SPATIAL ANALYSIS OF EMPLOYMENT AND POPULATION DENSITY: THE CASE OF THE AGGLOMERATION OF DIJON, 1999

by

Catherine Baumont, Cem Ertur, and Julie Le Gallo

REAL 03-T-6 March, 2003

SPATIAL ANALYSIS OF EMPLOYMENT AND POPULATION DENSITY: THE CASE OF THE AGGLOMERATION OF DIJON, 1999*

Catherine Baumont, Cem Ertur
LATEC UMR n°5118 and University of Burgundy, Pôle d'Economie et de Gestion, B.P. 26611, 21066 Dijon Cedex, France.
(e-mail: catherine.baumont@u-bourgogne.fr; certur@u-bourgogne.fr)

Julie Le Gallo
REAL, University of Illinois at Urbana Champaign, 220 Davenport Hall, 607 S. Mathews Av., Urbana, IL 61801-3671, USA.
(e-mail: jlegallo@uiuc.edu)

Abstract

The aim of this paper is to analyze the intra-urban spatial distributions of population and employment in the agglomeration of Dijon (regional capital of Burgundy, France). We study whether this agglomeration has followed the general tendency of job decentralization observed in most urban areas or whether it is still characterized by a monocentric pattern. In that purpose, we use a sample of 136 observations at the communal and at the IRIS (infra-urban statistical area) levels with 1999 census data and the employment database SIRENE (INSEE). First, we study the spatial pattern of total employment and employment density using exploratory spatial data analysis. Apart from the CBD, few IRIS are found to be statistically significant, a result contrasting with those found using standard methods of subcenter identification with employment cut-offs. Next, in order to examine the spatial distribution of residential population density, we estimate and compare different specifications: exponential negative, spline-exponential and multicentric density functions. Moreover, spatial autocorrelation, spatial heterogeneity and outliers are controlled for by using the appropriate maximum likelihood, generalized method of moments and Bayesian spatial econometric techniques. Our results highlight again the monocentric character of the agglomeration of Dijon.

Keywords: Bayesian spatial econometrics, exploratory spatial data analysis, outliers, population density, spatial autocorrelation, spatial heterogeneity, subcenters

JEL Classification: C11, C21, C52, R12, R14

---

* An earlier version of this paper was presented at the 49th North American Meeting of the RSAI, Puerto Rico. The authors would like to thank Françoise Bourdon for the help provided on the data and Don Lacombe, James LeSage, Daniel MacMillen and Antonio Paez for their helpful comments.

1. INTRODUCTION

Over the last decades, there has been considerable interest in the analysis of urban spatial structures. Indeed, urban growth has exhibited complex spatial patterns including both population spread and employment decentralization from the central city towards the suburbs. The validity of the monocentric model (Alonso, 1964; Muth, 1969) to explain urban patterns has therefore been questioned since employment decentralization has recently taken a polycentric form, with a number of employment subcenters influencing the spatial distribution of employment and population. The polycentric urban phenomenon has been extensively documented for many years. Most studies have been carried out on North American urban agglomerations: Chicago (McMillen and MacDonald, 1998a, b), Dallas-Fort Worth (Wadell and Shukla, 1993), Los Angeles (Gordon *et al.*, 1986; Heikkila *et al.*, 1989; Guiliano and Small, 1991; Small and Song, 1994; Sivitanidou, 1996), San Francisco (Cervero and Wu, 1997, 1998), Montréal (Coffey *et al.*, 1996)… This trend toward employment decentralization is not limited to North American areas (see, for example, Alperovitch and Deutsch, 1996 for Jerusalem; Chen, 1997 for Taipei; Wu, 1998 for Guangzhou; Gaschet, 2000 for Bordeaux, Boiteux-Orain and Guillain, 2002 for Ile-de-France).

Few studies have been carried out on middle-sized urban areas. Therefore, it is interesting to investigate whether these particular areas have experienced a similar trend toward employment decentralization or whether the monocentric model is still valid to explain employment and population spatial distributions. From an empirical point of view, studying polycentric rather than monocentric urban configurations raises a set of challenges (Anas *et al.*, 1998; Baumont and Le Gallo, 1999) which can be summarized as follows. How many economic subcenters can be identified apart from the traditional Central Business District (CBD)? What are their sizes and their boundaries? How do these multiple economic centers influence land values, population and employment distributions? …

In this paper, we are interested in this empirical challenge applied to the agglomeration of Dijon, which is the capital of Burgundy (France). The study covers the territory of the *Communauté de l'Agglomération Dijonnaise* (COMADI) in 1999, which is made up of 16 contiguous towns and has a total population of almost 250 000 inhabitants. We study the intra-urban employment and population distributions across the agglomeration using spatial statistic and econometric methods. First, we apply subcenter identification methods combined with spatial statistic techniques to study the characteristics of this agglomeration in 1999. Let us note that the identification of subcenters in this area using the standard methods suggested by Guiliano and Small (1991) has already been carried out in Baumont and Bourdon (2002) with the conclusion in favor of a multicentric urban configuration. In this paper, we underline the relevance of an alternative identification methodology, namely exploratory spatial data analysis (Anselin, 1995, 1996) in the spirit of Scott and Lloyd (1997). Indeed, these methods allow detecting both spatial autocorrelation, in the form of spatial clusters of population or employment, and spatial heterogeneity in the form of differentiated cluster patterns across space. Furthermore, they constitute an improvement over existing methodologies that don't assess the significance of their results and necessitate the definition of arbitrary cut-offs. Our results highlight a monocentric pattern based on the employment density distribution whereas some potential employment subcenters are detected using the total employment distribution. Second, we analyze whether the employment clusters detected in the previous step have a significant influence on the distribution of population using both monocentric and multicentric population density functions. It is well known that the presence of spatial autocorrelation yields inconsistent and inefficient OLS estimators but only a few studies have used spatial econometric techniques in the case of population density functions (Griffith, 1981; Anselin and Can, 1986; Stern, 1993; Griffith and Can, 1996; McMillen, 2002). Given the presence of spatial autocorrelation and spatial heterogeneity in the population density distribution of the COMADI, the appropriate spatial econometric methods are used (Anselin, 2001; LeSage, 1999, 2002). Our results show that the monocentric model explains well the population density provided that spatial autocorrelation, spatial heterogeneity and outliers are taken into account using maximum likelihood, generalized method of moments and Bayesian estimation techniques.

The paper is organized as follows. In the following section, we describe the data and the spatial weight matrices used in this study. In the third section, we study the spatial pattern of total employment and employment density using exploratory spatial data analysis. In the fourth section, we provide a spatial econometric analysis of monocentric and multicentric population density functions. Different specifications are estimated and compared: exponential negative, spline-exponential and multicentric density functions. The paper concludes with a summary of key findings.

2. DATA AND SPATIAL WEIGHT MATRIX

Our study focuses on a middle-sized urban area named COMADI (*Communauté de l'Agglomération Dijonnaise*), which is located in the French region of Burgundy. A community of agglomerations is a kind of large town council composed of several towns adjacent to a major city. More precisely, the COMADI is composed of 16 adjacent towns: the central city is Dijon, which is the capital of Burgundy, and 15 suburban towns around Dijon: Ahuy, Chenôve, Chevigny-Saint-Sauveur, Daix, Fontaine-lès-Dijon, Longvic, Marsannay-la-Côte, Neuilly-lès-Dijon, Ouges, Perrigny-lès-Dijon, Plombières-lès-Dijon, Quétigny, Saint-Apollinaire, Sennecey-lès-Dijon and Talant. In order to reduce edge effects present in spatial data analysis, we consider 22 additional suburban towns immediately surrounding the COMADI area. We will label these towns "urban fringe".

The spatial configuration of the COMADI area and its urban fringe is displayed in Map 1[1] and the main geographic and demographic characteristics of this area are shown in Table 1. From Table 1, it can be seen that the COMADI is a small area compared with the urban areas usually analyzed in urban studies (North American cities, major cities in Europe, Asia or in developing countries). With almost 250 000 inhabitants, the COMADI is the largest French community of agglomerations located between Paris, the largest French agglomeration with almost 10 million people, and Lyon, the second one with almost 2 million people. Moreover, the COMADI can be considered as the urban pole[2] structuring the metropolitan area of Dijon composed of 214 towns in 1999 for a total surface of 561 156 acres and 327 000 inhabitants.

[Map 1 about here]

[Table 1 about here]

We use data on employment and total population without double counting in the 1999 census. They are drawn from the RGP census and the SIRENE[3] data provided by INSEE (French national statistics institute) for all communes in the region of Burgundy. The employment data from the SIRENE files correspond to salaried employment. They relate primarily to private-sector employment. These data are collected at the communal level and at

---

[1] Maps are created using Arc-View©3.2 software on the basis of maps provided by the *Direction Régionale Bourgogne de l'INSEE.*

[2] The towns of the COMADI correspond to the towns within the *urban pole* of Dijon plus the town of Ahuy.

[3] RGP: *Recensement Général de la Population.* SIRENE: *Système Informatique pour le Répertoire des Entreprises et des Etablissements*.

a finer scale (IRIS-2000®) for places of more than 5000 inhabitants (cf. Appendix 1 for a description of IRIS-2000® level). Of the 16 towns that make up the COMADI, IRIS data are collected for nine of them: Chenôve (9 IRIS), Chevigny (5), Dijon (66), Fontaine (5), Longvic (4), Marsannay (3), Quétigny (5), Saint Apollinaire (4) and Talant (6) (see Map 2 for a picture of the IRIS scale). The other 7 towns of the COMADI and none of the 22 urban fringe communes have IRIS data. Finally, our sample comprises 136 spatial units for which different demographic, economic and geographic data are available.[4]

[Map 2 about here]

Spatial data analysis needs modeling the spatial interdependence between the observations by the mean of a spatial weight matrix W where each observation is connected to a set of neighboring observations according to a spatial pattern defined exogenously. The elements w_{ii} on the diagonal are set to zero whereas the elements w_{ij} indicate the way the unit i is spatially connected to the unit j. These elements are non-stochastic, non-negative and finite. In order to normalize the outside influence upon each unit, the weight matrix is standardized such that the elements of a row sum up to one.

Since various spatial weight matrices can be considered, the choice of some specific weight matrices depends on the geographical characteristics of the spatial area. The robustness of the results using different weight matrices must also be investigated. Given the specific geographical configuration of the spatial units in our sample, we choose not to consider inverse distance matrices. The analysis of the distance distribution between all pairs of spatial units reveals that the minimal distance cut-off for which each unit has at least one neighbor is above the first quartile and is very large (almost 5 km) compared to the size of the urban area (18 km by 16 km). If we consider such a distance cut-off, central IRIS will be connected to almost a quarter of the whole urban area, which is probably too much. This particular feature is a consequence of the important size heterogeneity since two spatial scales are used: the IRIS and the communal level. Another feature is that the IRIS situated in the centrally urbanized areas of some subdivided communes are very small (for example in Dijon, Chenôve or Saint-Apollinaire) whereas the communes that are not subdivided are much larger spatial units and are generally located in the periphery of the Dijon. Furthermore, residential

---

[4] A complete list of the observations and their codes is displayed in Table 3.

IRIS are generally much smaller than business IRIS.[5] For these reasons, distance-based spatial matrices with fixed cut-offs are not very relevant whereas simple binary contiguity and *k*-nearest neighbors matrices appear to be more appropriated. In a simple contiguity matrix, an element w_{ij} is one if the units *i* and *j* share a common border, and 0 otherwise. A *k*-nearest neighbors weight matrix is computed from the distance between the units' centroids and implies that each spatial unit is connected to the same number *k* of neighbors, wherever it is localized.

The general form of a *k*-nearest neighbors weight matrix *W*(*k*) is defined as following:

$$\begin{cases} w^*_{ij}(k) = 0 \ \text{ if } i = j, \forall k \\ w^*_{ij}(k) = 1 \ \text{ if } d_{ij} \le d_i(k) \\ w^*_{ij}(k) = 0 \ \text{ if } d_{ij} > d_i(k) \end{cases} \quad \text{and} \quad w_{ij}(k) = w^*_{ij}(k) / \sum_j w^*_{ij}(k) \tag{1}$$

where $w_{ij}(k)$ is an element of the standardized weight matrix and $d_i(k)$ is a critical cut-off distance defined for each unit *i*. More precisely, $d_i(k)$ is the k^{th} order smallest distance between unit *i* and all the other units such that each unit *i* has exactly *k* neighbors. Since the average number of neighbors in our sample is 5.76, we choose $k = 6$.

Finally, all our spatial data analysis and spatial econometric estimations have been carried out with the simple contiguity weight matrix and the 6 nearest-neighbors weight matrix.

3. IDENTIFICATION OF SUBCENTERS BASED ON EXPLORATORY SPATIAL DATA ANALYSIS

The identification of subcenters is often carried out using Giuliano and Small's (1991) methodology where a center is defined as a cluster of contiguous zones for which the total employment exceeds a predetermined cut-off and the employment density of each zone is higher than for all adjacent zones and is above a predetermined cut-off. The critical values chosen for these levels depend on the metropolitan area and may even vary over the metropolitan area if one observes strong variations in the employment or density employment distributions. For example, in the case of the Los Angeles region, Giuliano and Small (1991)

---

[5] The average surface of the residential IRIS is 209 acres whereas the average surface of business IRIS is 628 acres.

choose a level of 10 000 jobs for 2 counties (Los Angeles and Orange) and decrease the employment cut-off to 7 000 jobs for 3 other counties (Riverside, San Bernardino and Ventura) in order not to eliminate "definite peaks in employment density" (p. 167). In a recent study, Giuliano and Small (1999) use a lower level of 3000 jobs to analyze the polycentric pattern of Los Angeles in 1970. For other authors, the total employment level used for the cluster is replaced by a critical level for each zone included in the cluster, the value of the cut-off depending of the size of the urban area (Boiteaux-Orain and Guillain, 2002, choose a level of 7 000 jobs for Ile de France, Gaschet, 2000, chooses a level of 2 000 jobs for Bordeaux and Baumont and Bourdon, 2002, choose a level of 1 400 jobs for the COMADI). Therefore, this identification method depends heavily on the choice of arbitrary cut-offs. Different methods have been suggested to overcome these problems and to avoid the determination of arbitrary cut-offs. For example, Craig and Ng (2001) and McMillen (2001) use nonparametric techniques.

In this paper, we suggest an alternative method using exploratory spatial data analysis (ESDA) and define a potential employment subcenter as an area having significantly higher employment and employment density than neighboring sites. ESDA is a set of techniques aimed at describing spatial distributions in terms of spatial association patterns such as global spatial autocorrelation, local spatial autocorrelation and spatial heterogeneity. Since these patterns are associated to spatial weight matrices, where each unit is connected to a set of neighboring sites, then the way the employment or employment density of each unit is compared to those of its neighbors is directly taken into account. Moreover, using different spatial weight matrices, the notion of neighbors is large and is not limited to the notion of contiguity as in Giuliano and Small's method.[6] Finally, ESDA provides statistical tests aimed at indicating if the global and local spatial associations are significant.

The identification of subcenters in the COMADI is carried out applying ESDA both on total employment (Emp99) and gross employment density (Demp99) distributions in 1999. To reduce edge effects, we work with the sample “COMADI + Urban fringe” of 136 observations at IRIS and communal levels. Our results are compared to those obtained by Baumont and Bourdon (2002) using Giuliano and Small's method.

Since net employment densities are not available for our sample, the size heterogeneity problem previously mentioned affects the gross employment density distribution in two ways. First, at the communal level, urbanized areas in one town don't cover

---

[6] More precisely, Giuliano and Small (1991) consider that two zones are adjacent if they have at least 0.25 miles of common boundary.

the total surface of this town, especially for peripheral towns. Therefore, gross densities are smaller there than at the IRIS level defined for residential and business uses. Second, at the IRIS level, peripheral business IRIS are devoted to large industrial and commercial buildings and are often larger than central IRIS where offices and retail shops are mainly located. Therefore, we expect that gross employment densities decrease from central to peripheral spatial observations and that the communal scale measure reinforces the phenomena in peripheral areas.

First, if we consider global spatial autocorrelation, which is usually based on Moran's I statistic (Table 2)[7], it appears that total employment and employment density are positively spatially autocorrelated with at least 5% significance level for both variables.[8] These results indicate that similar values (high or low) of employment and employment density tend to be spatially clustered in the COMADI.

[Table 2 about here]

However, given our definition of a subcenter, we need to discriminate between a spatial clustering of high values and a spatial clustering of low values. Moreover, we need a measure of local spatial autocorrelation to compare each zone's total employment or employment density to that of its neighbors. In that purpose, Moran scatterplots and Local Indicators of Spatial Associations (Anselin, 1995, 1996) are used.[9] Moran scatterplots, which plot the spatial lag Wz against the original values z of a variable, aim at visualizing four types of local spatial association between an observation and its neighbors, each of them being localized in a quadrant of the scatterplot: quadrant HH refers to an observation with a high[10] value surrounded by observations with high values, quadrant LH refers to an observation with low value surrounded by observation with high values, etc. Quadrants HH and LL (resp. LH and HL) indicate positive (resp. negative) spatial autocorrelation indicating spatial clustering of *similar* (resp. *dissimilar*) values. In order to assess the significance of such spatial associations, Local Indicators of Spatial Associations (LISA) statistics are computed.

---

[7] All computations are done with SpaceStat 1.90 (Anselin, 1999)

[8] Inference is based on the permutation approach with 9999 permutations.

[9] The identification of subcenters can also be based on Getis and Ord's (1992) indicators of Gi(d) statistics. For applications, see Scott and Lloyd (1997) or Páez *et al.*(2001).

[10] High (resp. low) means above (resp. below) the mean.

Let us remind that the local version of Moran's *I* statistic for each observation i is written as:

$$I_i = \frac{(x_i - \mu)}{m_0} \sum_j w_{ij}(x_j - \mu) \qquad \text{with } m_0 = \sum_i (x_i - \mu)^2 / n \qquad (2)$$

where x_i is the observation in unit i; $n = 136$; μ is the mean of the observations across spatial units and where the summation over j is such that only neighboring values of j are included. A positive value for I_i indicates spatial clustering of similar values (high or low) whereas a negative value indicates spatial clustering of dissimilar values between a zone and its neighbors. Due to the presence of global spatial autocorrelation, inference must be based on the conditional permutation approach with 9 999 permutations. The *p*-values obtained for the local Moran's statistics are then pseudo-significance levels. Note that inference in this case is further complicated by the problem of multiple comparisons since the neighborhood sets of two spatial units contain common elements (Anselin, 1995; Ord and Getis, 1995; Le Gallo and Ertur, 2003). Therefore, the overall significance of 5% is not restricted enough and we also consider significance levels at 1% and 0.1%.

For each observation, the type and significance level of local spatial associations for both variables and both spatial weight matrices are displayed in Table 3. Moran significance maps for total employment (Map 3) and employment density (Map 4) combine the information in a Moran scatterplot and the significance of LISA by showing the IRIS with significant LISA and indicating by a color code the quadrants in the Moran scatterplot to which these IRIS belong. For both variables, significant HH spatial association may indicate an economic center covering several contiguous IRIS while the significant association of dissimilar values HL may indicate isolated economic centers.[11]

[Table 3 about here]

[Maps 3 and 4 about here]

For total employment, it appears that most of the observations are characterized by spatial positive association (51.5% in quadrant LL and 16.2% in quadrant HH) while the other IRIS are characterized by negative spatial association (8.1% in quadrant HL and 24.2% in quadrant LH). At the 5% pseudo-significance level, 6 IRIS are significantly HH, 3 being located in the center of the COMADI (D1, D4 and D6) and the other 3 in the south (Co8, D27

---

[11] Maps using the 6-nearest neighbors spatial weight matrix display a similar picture. They are not presented here due to space limitations but they are available from the authors upon request.

and Ma3). One eastern IRIS (Ch5), that is significantly HL, can be interpreted as an isolated employment pole. However, only 3 IRIS are significant at the 1% pseudo-significance level: central IRIS Monge (D1) in HH, southern IRIS Economic District of Marsannay (Ma3) in HH and the Economic District of Chevigny (Ch5) in HL. These results may indicate the existence of potential subcenters in the COMADI.

For employment density, positive spatial associations are even stronger (66.2% in quadrant LL and 14.6% in quadrant HH) than those of negative spatial associations (11% in quadrant LH and 8.8% in quadrant HL). All the IRIS having significantly local spatial association of HH type are centrally located. Among them 9 IRIS are significantly HH at the 1% pseudo-significance level: D1, D2, D5, D6, D8, D11, D40, D43 and D64. Finally, a quite different picture is obtained when considering employment density instead of total employment, possibly reflecting the bias introduced by gross employment density measure and pointing out the monocentric character of the agglomeration.

Concerning the identification of employment subcenters in the COMADI, these results can be interpreted as follows. First, both total employment and employment density distributions are characterized by significant local positive spatial autocorrelation. The local clusters of high employment are primarily located in the inner center of the agglomeration. If total employment distribution is considered, it appears that several southern IRIS located in the South of the agglomeration can be considered as employment subcenters (i.e. the economic district of Marsannay-la-Côte and its neighborhood). Moreover, one IRIS located in the East of the agglomeration has significantly more employments than its neighbors and may be considered as an isolated economic center. However, this latter result should be considered with caution since this IRIS is surrounded by a lot of open areas. When employment density is considered, only the central IRIS are found to be statistically significant. Finally, it appears that ESDA mainly detects the central business district of the COMADI and highlights the monocentric character of the agglomeration of Dijon.

It is worthwhile to compare these results to those obtained by Baumont and Bourdon (2002) where standard subcenter identification methods have been used. Following Giuliano and Small (1991), Baumont and Bourdon define a center as a zone or a set of contiguous zones where total employment of each zone is greater than a given level $\overline{E}$ and greater than total employment in surrounding zones and where employment density of each zone is greater than a given level $\overline{D}$ and greater than the density in surrounding zones. For the 114 contiguous zones composing the COMADI, the authors consider the employment level that allows taking into account a sufficient number of IRIS to include more than 50% of the total employment of the COMADI. This level is $\overline{E}$ = 1 400 jobs. The density employment level is $\overline{D}$ = 10 jobs per acre. Eleven IRIS containing more than 1 400 jobs are identified but among them, only 5 IRIS have the sufficient level for employment density. More precisely, we note that many peripheral IRIS have many jobs although they have low employment densities. On the contrary, central IRIS have high employment densities. This is a traditional feature due to the heterogeneity of the spatial scale that we have mentioned. Therefore, Baumont and Bourdon (2002) prefer not considering the employment density as a relevant indicator to define economic center and name "*Employment poles*" the zone or the set of contiguous zones that have more than 1 400 jobs. Five employment poles have been identified (Table 4). They are composed of central IRIS, like the CBD (9 644 jobs), or of IRIS from several contiguous different towns, like the "multi-towns" Poles in the South (11 540 jobs) and in the North (9 634 jobs). Others are single IRIS like the "isolated" Poles Quétigny (4 014 jobs) and Chevigny (1 421 jobs). Finally, these authors conclude that the COMADI exhibits a multicentric economic pattern for employment with the traditional CBD and four economic subcenters (South, North, Quétigny and Chevigny) as shown in Map 5. This result contrasts with that found with ESDA.

[Table 4 about here]
[Map 5 about here]

It is interesting to further analyze the spatial association characteristics of these economic centers with the exploratory spatial analysis that we have carried out above. Although a different sample has been used, we consider that we can compare the two studies for the following reasons. First, our sample is composed of the IRIS belonging to the COMADI and its urban fringe. If we analyze the employment composition of the towns belonging to the urban fringe (Table 1), only 3.54% of jobs are added to the analysis. None of these towns have more than 1 400 jobs and more than 10 jobs per acre. Second, the main statistical characteristics (mean and quartiles) of the total employment and employment density distributions are quite similar in the two samples. Local spatial association indicators associated to the employment poles are displayed in Table 4. We can easily note that the IRIS belonging to the CBD have significant LISA statistics. On the contrary, the peripheral employment poles are not significant according to ESDA except the isolated pole "Chevigny", which has to be considered with caution. Therefore, the multicentric pattern of employment highlighted by standard methods of employment subcenter identification is not fully confirmed by ESDA.

This contradictory result is indeed mainly explained by the statistical tests included in exploratory spatial data analysis. If only the Moran Scatterplot results are considered (Table 3), then all the subcenters identified in Baumont and Bourdon's study have a HH or HL local spatial association type. They can be considered as potential subcenters if no further analysis of significance of these clusters is carried out. However, since statistical test results do not confirm that these local spatial associations are significant, we cannot consider the "isolated" Pole Quétigny and the "multi-towns" Pole North as subcenters in the COMADI. Moreover, this evidence is consistent with statistical tests using the *6*-nearest-neigbors weight matrix.

Finally, ESDA, as a subcenter identification method, differs from Giuliano and Small's method in three ways. First, it doesn't need a priori cut-offs. Second, it allows formalizing the contiguity between spatial observations and also many other forms of proximity through the spatial weight matrices. Third, it assesses the significance of the results produced. In this framework, the detection of monocentric or multicentric employment patterns no more depends on exogenous cut-offs and/or on the personal judgment of the researcher. However, let’s note that they keep on depending, as other identification subcenters methods, on the employment indicator used: total employment or employment density for example.

Therefore, our study highlights the relevance of ESDA in this framework. However, if we consider that an economic center influences residential location choices, we must also estimate the effects of all subcenters on population density. In that purpose, different population density functions, with different functional forms and including one or more potential economic centers, are considered in the following section.

4. SPATIAL ECONOMETRIC ANALYSIS OF POPULATION DENSITY FUNCTIONS

The analysis of urban structures is usually conducted using population residential density functions including the distance from the CBD as an explanatory factor. The negative exponential density function defined by Clark (1951) is the most used theoretical specification. It has been largely improved in order to better capture the irregularities of the population density distribution in real urban areas.[12] For example, Anderson (1982, 1985)

---

[12] See Mills and Tan (1980) and McDonald (1989) for surveys of results and methodology.

suggests the use of cubic spline specifications when population densities do not homogeneously decrease as the distance from the CBD increases. Brueckner (1986) estimates distance-oriented density functions, with an unknown number of possible regimes, using switching regressions. Alperovich and Deutsch (2002) find evidence of two distinct regimes in the urban area of Tel-Aviv. In fact, all these studies take into account in different ways spatial heterogeneity: the estimated coefficients are different depending on their distance from the CBD or on the spatial regime they belong. Moreover, other economic centers than the CBD may influence the spatial population distribution on the urban area. In this case, the distance from several potential economic subcenters are added as explanatory variables. If more than one of the associated coefficients is statistically significant, then the urban pattern is considered as multicentric and it is considered as monocentric otherwise.

In this section, as for total employment and employment density, we first carry out an ESDA on population density in order to identify possible patterns of spatial heterogeneity and/or spatial autocorrelation. We then estimate population residential density functions for the COMADI using different specifications including the distance from the CBD and from the potential economic subcenters detected in the previous section and by Baumont and Bourdon (2002). The presence of spatial autocorrelation is systematically tested because if it is ignored, the OLS estimators are at best inefficient and statistical inference is biased, at worst they are biased and inconsistent. Wrong conclusions can therefore be drawn out of the results. However, except in some isolated studies (Griffith, 1981; Anselin and Can, 1986; Stern, 1993; Griffith and Can, 1996; McMillen, 2002), spatial autocorrelation is not taken into account in the estimation of population density functions.

The population density is named Dpop99 and is defined as the population per acre. It is measured for each observation of the sample of 136 observations of the COMADI and its urban fringe. We use the same spatial weight matrices as in the ESDA on employment: the contiguity weight matrix and the 6 nearest neighbors weight matrix.

Concerning the detection of global spatial autocorrelation, the value of the Moran's I statistic is positive and significant with $p = 0.001$ (Table 5). This result suggests that the IRIS with relatively high values (resp. low) of population density are surrounded by IRIS with relatively high values (resp. low) of population density.

[Table 5 about here]

If we look at the spatial distribution of LISA statistics (Table 6 and Map 6), two additional features appear. First, more than 82% of the spatial observations exhibit a positive spatial autocorrelation pattern (HH and LL types). In the case of the contiguity matrix, 51 IRIS are of type HH and 61 IRIS are of type LL. Only 8 IRIS are of type HL and 16 IRIS of type LH. Second, using a significance level of 5%, we detect clusters of high values in the central part of the COMADI and cluster of low values in the peripheral towns of the agglomeration. Note also that two clusters of high population density values are detected in the district of "Fontaine d'Ouche" (a neighborhood in Dijon) and in the town of "Chenôve". In conclusion, this spatial autocorrelation pattern is not in contradiction with the standard theoretical distribution of residential density associated with the monocentric assumption, except for the two local peaks located at the western part of the city of Dijon that may reflect a form a spatial heterogeneity. We investigate these issues further with population density functions.

[Map 6 about here]
[Table 6 about here]

Let us take as a starting point the following negative exponential function:

$$D(u) = D(0)e^{-\gamma u+\varepsilon} \tag{3}$$

where $D(u)$ is the population density at distance u from the CBD; $D(0)$ is the population density at the CBD; γ is the density gradient and measures the proportional rate at which population density falls with distance; ε is the error term with the usual properties. Note that this particular form can be derived from the monocentric model with several restrictive assumptions, i.e. constant returns Cobb-Douglas production function for housing, consumer with identical tastes and incomes and unit price elasticity of demand of housing. The function is then estimated by taking logs on both sides:

$$\ln D(u) = \ln D(0) - \gamma u + \varepsilon \tag{4}$$

All distances are measured in straight-line km from the centroid of the IRIS Monge (D1). The results of the estimation by OLS of this model are given in the first column of Table 7. The density gradient is strongly significant and negative, $\hat{\gamma} = -0.535$, which confirms the decay of population density from the center. Note also that the model fit is quite good: R^2-adjusted $\simeq 60\%$. However, using the first order contiguity spatial weight matrix[13], the Moran's *I* test adapted to regression residuals (Cliff and Ord, 1981) indicates the presence of spatial autocorrelation. It is well known that OLS estimators are in this case at best inefficient and at worst biased and inconsistent. A classical "specific to general" specification search approach[14] outlined in Anselin and Rey (1991) or Anselin and Florax (1995a) in the context of spatial econometric modeling can then be applied to discriminate between the two forms of spatial dependence – spatial autocorrelation of errors or endogenous spatial lag. The performance of such an approach is experimentally investigated in Florax and Folmer (1992). Furthermore, Florax *et al.* (2003) show by means of Monte Carlo simulation that this classical approach outperforms Hendry's "general to specific" approach. Here, the Lagrange Multiplier tests, LMERR, LMLAG and their robust versions (Anselin, 1988; Anselin *et al.*, 1996) indicate the presence of spatial error autocorrelation rather than a spatial lag.

[Table 7 about here]

Before turning to the spatial error exponential density function, we also investigate the presence of some form of spatial heterogeneity given the presence of local population density peaks detected in the ESDA. In that purpose, we use the spline-exponential function as suggested by Alperovitch (1995). This specification is an extension of the negative exponential function and is adapted when the population density does not decrease monotonously with distance from the CBD. One or more "knots" are specified defining distance intervals. The function is then exponential between knots and the gradient of the function is allowed to vary along different distance intervals. Note that cubic-spline functions are often used in empirical analyses. However, we prefer the use of a spline-exponential since it avoids the use of high-order distance terms and therefore limits the amount of multicollinearity (Alperovitch, 1995). We define one knot located at 4 km of the CBD because it is approximately the distance at which the three local population density peaks are situated. Moreover, this distance corresponds more or less to the boundaries of Dijon where several high density housing projects were realized in the seventies. Until this distance,

---

[13] We used also the 6-nearest neighbors weight matrix, but the results are not presented here due to space limitations. All the estimation results using this latter matrix are available from the authors upon request.

[14] Nevertheless it must be stressed that this classical approach has three main drawbacks: the significance levels of the sequence of tests are unknown, every test is conditional on arbitrary assumptions, it does not always lead to the "best model". Some authors prefer to pre-whiten or filter the variables to get rid of spatial autocorrelation (e.g., Getis and Griffith, 2002 among others). Conley (1999) proposes yet an interesting alternative approach based on nonparametric estimation of covariance matrices yielding standard error estimates for coefficients that are robust versus spatial autocorrelation and heteroscedasticity. His approach is the spatial analog of that followed in time-series by e.g., Newey and West (1987) or Andrews (1991).

population densities are measured at the IRIS scale. On the contrary, beyond this distance are located peripheral residential towns often characterized by low population density levels and measured at the communal level.

The spline-exponential function can then be written as:

$$\ln D(u) = \ln D(0) - \gamma u + \beta u_a + \varepsilon \qquad (5)$$

where $D(u)$, $D(0)$, γ and u are defined as before; β is the parameter describing the change of the gradient of the function occurring within the distance interval defined by the 4 km knot. Finally, u_a is defined by:

$$u_a = \begin{cases} 0 & \text{if } u \leq a \\ u - a & \text{if } u > a \end{cases} \qquad \text{with } a = 4 \text{ kms} \qquad (6)$$

The results of the estimation by OLS of this model are given in the first column of Table 8. While the gradient is still significant and negative, the coefficient β is not significant. Therefore it seems that the gradient does not change after 4 kms of the CBD. The model fit is approximately the same as in the simple exponential density function and the spline-exponential does not perform better in terms of information criteria. However, this model is misspecified since the spatial autocorrelation tests indicate the presence of spatially autocorrelated errors and therefore statistical inference based on it is not reliable.

[Table 8 about here]

Both models must therefore be modified to integrate spatial autocorrelation explicitly in the form of a spatial error model in order to achieve reliable inference. In equation (4) and (5), the following error structure is added:

$$\varepsilon = \lambda W \varepsilon + u \qquad u \sim N(0, \sigma_u^2 I) \qquad (7)$$

The estimation results for the exponential density function by maximum likelihood (ML) and iterated general method of moments (GMM) are displayed in the second and third columns of table 7. Using ML, it appears that all coefficients are strongly significant. The density gradient is slightly lower than in the model estimated by OLS, population density declines by 51.9% for each km from the CBD compared to 53.5% obtained previously. It is as well important to note that a significant positive spatial autocorrelation of the errors is found

($\hat{\lambda} = 0.439$). Furthermore, the LMLAG* test does not reject the null hypothesis of the absence of an additional autoregressive lag variable in the spatial error model. Estimation of this model by GMM leads to practically the same results on the parameters of interest.

The estimation results for the spline exponential density function by ML and GMM are given in the second and third columns of Table 8. Compared to the model estimated by OLS, it appears that the estimated gradient is higher and that β is positive and now significant, highlighting a change of the gradient after 4 kms. For example, using ML, others things being equal, the value of the constant term and of the absolute value of the density gradient are lower for units located farther than 4 kms from the CBD than for areas located closer than 4 kms from the CBD. Population density is estimated to decline by 86.7% for each km from the CBD for areas located closer than 4 kms from the CBD, while it is estimated to decline by 38.8% for each km from the CBD for areas located farther then 4 kms from the CBD. A significant positive spatial autocorrelation of the errors is also found ($\hat{\lambda} = 0.509$) while the LMLAG* test does not reject the null hypothesis of the absence of an additional autoregressive lag variable in the spatial error model. The spline-exponential model seems then appropriate to capture the spatial heterogeneity in middle-sized agglomerations where the downtown still dominates.

Finally, in order to deal with potential outliers, which could exert a substantial impact on inference regarding the density gradient in the context of our relatively small sample of 136 observations together with heteroscedasticity, we estimate a Bayesian heteroscedasticity robust model using the method described in LeSage (1999, 2002). This model allows the disturbances to take the form $\varepsilon \sim N(0, \sigma^2 V)$, where $V = diag(v_1, v_2, ..., v_n)$ and is estimated using MCMC methods. A prior distribution is assigned to the v_i terms taking the form of a set of n independent, identically distributed, $\chi^2(r)/r$ distributions, where r represents the single parameter of the χ^2 distribution. This allows us to estimate the additional n variance scaling parameters v_i by adding only a single parameter r to the model (see Geweke, 1993).

The χ^2 prior assigned to the v_i terms can be motivated by considering that the prior mean equals unity and the prior variance is $2/r$. This implies that as our prior assignment of a value for r becomes very large, the terms v_i will all approach unity, resulting in $V = I_n$, the traditional assumption of constant variance across space. On the other hand, assigning small prior values to r leads to a skewed distribution permitting large values of v_i that deviate greatly from the prior mean of unity. The role of these large v_i values is to accommodate outliers or observations containing large variances by down-weighting these observations. In the context of spatial modeling, outliers arise due to "enclave effects", where a particular observation exhibits divergent behavior from nearby observations. Geweke (1993) shows that this approach to modeling the disturbances is equivalent to a model that assumes a Student-*t* distribution for the errors. This type of distribution has frequently been used to deal with sample data containing outliers, (e.g., Lange *et al.* 1989). In practice, one can either assign an informative prior for the parameter r based on the exponential distribution centered on a small value, or treat this as a hyper-parameter in the model, set to a small value, say 4 to 7. Our estimates presented in the last columns of Table 7 and Table 8 are based on $r = 4$.[15]

Concerning the negative exponential density function, they are close to those obtained earlier, only three areas exhibit posterior means for the variance scalars v_i greater than 4 (Figure 1). These potential outliers are D62 (a peripheral new commercial district), D66 (a large green area) and Lo4 (an industrial district surrounded with highly populated districts), but they are not influencing much the results. For the spline exponential density function, β is no more significant. Even though only two areas D62 and Lo4 are potential outliers, down-weighting these observations results in a noticeable change in the β estimate relative to those based on ML where constant variance was assumed. The density gradient can no more be considered as different before and after a distance of 4 kms from the CBD.

[Figure 1 about here]

A comparison between ML and Bayesian heteroscedastic β estimates is presented in Figure 2 where the simulated normal distribution for β from the ML estimation and the posterior distribution for β from the Bayesian estimation are plotted. The posterior distribution for β appears to be skewed to the left ($\text{mode} = 0.176$, $\text{median} = 0.202$, $\text{mean} = 0.208$) compared to the simulated normal distribution, most likely because of the outliers or non-constant variances, as an illustration of our point.

[Figure 2 about here]

All these results indicate that the monocentric model explains well the spatial pattern of population density, provided that spatial autocorrelation, spatial heterogeneity and outliers

---

[15] All computations are carried on by means of the Matlab Spatial Econometrics Toolbox developed by James LeSage (http://www.spatial.econometrics.com).

are taken into account. However, it is interesting to study how multicentric density functions perform compared to the monocentric model. In that purpose, we consider the following model:

$$\ln D(u) = \ln D(0) - \gamma u + \delta D_{SUB}^{-1} + \varepsilon \tag{8}$$

where $D(u)$, $D(0)$, γ are defined as before; D_{SUB}^{-1} is the distance from the nearest subcenter site; δ is the associated parameters to be estimated. Since only the distance from the nearest subcenter is included, this specification expresses the local influence of subcenters, compared to the CBD that we assume to have a global influence. Note furthermore that the distance from the nearest subcenter is expressed in inverse form because this specification allows the effect of distance from the nearest subcenter to decline rapidly with distance. Furthermore, it limits the amount of multicollinearity (McMillen and McDonald, 1998a, b). Indeed, the maximum condition number found in the following regressions is 7.388. In this specification, a positive coefficient δ indicates that population density falls with distance from nearest subcenter.

We consider three sets of regressions. In the first one, D_{SUB}^{-1} only contains the inverse minimum distance from the 2 IRIS that were significant in ESDA on total employment: Ch5 and Ma3 with the associated parameter δ_1. In the second one, it includes the inverse minimum distance from the subcenters that were detected in Baumont and Bourdon (2002): Ma3[16], Qu5 and D63[17] with the associated parameter δ_2. In the last one, D_{SUB}^{-1} contains the inverse minimum distance from all four potential subcenters with the associated parameter δ_3.

The estimation results by OLS of these three specifications are displayed in Table 9. It appears that the CBD gradient is always strongly significant and negative, highlighting the importance of the CBD, even when other subcenters are included. However, the distances from the various subcenters included are either not significant (δ_1 and δ_3 estimates), or are weakly significant but do not have the expected sign (δ_2). The model fits are comparable to monocentric models estimated by OLS (R^2-adjusted $\simeq 60\%$) and these specifications do not perform much better in terms of information criteria. However, all three specifications are misspecified due to the presence of spatial autocorrelation, as indicated by the Moran's *I* test. The inclusion of distance to the nearest subcenter did not remove spatial autocorrelation. The

---

[16] Its centroid is taken as the centroid of the southern employment pole.
[17] Its centroid is taken as the centroid of the northern employment pole.

classical “specific to general” specification search approach again points towards the spatial error model rather than the spatial lag model.

[Table 9 about here]

The estimation results of these spatial error models by ML and GMM are also displayed in Table 9. It appears that whatever the multicentric specification estimated, the CBD gradient is again strongly significant and negative and its estimated value remains close to that obtained in monocentric specifications. A strongly significant positive spatial autocorrelation of the errors is also found for each specification. However, the δ_2 and δ_3 estimates are now significant, but negative. Each additional km from the nearest employment subcenter increases population density by respectively 115.7% and 101%. These findings are furthermore confirmed by the Bayesian heteroscedasticity and outliers robust estimations (respectively 105% and 108.2%), where again D62, D66 and Lo4 exhibit posterior means for the variance scalars v_i greater than 4 and appear thus as potential outliers for all the various specifications. Figure 2 displays the posterior means for the v_i estimates for the third specification (the picture is quite similar for the other specifications). However, these potential outliers do not seem to influence much the results.

[Figure 3 about here]

These results are in contradiction with the standard conclusions of residential choice models since it indicates that population densities increase with the distance from the nearest employment pole. Therefore, such results question the nature of the attractiveness of these employment poles for residential choice. One possible explanation is the small size of the COMADI where locations farther away from the potential subcenters are in fact located closer to the CBD, implying an increase in the population density. Note that this result has also been found for Chicago by McMillen (2003), which provide two explanations that are also relevant to our case. First, using gross densities rather than net densities implies that densities near subcenters are low since by definition most of the land area in subcenters is in non-residential use. Second, these subcenters may not be yet large enough to influence the distribution of population density. Furthermore, in our case, subcenters are mainly located at the border of the COMADI and their spatial influence is indeed limited within that strongly urbanized area. Rather, subcenters may first have an effect on land values or employment density within the

COMADI or may have an effect on population density in outlying suburbs where a lot of vacant land is available.

Finally, these results show few evidence of a multicentric urban pattern in the COMADI and are in conformity with those found in the ESDA on total employment and employment density. Rather, the monocentric model appears as the most appropriate specification.

5. CONCLUSION

In this paper, we have analyzed the intra-urban spatial distributions of population and employment in the agglomeration of Dijon (regional capital of Burgundy, France). Our aim was to study whether this agglomeration has followed the general tendency of job decentralization observed in most urban areas or whether it is still characterized by a monocentric pattern. In that purpose, a sample of 136 observations at the communal and at the IRIS (infra-urban statistical area defined by INSEE) levels with 1999 census data and the employment database SIRENE (INSEE) was used.

First, the spatial pattern of total employment and employment density using Exploratory Spatial Data Analysis (ESDA) has been studied. Contrary to standard methods of employment subcenter identification, ESDA does not require the determination *a priori* of arbitrary cut-offs and it allows assessing the significance of clusters of high employment. The application of these procedures to total employment and employment density shows that, apart from the CBD, few IRIS are found to be statistically significant. These results contrast with those found using standard methods, where potential employment subcenters were detected in the North, South and East of the COMADI.

Second, in order to examine the spatial distribution of residential population density, different specifications have been estimated and compared. On the one hand, an exponential negative and a spline-exponential density function have been considered. The latter has been estimated due to the presence of local clusters of high population density located in the western part of the COMADI. The exponential negative density function has been found to perform quite well. On the other hand, multicentric density functions including various subcenters yield results with distances from the subcenters that are significantly negative, similar to those obtained by McMillen (2003) for Chicago, indicating that proximity to these subcenters *reduces* population density. In each case, spatial autocorrelation, spatial heterogeneity and outliers are controlled for by using maximum likelihood, generalized method of moments and Bayesian estimation techniques. Applying the classical “specific to general” specification search approach, the spatial error model appeared as the most appropriate specification in each case. However, further investigation of the spatial lag model, the spatial Durbin model and spatial autoregressive local estimation (Pace and LeSage, 2002) could also be interesting and is left for further research.

Finally, all the results highlight the monocentric character of the agglomeration of Dijon. Although some job decentralization, following urban policies, has taken place in the last years, there are no clusters of employment having a significant impact on the distribution of population density. These findings could be extended by considering the distribution of land and housing values within the COMADI and/or by studying a larger area surrounding the COMADI in order to analyze the spatial influence of subcenters on the distribution of population density in peri-urban areas.

References

Akaike H. (1974) A new look at the statistical model identification, *IEEE Transactions on Automatic Control,* AC-19, 716-723.

Alonso W. (1964) *Location and Land Use: Toward a General Theory of Land Rent*, Harvard University Press, Cambridge Massachusetts.

Alperovich G. (1995) The effectiveness of spline urban density functions: an empirical investigation, *Urban Studies*, 32, 1537-1548.

Alperovich G., Deutsch J. (1996) Urban structure with two coexisting and almost completely segregated populations: the case of East and West Jerusalem, *Regional Science and Urban Economics*, 26, 171-187.

Alperovich G., Deutsch J. (2002) An application of a switching regimes regression to the study of urban structure, *Papers in Regional Science*, 81, 83-98.

Anas A., Arnott R., Small K.A. (1998) Urban spatial structure, *Journal of Economic Perspectives*, 36, 1426-1464.

Anderson J. (1982) Cubic spline urban density functions, *Journal of Urban Economics*, 12, 55-67.

Anderson J. (1985) The changing structure of a city: temporal changes in cubic spline urban density patterns, *Journal of Regional Science*, 25, 413-425.

Andrews D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, 59, 817-859.

Anselin L. (1988) *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.

Anselin L. (1995) Local indicators of spatial association-LISA, *Geographical Analysis*, 27, 93-115.

Anselin L. (1996) The Moran scatterplot as an ESDA tool to assess local instability in spatial association, in: Fisher M., Scholten H.J., Unwin D. (eds.), *Spatial Analytical Perspectives on GIS*, Taylor & Francis, London.

Anselin L. (1999) SpaceStat, a software package for the analysis of spatial data, Version 1.90, Ann Arbor, BioMedware.

Anselin L. (2001) Spatial econometrics, in: Baltagi B. (ed.), *Companion to Econometrics*, Basil Blackwell, Oxford.

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

Anselin L., Can A. (1986) Model comparison and model validation issues in empirical work on urban density functions, *Geographical Analysis*, 18, 179-197.

Anselin L., Florax R. (1995) Introduction, in: Anselin L., Florax R. (eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.

Anselin L., Florax R. (1995) Small sample properties of tests for spatial dependence in regression models, in Anselin L., Florax R. (eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.

Anselin L., Rey S. (1991) Properties of tests for spatial dependence in linear regression models. *Geographical Analysis*, 23, 110-131.

Baumont C., Bourdon F. (2002) Centres secondaires et recomposition économique des espaces urbain, le cas de la Communauté de l'Agglomération Dijonnaise, (1990 ;1999), *LATEC Working Paper,* n° 2002-04, Université de Bourgogne, Dijon (english version available upon request from the authors)

Baumont C., Le Gallo J. (1999) Empirical foundations of multicentric urban models, paper presented at the *46th North American Congress of the Regional Science Association International,* Montréal (Canada), November 11-14, 1999.

Boiteux-Orain C., Guillain R. (2002) Information technology and evolution of the producer services geography in Ile-de-France (1978-1997), paper presented at the *IX National Meeting of APDR*, Lisbon (Portugal), June 27-29, 2002.

Brueckner J.K. (1986) A switching regression analysis of urban population densities, *Journal of Urban Economics*, 19, 174-189.

Cervero R., Wu K.-L. (1997) Polycentrism, commuting, and residential location in the San Francisco Bay area, *Environment and Planning A*, 29, 865-886.

Cervero R., Wu K.-L. (1998) Sub-centring and commuting: evidence from the San Francisco Bay area, *Urban Studies*, 35, 1059-1076.

Chen H.-P. (1997) Models of urban population and employment density: the spatial structure of monocentric and polycentric functions in Greater Taipei and a comparison to Los Angeles, *Geographical & Environmental Modelling*, 1, 135-151.

Clark C. (1951) Urban population densities, *Journal of the Royal Statistical Society (Series A*), 114, 490-496.

Cliff A.D., Ord J.K. (1981) *Spatial Processes: Models and Applications*, Pion, London.

Coffey W.J., Polèse M., Drolet R. (1996) Examining the thesis of Central Business District decline: evidence from the Montreal metropolitan area, *Environment and Planning A,* 28, 1795-1814.

Conley T.G. (1999) GMM estimation with cross sectional dependence, *Journal of Econometrics*, 92, 1-45.

Craig S., Ng P. (2001) Using quantile smoothing splines to identify employment subcenters in a multicentric urban area, *Journal of Urban Economics*, 49, 100-120.

Florax R., Folmer H. (1992) Specification and estimation of spatial linear regression models: Monte-Carlo evaluation of pre-test estimator, *Regional Science and Urban Economics*, 22, 405-432.

Florax R., Fomer H., Rey S. (2003) Specification searches in spatial econometrics: The relevance of Hendry's methodology, forthcoming in *Regional Science and Urban Economics*, 33.

Gaschet F. (2000) La structure d'un espace urbain polycentrique: la métropole bordelaise, in: Derycke P.-H. (ed.), *Structure des villes, entreprises et marchés urbains*, L'Harmattan, Paris.

Gelman A., Carlin J.B., Stern H.S., Rubin D.B. (1995), *Bayesian Data Analysis*, London, Chapman & Hall.

Getis A., Griffith D.A. (2002) Comparative spatial filtering in regression analysis, *Geographical Analysis*, 34, 130-140.

Getis A., Ord J.K. (1992) The analysis of spatial association by use of distance statistics, *Geographical Analysis*, 24, 189-206.

Geweke (1993) Bayesian treatment of the independent Student-*t* linear model, *Journal of Applied Econometrics*, 8, 19-40.

Giuliano G., Small K.A. (1991) Subcenters in the Los Angeles region, *Regional Science and Urban Economics,* 21, 163-182.

Giuliano G., Small K.A (1999) The determinants of growth of employment subcenters, *Journal of Transport Geography*, 7, 189-201.

Gordon P., Richardson H.W., Wong H.L. (1986) The distribution of population and employment in a polycentric city: the case of Los Angeles, *Environment and Planning A,* 18, 161-173.

Griffith D.A. (1981) Modelling urban population density in a multi-centered city, *Journal of Urban Economics*, 9, 298-310.

Griffith D.A., Can A. (1995) Spatial statistical/econometric versions of simple urban population density models, in: Arlinghaus S.L. (ed.) *Practical Handbook of Spatial Statistics*, CRC Press, Boca Raton.

Heikkila E., Gordon P., Kim J.I., Peiser R.B., Richardson H.W. (1989) What happened to the CBD-distance gradient? Land values in a policentric city, *Environment and Planning A*, 21, 221-232.

INSEE (2000) IRIS-2000 : un nouveau découpage pour mieux lire la ville, *Chiffres pour l'Alsace*, 43, 5.

Kelejian H.H., Prucha I.R. (1999) A generalized moments estimator for the autoregressive parameter in a spatial model, International Economic Review, 40, 509-533.

Lange K.L., Little R.J.A., Taylor J.M.G. (1989) Robust statistical modeling using the *t* distribution, *Journal of the American Statistical Association*, 84, 881-896.

Le Gallo J., Ertur C. (2003) Exploratory Spatial Data Analysis of the distribution of regional per capita GDP in Europe, 1980-1995, *Papers in Regional Science*, forthcoming.

LeSage J. (1999) *Spatial Econometrics*, in *The Web Book of Regional Science* (www.rri.wvu.edu/regscweb.htm), Jackson R.W. (ed.), Morgantown, WV: Regional Research Institute, West Virginia University.

LeSage J. (2002) Application of Bayesian methods to spatial econometrics, Working Paper, University of Toledo, Ohio, presented at the 49th Regional Science Association North American Meeting, November 14-16, San-Juan, Puerto-Rico.

McDonald J.F. (1989) Econometric studies of urban population density: a survey, *Journal of Urban Economics*, 26, 361-385.

McMillen D.P. (2001) Nonparametric employment subcenter identification, *Journal of Urban Economics*, 50, 448-473.

McMillen D.P. (2002) Employment densities in large metropolitan areas: spatial autocorrelation versus functional form misspecification, paper presented at the *49th North American Congress of the Regional Science Association International,* Porto Rico (USA), November 14-16, 2002.

McMillen D.P. (2003) Employment subcenters in Chicago: past, present, and future, forthcoming in *Economic Perspectives.*

McMillen D.P., McDonald J.F. (1998a) Population density in Chicago: a bid rent approach, *Urban Studies,* 7, 1119-1130.

McMillen D.P., McDonald J.F. (1998b) Suburban subcenters and employment density in metropolitan Chicago, *Journal of Urban Economics,* 43, 157-180.

Mills E.S., Tan J.P. (1980) A comparison of urban population density functions in developed and developing countries, *Urban Studies*, 17, 313-321.

Muth R.F. (1969) *Cities and Housing: The Spatial Pattern of Urban Residential Land-Use,* The University of Chicago Press, Chicago and London.

Newey W.K., West K.D. (1987) A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703-708.

Ord J.K., Getis A. (1995) Local spatial autocorrelation statistics: distributional issues and an application, *Geographical Analysis*, 27, 286-305.

Pace R.K., Le Sage J.P. (2002) Spatial autoregressive local estimation, forthcoming in Recent Advances in Spatial Econometrics, Mur J., Zoller H., Getis A. (eds.), Palgrave Publishers.

Páez A., Uchida T., Miyamoto K. (2001) Spatial association and heterogeneity issue in land price models, *Urban Studies*, 38, 9, 1493-1508.

Schwarz G. (978) Estimating the dimension of a model, *The Annals of Statistics*, 6, 461-464.

Scott, L., Lloyd W.J. (1997) Spatial Analysis in a GIS environment: employment patterns in greater Los Angeles, 1980 - 1990, *Proceedings of the University Consortium for Geographic Information Science*, Orono, Maine.

Sivitanidou R (1996) Do office-commercial firms value access to service employment centers? A hedonic value analysis within polycentric Los Angeles, *Journal of Urban Economics,* 40, 125-149.

Small K., Song S. (1994) Population and employment densities: structure and change, *Journal of Urban Economics,* 36, 292-313.

Stern D.I. (1993) Historical path-dependence of the urban population density gradient, *Annals of Regional Science*, 27, 259-283.

Waddell P., Shukla V. (1993) Employment dynamics, spatial restructuring and the business cycle, *Geographical Analysis,* 25, 35-52.

White H. (1980) A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, 48, 817-838.

Wu F. (1998) Polycentric urban development and land-use change in a transitional economy: the case of Guangzhou, *Environment and Planning A*, 30, 1077-1100.

Appendix 1: IRIS-2000® zoning

The acronym **IRIS** stands for *Ilots Regroupés pour l'Information Statistique* (blocks clustered for statistical information)

IRIS-2000® is an infra-communal level division available for all urban communes of at least 10 000 inhabitants and most communes of 5000 to 10 000 inhabitants. It is a small district, defined as a group of adjacent blocks of houses. IRIS-2000® are subdivided into three types of zone (INSEE. 2000):

- residential IRIS: IRIS-2000® with populations of 1800 to 5000 inhabitants. They are homogeneous in respect of types of housing.
- business IRIS: IRIS-2000® clustering more than 1000 employees and with twice as many salaried jobs as resident inhabitants.
- miscellaneous IRIS: IRIS-2000® covering large areas and for special purposes (woods, parkland, docklands, etc.).

In data bases covering several communes, IRIS data correspond either to IRIS-2000® for subdivided communes or to the entire commune for small non-subdivided communes. The population and employment data are collected both at the communal level and at the IRIS level. Note that INSEE's procedures for breaking down employment or population data in the IRIS system are relatively recent and they require relevant information to be available on the localization of individuals and firms. Sometimes these localizations cannot be identified or assigned to an IRIS zone. There is therefore some discrepancy between the data provided for the whole town and the data for the town computed from the IRIS data. Such discrepancies are relatively minor, however, both for population and for employment data in our sample.

The employment data from the SIRENE files correspond to salaried employment. They relate primarily to private-sector employment and are far from complete when it comes to public-sector employment and agricultural employment as well as employment in some major financial organizations. There are a number of reasons for this: agricultural employment figures come from the farming census, public-sector employment and employment in large financial organizations cannot be broken down in the IRIS system because it is assigned for the regional head office and not for the actual place of business.

Tables, maps and figures

Table 1: Main geographic and demographic characteristics of study area

Characteristics	**COMADI**	**COMADI + Urban Fringe**	**Urban fringe/ COMADI (%)**	**Metropolitan Area (% COMADI/ M.A.)**	**Burgundy Region** (% COMADI/ Burgundy)
Total area (acres) (Width×Length)	172.4 km^2 (42 600 acres) (16km×18km)	412.1 km^2 (101 833 acres)		2 271 km^2 (8%)	31 581.96 km^2 (0.5%)
Population 1999*	238 309	257 844	8.08%	326 887 (73%)	1 608 262 (14.8%)
Working pop. 1999*	111 028	120 587	8.6%	153 401 (72%)	708 174 (15.7%)
Employment 1999**	70 770	73 277	3.54%	82 958 (85%)	387 183 (18.3%)

*Sources: * Recensement Général de la Population 1999 (RGP Census)*
*** SIRENE (INSEE)*

Table 2: Moran's *I* statistics for total employment and employment density in 1999

	Contiguity weight matrix			6 nearest neighbors weight matrix		
Variable	Moran's *I*	St. dev.	St. value	Moran's *I*	St. dev.	St. value
Emp99*	0.117	0.049	2.255	0.136	0.042	3.365
Demp99*	0.378	0.045	8.467	0.436	0.039	11.249

* The expected value for Moran's *I* statistic is -0.007 for Emp99 and Demp99. All statistics are significant at 5% level.

Table 3: LISA statistics for total employment and employment density in 1999

		Total employment 1999		***Employment density 1999***	
		Contiguity	***6 near. neighbors***	***Contiguity***	***6 near. neighbors***
Ahuy	***Ahuy***	LH	LH	LL	LL
	Chenôve				
Co1	Piscine-Valendons	LL	LL	LL	LL
Co2	Petignys-Chaufferies	LL	LL	LL	LL
Co3	Chapitre-Bibliothèque	LL	LL	LL	LL
Co4	Saint-Exupery	LL	LL	HL	HL
Co5	Vieux Bourg-Grands Crus	LH	LL	LL	LH
Co6	Atliers SNCF	HH	HH	HL	HL
Co7	Mairie	LH	LH	LH	LH
Co8	Zone industrielle	HH*	HL	HH	HH
Co9	STRD	HH	HH	HL	HL
Chevigny-Saint-Sauveur					
Ch1	Breuil	LL	LL	LL*	LL
Ch2	Corcelles	LL	LL	LL**	LL*
Ch3	Centre-Ville	LL	LL	LL*	LL
Ch4	Château	LH	LH	LL	LL
Ch5	Zone Economique	HL**	HL*	LL***	LL**
Daix	***Daix***	HL	HL	LL**	LL
	Dijon				
D1	Monge	HH**	HH**	HH***	HH***
D2	Cordeliers	HH	HH**	HH**	HH***
D3	Saint Michel	LH	LH	HH	HH
D4	Grangier	HH*	HH*	HH***	HH***
D5	J-J Rousseau	HH	HH*	HH**	HH**
D6	Darcy	HH*	HH*	HH***	HH***
D7	Les Roses	LH	LH	HH	HH**
D8	République	HH	HH	HH**	HH**
D9	Clémenceau	HL	HH	HH	HH
D10	Davout	HH	HH	HH*	HH*
D11	Petit Citeaux	LH	LH*	HH**	HH**
D12	Saint Pierre	LL	LH	LH	LH
D13	Docteur Lavalle	LH	LL	HH	HH
D14	Voltaire	HL	HL	HH	HH
D15	Lyautey	HL	HL	HL	HL
D16	Parc des Sports	LH	LL	LL	LL
D17	Champmaillot	LL	LL	LH	LH
D18	Universités	HH	HL	LL	LL
D19	Lentillères	LL	LL	LL	LL
	Dijon (continued)				
D20	Petites Roches	LL	LL	LL	LH
D21	Mansart	LL	LL	LL	LL
D22	Abattoirs	LH	LH	LL	LL
D23	Castel	LH	LL	LH*	LL
D24	Stearinerie	LL	LL	LL	LL
D25	Carrousel	LL	LL	LL	LL
D26	Greuze	LH*	LL	LL	LL
D27	Arsenal	HH*	HL	LL	LL
D28	Bel Air	LL	LL**	LL	LL
D29	Larrey	LL	LL	LL	LH
D30	Bourroches Ouest	LL	LL	LL	LL
D31	Bourroches Est	LL	LL	LH	LL
D32	Trois Forgerons	HL	HL	HL	HL
D33	Les Valendons	LL	LL*	LL	LL
D34	La Montagne	LL	LL	LL*	LL
D35	Tire Pesseau	LL	LL	LL	LL
D36	Le Lac	LL	LL	HL	HL
D37	E. Belin	LL	LL	LL	LL
D38	Champ Perdrix	LL	LL	HL	HL
D39	Chartreuse	HH	HH	HH*	HH**
D40	Arquebuse	HH	HH**	HH**	HH***
D41	Tanneries	LH	LH	LH*	LH*
D42	Providence	LH	LL	LH	LL
D43	Carrières Basquin	HH	HL	HH**	HH
D44	F. Pompom	LL	LL	LL	LL
D45	Hauts Montchapet	LL	LL	LL	LL
D46	E. Spuller	LL	LL	LH	LH
D47	La Charmette	LL	LL	LL	LL
D48	Fauconnet	LL	LL	LL	LL
D49	Jouvence Ouest	LL	LL	LH	LH
D50	Jouvence Est	LL	LL	HH	HH
D51	Balzac	LH	LL	LL	LL
D52	Stalingrad	LH**	LL	LH	LL
D53	Casernes	LL	LL	LH	LH
D54	Sacré Cœur	LH	LL	LH	LH
D55	York	LH	LL	LH	LL
D56	Lochères	LH	LL	LH	LL

***Notes**:* *5% pseudo-significance level; **1% pseudo-significance level; ***0.1% pseudo-significance level; inference based on 9999 permutations.

Table 3 (continued): LISA statistics for total employment and employment density in 1999

		Total employment 1999		Employment density 1999	
		Contiguity	*6 near. neighbors*	*Contiguity*	*6 near. neighbors*
Dijon (continued)					
D57	Grésilles Centre	LH*	LL	LL	LL
D58	Castelnau	LL	LL	LH	LL
D59	Charles de Gaulle	LH	LL	LL	LL
D60	Concorde	HH	HL	HL	HL
D61	Clos de Pouilly	LH	LH	HL	HL
D62	La Toison d'Or	HH	HL	LL	LH
D63	ZI Nord Est	HH	HL	HL	HL
D64	La Gare	HH	HH*	HH***	HH***
D65	Le Bocage	LH	LL	LL	LL
D66	Combe à la Serpent	LL*	LL**	LL	LL
Fontaine-lès-Dijon					
Fo1	Vieux Village	LL	LL	LL	LL
Fo2	Saverney	LL	LL	LL	LL
Fo3	Saint Martin	LL	LL	LL	LL
Fo4	Majnoni	LL	LL	LL	LL
Fo5	Activités économiques	HL	HL	HL	HL
Longvic					
Lo1	Bief du Moulin	LH*	LH	LL	LL
Lo2	Bourg	LH	LH	LL	LL
Lo3	Parc Poussot	LH	LL	LL	LL
Lo4	Zone indust. Colombières	HH	HH	LL	LL
Marsannay-la-Côte					
Ma1	Le Bourg	LH*	LH	LL	LH
Ma2	Champagne Haute	LL	LL	LL*	LH
Ma3	ZAC	HH**	HH**	LL	LH
Neuil	***Neuilly-lès-Dijon***	LL	LL	LL***	LL**
Ouges	***Ouges***	LH	LH	LL*	LL
Perry	***Perrigny-lès-Dijon***	HL	HH**	LL*	LH
Plomb	***Plombières-lès-Dijon***	LL	LL	LL***	LL
Quétigny					
Qu1	Atrias-Vieux Village	LH	LH	LL	LL
Qu3	La Fontaine aux Jardins	LH	LH	LL	LL
Qu3	Les Huches	LL	LH	LL	LL
Qu4	Place Centrale	LH	LH	LL	LL
Qu5	Zone Activités Cap Vert	HL	HL	HL	HL

		Total employment 1999		Employment density 1999	
		Contiguity	*6 near. neighbors*	*Contiguity*	*6 near. neighbors*
Saint-Apollinaire					
Sapo1	Nord Village	LH	LH	LL	LL
Sapo2	Sud Village	LH	LH	LL	LL
Sapo3	Nord-Est	HH	HH	LL	LL
Sapo4	Sud-Est	HH	HH	LL	LL
Senec	***Sennecey-lès-Dijon***	LL	LL	LL**	LL
Talant					
Ta1	Vieux Talant-Clinique	LL	LL*	LL	LL
Ta2	Maronniers-Neruda	LL	LL	LL	LL
Ta3	Mail-Canzio-Jouvet	LL	LL	LL	LL
Ta4	Prévert-plein ciel	LL	LL	LL	LL
Ta5	Boris Vian-Triolet	LL	LL	LL	LL
Ta6	Montoillots- ZA	HL	HL	LL*	LL
Urban fringe of COMADI					
Asn	***Asnières-lès-Dijon***	LH	LH	LL	LL
Belf	***Bellefond***	LH	LH*	LL	LL
Bres	***Bressey-sur-Tille***	LL	LL	LL	LL*
Bret	***Bretenière***	LL	LL*	LL**	LL***
Corc	***Corcelles-les-Monts***	LL	LL	LL***	LL***
Couch	***Couchey***	LL	LL	LL**	LL**
Cout	***Couternon***	LL	LL	LL**	LL**
Crim	***Crimolois***	LL	LL	LL**	LL**
Darois	***Darois***	LL	LL	LL**	LL***
Fauv	***Fauverney***	LL	LL*	LL*	LL**
Fenay	***Fénay***	LH*	LH*	LL	LL
Fixin	***Fixin***	LL	LL	LL*	LL*
Flav	***Flavignerot***	LL	LL	LL***	LL***
Hautv	***Hauteville-lès-Dijon***	LL	LL	LL**	LL*
Lant	***Lantenay***	LL	LL*	LL**	LL***
Magny	***Magny-sur-Tille***	LL	LL	LL*	LL**
Mess	***Messigny-et-Vantoux***	LL	LL	LL*	LL***
Pren	***Prenois***	LL	LL	LL**	LL***
Rouv	***Rouvres-en-Plaine***	LL	LL*	LL***	LL***
Ruff	***Ruffey-lès-Echirey***	LH*	LH*	LL	LL
Varois	***Varois-et-Chaignot***	LL	LL	LL*	LL*
Velars	***Velars-sur-Ouche***	LL	LL	LL***	LL***

***Notes**:* [*] 5% pseudo-significance level; [**]1% pseudo-significance level; [***]0.1% pseudo-significance level; inference based on 9999 permutations.

Table 4: Employment poles in 1999 detected in Baumont and Bourdon (2002) compared to LISA statistics

Employment poles			Emp99	**Total employment 1999**		Demp99a	**Employment density 1999**	
				Contiguity	***6 near. neighbors***		***contiguity***	***6 near. neighbors***
Name	**Characteristics**	IRIS Composition						
Inner City of Dijon	**CBD**	Monge D1	1501	HH**	HH**	28,5	HH***	HH***
		Grangier D4	4080	HH*	HH*	62,7	HH***	HH***
Total jobs: 9 644		La Gare D64	4063	HH	HH*	74,5	HH***	HH***
South	**Multi-towns**	Zone industrielle Co8	4776	HH*	HL	9,1	HH	HH
		ZAC Ma3Activités économiques	1505	HH**	HH**	2,7	LL	LH
Total jobs: 11 540		Zone indust. Colombières Lo4	5259	HH	HH	3,5	LL	LL
North	**Multi-towns**	La Toison d'Or D62	2670	HH	HL	3,1	LL	LL
		ZI Nord Est D63	5558	HH	HL	10,5	HL	HL
Total jobs: 9 634		Nord-Est Sapo3	1406	HH	HH	1,2	LL	LL
Quétigny	**Isolated pole**	Zone Activités Cap Vert Qu5	4014	HL	HL	8,8	HL	HL
Chevigny	**Isolated pole**	Zone Economique Ch5	1421	HL**	HL*	1,2	LL***	LL**

Table 5: Moran's *I* statistics for population density in 1999

	Contiguity weight matrix			6 nearest neighbors weight matrix		
Variable	Moran's *I*	St. dev.	St. value	Moran's *I*	St. dev.	St. value
Dpop99*	0.474	0.0494	9.734	0.367	0.043	8.771

* The expected value for Moran's *I* statistic is -0.008 for Dpop99. All statistics are significant at 5% level.

Table 6: LISA statistics for population density in 1999

		Contiguity	*6 near. neighbors*
Ahuy	***Ahuy***	LL**	LL
	Chenôve		
Co1	Piscine-Valendons	HH	HH*
Co2	Petignys-Chaufferies	HH**	HH*
Co3	Chapitre-Bibliothèque	HH*	HH*
Co4	Saint-Exupery	HH	HH
Co5	Vieux Bourg-Grands Crus	LH	LH*
Co6	Atliers SNCF	LH*	LH*
Co7	Mairie	HH	HH
Co8	Zone industrielle	LL*	LH
Co9	STRD	LL	LH
	Chevigny-Saint-Sauveur		
Ch1	Breuil	LL*	LL
Ch2	Corcelles	LL*	LL*
Ch3	Centre-Ville	HL	HL
Ch4	Château	LL**	LL
Ch5	Zone Economique	LL**	LL*
Daix	***Daix***	LL***	LH
	Dijon		
D1	Monge	HH	HH
D2	Cordeliers	HH**	HH
D3	Saint Michel	HH*	HH*
D4	Grangier	HH*	HH*
D5	J-J Rousseau	HH*	HH*
D6	Darcy	HH	HH*
D7	Les Roses	HH	HH*
D8	République	HH**	HH**
D9	Clémenceau	HH	HH*
D10	Davout	HH*	HH**
D11	Petit Citeaux	HH	HH
D12	Saint Pierre	HH	HH*
D13	Docteur Lavalle	HH	HH
D14	Voltaire	HH	HH*
D15	Lyautey	HH	HH
D16	Parc des Sports	LL	LH
D17	Champmaillot	HH	HH
D18	Universités	LL	LL
D19	Lentillères	HH	HH
	Dijon (continued)		
D20	Petites Roches	HH	HH
D21	Mansart	HL	HH
D22	Abattoirs	LL	LL
D23	Castel	LH	LH
D24	Stearinerie	HL	HH
D25	Carrousel	LH	LH
D26	Greuze	LL	LH
D27	Arsenal	LH	LH
D28	Bel Air	LH	LH
D29	Larrey	HL	HH
D30	Bourroches Ouest	HL	HH
D31	Bourroches Est	HH	HH
D32	Trois Forgerons	HH	HH
D33	Les Valendons	HL	HH**
D34	La Montagne	LL	LH**
D35	Tire Pesseau	HH*	HH**
D36	Le Lac	HH*	HH**
D37	E. Belin	HH*	HH
D38	Champ Perdrix	HH**	HH**
D39	Chartreuse	LL	LH
D40	Arquebuse	HL	HH
D41	Tanneries	HH	HH
D42	Providence	LL	LH
D43	Carrières Basquin	HH	HH
D44	F. Pompom	HL	HH
D45	Hauts Montchapet	HH	HH
D46	E. Spuller	HH*	HH*
D47	La Charmette	HH	HH
D48	Fauconnet	HH	HH
D49	Jouvence Ouest	HH*	HH*
D50	Jouvence Est	HH	HH*
D51	Balzac	HL	HH
D52	Stalingrad	HL	HH
D53	Casernes	HH	HH*
D54	Sacré Cœur	LH	LH*
D55	York	HH	HH
D56	Lochères	HH	HH

Notes: * 5% pseudo-significance level; **1% pseudo-significance level; ***0.1% pseudo-significance level; inference based on 9999 permutations.

Table 6 (continued): LISA statistics for population density in 1999

		Contiguity	***6 near. neighbors***
	Dijon (continued)		
D57	Grésilles Centre	LL	LH
D58	Castelnau	HH	HH
D59	Charles de Gaulle	LL	LL
D60	Concorde	HL	HL
D61	Clos de Pouilly	LL	LL
D62	La Toison d'Or	LL*	LL
D63	ZI Nord Est	LL	LL
D64	La Gare	LH	LH*
D65	Le Bocage	LL	LH
D66	Combe à la Serpent	LH	LH***
	Fontaine-lès-Dijon		
Fo1	Vieux Village	LL	LL
Fo2	Saverney	LH	LH
Fo3	Saint Martin	LH	LH
Fo4	Majnoni	HL	HL
Fo5	Activités économiques	LL	LL
	Longvic		
Lo1	Bief du Moulin	HL	HL*
Lo2	Bourg	LL*	LL
Lo3	Parc Poussot	LL	LL
Lo4	Zone indust. Colombières	LL*	LL
	Marsannay-la-Côte		
Ma1	Le Bourg	LL*	LH
Ma2	Champagne Haute	LL**	LH
Ma3	ZAC	LL**	LL*
Neuil	***Neuilly-lès-Dijon***	LL**	LL*
Ouges	***Ouges***	LL**	LL*
Perry	***Perrigny-lès-Dijon***	LL***	LL**
Plomb	***Plombières-lès-Dijon***	LL***	LH
	Quétigny		
Qu1	Atrias-Vieux Village	LL	LL
Qu3	La Fontaine aux Jardins	LL	LL
Qu3	Les Huches	HL	HL
Qu4	Place Centrale	HL	HL
Qu5	Zone Activités Cap Vert	LL	LL

		Contiguity	***6 near. neighbors***
	Saint-Apollinaire		
Sapo1	Nord Village	LL	LH
Sapo2	Sud Village	LL	LH
Sapo3	Nord-Est	LL*	LL
Sapo4	Sud-Est	LL	LL
Senec	***Sennecey-lès-Dijon***	LL*	LL
	Talant		
Ta1	Vieux Talant-Clinique	LL	LH
Ta2	Maronniers-Neruda	LH	LH**
Ta3	Mail-Canzio-Jouvet	HH	HH**
Ta4	Prévert-plein ciel	HH	HH*
Ta5	Boris Vian-Triolet	HH	HH**
Ta6	Montoillots- ZA	LL	LH
	Urban fringe of COMADI		
Asn	***Asnières-lès-Dijon***	LL**	LL*
Belf	***Bellefond***	LL*	LL*
Bres	***Bressey-sur-Tille***	LL*	LL*
Bret	***Bretenière***	LL**	LL***
Corc	***Corcelles-les-Monts***	LL***	LL**
Couch	***Couchey***	LL***	LL**
Cout	***Couternon***	LL**	LL*
Crim	***Crimolois***	LL**	LL
Darois	***Darois***	LL**	LL***
Fauv	***Fauverney***	LL**	LL**
Fenay	***Fénay***	LL***	LL*
Fixin	***Fixin***	LL*	LL**
Flav	***Flavignerot***	LL**	LL***
Hautv	***Hauteville-lès-Dijon***	LL**	LL
Lant	***Lantenay***	LL*	LL***
Magny	***Magny-sur-Tille***	LL*	LL**
Mess	***Messigny-et-Vantoux***	LL*	LL***
Pren	***Prenois***	LL**	LL***
Rouv	***Rouvres-en-Plaine***	LL**	LL***
Ruff	***Ruffey-lès-Echirey***	LL***	LL**
Varois	***Varois-et-Chaignot***	LL**	LL**
Velars	***Velars-sur-Ouche***	LL**	LL***

Notes: * 5% pseudo-significance level; **1% pseudo-significance level; ***0.1% pseudo-significance level; inference based on 9999 permutations.

Table 7: Estimation results for the monocentric negative exponential density function

	Negative exponential			
	OLS-White	**SEM-ML**	**SEM-GMM**	**SEM-Hetero-Bayesian**
Ln*D*(0)	4.038	4.178	4.273	4.338
	(0.000)	(0.000)	(0.000)	(0.000)
γ [*d(CBD)*]	-0.535	-0.519	-0.510	-0.516
	(0.000)	(0.000)	(0.000)	(0.000)
λ	-	0.561	0.561	0.491
		(0.000)		(0.000)
R^2	0.5997	-	-	0.6531
R^2-adj	0.5967	-	-	0.6505
Sq.corr	-	0.600	0.600	-
LIK	-232.159	-225.490	-	-
AIC	468.318	454.981	-	-
BIC	474.143	460.806	-	-
σ^2	1.806	1.552	1.523	1.542
Condition number	3.081	-	-	-
MORAN	4.372	-	-	-
	(0.000)			
LMERR	15.734	-	-	-
	(0.000)			
R-LMERR	1.295	-	-	-
	(0.255)			
LMLAG or LMLAG*	12.945	1.636*	-	-
	(0.000)	(0.201)		
R-LMLAG	0.129	-	-	-
	(0.720)			
LR-error	-	13.337		
		(0.000)		
SARMA	15.862	-	-	-
	(0.000)			

Notes: *p*-values are in parentheses. OLS-White indicates the use of the White (1980) heteroskedasticity consistent covariance matrix estimator for statistical inference in the OLS estimation. SEM-ML indicates maximum likelihood estimation of the spatial error model. SEM-GMM indicates iterated generalized moments estimation (Kelejian and Prucha, 1999). SEM-Hetero-Bayesian indicates Bayesian estimation of the spatial error model robust versus heteroscedasticity and outliers (LeSage, 1999). For Bayesian estimations, Bayesian p-values suggested by Gelman *et al*. (1995) as an analogue to conventional *p*-values are in parenthesis. Sq. Corr. is the squared correlation between predicted values and actual values. LIK is the value of the maximum likelihood function. AIC is the Akaike (1974) information criterion. BIC is the Schwarz information criterion (1978). MORAN is the Moran's *I* test adapted to OLS residuals (Cliff and Ord, 1981). LMERR is the Lagrange multiplier test for residual spatial autocorrelation and R-LMERR is its robust version. LMLAG is the Lagrange multiplier test for spatially lagged endogenous variable and R-LMLAG is its robust version (Anselin and Florax, 1995b; Anselin *et al*., 1996). LMLAG* is the Lagrange multiplier test for an additional spatially lagged endogenous variable in the spatial error model (Anselin, 1988). LR-error is the likelihood ratio test for the spatial error parameter. SARMA is the joint test of residual spatial autocorrelation and spatially lagged endogenous variable.

Table 8: Estimation results for the monocentric spline-exponential density function

	Spline-exponential			
	OLS-White	**SEM-ML**	**SEM-GMM**	**SEM-Hetero-Bayesian**
Ln*D*(0)	4.368 (0.000)	5.062 (0.000)	5.476 (0.000)	4.666 (0.000)
γ **[*d(CBD)*]**	-0.678 (0.000)	-0.867 (0.000)	-0.977 (0.000)	-0.670 (0.000)
β **[*d(>4km)*]**	0.208 (0.238)	0.479 (0.028)	0.622 (0.010)	0.208 (0.186)
λ	-	0.509 (0.000)	0.622	0.494 (0.000)
R²	0.6050	-	-	0.6618
R²-adj	0.5991	-	-	0.6567
Sq.corr	-	0.600	0.587	-
LIK	-231.247	-219.656	-	-
AIC	468.494	445.313	-	-
BIC	477.232	454.051	-	-
σ^2	1.795	1.481	1.445	1.503
Condition number	10.388	-	-	-
MORAN	4.702 (0.000)	-	-	-
LMERR	17.161 (0.000)	-	-	-
R-LMERR	4.850 (0.028)	-	-	-
LMLAG or LMLAG*	13.110 (0.000)	1.877* (0.349)	-	-
R-LMLAG	0.798 (0.372)	-	-	-
LR-err	-	15.863 (0.000)	-	-
SARMA	17.960 (0.000)-	-	-	-

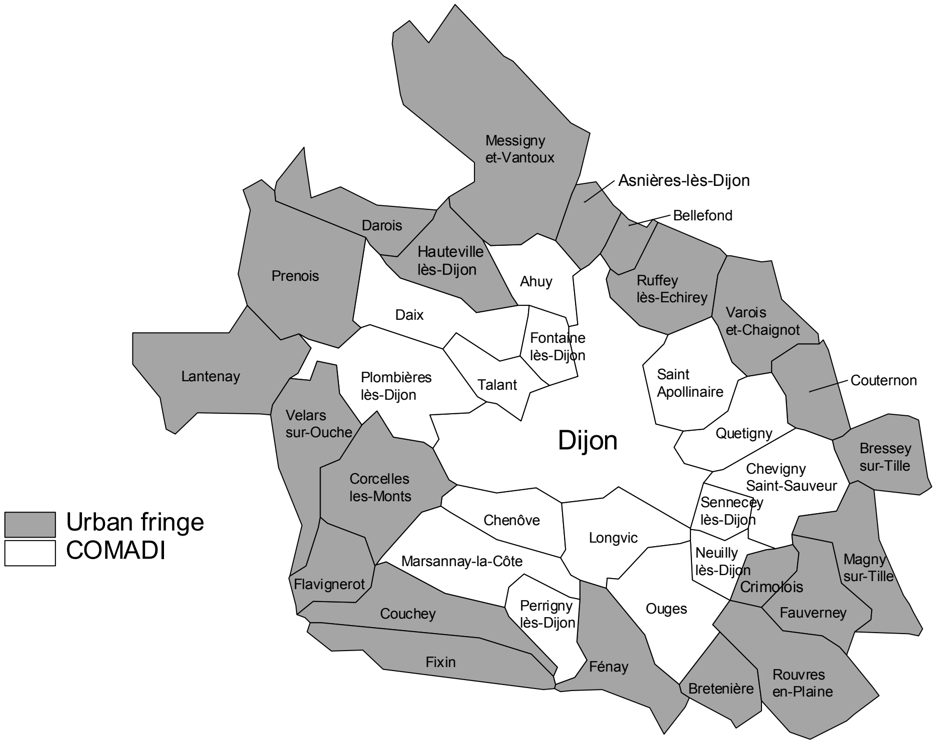
Notes: *p*-values are in parentheses. OLS-White indicates the use of the White (1980) heteroskedasticity consistent covariance matrix estimator for statistical inference in the OLS estimation. SEM-ML indicates maximum likelihood estimation of the spatial error model. SEM-GMM indicates iterated generalized moments estimation (Kelejian and Prucha, 1999). SEM-Hetero-Bayesian indicates Bayesian estimation of the spatial error model robust versus heteroscedasticity and outliers (LeSage, 1999). For Bayesian estimations, Bayesian p-values suggested by Gelman *et al.* (1995) as an analogue to conventional *p*-values are in parenthesis. Sq. Corr. is the squared correlation between predicted values and actual values. LIK is the value of the maximum likelihood function. AIC is the Akaike (1974) information criterion. BIC is the Schwarz information criterion (1978). MORAN is the Moran's *I* test adapted to OLS residuals (Cliff and Ord, 1981). LMERR is the Lagrange multiplier test for residual spatial autocorrelation and R-LMERR is its robust version. LMLAG is the Lagrange multiplier test for spatially lagged endogenous variable and R-LMLAG is its robust version (Anselin and Florax, 1995b; Anselin *et al.*, 1996). LMLAG* is the Lagrange multiplier test for an additional spatially lagged endogenous variable in the spatial error model (Anselin, 1988). LR-error is the likelihood ratio test for the spatial error parameter. SARMA is the joint test of residual spatial autocorrelation and spatially lagged endogenous variable.

Table 9: Estimation results for the multicentric density functions

	Multicentric 1				Multicentric 2				Multicentric 3			
	OLS-White	SEM-ML	SEM-GMM	SEM-Hetero-Bayesian	OLS-White	SEM-ML	SEM-GMM	SEM-Hetero-Bayesian	OLS-White	SEM-ML	SEM-GMM	SEM-Hetero-Bayesian
ln *D*(0)	4.097	4.392	4.543	4.573	4.313	4.676	4.838	4.770	4.219	4.675	4.868	4.810
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
γ **[*d(CBD)*]**	-0.531	-0.516	-0.512	-0.522	-0.542	-0.534	-0.529	-0.533	-0.533	-0.525	-0.524	-0.532
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
δ_1	-0.311	-0.810	-0.910	-0.917	-	-	-	-	-	-	-	-
	(0.491)	(0.083)	(0.053)	(0.080)								
δ_2	-	-	-	-	-0.652	-1.054	-1.157	-1.050	-	-	-	-
					(0.069)	(0.008)	(0.004)	(0.033)				
δ_3	-		-	-	-	-	-	-	-0.466	-0.995	-1.010	-1.082
									(0.148)	(0.003)	(0.001)	(0.013)
λ	-	0.477	0.595	0.514	-	0.490	0.603	0.519	-	0.523	0.634	0.548
		(0.000)		(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
R²	0.601	0.591	0.588	0.663	0.610	0.596	0.589	0.673	0.606	0.615	0.619	0.678
R²-adj	0.595	-	-	0.658	0.604	-	-	0.668	0.600	-	-	0.673
Sq. corr	-	0.597	0.596	-	-	0.605	0.603	-	-	0.598	0.594	-
LIK	-231.916	-224.06	-	-	-230.462	-222.169	-	-	-231.083	-221.630	-	-
AIC	469.831	454.119	-	-	466.925	450.338	-	-	468.166	449.260	-	-
BIC	478.569	462.857	-	-	475.663	459.076	-	-	476.904	457.998	-	-
σ^2	1.813	1.506	1.477	1.499	1.775	1.462	1.4316	1.454	1.791	1.440	1.407	1.432
Cond. num.	3.63	-	-	-	4.18	-	-	-	3.92	-	-	-
MORAN	4.687	-	-	-	4.858	-	-	-	5.026	-	-	-
	(0.000)				(0.000)				(0.000)			
LMERR	17.555	-	-	-	18.756	-	-	-	20.307	-	-	-
	(0.000)				(0.000)				(0.000)			
R-LMERR	2.823	-	-	-	5.473	-	-	-	6.375	-	-	-
	(0.093)				(0.019)				(0.012)			
LMLAG or LMLAG*	14.756	0.273	-	-	13.73	0.022	-	-	14.653	0.027	-	-
	(0.000)	(0.601)			(0.000)	(0.882)			(0.000)	(0.869)		
R-LMLAG	0.0242	-	-	-	0.448	-	-	-	0.721	-	-	-
	(0.876)				(0.503)				(0.396)			
LR-err	-	15.712	-	-	-	16.587	-	-	-	18.906	-	-
		(0.000)				(0.000)				(0.000)		
SARMA	17.579	-	-	-	19.204	-	-	-	21.028	-	-	-
	(0.000)				(0.000)				(0.000)			

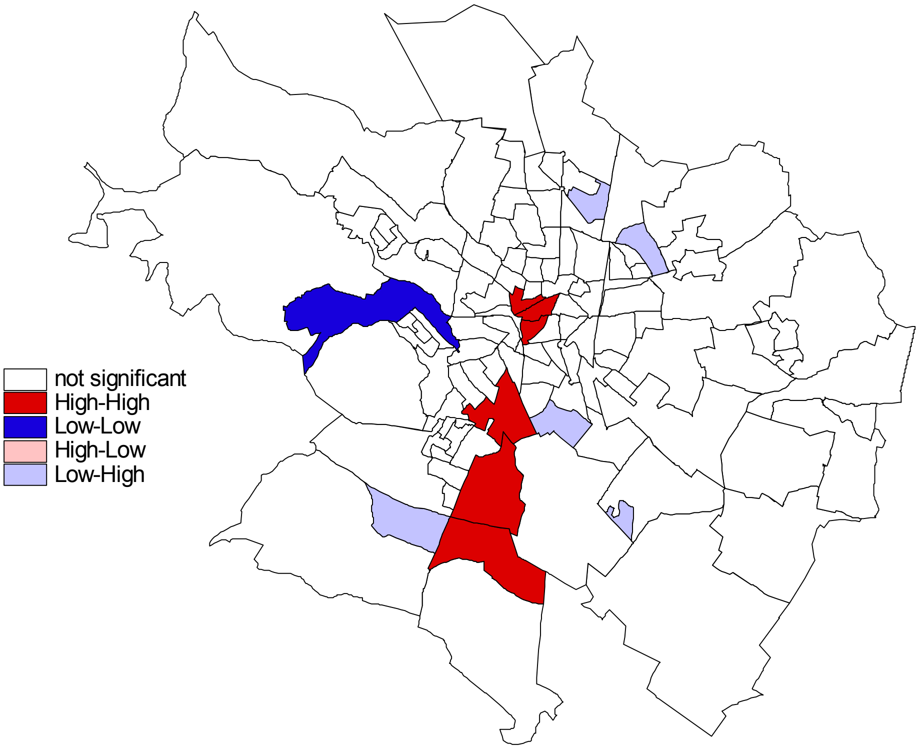
Notes: see Notes Table 8.

Map 1: The COMADI and its urban fringe

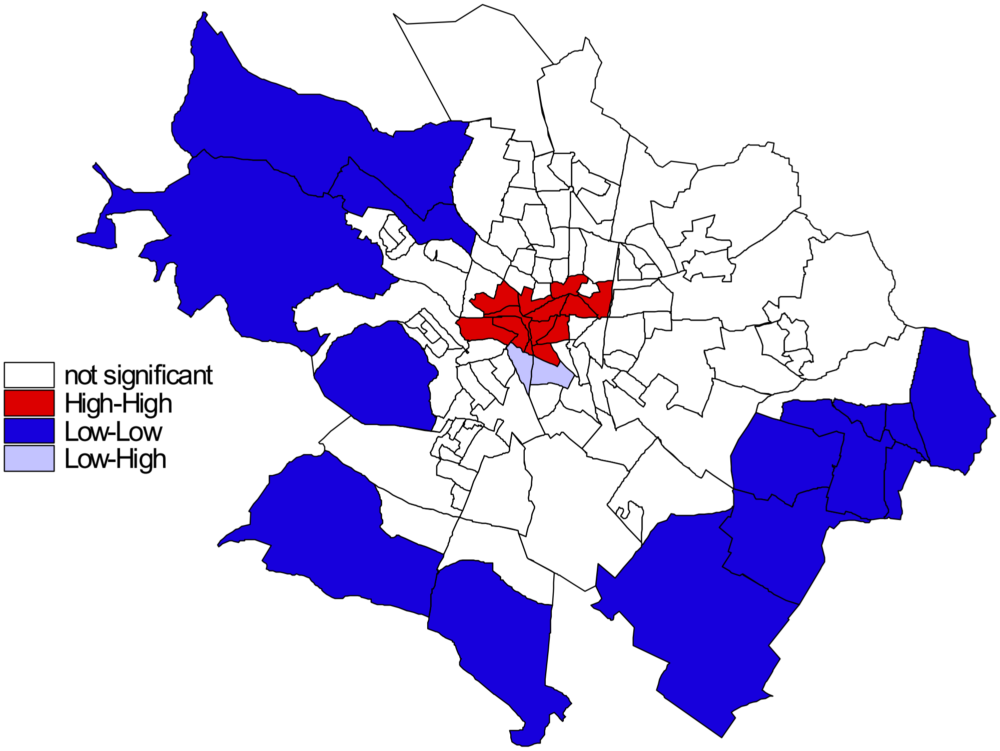


Map 2: The 114 IRIS of the COMADI

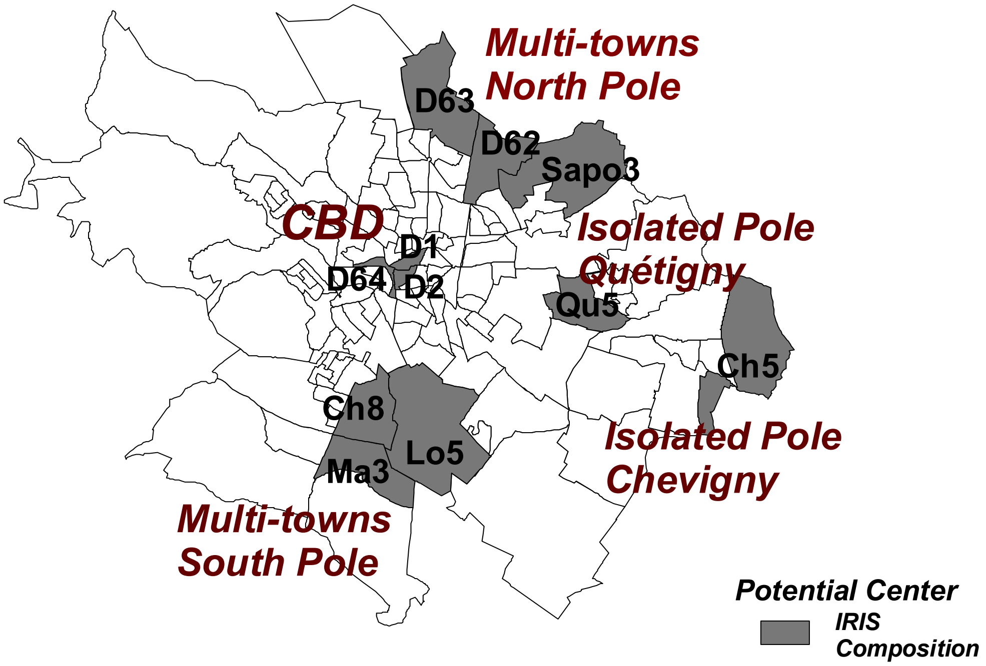
Map 3: Moran significance map for total employment 1999 (contiguity weight matrix)



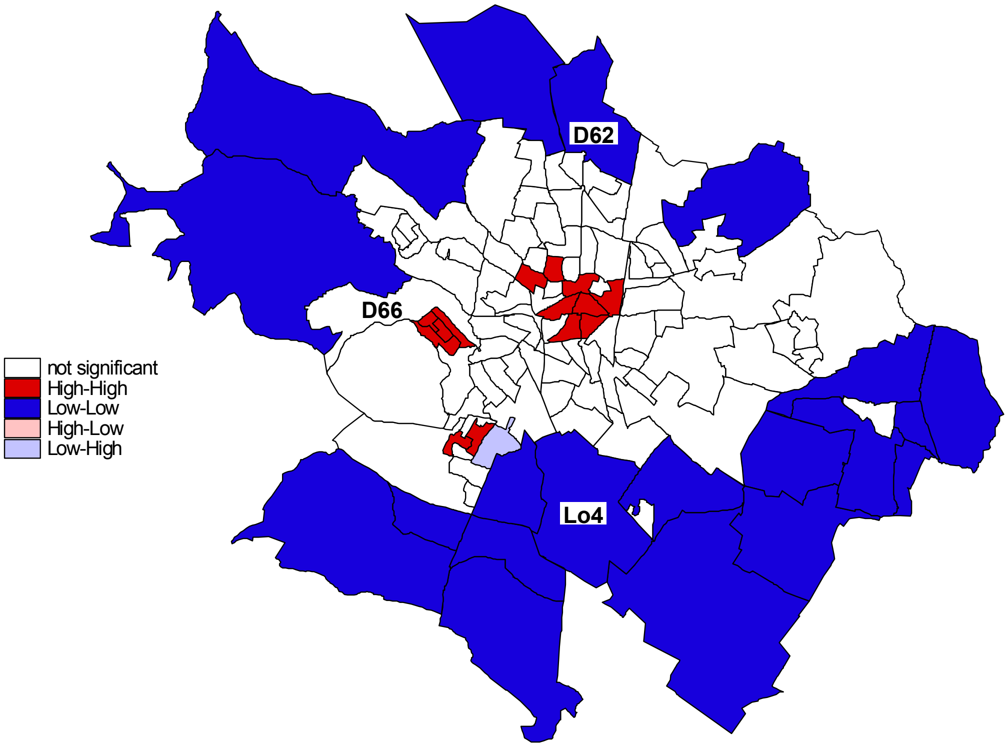
Map 4: Moran significance map for employment density 1999 (contiguity weight matrix)



Map 5: Employment poles (Baumont and Bourdon, 2002)



Map 6: Moran significance map for population density 1999 (contiguity weight matrix)



Notes: D62, D66 and Lo4 are potential outliers detected by Bayesian heteroscedastic estimation

Figure 1: monocentric spline exponential density function : posterior means of v_i estimates

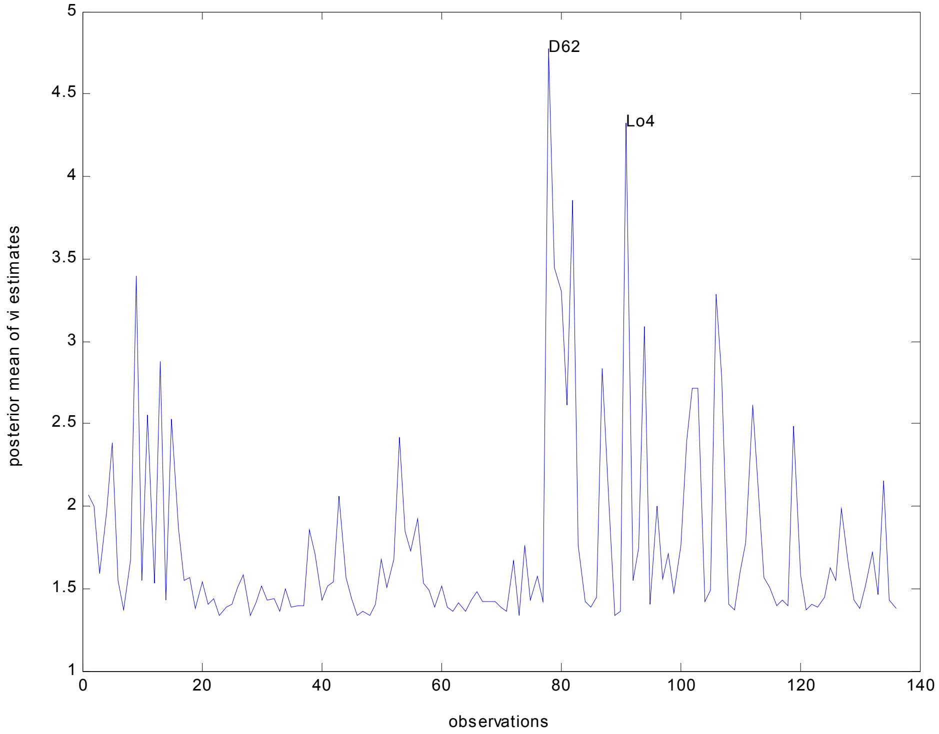


Figure 2: ML simulated normal distribution versus Bayesian heteroscedastic posterior distribution for β in the monocentric spline exponential density function

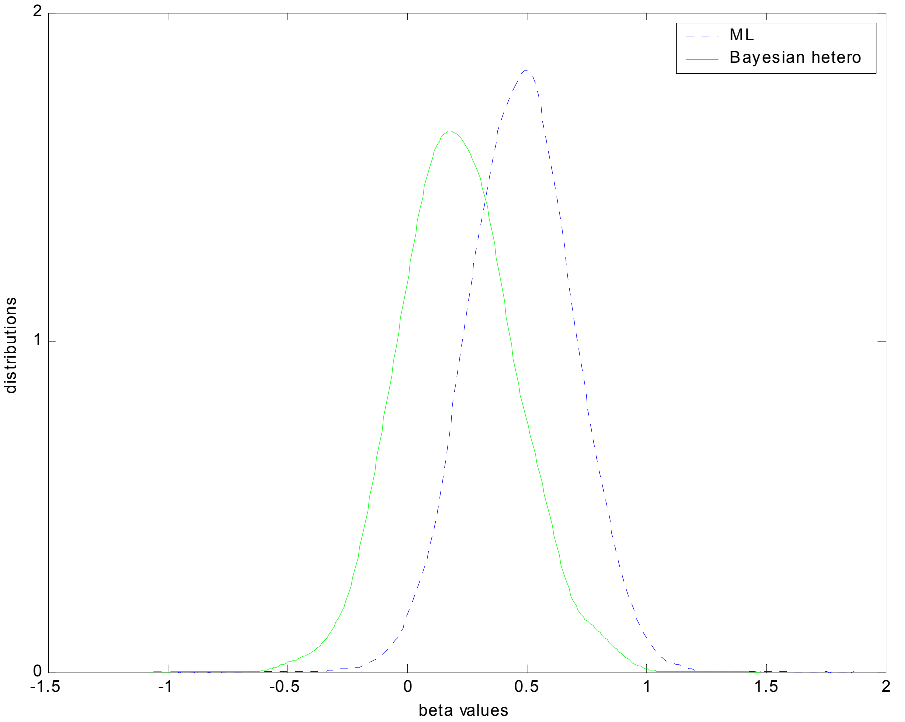


Figure 3: multicentric density function (3): posterior means of v_i estimates