *Cultural barriers* have to do with problems such as consumers' lack of trust in the financial system; and *Location barriers* are primarily geographic barriers to access resulting from the spatial localization of bank branches. Location barriers are very important even at an intrametropolitan scale: nearly 10% of those unbanked in Mexico City with enough income to use banking services report that *location* of bank branches is a key factor impeding their access to the banking system (WB, 2005; p. 22).

Nevertheless, except in a few cases (e.g. Ruiz, 2004; WB, 2005), geographic access to banking services at an intrametropolitan scale has been studied very little in Mexico, and studies conducted have not taken into account key issues of spatial organization which turn out to be critical in understanding unbaking at an intrametropolitan scale, even though the importance of the spatial variable in access to banking is widely recognized (Ruiz, 2004: p. 568; WB, 2005: p. xvii). For the purposes of this work, two central issues are emphasized: i. the influence of existing *supply* of banking services on localization of new bank branches (e.g. do existing bank branches attract, reject or have no effect, in spatial terms, on new bank branches entering the intrametropolitan market?) and ii. the influence of *demand* characteristics on the spatial localization of banking services (e.g. which population or employment characteristics, heterogeneously localized in the territory, have more influence on the localization of bank branches?).

The objective of this work is, therefore, to explore the influence of these two key determinants of spatial organization of banking systems in a specific city, to better understand the logic of geographic behavior of a banking system in an intrametropolitan space. This will advance key learnings and lead to public policy designs aimed at creating more opportunity of access to financial services for unbanked populations and businesses in large Mexican cities.

This paper comprises six sections. First, we summarize the principal findings, methods and indicators reported in the literature to analyze the spatial organization of banking systems at an intrametropolitan scale. In section two we briefly describe our methodology, study zone, information sources and spatial analysis software used for this work. In section three we examine in detail the spatial pattern of bank branches via the *Planar K-function* and in section four we evaluate the spatial attraction, rejection or indifference between branches of different banking firms, estimated via the *Cross K-function*. Having analyzed the influence of *supply* on spatial organization of a banking system, in section five we complete the analysis by examining the spatial effects of the principal segments of *demand* (population and employment). Finally, in section six we summarize our findings and the primary learnings from our study.

1. **State of the Art:** *Agglomeration, Dispersion or Randomness of the Spatial Pattern of Bank Branches, and Key Location Factors.*

The international literature contains very few examples of intrametropolitan location pattern analysis for bank branches. Frequently cited is the famous article by Avery (1991) which analyzes the supply of bank branches in cities, but which treats the urban zones as *non-dimensional* points, rather than as areas. More pertinent examples are those that treat cities as areas: Ashcroft, 1981 (which studies the spatial organization of the banking system in Ottawa, Canada); Caskey, 1992 (which studies the US cities of Atlanta, Denver, San Jose, New York and Washington, DC); Chang et al, 1997 (New York); Lee and Fukuy, 2003 (Tokyo, Japan); Lord and Wright, 1981 (Charlotte, U.S.); Medina and Núñez, 2006 (Bogotá, Colombia); Shearmur and Alvergne, 2002 (Paris, France); Topçu, 2001 (Istanbul, Turkey); and Zhou, 1998 (Los Angeles, U.S.). However, for Mexico there are no published studies of this type.

There is evidence that the intrametropolitan location of banking services, at least in certain US cities, tends towards spatial concentration, although the precise reasons for this locational behavior are only known within the banking firms (Chang et al, 1997). Nevertheless, it is reasonable to suppose that this spatial behavior is not coincidental, but rather is due to a number of advantages for firms when they agglomerate in the territory, generically called *economies of agglomeration* (Fujita and Krugman, 2004; Fujita and Thisse, 2002; O´Sullivan, 2007). However, a better understanding of the agglomeration processes of economic units, including banks, is complicated because of an important initial obstacle: the lack of widely accepted methods to correctly measure spatial concentration, dispersion or randomness (Guillain and Le Gallo, 2007; Holmes and Stevens, 2004; Marcon and Puech, 2003; Quah and Simpson, 2003). It is an accepted fact that if the spatial pattern of economic units cannot be reliably measured, it becomes very difficult to understand their spatial behavior (Duranton and Overman, 2005).

The literature reports a wide variety of methods and indicators used for characterizing the spatial patterns of economic units and activities; however, there is no consensus on which of them are more appropriate (Guillain and Le Gallo, 2007). Duranton and Overman (2005) identify three *generations* of methods for estimating spatial concentration, dispersion or randomness of firms. The first generation includes, paradoxically, *non-spatial* indicators such as those derived from Gini's Index, which in reality are statistical exercises not tied to territory (e.g. Krugman, 1991b). The second generation of indicators introduces some spatial elements and describes the underlying economic concentration, but like the first generation these indicators consider space in a *discrete* manner – organized in spatial administrative units chosen arbitrarily: countries, states, municipalities. This is driven by the way in which official statistical information is made available, not because they are more adequate for studying firms' spatial agglomeration or as a response to any spatial configuration of the economic processes in the territory (e.g. Ellison and Glaeser, 1997; Devereuex et al, 2004). Finally, the third generation of methods and indicators is derived for the most part from the *K-function* proposed by Ripley (1976), and is based on the concept of *continuous* spaces (i.e. surfaces not fragmented by artificial or arbitrary political-administrative units) enabling measurement of point concentrations (i.e. bank branches) *simultaneously* on diverse spatial scales (e.g. Arbia, 2001; Arbia et al, 2007; Duranton and Overman, 2005; 2006; Marcon and Puech, 2003). [1]

From a *geographic perspective*, the methodologies used to estimate the spatial agglomeration, dispersion or randomness of economic units in a territory can be divided into two major categories: those which treat the territory as a *discrete space* and those which treat it as a *continuous space*. The *discrete space* methodologies have been widely used (Greenstein, 2005; Maurel and Sédillot, 1999; Rysman and Mori et al, 2005), but are recognized to have important limitations: i. they rely

[1] Perhaps the methods and indicators called third generation by Duranton and Overman (2005) should be considered *fourth generation*, given the earlier appearance of methods derived from the closest-neighbor technique (Clark and Evans, 1954; Lee, 1979). The difference between these methods and those derived from Ripley’s K-function is that they don’t allow, among other things, simultaneous analysis over multiple spatial scales or a comparison with points from different study areas, as is possible with the more sophisticated methods based on the *K-function*.

on data artificially aggregated in *discrete* spatial units whose arbitrary limits do not necessarily correspond to the real behavior of the economy; ii. the discrete spatial units chosen are considered internally homogeneous (which implies a serious issue of *ecological fallacy*); iii. the definition (form) and scale (size) of the spatial units directly impact the analysis results (i.e. what geographers call the *Modifiable Areal Unit Problem*); iv. these are *non-spatial* methodologies since they do not consider territory localization of the discrete geographic units in which the information is organized (Bertinelli and Decrop, 2005; Guillain and Le Gallo, 2007); v. they use a general indicator of concentration/dispersion, but do not indicate its *level of significance* (so that it's not known exactly whether the results were obtained by chance or whether they are statistically significant). On the other hand, these methodologies have important advantages: they are relatively easy to apply, they rely on readily available data, and their interpretation is usually simple (Ellison and Glaeser, 1997; Maurel and Sédillot, 1999).

The limitations of these *discrete space* methodologies are overcome using *continuous space* methodologies (Arbia et al, 2007; Duranton and Overman, 2005; Marcon and Puech, 2003). However, these methodologies suffer from at least two important *operational* problems: i. they are very demanding in terms of the data necessary to conduct the analyses, since they are based on precise locations (geographic coordinates) of the economic units, information not usually recorded in readily available data sources (and when available, its use is typically restricted by confidentiality, e.g. Guillain and Le Gallo, 2007; Kosfeld et al, 2009), therefore requiring collection in the field; and ii. the necessary numerical calculations are compex and labor-intensive, requiring the use of adequate specialized *software* (Mitchell, 2005).

Methodologies for measuring agglomeration (dispersion) of firms in continuous spaces have relied primarily on functions derived from *Ripley's K-function* (Ripley, 1976), considered one of the most robust for simultaneously analyzing point patterns at various spatial scales (Bailey and Gatrell, 1995; Mitchell, 2005). The K-function has been utilized for some time to analyze territory patterns for various types of natural phenomena (Getis and Franklin, 1987). It is more recently being tried for exploring the spatial behavior of social and economic phenomena at a regional scale (Duranton and Overman, 2005; Yamada and Thill, 2004) and, much less frequently, in intrametropolitan spaces (Cuthbert and Anderson, 2004; Garrocho et al, 2010; Myint, 2008).

In terms of the key localization factors for bank branches, it is accepted that their location strategy follows, in general, the same localization logic as other private services firms (Birkin et al, 2002). That is, banks recognize that the spatial distribution of their markets (i.e. their customers, their branches and competitors' branches) is not homogeneous in the territory and therefore there are strategic locations in a city advantageous to their branches (Chang et al, 1997). Thus the challenge for banks in developing their location strategy is to identify these locations to maximize three key interrelated factors: i. access to potential customers; ii. sales of banking services; iii. overall bank earnings. The first factor is critical, as the other two factors depend on it (Beery, 2002). Therefore these key demand-related localization factors for bank branches are associated with certain key characteristics, including the spatial distribution both of the overall population and of the employed population.

2. Methodology, Study Zone, Sources of Information and Analysis Software

2.1. Methodology

The *planar K-function* was applied to reveal the location pattern of bank branches in the study area, that is, to establish if the pattern is predominantly agglomerated, dipersed or random. While the results showed a strong general tendency towards spatial agglomeration, the *K-function* does not produce information on whether this agglomeration is homogeneous across all branches of competing banking firms in the city, or if it is more or less intense between certain firms. Therefore to explore the spatial relationship between branches of different banks we estimated the *Cross K-function*. We conclude that there are banking firms whose branches have higher spatial interrelationship. Having established the locational influence of *supply*, we proceeded to characterize the effects of *demand* in the spatial pattern of bank branches. We conducted a multivariate analysis in which the dependent variable was the *accessibility* of the bank branches to potential consumers (population and employment) and the independent variables were the principal characteristics of the demand, including its spatial distribution. The analysis results enabled us to identify the demand characteristics that most strongly influence the location of bank branches in the intrametropolitan space.

2.2. Study Zone

The Toluca Metropolitan Area (TMA) is located a half hour from Mexico City, 40km to the southwest. It has an area of 269.6 square kilometers; its longest east-west axis measures 31.6km and its north-south axis measures 21.1km. With a population of 1.6 million, it is one of the five largest cities in Mexico. The TMA is one of the most dynamic urban areas in the country and has undergone large-scale metropolitan growth. This growth has affected its urban structure, transforming the single-center city at the beginning of the 1980s to a huge multi-centric metropolitan area at the start of the 21st century. In terms of the supply of banking services, by the end of 2009 ten banking firms were operating in the TMA with a total of 107 branches (Figure 1).

(Figure 1)

2.3. Sources of Information and Analysis Software

We obtained information on the location of bank branches from telephone directories and supplemented it with field data collected during 2009. We geographically referenced the bank branches with GPS instruments and integrated the data into ArcGIS, which facilitated management of the statistical and mapping data as well as the interface to *Spatstat* (Baddeley and Turner, 2005), the sofware utilized in our study to estimate both the Planar K-function and the Cross K-function. Spatsat automatically corrects for border effects in the study area. In addition, we utilized aerial photographs as supplementary mapping sources.

3. Spatial Pattern of Bank Branches: *Agglomerated, Dispersed or Random?*

3.1. The Planar K-function

The mathematical details of the planar K-function are widely reported in the international literature (i.e. Bailey and Gatrell, 1995; Diggle, 2003; Lu and Chen, 2006; Ripley, 1976; Ripley, 1981; Yamada and Thill, 2004; among many others). However, it is necessary to present its primary characteristics to clearly show the calculation method followed in this study.

Given a point distribution $P = \{p1,...,pn\}$ in a specific area of study, the planar K-function compares the observed value of K at a certain distance (K_{obs}) with the expected value of K at the same distance. Thus, the expected planar K-function can be expressed as:

$$K(r) = \frac{1}{\lambda} E \begin{pmatrix} \textit{number of points in P} \\ \textit{within a Euclidean distance r} \\ \textit{from an arbitrary point in P} \end{pmatrix} \qquad (1)$$

where:

- r: distance at which K is calculated
- λ: estimate of the distribution of the density of points in P (i.e. number of points per unit area)
- $E(\ )$: expected point value.

When the pattern of points follows a homogeneous Poisson process, that is, it has complete spatial randomness (CSR), the function can be expressed (Dixon, 2002; Okabe and Yamada, 2001; Smith, 2009) as:

$$K(r) = \pi r^2 \qquad (2)$$

For an observed set of n points, distributed over a region of area A the observed K-function takes the form:

$$K_{obs}(r) = \frac{1}{\lambda' n} \sum \sum_{i \neq j} i_r(d_{ij}) \qquad (3)$$

where:

- n: number of points observed
- d_{ij}: Euclidean distance between points p_i and p_j

$$I_r(d_{ij}) = \begin{cases} 1 \textit{ if } d_{ij} \leq r \\ 0 \textit{ otherwise} \end{cases} \qquad (4)$$

$$\lambda' = \frac{n}{A} \qquad (5)$$

Substituting λ' into the formula K_{obs} (3), K_{obs} (r) can be defined as

$$K_{obs}(r) = \frac{A}{n^2} \sum \sum_{i \neq j} i_r(d_{ij}) \qquad (6)$$

To simplify the calculations (Diggle, 2003), λ' can be redefined as

$$\lambda' = \frac{n-1}{A} \qquad (7)$$

which is practically the same as the previous definition of λ' for sufficiently large values of *n*, and therefore the function $K_{obs}(r)$ can now be represented as:

$$K_{obs}(r) = \frac{A}{n(n-1)} \sum \sum_{i \neq j} i_r(d_{ij}) \qquad (8)$$

Correction factors for border effects can be included as an additional term within the sums in equation (8) (Dixon, 2002).

As the value of *r* increases to cover the entire distribution of points *P* in the study area, a series of values $K_{obs}(r)$ can be obtained. These values can be graphed on the vertical axis of a Cartesian plane against values of *r* along the horizontal axis. Comparing the resulting curve for $K_{obs}(r)$ against the curve for K(r), it is possible to determine whether the observed distribution of points *P* follows a clustered (agglomerated), dispersed or random pattern that is statistically significant on various spatial scales (Levine, 2004; Lu and Chen, 2006).

3.2. Results of the Planar K-function

To determine the upper and lower confidence limits for our sample of 107 bank branches, we followed the standard practice of conducting 999 Monte Carlo simulations to a maximum distance of 6.0km, the distance recommended to calculate the K-function for the given dimensions of the TMA, that is, about 0.25 of the maximum distance between the furthest points in the study area (Baddeley and Turner, 2005). We calculated the planar K-function for 513 distance intervals (every 12 meters) and the results are shown on Figure 2.

(Figure 2)

From Figure 2 it is clear that the spatial pattern of the banks is strongly agglomerated at any distance, since the observed K-function is above the upper confidence limit determined by the Monte Carlo simulations (the so-called *theoretical K-function*) for all values of *r*.

To determine the *maximum agglomeration distance* as well as the *agglomeration intensity* we calculated three indicators that involve the relationship between the observed value and the upper confidence limit. Maximum agglomeration distance is obtained from equation (9) and is the distance at which the planar K-function reaches is maximum value (Duranton and Overman, 2006):

$$M(r) = Max\ [K_{obs}(r) - K_{ls}(r)] \qquad (9)$$

Using equation (9) the *maximum agglomeration distance* of the bank branches occurs at 5.2km. Figure 3 shows the results of the function *M* for all values of *r*.

(Figure 3)

A second way of calculating maximum agglomeration distances is via equation (10) (Kosfeld et al, 2009) which transforms the planar K-function into a standardized linear form with scaled variance, emphasizing the *maximum local agglomeration*:

$$L'(r) = \sqrt{\frac{K_{obs}(r)}{\pi}} - \sqrt{\frac{K_{ls}(r)}{\pi}} \qquad (10)$$

Results of the function (*L'*) are shown on Figure 4, where we can identify three agglomeration peaks, the highest of which occurs at 4.6km, the second highest at 2.5km and the third at 768 meters. These results are interpreted to mean that the bank branches in the TMA are agglomerated at three scales: i. high local agglomeration, for example a number of branches on the same city block, street or shopping center; ii. agglomeration at the scale of a TMA service subcenter; and iii. agglomeration at a metropolitan scale, where we record the maximum agglomeration levels.

(Figure 4)

To determine agglomeration intensity at the distances of maximum agglomeration we calculated the Index of Concentration (IC) which calculates the excess of bank branches with respect to the upper confidence limit at a particular distance (Dixon 2002a; 2002b; Fajardo, et al, 2006), as shown in equation (11):

$$IC(r) = \frac{K_{obs}(r)}{K_{ls}(r)} \qquad (11)$$

Thus, at the distance of absolute maximum agglomeration we estimated an IC of 1.87, which means that at a distance of 5.2km we observe 87% more bank branches than we would expect with *complete spatial randomness* (CSR). Using the standardized indicator *L'* we obtain an IC of 2.0 at 4.6km, an IC of 2.8 at 2.5km and an IC of 5.8 at 768 meters. These results demonstrate without a doubt the strong agglomeration of bank branches at all of the geographic scales in the intrametropolitan scale.

4. Co-Location of Bank Branches: *Spatial Attraction, Rejection or Indifference Between Branches of Competing Firms?*

The planar K-function only considers location of each of the points and ignores any other relevant information about the points. However, when the points are not identical but instead have different attributes (e.g. bank branches belonging to different firms) it is possible to conduct *multivariate* analysis of the point patterns. Each distinct attribute associated with the coordinates of a point is called a *mark* (Dixon 2002a). The marks can be discrete (e.g. belonging or not to a certain banking firm) or continuous (e.g. number of teller stations or size of the service areas). When the point patterns have marks, the analysis can extend beyond the general or *univariate* pattern characterization of the planar K-function, and examine if a spatial dependency exists between point patterns with different attributes (marks) (Rot, 2006).

This can be done by applying the *Cross K-function* to evaluate the spatial relationship between units of different firms (e.g. branches of different banks), to see if they *agglomerate*, *reject*, or distribute themselves *at random*. In other words, the Cross K-function proves whether or not the spatial distribution of a set of points (e.g. branches of a particular bank) is related to the distribution of another set of points (e.g. branches of other banks) (Mynet, 2008).

4.1. The Cross K-Function

The formula for the univariate planar K-function can be extended to consider sets of points with different marks: given a universe of points P located within the same polygon, and two marks, *i* and *j*, which generate two subsets of points, $P_i = \{P_1..P_n\}$ and $P_j = \{P_1..P_m\}$; the expected value of the cross K-function of *i* with respect to *j* is obtained via equation (12):

$$K_{ij}(r) = \frac{1}{\lambda_j} E\begin{pmatrix} \textit{number of points in } P_j \\ \textit{within a Euclidean distance } r \\ \textit{from an arbitrary point in } P_i \end{pmatrix} \qquad (12)$$

where:

- r: distance at which K_{ij} is calculated
- λ_j: estimate of the density distribution of points in P_j (i.e. number of points of type *j* per unit area)
- E: expected value of points

If there are *t* sets of points, the number of K-functions that can be calculated is t^2, that is, $K_{11}(r)$, $K_{12}(r)$ ... $K_{t1}(r), K_{t2}(r) \ldots$ $K_{tt}(r)$. However, considering that the functions $K_{11}(r)$, $K_{22}(r) \ldots K_{tt}(r)$ are univariate K-functions that would not show dependency – and therefore are similar to the traditional planar K-function, and considering also that the function K_{12} is equal to the function K_{21} (stationary spatial process without border-correction factors), the resulting number of K-functions to calculate is equal to $\frac{t^2-t}{2}$.

Assuming independence, the expected theoretical value of the cross K-function is similar to that of equation (2), that is:

$$K_{ij}(r) = \pi r^2 \qquad (13)$$

As in the univariate function, the confidence levels for the cross K-function are obtained via Monte Carlo simulations, although this is complicated by the fact that the simulations must maintain a spatial pattern in each individual process, breaking any dependence between them. Thus, for an observed set of points with marks *i* and *j* distributed in a region with area *A*, the observed cross K-function takes the form:

$$K_{ij}^{obs}(t) = \frac{1}{\lambda'_i \lambda'_j A} \sum_k \sum_l I(d_{i_k j_l}) \qquad (14)$$

where:

$\lambda'_i \lambda'_j$: Density of points of type *i* and type *j*

$d_{i_k j_l}$: Distance between the *k-th* location of type *i* and the *l-th* location of type *j*

$$I(d_{i_k j_l}) = \begin{Bmatrix} 1 \ if \ d_{i_k j_l} < r \\ 0 \ otherwise \end{Bmatrix} \qquad (15)$$

The border correction factors can be included as an additional term within the sums in equation (14).

Comparing the function K_{ij}^{obs} with the function K_{ij} (*theoretical)* we can determine if there is special agglomeration, rejection or independence between sets of points with the marks *i* and *j*. The spatial pattern of points *i* and *j* are agglomerated if K^{obs} is above the upper limit of significance of the theoretical K-function calculated via the Monte Carlo process. The sets of points i and j are spatially independent if the function K_{ij}^{obs} is within the confidence levels of the theoretical K-function, and represent spatial rejection if the function K^{obs} is below the lower limit of significance of the theoretical K-function.

4.2. Results of the Cross K-Function

We calculated cross K-functions for the five most important banks operating in the TMA (Table 1)[2] and the results are shown in Figure 5.

(Table 1)

Given the relatively small number of branches per bank, the Monte Carlo simulations generated relatively broad significance bands. Nevertheless we can draw statistically reliable conclusions.

(Figure 5)

First, the results of the cross K-function show a clear tendency towards agglomeration, even when considering sets of bank branches differentiated by the banking firm to which they belong. This is consistent with results from the planar K-function, showing the same behavior for the entire group of bank branches. However, we notice certain behavior differences (Graphs 2 and 3).

Banamex is the bank that tends to agglomerate the least with its competitors. Its average Index of Concentration is 1.36, implying that its branches are located with more spatial independence than branches from the rest of the firms. This suggests that Banamex follows a location strategy with some tendency towards spatial rejection relative to branches of its competitors (especially with those of BBVA and Banorte, which comprise with Banamex the three most important banks in the TMA).

[2] The five banks analyzed concentrate 83% of the branches they operate in the TMA; the rest of the banks have small presence in the zone (1 or 2 branches), for example Banjercito, IXE, Afirme, Banco del Bajio or ScotiaBank (with a presence of 7 branches).

(Table 2)

(Table 3)

The graphs in Figure 5 confirm the previous assertion, showing for example how Banamex avoids agglomerating at short distances from BBVA and Banorte, although the branches are spatially dependent at medium distances. This same behavior occurs between Banamex and the less important banking firms. Thus Banamex looks for locations that are independent from the rest of the firms, but are not too isolated. In general, Banamex will not locate a branch on the same block as its competitors, but will locate the branch in the same zone or intrametropolitan subcenter.

In contrast, Santander, HSBC and Banorte had the highest average Indices of Concentration (1.53, 1.49 and 1.50 respectively), which implies that these firms tend to co-locate more readily with their competitors in the intrametropolitan space. BBVA had intermediate agglomeration with its competitors, with an Index of Concentration right at the average of 1.45.

The spatial relationship between pairs of banks contributes key information about their location patterns. Banamex and Banorte had the lowest index of spatial concentration and the longest agglomeration distance of all the firms (1.22). A value over 1.0 denotes a certain degree of attraction between this pair of banks, but it was the weakest among pairs of firms considered. In contrast, Banorte and Santander had the highest Index of Concentration (1.72), followed by Banorte and HSBC (1.61).

In light of the importance and level of experience of Banorte as compared to HSBC and Santander in the TMA, it's possible to surmise that the latter two banks employ a "follow the leader" location strategy, even though this can lead to the so-called *rational herding effect* (Chang et al, 1997).[3] BBVA also shows a tendency to co-locate with competitors other than

[3] "The term rational-herding has been used to describe situations in which it is individually rational for agents/firms to mimic the actions of others even though such mimicry can potentially lead to aggregate outcomes that are sub-optimal" (Chang et al, 1997; p. 2).

Banamex, especially with Santander and HSBC, the two less important banks in the TMA. This seems to confirm the “follow the leader” location strategy.

5. Demand as a Determinant of Location Strategy: *Population and / or Employment?*

So far two conclusions emerge: i. banking services agglomerate in the territory of the TMA; and ii. some firms attract each other spatially more than others. But we haven’t related the spatial distribution of the supply with the location of the demand. This section does so, by examinining one of the key variables impacting not only the design of policies to combat unbanking but also the definition of a banking firm’s location strategy: *accessibility* (WB, 2005). Accessibility is so important for bank branches in the TMA that more than 95% of them are located on the main roads of the city (Figure 1). However, the inherent tendency of the banking system to function in an agglomerated manner in the TMA acts as a detriment to accessibility for a large part of the potential demand.

Before going any further, we must point out that *accessibility* of a service is very different than actual *provision* of the service. While service provision depends simply on the relative magnitude of supply and demand, accessibility also considers the spatial location of supply and demand (Unal et al., 2007). This is clearly the case in the TMA. Comparing the breadth of the TMA (270.0 km^2) against the number of bank branches, we can conclude that the 107 branches in the city are more than enough to cover the entire demand, with at least one branch every 1.0km.

However, as we saw, the bank branches are not homogeneously distributed in the TMA, but rather they tend to agglomerate in the intrametropolitan space. Therefore, there must be differences of accessibility to banking services in the TMA, and to explain these differences we must look beyond service provision, to the location strategy of the banks (Chang et al, 1997).

To explore the determinants of location strategy for bank branches as related to *demand*, we conducted a multiple linear correlation analysis, where the dependent variable was the accessibility of each AGEB (Area Geoestadística Básica, or Basic Geostatistical Area, the smallest spatial unit reported in the census in Mexico) to the bank branches, estimated via a gravity accessibility index in the following manner (Bath et al, 2002):

$$I_i = \sum_j \left(\frac{S_j}{O_{tot}} \right) C_{ij}^{-b} \qquad (16)$$

where S_j is the service supply in the service unit "j" (as defined by the number of teller stations in each bank branch, rather than by using the branch size); O_{tot} is the population demanding banking services in the study zone (total and employed population in each AGEB); C_{ij} are the transportation costs between the origin i (the centroid of each AGEB) and the destination bank branch j; and $-b$ is the friction parameter of the distance.

The independent variables were *population* (where population lives) and *employment* (where population works). The population variables were: i. population earning from two to five times minimum salary; ii. population earning over five times minimum salary; iii. total population 15 years and older. The employment variables correspond to 19 of the 20 sectors defined in the North American Industry Classification System (wich is used by Mexico, Canada and the United States). The agricultural sector was not included because it is not significant in the TMA.

To identify the variables most closely associated with the changes in accessibility to the bank branch system, the *Stepwise* method was utilized. We calcualted the correlation coefficients (adjusted R^2) and the partial coefficients, both standardized and non-standardized, for each variable selected in the final model. The following tests were conducted to ensure the significance of the results and prevent colinearity between the independent variables: correlations between the independent variables, variance analysis (ANOVA), F and t tests, and calculation of both Tolerance and Variance Inflation Factors (VIF) for each estimated model. The statistical package used was SPSS.

5.1. *Results and Interpretation*

The variables statistically most strongly correlated with accessibility to the banking system in the TMA are, in order of importance:[4] i. Government Activities (0.345); ii. Population with Income above Five Times Minimum Salary (0.339); iii. Total Population 15 years and older (-0.303); iv. Other Services Excluding Government Activities (0.277); and v. Retail Commerce (0.146). Together these variables explained 0.625 (Adjusted R^2) of the behavior of accessibility to the city's banking system. The only independent variable with a *negative* value was Total Population 15 years and older. This can be explained, since the larger the total population, the larger the low income segment (the direct correlation between these two variables is 0.851) and therefore, the lower the interest on the part of bank branches to locate in these low income spatial markets. The other variables all had positive values.

The correlation between accessibility to bank branches and the presence of government activity can be explained in two primary ways: a. public sector entities are intensive users of banking services; and b. a large proportion of state and local government offices is located in the central business district (agglomerated around the state government house and municipal headquarters, historically located in the center of town in Mexican cities). This area also provides advantages to bank branches for coming into contact with the rest of the demand in the city. Likewise, it is easy to understand the correlation between accessibility to bank branches and the high-income population, which constitutes, along with the employed population, a *strategic market segment* due to their high demand for banking services.

The other two variables with positive correlation are very interesting, because this correlation can be greatly explained by the economic characteristics of the TMA and by the new location patterns of the retail business sector. Thus the variable *Other Services Excluding Government Activities* likely turns out to be significant because it includes the subsegment *Repair and Maintenance Services*, a very strong economic activity in the city, especially as it relates to trucks, cars and agricultural, industrial, commercial and service equipment. Finally, the variable *Retail Commerce*

[4] The order of importance was determined according to the Standardized Beta Coefficients, which coincided with the Non-Standardized Beta Coefficients (the latter are shown between the parentheses).

is correlated with accessibility to bank branches because it includes the subsegment *Retail Commerce in Self-Service and Department Stores*, which not only has enjoyed very fast growth in the TMA in recent years, but also because the new resulting strategies of spatial organization of retailing in the city (shopping centers and business centers, for instance) provide location options highly valued by bank branches (in 2009, 17 bank branches, 16% of the total, were located in shopping and business centers). Additionally, this new retailing formats includes a variety of business activities that not only make extensive use of banking services but require to be highly accessible to their potential demand, which coincides with the interests of banking firms.

Thus, from the demand side, the results of the multiple correlation analysis shows that access to the employed population (especially government employees and workers in *key* private sector activities) as well as access to the high-income population, are at the heart of bank branch location strategies.

6. Conclusions

The evidence from our application of the planar K-function clearly shows the tendency of bank branches to agglomerate within the TMA at any scale of observation. In contrast, results from our estimation of the cross K-function also clearly show that not all banks attract each other equally in the intrametropolitan space. Some banks strongly attract each other while others less so, adopting to some degree a “follow the leader” strategy.

In other words, in most cases wherever one branch is located another branch will *co-locate* with it. This spatial behavior necessarily implies unequal accessibility of banking services within the TMA, which can worsen, at least partially, the unbanking of the population, micro-enterprises and small business. While the specific reasons for spatial agglomeration of bank branches is part of the carefully guarded internal *know-how* of banking firms (Chang et al, 1997), for the TMA we can propose additionally the following explanations: i. bank branches achieve economies of agglomeration, enabling them to lower customer costs of finding and obtaining banking services; ii. agglomeration economies for bank branches also include the ability to share services and costs, in increasingly frequent cases of agglomeration in shopping and business centers (e.g. security, exterior maintenace, lighting); iii. agglomeration is also a result of “follow the leader” location strategies, enabling smaller firms to take advantage of the experience, know-how and decision-making capability of the leaders (which can in turn create a *spatial rational herding effect*); and iv. demand for banking services is agglomerated in the territory, causing firms to prioritize location decisions that maximize their access to strategic demand segments: a. government employees and workers in *key* private sector activities and b. the high-income population.

Bibliography

Arbia G, 2001, “Modelling the Geography of Economic Activities in a Continuous Space” *Papers in Regional Sciences* **80** 411-424

Arbia, G, Espa, G, Quah, D, 2007, “A Class of Spatial Econometric Methods in the Empirical Analysis of Clusters of Firms in the Space”, Discussion Paper No. 5, University of Trento, Department of Economics, Trento, Italy.

Ashcroft R, 1981, “Intrametropolitan Branch Banking: Ottawa, 1930-1979” *Ontario-Geography* (17) 1-18

Avery, R, 1991, “Deregulation and the Location of Financial Institution Offices” *Economic Review* Federal Reserve Bank of Cleveland Third Quarter 30-42

Baddeley A, Turner, R, 2005, “Spatstat: an R Package for Analyzing Spatial Point Patterns” *Journal of Statistical Software* **12**(6) 1-42

Bailey T, Gatrell A, 1995 *Interactive Spatial Data Analysis* (Longman Press, Harlow, UK)

Bath C, Handy, S, Kockelman, Kara, Mahmassani, Hani, Gopal A, Srour I, Weston L, 2002, “Development of an Urban Accessibility Index: Formulations, Aggregation, and Application”, Research Report No. FHWA/TX-02-4938-4, Center for Transportation Research, The University of Texas at Austin, Austin, USA, http://www.utexas.edu/research/ctr/pdf_reports/4938_4.pdf

Beery A, 2002, "Site Selection Insight" *American Bankers Association Bank Marketing*, http://findarticles.com/p/articles/mi_hb4838/is_4_34/ai_n28916322/

Bertinelli L, Decrop J, 2005, "Geographical Agglomeration: Ellison and Glaeser's Index Applied to the Case of Belgian Manufacturing Industry" *Regional Studies* **39**(5) 567-583

Birkin Mark, Clarke G, Clarke M, 2002, *Retail Geography and Intelligent Network Planning* (John Wiley and Sons, Sussex, UK)

Caskey J, 1992, "Bank Representation in Low-Income and Minority Urban Communities", Research WP 92-10, Federal Reserve Bank of Kansas City, Kansas, USA.

Chang A, Chaudhuri S, Jayaratne J, 1997, *"Rational Herding and the Spatial Clustering of Bank Branches: an Empirical Analysis"*, Discussion Paper Series No. 9697-24, Federal Reserve Bank of New York, NY, USA.

Clark P, Evans F, 1954, "Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations" *Ecology* **35**(4) 445-453

Connolly C, Hajaj K, 2001, "Financial Services and Social Exclusion" Financial Services Consumer Policy Centre, University of New South Wales, Sydney, Australia, http://www.fscpc.org.au/publ/publications.htm

Cuthbert A, Anderson W, 2004, "Using Spatial Statistics to Examine the Pattern of Urban Land Development" *The Professional Geographer* **54**(4) 521-532

Devereux, M, Griffith R, Simpson, H, 2004, "The Geographic Distribution of Production Activity in the UK" *Regional Science and Urban Economics* **34**(5) 533-564

Diggle P, 2003 *Statistical Analysis of Spatial Point Patterns* (Oxford University Press, London, UK)

Dixon P, 2002a, “Ripley´s K-Function”, in *Encyclopedia of Environmetrics* **3** Eds A El-Shaarawi, W Piegorsch (John Wiley & Sons, Chichester, UK) pp 1796-1803

Dixon P, 2002b, “Nearest Neighbour Methods”, in *Encyclopedia of Environmetrics* **3** Eds A El-Shaarawi, W Piegorsch (John Wiley & Sons, Chichester, UK) pp 1370-1383

Duranton G, Overman H, 2005, “Testing for Localization Using Micro-Geographic Data” *Review of Economic Studies* **72**(4) 1077-1106

Duranton G, Overman H, 2006, “Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data”, CEP Discussion Paper No. 756, Centre for Economic Performance, London School for Economics and Political Science, London, UK.

Ellison G, Glaeser E, 1997, “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach” *Journal of Political Economy* **105**(5) 889-927

Fajardo A, Goodburn J, Graham J, 2006, “Spatial Patterns of Regeneration in Managed Uneven-Aged Ponderosa Pine/Douglas Fir Forests of Western Montana, USA” *Forest Ecology and Management* **223**(2) 255-266

Fujita M, Thisse JF, 2002, *The Economics of Agglomeration: Cities, Industrial Location and Regional Growth* (Cambridge University Press, Cambridge, UK)

Fujita M, Krugman P, Venables A, 1999, *The Spatial Economy* (MIT Press, Cambridge, MA, USA)

Garrocho C, Álvarez JA, Chávez T, (2010), “Patrón Espacial de Firmas Terciarias en el Centro Tradicional de Negocios: Del Espacio Continuo al Espacio Red”, *Economía Mexicana* (in print)

Getis A, Franklin J, 1987, "Second-Order Neighborhood Analysis of Mapped Point Patterns" *Ecology* **68**(3) 473-477

Graves S, 2003, "Landscapes of Predation, Landscapes of Neglect: A Location Analysis of Payday Lenders and Banks" *The Professional Geographer* **55**(3) 303-317

Guillain R, Le Gallo J, 2007, "Agglomeration and Dispersion of Economic Activities in Paris and its Surroundings: An Exploratory Spatial Data Analysis", Discussion Paper 06-T-10, Regional Economics Applications Laboratory, University of Illinois -The Federal Reserve Bank of Chicago, Urbana, IL, USA.

Holmes T, Stevens J, 2004, "Spatial Distribution of Economic Activities in North America", in *Handbook of Urban and Regional Economics: Cities and Geography* Eds V Henderson, JF Thisse http://www.core.ucl.ac.be/staff/thisseHandbook/holmes%3Astevens.pdf

Kosfeld R, Eckey HF, Lauridsen J, 2009, "Spatial Point Pattern Analysis and Industry Concentration", Joint Discussion Paper Series in Economics 16-2009, Universities of Aachen - Gießen - Göttingen – Kassel - Marburg - Siegen, Germany http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/16-2009_Kosfeld.pdf

Krugman P, 1991, *Geography and Trade* (MIT Press, Cambridge, MA, USA)

Kumar A, Beck T, Campos C, Chattopadhyay, S, 2005, "Assessing Financial Access in Brazil", World Bank WP 50, World Bank, Washington, USA http://siteresources.worldbank.org/EXTINCLUSIVEFINSYS/Resources/AssessingFinancialAccessBrazil.pdf

Lee Y, Fukui H, 2003, "Identifying the Relationship Between Area Characteristics by Social Class and Bank Branches Distribution using GIS: Case Study of Tokyo 23 Wards", Poster Session at Map Asia 2003, 13-15 October, Kuala Lumpur, Malaysia http://www.gisdevelopment.net/application/business/ma03177pf.htm

Lee Y, 1979, "A Nearest-Neighbor Spatial-Association Measure for the Analysis of Firm Interdependence" *Environment and Planning A* **11**(2) 169-176

Levine N, 2004, *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations V.3.0* (Ned Levine & Associates, Houston, TX - The National Institute of Justice, Washington, DC, USA)

Lord J, Wright D, 1981, "Competition and Spatial Location Strategy in Branch Banking: Spatial Avoidance or Clustering" *Urban Geography* **2**(3) 189-200

Lu Y, Chen X, 2006, "On the False Alarm of Planar K-Function When Analyzing Urban Crime Distributed Along Streets" *Social Science Research* **36**(2) 611-632

Marcon Eric, Puech F, 2003, "Evaluating the Geographic Concentration of Industries Using Distance-Based Methods" *Journal of Economic Geography* **3**(4) 409-428

Maurel F, Sédillot B, 1999, "A Measure of the Geographic Concentration in French Manufacturing Industries" *Regional Science and Urban Economics* **29**(5) 575-604

Medina C, Núñez J, 2006, *La Oferta de Servicios del Sector Financiero Formal en Bogotá*, Documento Centro de Estudios sobre Desarrollo Económico (CEDE) 2006-05, Universidad de los Andes, Bogotá, Colombia http://ideas.repec.org/p/col/000089/003043.html

Mitchell A, 2005, *The ESRI Guide to GIS Analysis, Vol. 2: Spatial Measurement & Statistics* (ESRI Press, Redlands, CA, USA)

Myint S, 2008, "An Exploration of Spatial Dispersion, Pattern and Association of Socio-Economic Functional Units in an Urban System" *Applied Geography* **28**(3) 168-188

Okabe A, Yamada I, 2001, "The K-Function Method on a Network and its Computational Implementation" *Geographical Analysis* **33**(3) 271-290

O´Sullivan A, 2007, *Urban Economics* (McGraw-Hill, Boston, USA)

Quah D, Simpson H, 2003, "Spatial Cluster Empirics", Monograph, Economics Department - Institute for Fiscal Studies, London School of Economics and Political Science, London, UK.

Ripley B, 1976, "The Second-Order Analysis of Stationary Point Process" *Journal of Applied Probability* **13**(1) 255-266

Ripley B, 1981, *Spatial Statistics* (John Wiley & Sons, Chichester, UK)

Rojas L, 2006, "El Acceso a los Servicios Bancarios en América Latina: Identificación de Obstáculos y Recomendaciones", WP 124, Center for Global Development**,** 1800 Massachusetts Avenue NW, Third Floor, Washington, USA.
http://www.cgdev.org/doc/experts/Serv_Bancarios_AL.pdf

Rot M, 2006, "Introducción al Análisis de Datos Mapeados o Algunas de las (Muchas) Cosas que Puedo Hacer Si Tengo Coordenadas", *Ecosistemas* **15**(3) 20-40

Ruiz C, 2004, "Los Desbancarizados: El Problema de los Mercados Financieros Segmentados" *Comercio Exterior* **54**(7) 566-574

Rysman M, Greenstein S, 2005, "Testing for Agglomeration and Dispersion" *Economics Letters* **86**(3) 405-411

Shearmur R, Alvergne, C, 2002, "Intrametropolitan Patterns of High-order Business Service Location: A Comparative Study of Seventeen Sectors in Ile-de-France" *Urban Studies* **39**(7) 1143-1163

Smith T, 2009, *Notebook on Spatial Data Analysis*, Web Book, http://www.seas.upenn.edu/~ese502/#notebook

Solo T, 2008, “Financial Exclusion in Latin America or the Social Costs of Not Banking the Urban Poor” *Environment and Urbanization* **20**(1) 47-66

Topçu Ü, 2001, “Spatial Distribution of Banks in Istanbul”, paper presented at the 41st Congress of the European Regional Science Association, Zagreb, Croatia, August 29 – September 1, 2001 http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa01/papers/full/222.pdf

Unal E, Chen, S, Waldorf B, 2007, “Spatial Accessibility of Health Care in Indiana”, Working Paper # 07-07, Department of Agricultural Economics, Purdue University, West Lafayette, IN, USA, http://ageconsearch.umn.edu/bitstream/7329/2/wp070007.pdf

WB, 2005, “Broadening Access to Financial Services Among the Urban Population: Mexico City´s Unbanked” Vol. 1: Main Report 32418-MX (World Bank, Washington, USA) http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&theSitePK=523679&entityID=000160016_20060111111546&searchMenuPK=64187283&theSitePK=523679

Yamada I, Thill JC, 2004, “Comparison of Planar and Network K-Functions in Traffic Accident Analysis” *Journal of Transport Geography* **12**(2) 149-58

Zhou Y, 1998, “Beyond Ethnic Enclaves: Location Strategies of Chinese Producer Service Firms in Los Angeles” *Economic Geography* **74**(3) 228-251

Figure 1. Study area: bank branches and main streets, 2009

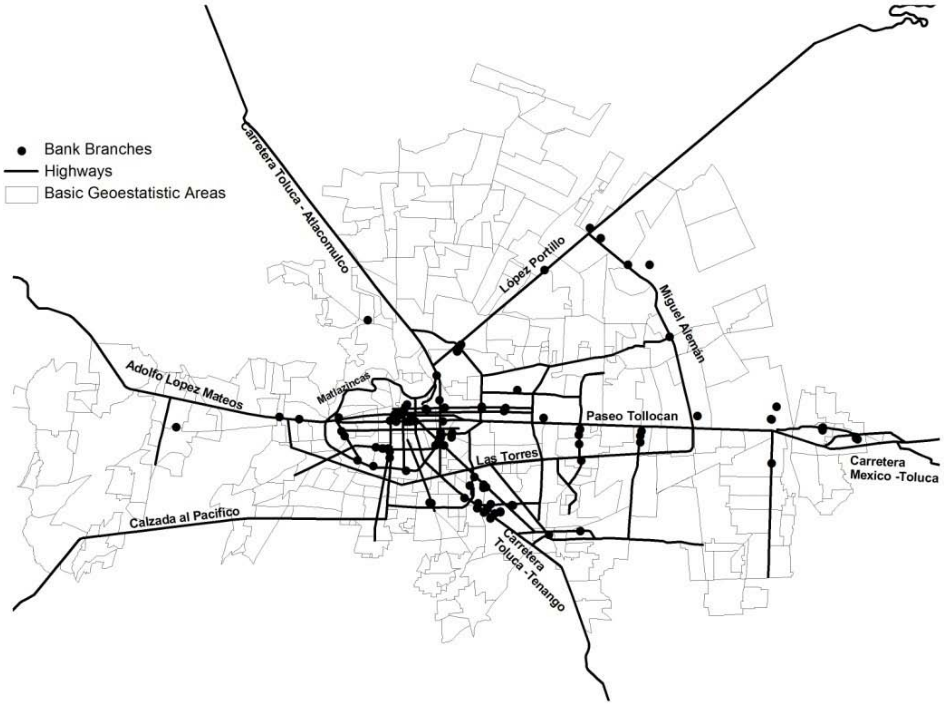


Figure 2. K-planar Function

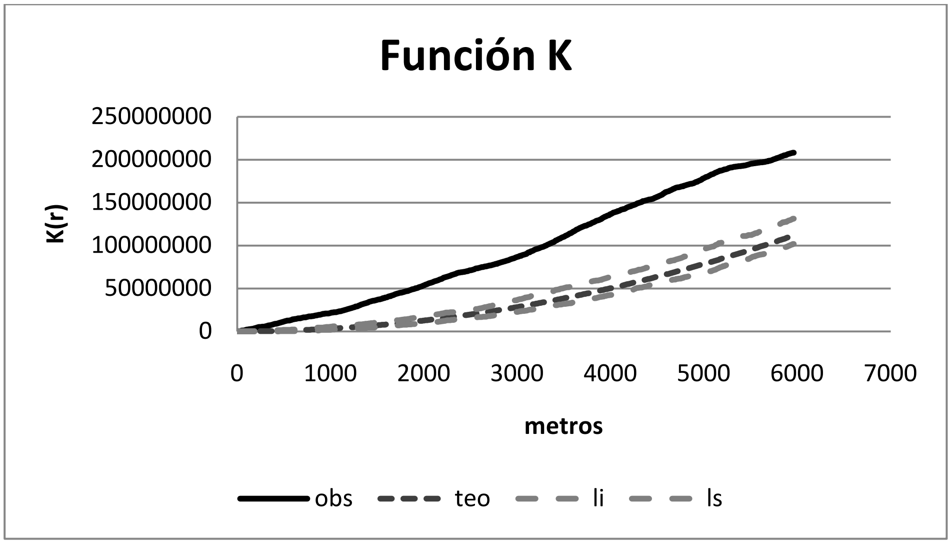


Figure 3. “M” Index

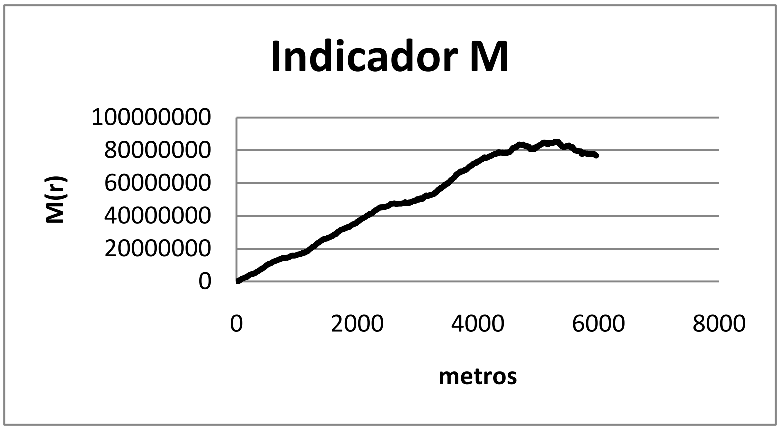


Figure 4. L’ Function

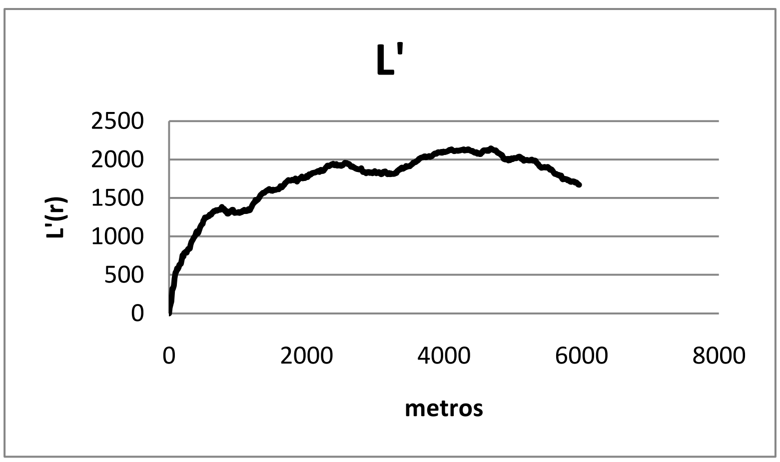
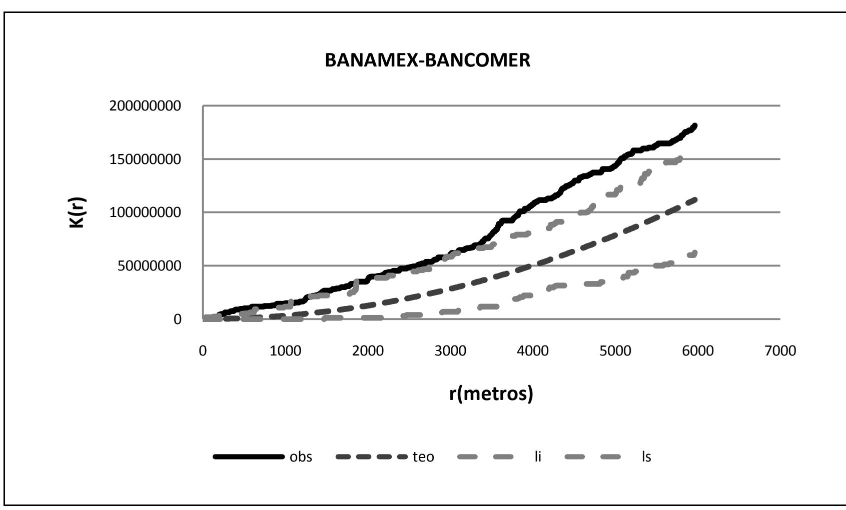
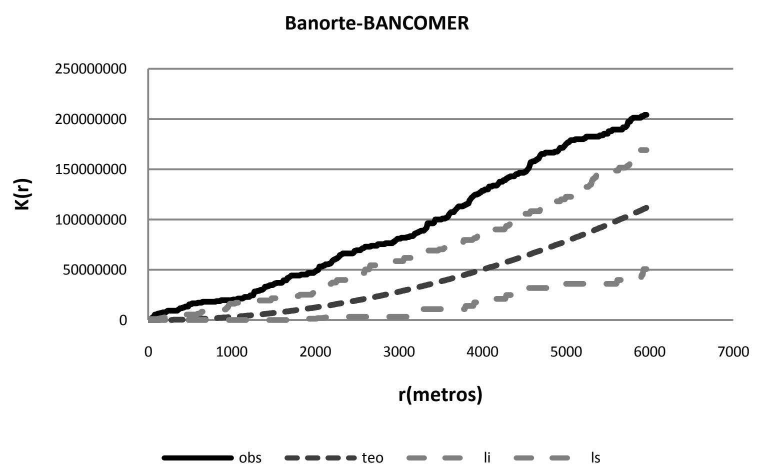
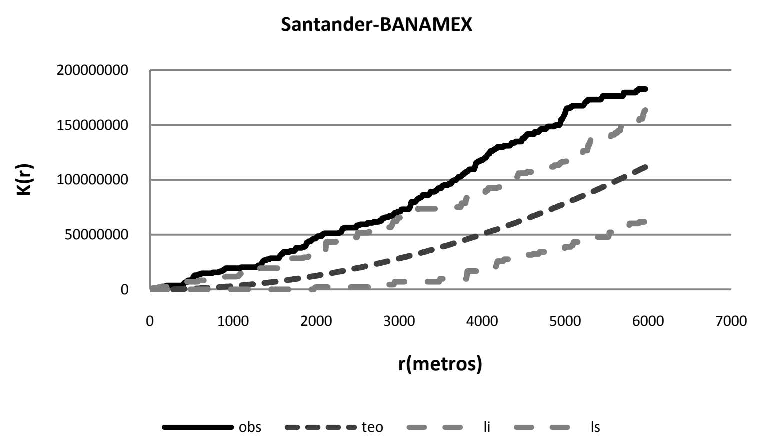
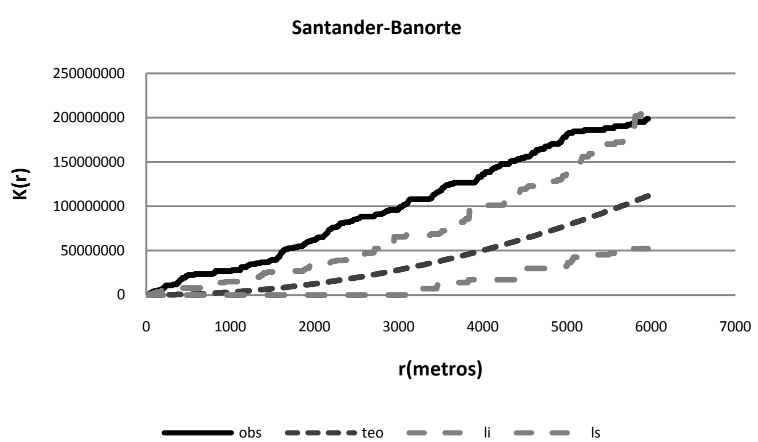
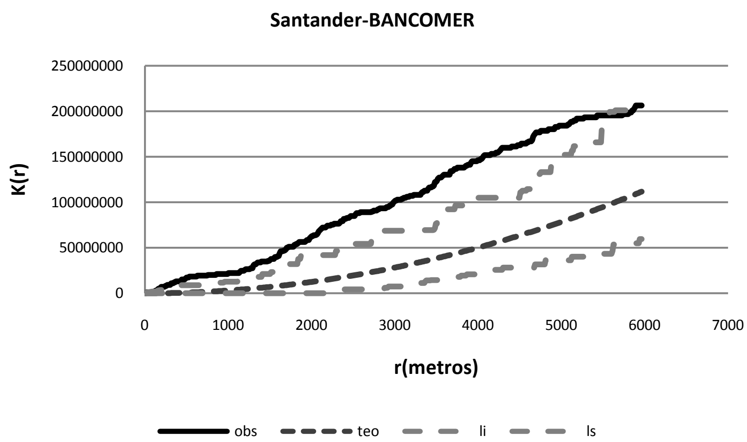
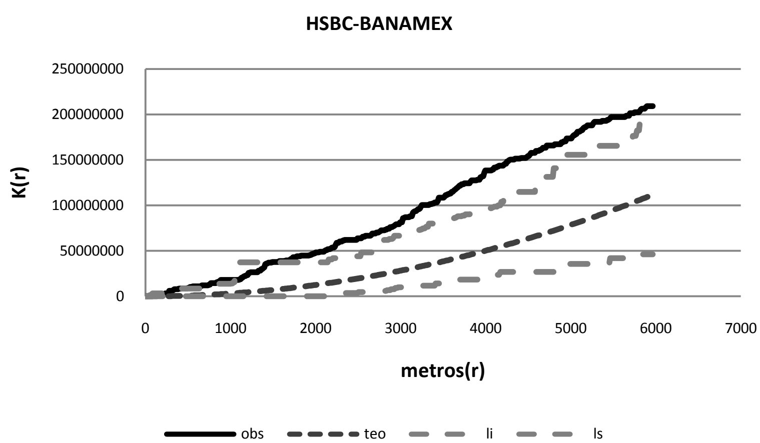
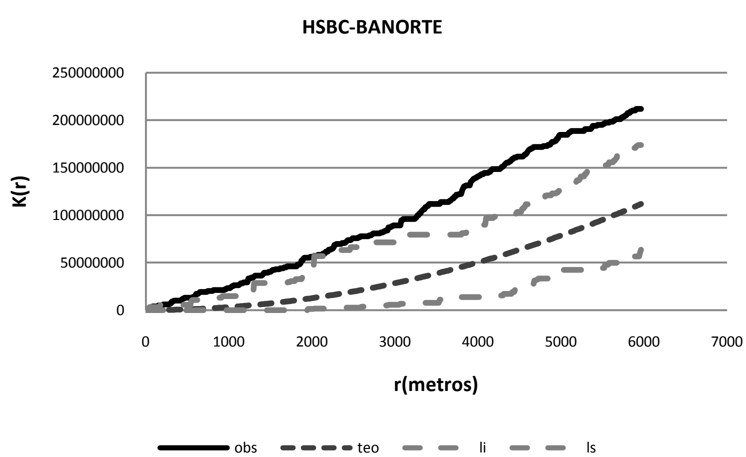


Figure 5. Dependency Pattern Between Banks in the TMA (Cross K-function)



- Oscillation between light agglomeration and independence at distances smaller than 3km.
- Agglomeration at the remaining distances.

	• Agglomeration at all distances.
	• Agglomeration, particularly at medium distances from 3.1km.
	• Agglomeration with tendency to independence from 5.8km.

	• Agglomeration with tendency to independence from 5.9km.
	• Agglomeration from 1.5km. There is no dependency at smaller distances.
	• Agglomeration, particularly from 2.0km.

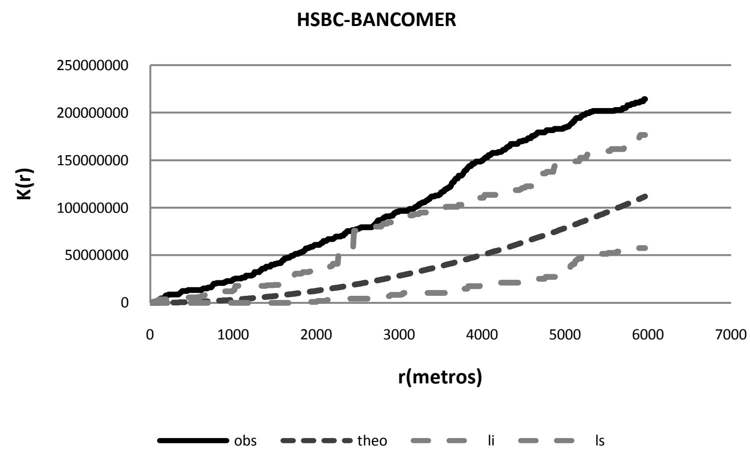
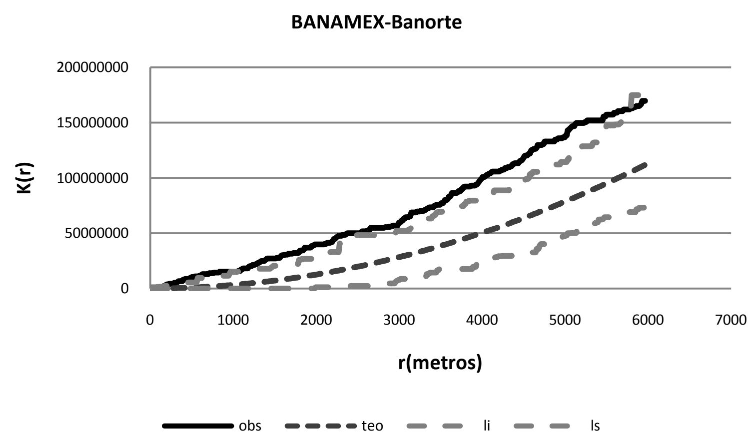
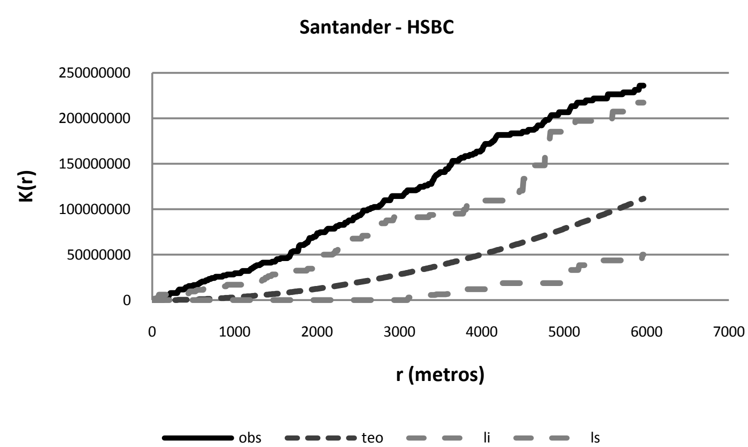
	• Agglomeration with tendency to independence from 5.8km.
	• Slight agglomeration with tendency to independence from 5.8km.
	• Agglomeration from 200 meters.

Table 1. Bank branches in the MAT, 2009

Bank	Number of Branches	%
Banamex	22	10.6
Banorte	19	17.8
BBVA	19	17.8
HSBC	15	14.0
Santander	14	13.1
Otros	18	16.7
Total	**107**	**100.0**

Table 2. Distance of maximum agglomeration (Km)

Banks	Banamex	BBVA	Banorte	HSBC	Santander
Banamex	--	4.5	5	3.9	5
BBVA	4.5	--	5	4.6	4.4
Banorte	5	5	--	4.4	3.6
HSBC	3.9	4.6	4.4	--	4.6
Santander	5	4.4	3.6	4.6	--

Table 3. Concentration Index at distance of maximum agglomeration

Banks	Banamex	BBVA	Banorte	HSBC	Santander
Banamex	--	1.37	1.22	1.47	1.41
BBVA	1.37	--	1.44	1.45	1.55
Banorte	1.22	1.45	--	1.61	1.72
HSBC	1.47	1.45	1.61	--	1.45
Santander	1.41	1.55	1.72	1.45	--
Sum	**5.47**	**5.82**	**6.00**	**5.98**	**6.13**
Average	**1.368**	**1.455**	**1.500**	**1.495**	**1.533**