51st European Congress of the Regional Science Association International
Barcelona, 30th August - 3rd September 2011

EXPLORING SPATIAL CONTAGION IN SPAIN'S INTERNATIONAL MIGRATION DISTRIBUTION*

María Hierro
Department of Economics, University of Cantabria, Avda. Los Castros s/n, 39005 Santander (Spain)

Abstract. This paper explores the Spain's international migration distribution (SIMD) for the 1998-2009 period. Beyond a general depiction of the distribution, the study pays special attention to the role played by space and, particularly, to the possibility of geographical contagion effects. For this latter, and using a spatial Markov chain approach, two new measures of positive and negative contagion are proposed. The results do identify space as key determinant of the SIMD. Furthermore, results reveal that there are contagion effects, positive contagion among provinces surrounded by high-immigration provinces being the most significant.

Key words: International migration; spatial Markov chain approach; contagion effect; Spanish provinces.

1 Introduction

The new international migration reality in Spain has come at the forefront of the regional inquiry since the turn of the 21st century or even earlier. In scarcely a decade and a half, Spain has evolved into one of the major immigration receiving countries in Europe (Carling, 2007; Arango and Finotelli, 2009). At the late 1990s the widely embraced thesis was that international migration provided a unique opportunity for tackling cross-cutting challenges, such as labour market and demographic imbalances and the sustainability of the social security system. Nevertheless, hardly a few years later the wave of immigration overwhelmed any rational prevision, its share in the total population escalating from 1.6% in

* The author is grateful to Professor Adolfo Maza for his valuable comments and continuing encouragement and support. However, responsibility for any error is solely mine.

1998 to 6.2% in 2003. More recently, although fears of a massive wave of immigration from future European Union (EU) enlargements seem unfounded and intensity of international migration flows to Spain is somewhat lower than in the first half of 2000s,[1] potential effects of current levels of immigration on the welfare system, employment opportunities of native workers and national security have ignited a major debate on the current immigration paradigm.

Thus, at the present stage of the debate it seems mandatory to build a more in depth understanding of some intrinsic characteristics of the SIMD that permits precise diagnosis of the real state of the international migration phenomena. This is particularly so because, despite a remarkable upsurge in the volume of empirical studies on international migration in Spain over the last few years (Bover and Velilla, 2002; Delgado, 2002; Recaño, 2002; Arango, 2003; Recaño, 2004; Izquierdo and Carrasco, 2005; Pumares et al., 2006; Recaño and Domingo, 2006; Hierro, 2007; Fernández and Ortega, 2008; Izquierdo et al., 2009; Hierro and Maza, 2010a,b; Hierro, 2011), a thorough analysis of the main characteristics of Spain's international migration distribution is still a pending question that has to be looked at squarely.

An interesting issue to be analyzed relates to the spatial pattering in Spain's international migration. More directly, the appealing question is whether levels of international migration are not randomly distributed across the Spanish provinces, so much so that some kind of distribution clustering is happening. Furthermore, in the case that space mattered, we must question whether high (low) incidence of international migration in nearby provinces might be transmitted to a province with initially low (high) levels of international migration; roughly speaking, might we expect some kind of contagion effect?. These premises lead us to the main motivation of this paper. Along with a detailed description of main characteristics of the SIMD, this paper is aimed at shedding light on the role played by geographic space in the distribution of international migration across the Spanish provinces and on the presence of potential contagion effects.

Methodologically, and regarding spatial questions, needless to say that Exploratory Spatial Data Analysis (ESDA) is the most common approach to examine spatial dependence, and also to elucidate phenomena lying behind observed spatial patterns, such as contagion. However, concerning the last issue, the application of ESDA only allows static comparisons,

[1] Some factors that might explain the reduction of international migration inflows to Spain since the second half of 2000s are the imposition of visa requirements for nationals from Colombia, Ecuador and Bolivia in 2002, 2003 and 2007, respectively, and reduced job opportunities associated to recent economic crisis.

being, therefore, no possible by means of this analysis to distinguish true contagion from apparent contagion (Messner and Anselin, 2004). In order to resolve this drawback, in this paper we propose two new measures of positive and negative contagion defined on the ground of a spatial Markov chain approach.

In order to accomplish the study, we use annual data on officially registered per capita foreign-born population[2] for the period 1998-2009 originate from the Spanish National Statistics Institute; more specifically, we employ data from the "Municipal Register" (*Padrón Municipal de Habitantes*). Some justification for the use of this database seems mandatory. Along with its ever-growing quality and coverage, data provided are annual, what makes possible a thorough dynamic analysis of the distribution. Additionally, our data set encompasses both regular and irregular foreign-born population. This represents a clear advantage as irregular immigration is usually hidden from view for statistics on immigration.[3] As for the level of territorial disaggregation, we have opted to use the 50 Spanish provinces that correspond to the Nomenclature of the Territorial Units for Statistics (NUTS) level III, because this allows one to take into account movements across provinces belonging to a region that go unnoticed when using NUTS-2 regions.

The remainder of this paper is as follows. In Section 2 we provide a general overview of main international migration patterns in Spain since the late 1990s. This is followed, in Section 3, by an examination of main characteristics of the international migration distribution in Spain, with especial attention to inequality, external shape, polarization and spatial dependence. Next, in Section 4 an appraisal of possible contagion effects is accomplished by means of the proposal of two new measures of positive and negative contagion defined on the ground of a spatial Markov chain approach. Finally, some concluding remarks are presented in Section 5.

[2] This paper focuses on international migration stock residing in a country. Notwithstanding that, and not denying the distinction between "immigration" (migration inflow) and "foreign-born population" (migration stock), both terms are used synonymously in this paper.

[3] Following the extension of some social rights to immigrants by the *Ley Orgánica 4/2000*, and particularly the provision of health assistance to undocumented immigrants registered in a Municipal Register, a large number of undocumented immigrants residing in Spain has become "visible", this yielding to a more precise count of international migration numbers through this official statistic. In addition, since December 2003 obligation of non-EU foreign-born residents without permanent legal permit to renew their registration in the municipal registry every two years might have contributed to improve reliability of the statistic (Arango and Finotelli, 2009).

2 Recent international migration trends in Europe: A focus on Spain

Challenges and opportunities concerning international migration have become an EU-wide priority. The dissemination of international migration flows into areas that have not traditionally been immigrant magnets and increasing interdependencies across European Member States provide, as outlined in the 2000 "Lisbon Agenda", strong foundations for cooperation among the EU's Member States to tackle these challenges. Although the formulation of a community immigration policy seems to be the path to be followed, different perceptions across Member States about international migration, different national circumstances and some political reluctance among European Member States to lose sovereignty over immigration-related issues have hindered this objective so far (Delgado, 2002).

On the other hand, in the last decades profound changes have occurred in international migration patterns in Europe. Until the mid 1970s, Europe's migration was predominantly marked by migration from Southern to Western European countries. Following the 1973 oil price shock, halted recruitment and restricted immigration policies in Western Europe led promptly to a reduction of these flows (Fassmann and Münz, 1992; Hierro, 2011). Afterwards, liberalization of emigration restrictions that followed the fall of the 'Iron Curtain' in the late 1989 fuelled intense East-West migration flows (Dietz, 2000). Since the late 1990s one of the foremost changes in Europe's migration patterns was the rapid and unexpected turnaround in Southern Europe, with some Southern European countries traditionally exporters of manpower (like Italy, Spain, Portugal and Greece) shifting to also become areas of attraction of international migration headed to the EU (Arango and Finotelli, 2009).[4] Table 1 reports data on international migration taken from the United Nation's Population Prospects for the EU-27 between years 1995 and 2010. A glimpse at this table shows that the annual growth rate in international migration stock for Southern European countries, of 8.1%, is, by far, the highest recorded throughout the period 1995-2010.

[4] Other relevant changes in migration patterns in Europe include, for instance, the diversification of migration origins, the increase of undocumented migration, increasing importance of family reunification, high demand for immigrant labour in service sector activities and the role of female immigrants. For more details, see for instance Delgado (2002).

Table 1 International migration in the EU-27

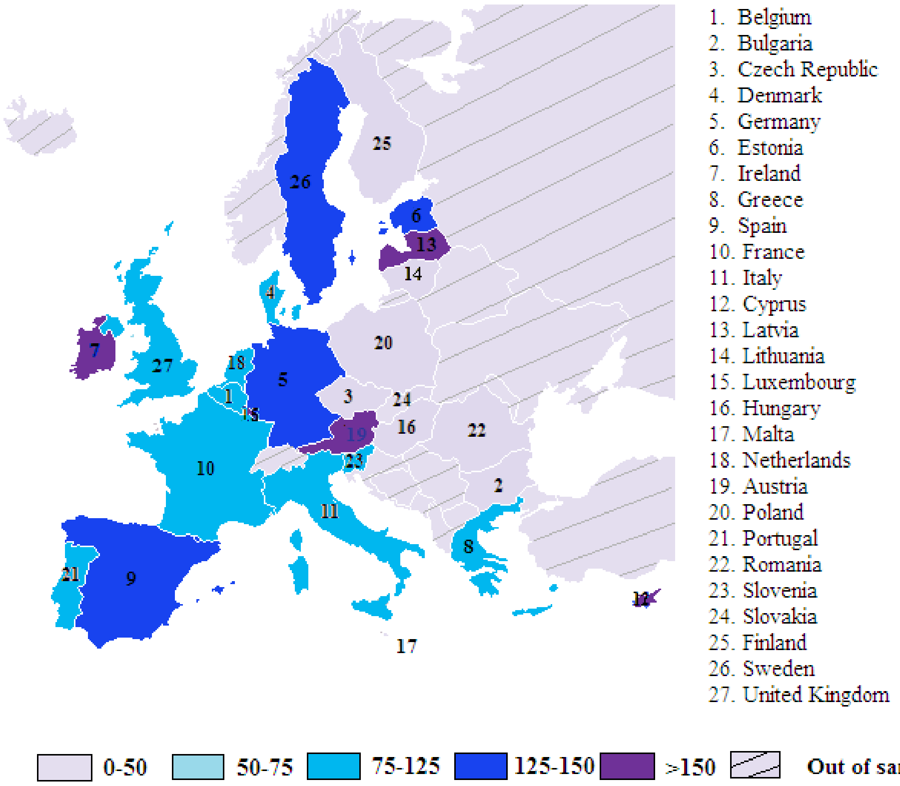
	1995		2010		Annual growth rate (%)
	Thousands	Per capita (%)	Thousands	Per capita (%)	
Northern	6,869	7.8	10,012	10.7	2.5
Southern	4,104	3.4	13,226	10.1	8.1
Eastern	2,006	2.1	2,019	2.2	0.0
Western	18,495	10.7	21,654	12.0	1.1
EU-27	**31,474**	**6.6**	**46,911**	**9.4**	**2.7**

Source: *United Nations Population Division*, 2009.

Note: Northern EU-27 Member States: Denmark, Estonia, Ireland, Latvia, Lithuania, Finland, Sweden, The United Kingdom. Southern EU-27 Member States: Cyprus, Greece, Italy, Malta, Portugal, Slovenia, Spain. Eastern EU-27 Member States: Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovakia. Western EU-27 Member States: Austria, Belgium, France, Germany, Luxembourg, Netherlands.

Even so, international migration numbers in Southern European countries, around 13.2 million persons in 2010, are still far from those recorded in Western Europe, around 21.6 million persons, albeit they already exceed numbers for Northern European countries, of around 10 millions. In per capita terms, however, we can observe that differences are not so large (see, also Figure 1); per capita international migration represents 10.1% of the total population in Southern European countries, while in Western Europe this percentage reaches 12%.

Figure 1 Per capita international migration in the EU-27's Member States, year 2010 (EU-27=100)



As far as this issue is concerned, Spain, the European country in which the rest of the paper will have its focus, is by far the Southern European country where international migration has reached really massive proportions. Throughout most of the 20th century, international emigration was one of the main factors determining the Spain's population dynamics. However, the late 1990s ushered a new demographic phase for Spain. Massive immigration, coming largely from Europe, Latin America and North Africa,[5] transformed Spain from sending to receiving country (Hierro, 2011). Table 2 reports the evolution of foreign resident population in Spain between 1998 and 2009, both in absolute and per capita numbers (relative to population size). According to this table, foreign-born population has increased strongly, from 637 to 5,649 thousand persons. This increase was especially marked between 2001 and 2004; according to Hierro (2011), large migration inflows coming from some South American countries benefited from visa-free entry into Spain, mainly Ecuador and Colombia, propelled the increase. It is also worth noting that its share in the total

[5] According to INE data, in 2009 44.2% of the foreign-born population in Spain came from Europe, 32.1% from Latin American countries and 17.9% from Africa.

population has experienced an even more impressive growth, from 1.6% in 1998 to 12.1% in 2009.

Table 2 International migration in Spain

Year	Foreign-born population (thousands)	Per capita (%)
1998	637	1.6
1999	749	1.9
2000	924	2.3
2001	1,371	3.3
2002	1,978	4.7
2003	2,664	6.2
2004	3,034	7.0
2005	3,731	8.5
2006	4,144	9.3
2007	4,520	10.0
2008	5,269	11.4
2009	5,649	12.1

Source: INE data.

Among the main reasons for the sharp rise in international migration in Spain are: (1) economic opportunities derived from Spain's entry into the European Community in 1986 and vigorous economic growth during the late 1990s and early 2000s, (2) a rather lenient immigration policy (including visa exemption accords for tourists from some non-European countries and relatively easy channels for regularization and naturalization), (3) high demand for migrant labour resulting from a scarcity of local labour in low-skilled and low-paid activities (largely in agriculture, construction and service sectors) and (4) a reduction of the salary gap between Northern and Southern European countries (Fassmann and Münz, 1992; King et al., 1997; Bruquetas et al., 2008; Hierro, 2011). However, in dealing with main reasons explaining the high increase observed in international migration numbers, we must not disregard the effect of periodical regularization programs conducted in Spain in 1986, 1991, 1996, 2000, 2001 and 2005. As indicated by Bruquetas et al. (2008), regularizations have been the primary avenue for providing legal status to immigrants residing illegally in Spain. In this vein, a large number of undocumented immigrants residing in Spain have become legally residents and, therefore, "visible" for official statistics. As can be seen in Table 3, over the 1986-2005 period the Spanish authorities received almost 1.5 million applicants for regularization, the coverage ratio reaching 74 per cent. In addition, high numbers for the regularization of 2005 reflect clearly that, as indicated by Arango and Finotelli (2009), irregular immigration has become a structural characteristic of the Spanish migration regime.

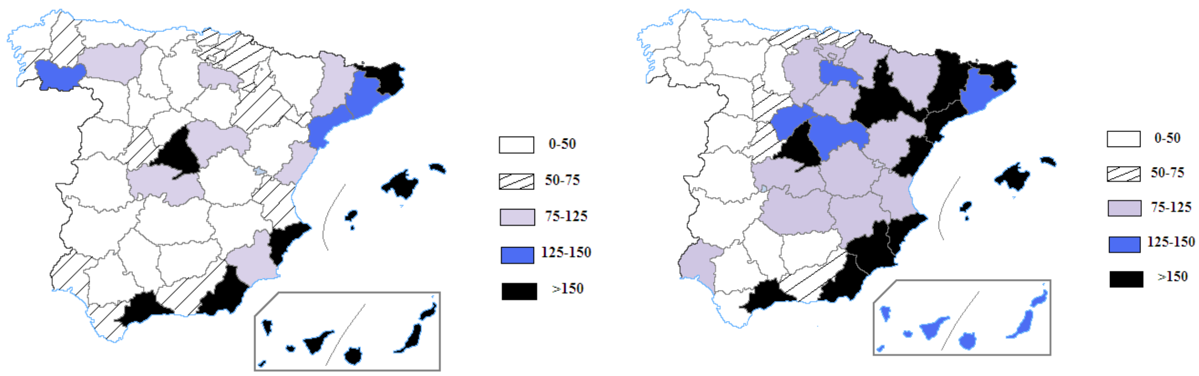
Table 3 Regularization programs in Spain

Regularization campaign	Appliances (thousands)	Acceptance (thousands)	Coverage ratio %
1986	44	38	87.1
1991	130	108	83.0
1996	25	21	84.7
2000	246	137	56.2
2001	350	216	61.8
2005	692	578	83.6
1986-2005	**1,487**	**1,100**	**74.0**

Source: Ministry of Labour and Social Affairs.

Concerning the settlement patterns of international migration across the Spanish provinces, a look at Spain's map in both 1998 and 2009 years (Figure 2) reveals that it has been prone to cluster in areas characterized by economic dynamism and high demand for migrant labour, such as the Mediterranean and South-Eastern coast, the Ebro Valley, as well as Madrid and its area of influence. Besides this, concentration of foreign-born residents in both Balearic and Canary Islands also deserves to be mentioned; in this case, however, the bulk of them are non-labour immigrants (i.e. people retiring to Spain for climatic and lifestyle reasons) coming largely from Northern and Western Europe. Interestingly, international migration has also tended to concentrate, although in a lesser extent, in some provinces with low economic dynamism but located in the area of influence of the Mediterranean coast; among some reasonable reasons, we might think, for instance, in factors like geographic closeness, less-saturated labour markets and lower residential costs.

Figure 2 Per capita international migration in the Spanish provinces (Spain=100) (%)



a) Year 1998 b) Year 2009

In view of this figure, the possibility that international migration is not randomly distributed across the Spanish provinces can not be ruled out. In the following sections we will further explore this point by providing new insights into the spatial pattering of international migration in Spain.

3 Characterizing the SIMD

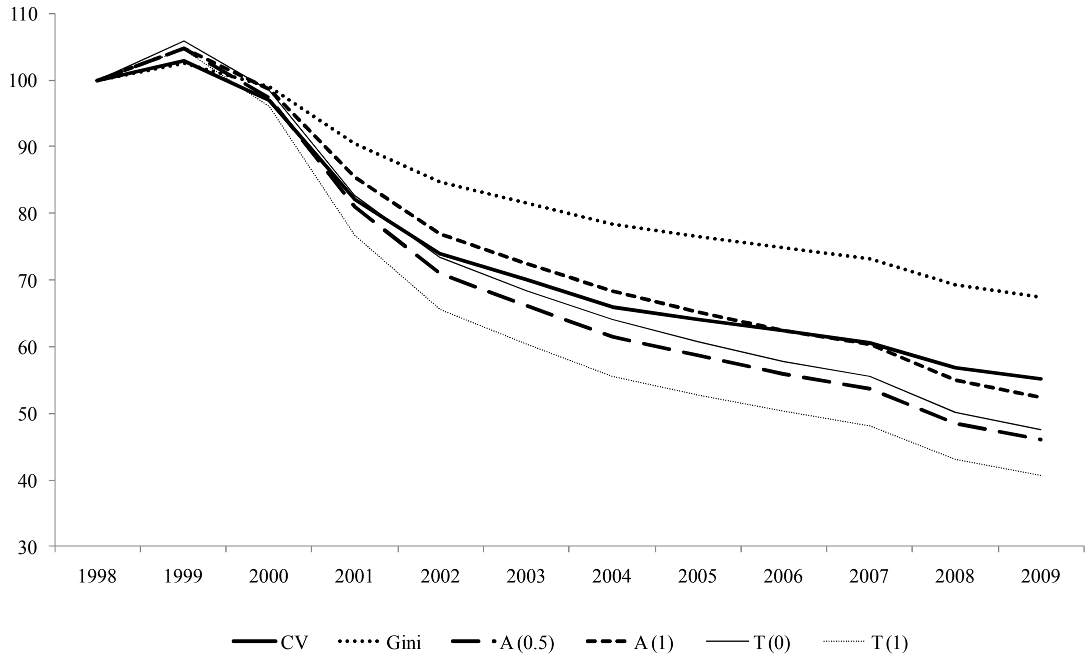
In this section we set the stage by examining, for the sample period 1998-2009, some general characteristics of the per capita international migration distribution in Spain concerning inequality levels, external distribution shape, polarization degree and spatial association.

3.1 Inequality

We start our analysis by examining the evolution of per capita international migration disparities across the Spanish provinces. For the sake of comparison, we have considered, among the whole array of inequality indicators, the Coefficient of Variation (CV), the Gini coefficient, the Atkinson indexes $A(0.5)$ and $A(1)$, and two Theil indexes $T(0)$ and $T(1)$.

Figure 3 depicts, taking the value of all the measures in 1998 equal to 100, the evolution of inequality over the 1998-2009 period. It is first worth noting that all the scalar indicators follow the same trend, indicating a small rebound in 1999 followed by a declining tendency thereafter. Particularly, inequality has declined, depending on the inequality measure considered, in the range of 41-67% between 1998 and 2009. In addition, it can be seen that reduction of inequality has been a bit more pronounced in both the early and late 2000s.

Figure 3 Per capita international migration inequality in Spain (1998=100)



3.2 External shape

Additional insights into per capita international migration distribution in Spain can be obtained from the construction of a density function. This graphical tool, which can be understood as being a smoothed version of a histogram, provides a good approximation to the external shape and, furthermore, on some relevant characteristics of the distribution, such as possible distribution clustering. Comparison of a density function at different points in time allows us, in addition, to obtain a general idea on the law of motion of the distribution, i.e., how it evolves over time.

Following Silverman (1986), a density function can be estimated by a sum of kernel functions at each data points of the sample, that is:

$$\hat{f}(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{y-y_i}{h}\right) \quad (1)$$

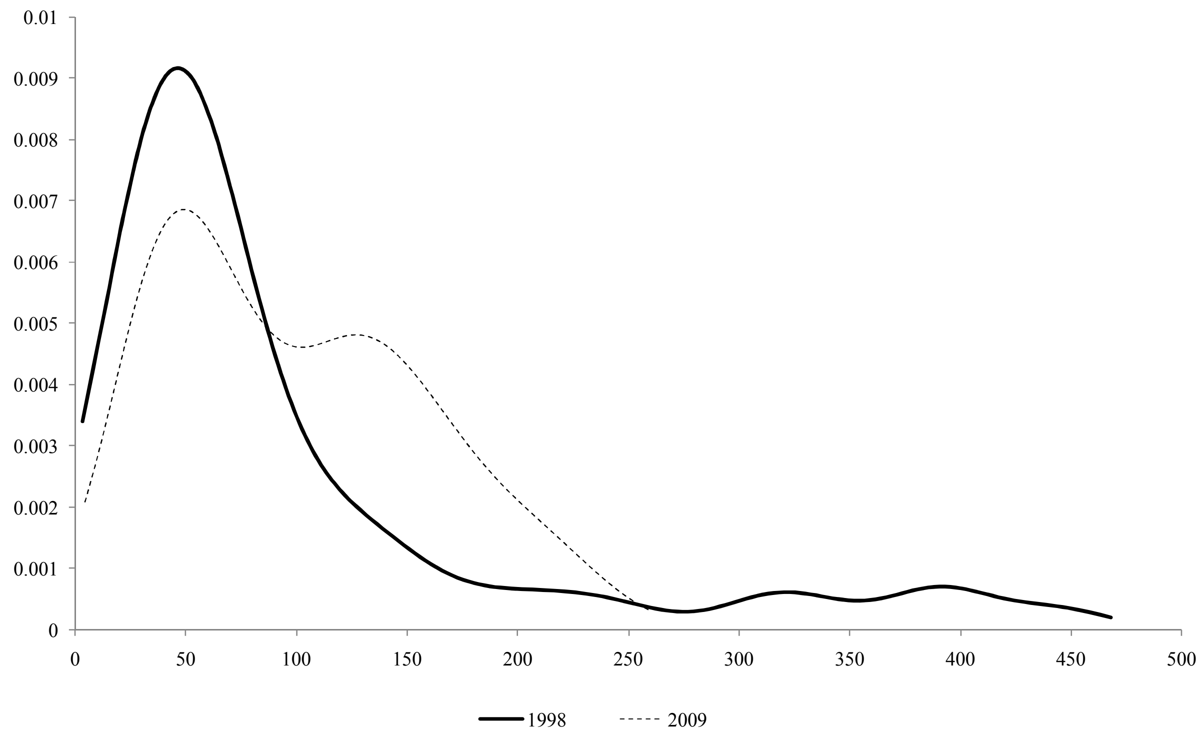
where y is the point at which the density function is being evaluated; y_i is province i's per capita international migration; n is the number of data points (in our case, provinces); h is the bandwidth or smoothing parameter that controls the smoothness of the estimated densities; and $K(\cdot)$ is the Kernel function. In this study we have used a Gaussian kernel function[6] with optimal bandwidth according to Silverman's rule-of-thumb (Silverman, 1986) defined as:

$$K(y) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}y^2\right) \tag{2}$$

Figure 4 plots the SIMD for the initial and final years of the sample. Results are based on values for each provinces expressed in relative terms with the national average as 100. The figure shows that some significant shifts have occurred in the external shape of the distribution between 1998 and 2009. Firstly, the distribution has gradually moved to the average values, this being a clear sign of reduction of disparities in per capita international migration across the Spanish provinces. Secondly, another noteworthy characteristic is that the prominent upper tail visible at the beginning of the sample period, representing high-immigration provinces, has progressively vanished, the distribution evolving to a bimodal distribution characterized by twin peaks. Following the proposal by Salgado-Ugarte et al. (1997) for non-parametric assessment of multimodality, the position of the first peak (located at 47.9 in 1998 and at 47.7 at 2009) has hardly changed, while the second one has approached to the average, this one being located at 124.1 in 2009.

[6] Although multiple options for the kernel selection exist (uniform, Gaussian, Epanechnikov, triangular, rectangular, bi-weight, etc.), the Gaussian kernel is probably the most common choice.

Figure 4 Density plots for relative per capita international migration in Spain



3.3 Polarization

While a glance at density functions seems to point to some polarization in the distribution, it seems appropriate to complement this graphical tool, if possible, by some scalar indicator that provides direct evidence on the degree to which provinces cluster around a set of migration poles. For it, and among the proposals made in the field of polarization measurement, we resort to the polarization measure proposed by Esteban et al. (2007) –henceforth EGR–, as being undoubtedly the most widely used in empirical analysis.

For it, let us consider an international migration distribution defined by f, and a partition for it that defines k non-overlapping groups as intervals of per capita international migration $[z_{i-1}, z_i]$, with $i = 1, \dots, k$. The EGR polarization measure can be expressed as:

$$EGR(\alpha, \beta) = \sum_{i=1}^{k} \sum_{j=1}^{k} p_i^{1+\alpha} p_j \left| \frac{y_i}{\mu} - \frac{y_j}{\mu} \right| - \beta(G - G_S) \quad (3)$$

where, for the purpose of the present study, p_i and p_j denote population shares for groups i and j; y_i and y_j are per capita international migration for groups i and j; μ is the Spanish average per capita international migration; α is a parameter measuring the degree of sensitivity of the index to polarization that falls in the interval $[1, 1.6]$; and G and G_S are the Gini coefficients of the original (i.e. ungrouped) and the grouped distribution, respectively. The first part in Equation (3) is commonly referred to as *simple polarization* $ER(\alpha, \beta)$, while the second part represents the *specification error*[7] ε –or *lack of identification*– modulated by a parameter $\beta \geq 0$ that reflects the sensitivity of the index to the groups' level of cohesion. Accordingly, polarization in a distribution may increase either because increasing heterogeneity between groups, i.e. higher simple polarization, or because, as a result of higher homogeneity within groups, these ones become more identified.

Before going any further, it is necessary to clear up some points concerning the computation of the EGR measure. First, we only consider the case of 2 and 3 groups. Second, optimal partition of the distribution in a given number of groups has been obtained following the methodology proposed by Davies and Shorrocks (1989).[8] Third, as the choice of values of α being somewhat arbitrary, we have considered $\alpha = 1$ and $\alpha = 1.5$. Finally, regarding the value of β, as there is general agreement that it must be close to 1, we have chosen $\beta = 1$.

Bearing these considerations in mind, a substantial decrease in polarization degree has occurred over the sample period irrespective of the number of groups considered (Table 4). This decline has been more intense between 2003 and 2009. Table 5 provides further insights into the role played by polarization between groups and homogeneity within them. We observe that, whereas polarization between groups has decreased, albeit at a slower rate than polarization as a whole, groups have become more identified, this being especially so between 1998 and 2003.

[7] Definition of this measure needs the distribution to be previously pre-arranged in k groups and then replacing migration data within a group by the group mean. As some intra-group dispersion is to be expected, partition of the distribution is likely to cause a loss of distributional information and therefore to induce an approximation error.

[8] Following Davies and Shorrocks (1989), optimal grouping is that partition which minimizes the specification error and therefore intra-group dispersion.

Table 4 Polarisation between Spanish provinces $EGR(\alpha = 1, \beta = 1)$

	$EGR(\alpha = 1, \beta = 1)$		$EGR(\alpha = 1.5, \beta = 1)$	
	$k = 2$	$k = 3$	$k = 2$	$k = 3$
1998	0.191	0.203	0.095	0.104
2003	0.171	0.153	0.088	0.062
2009	0.111	0.107	0.049	0.034

Source: Own elaboration from INE data.

Table 5 Polarization by components: simple polarization (ER) and lack of identification (ε)

	$ER(\alpha = 1)$		$ER(\alpha = 1.5)$		ε	
	$k = 2$	$k = 3$	$k = 2$	$k = 3$	$k = 2$	$k = 3$
1998	0.336	0.283	0.240	0.183	0.145	0.080
2003	0.282	0.222	0.199	0.131	0.111	0.069
2009	0.217	0.174	0.156	0.101	0.107	0.067

Source: Own elaboration from INE data.

3.4 Spatial association

Since the 1990s or even earlier, an expanding body of work in regional economics has shown that space matters when analyzing many economic phenomena (Anselin, 1995; Fingleton, 2001; Rey, 2001; Arbia and Paelinck, 2003; Maza and Villaverde, 2009; Dall'erba, 2005; Durlauf et al., 2005; Rey and Janikas, 2005; Fingleton and López-Bazo, 2006; Bosker, 2009), this encouraging many researchers to change their view of geographical units as isolated areas and see themselves as interdependent geographical units conditioned by the spatial context.

From a statistical standpoint, the existence of spatial association implies that data are not randomly distributed across geographical units. This obviously may have serious implications for economic modeling, as if spatial dependence exists and the model does not explicitly account for it, results might be unbiased by spatial dimension of data. As for the case of international migration, if we assume spatial association between provinces, the relative location of a province may affect its international migration performance. Taking this in mind, it seems mandatory to have a detailed look at spatial association. In order to do it, we have turned to some traditional measures of spatial association provided by the ESDA: the Moran's I statistic (Moran, 1948) and the Geary's C statistic (Geary, 1954).

To perform this analysis, it is necessary to pre-define the spatial connectivity structure of the sample (LeSage et al., 2007). For it, and following Le Gallo (2004), we define a spatial weights matrix W (in our case of dimension 50×50) with elements w_{ij}, called spatial weights, defined as:

$$w_{ij} = \begin{cases} 0 & if\ i = j \\ s_{ij}^{-2} & if\ i \neq j \quad d_{ij} \leq S_1 \\ 0 & if\ i \neq j \quad d_{ij} > S_1 \end{cases} \tag{4}$$

where s_{ij} represents the distance between the centroids of provinces i and j, and S_1 is a cutt-off point defined as the lower quartile of the great circle distance distribution.[9] As the way in which spatial weights are defined, spatial interactions between provinces decay inversely with

[9] While the definition of a cut-off point is somewhat arbitrary, it permits to restrict neighbouring influence to a reasonable great circle distance and therefore elude the neighbouring influence of far enough provinces. In our case, the cutt-off point is equal to 275.729 kilometres.

the quadratic distance. By convention, each element of the matrix is divided by its row sum, giving rise to a standardized spatial weights matrix W^*, with elements w^*_{ij}. [10]

Then, we start the analysis by computing the Moran's I statistic. Specifically, the Moran's I statistic can be expressed for a period t as (Anselin, 1988):

$$I(t) = \frac{n}{\sum_i \sum_j w^*_{ij}} \frac{\sum_i \sum_j w^*_{ij}[y_i(t)-\mu(t)][y_j(t)-\mu(t)]}{\sum_i [y_i(t)-\mu(t)]^2} \tag{5}$$

where, in the context of this paper, $y_i(t)$ and $y_j(t)$ are the per capita international migration of provinces i and j at time t; $\mu(t)$ is the Spanish average per capita international migration; and w^*_{ij} are the standardized spatial weights. In order to facilitate the interpretation of the statistic, the standardized value (z −value) is obtained.[11] Accordingly, a significant positive (negative) value for the Moran's I statistic will imply positive (negative) spatial association, herein interpreted to imply similar (dissimilar) values of per capita international migration being clustered together in space.

In addition, and for the sake of robustness, we have computed the Geary's C statistic:

$$C(t) = \frac{n-1}{2\sum_i \sum_j w^*_{ij}} \frac{\sum_i \sum_j w^*_{ij}[y_i(t)-y_j(t)]^2}{\sum_i [y_i(t)-\mu(t)]^2} \tag{6}$$

For interpretation purposes, the standardized value of the Geary's C statistic is also obtained, so that a significant negative (positive) z −value for the statistic will imply positive (negative) spatial association.

Table 6 displays the results of computing both Moran's I and Geary's C statistics for the years 1998 to 2009. As can be seen, the sign of the z −value is positive for Moran's I and significant at the 5% level for all periods. As for the Geary's C statistic, we obtain negative

[10] Accordingly, neighbouring provinces are those sharing a border, but also those within a critical distance of each other. As such, we have discarded contiguity-based spatial weights, as some provinces, despite not sharing a border, can exert influence on another (see Le Gallo, 2004). In addition, that kind of criteria is unfeasible when islands are included into the analysis, as that is the case. For alternative definitions of distance-based spatial weights, see Anselin (1988) and Florax and Rey (1995).

[11] The standardized *z*-value is obtained by subtracting the expected value for the statistic, and then dividing the result by the corresponding standard deviation.

and highly significant values for all the years of the sample period.[12] Therefore, the results indicate positive spatial association regardless of the statistic used. We can also observe increasingly higher values for both statistics over the sample period, this piece of evidence suggesting that the geographical context has played an increasing role in the distribution.

Table 6 Moran's I and Geary's C statistics

t	Moran's I statistic			Geary's C statistic		
	Value	z −value	p −value	Value	z −value	p −value
1998	0.163	2.313	0.021	0.686	-3.847	0.000
1999	0.154	2.205	0.027	0.696	-3.729	0.000
2000	0.186	2.602	0.009	0.667	-4.076	0.000
2001	0.253	3.455	0.001	0.620	-4.652	0.000
2002	0.312	4.193	0.000	0.578	-5.168	0.000
2003	0.363	4.844	0.000	0.540	-5.639	0.000
2004	0.404	5.354	0.000	0.519	-5.894	0.000
2005	0.423	5.598	0.000	0.504	-6.083	0.000
2006	0.438	5.783	0.000	0.494	-6.198	0.000
2007	0.454	5.989	0.000	0.477	-6.405	0.000
2008	0.463	6.106	0.000	0.473	-6.460	0.000
2009	0.467	6.154	0.000	0.470	-6.494	0.000

Source: Own elaboration from INE data.

4 Exploring geographic contagion in SIMD: A spatial Markov chain approach

Results obtained in the previous section provided evidence on the existence of positive and significant spatial association. This being so, we might question whether some type of contagion force underlies spatial interactions in international migration levels across provinces; namely, whether high (low) per capita international migration levels to be located in neighbouring provinces increases the likelihood of having higher (lower) levels in a province. It is undoubtedly that the existence of some kind of "imitation" is a question of no minor importance, as it might provide some hint on the influence of social networks on

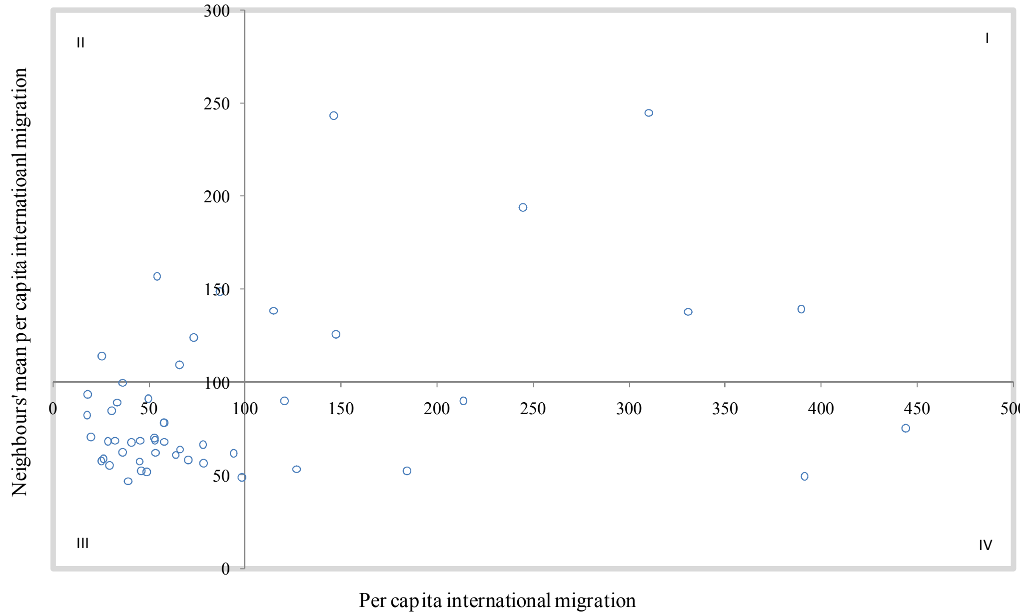
[12] Significance of both Moran's I and Geary's C statistics was based on the assumption that, following Anselin (1992), the standardized statistic follows a normal distribution. For the sake of robustness, we additionally applied two alternative approaches (the randomized and permutation approaches), but the results were roughly the same.

mapping international migration in Spain, as well as on the strength of inter-provincial ties pertaining to immigration issues.

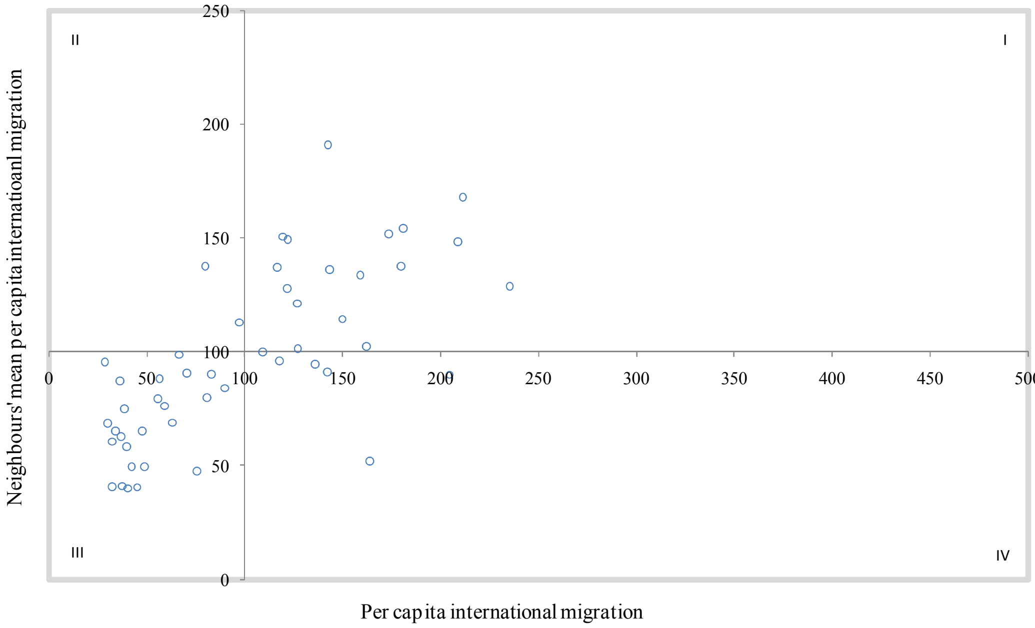
A straightforward way to detect some kind of contagion process consists on direct comparison, in two different time periods, of provinces' locations in the Moran scatter plot. This representation originated as a way of visualizing more intuitively spatial association derived from the computation of the Moran's I statistic (Anselin, 1995). In order to construct it, it is necessary to previously define the concept of *spatial lag*. For it, let suppose that per capita international migration for a province i at time t is given by $y_i(t)$. Its spatial lag is then defined as the weighted sum of each province's per capita international migration, with weights the standardized spatial weights previously defined, that is, $y_i^*(t) = \sum_j w_{ij}^* y_j(t)$. Accordingly, the Moran scatter plot is a plot with the variable of interest –per capita international migration– on the $x-$axis and the spatial lag on the $y-$axis, the slope of the linear regression line providing the Moran's I statistic. Its interpretation is straightforward: quadrants I and III (the upper right and lower left, respectively) represent positive spatial association, whereas quadrants II and IV (the upper left and lower right, respectively) negative spatial association.

By way of illustration, Figure 5 displays the Moran scatter plot obtained for the initial and final years of the sample period. The scatter plot is centered on the mean, so that the position of each point makes sense. As it is clear from the figure, positive spatial association exists and its intensity is higher in 2009 than in 1998.

Figure 5 Moran scatter plots



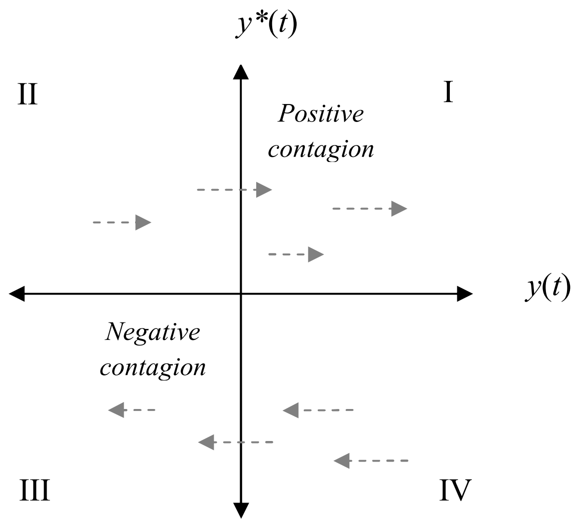
a) Year 1998



b) Year 2009

Intuitively, when comparing Moran's scatter plots depicting spatial association patterns at the initial and final period we might distinguish between two kinds of contagion effect: *positive contagion* and *negative contagion*. As illustrated in Figure 6, positive contagion is likely the most reasonable to expect and occurs when provinces' location changes through quadrant II to I between the two periods considered, that is, provinces with initially low per capita international migration, but surrounded by provinces with high per capita international migration, move towards higher per capita international migration levels. Conversely, negative contagion happens when provinces' location changes through quadrant IV to III, that is, provinces with initially high per capita international migration, but surrounded by provinces with low per capita international migration, move towards lower per capita international migration levels.

Figure 6 Positive and negative contagion



Although comparison of Moran scatter plots in Figure 5 invites to think in some kind of contagion effect, a simple comparison of Moran scatter plots is not enough to draw definite conclusions concerning contagion in the distribution. This is because this approach has several drawbacks. Firstly, as indicated by Messner and Anselin (2004), it only provides a static comparison between two periods, so that the dynamic nature of a contagion process is not captured directly in the analysis. Secondly, and this is probability the main drawback, it fails to provide a quantitative measure of contagion. In order to overcome such shortcomings, in this paper we resort to the spatial Markov chain approach, and more specifically to the proposal by Rey (2001). This consists basically on an extension of the classical Markov chain framework, in which current per capita international migration for a province is doubly conditioned on its per capita international migration in a previous period and the per capita international migration of its neighbouring provinces. Thereby, the researcher can analyze not only how geographical areas move over time across different per capita international migration levels, commonly known as states and hereafter referred to as migration states, but also the geographical dimension of these movements (Rey, 2001; Bosker, 2009).

To perform the analysis, let us suppose that provinces are first grouped into two groups based on the spatial lag variable $y^*(t)$, herein after referred to as spatial lags and denoted by k: provinces with neighbours' mean per capita international migration below the national average ($k = 1$) and provinces with neighbours' mean per capita international migration above the national average ($k = 2$). For each spatial lag, let suppose that the original international migration distribution y is divided into an exhaustive finite set of m mutually exclusive migration states and that $X_k(t)$ indicates the migration state occupied at time t given the spatial lag k.[13] Then, for a given time period $(t, t + s)$, we can define two spatial transition matrices, $P_k(t, t + s)$, with dimension $m \times m$ and entries:

$$p_{ij|k}(t, t + s) = Pr[X_k(t + s) = j | X_k(t) = i] \quad (7)$$

for all i, j, k, representing the spatially conditioned transition probabilities or spatial transition probabilities, namely the probability of transition from a migration state i to another j between t and $t + s$, given that the spatial lag was k at t. Accordingly, these probabilities measure how a province's position in the relative international migration distribution changes over time depending on the spatial lag considered. This means, for instance, that if we consider the migration states i and j representing low-immigration and high-immigration provinces, respectively, $p_{ij|2}(t, t + s)$ will define the probability that low-immigration provinces surrounded by above-average immigration provinces have to move up towards high-immigration provinces.

The main aspect that makes a spatial Markov chain approach especially appealing in this context is that it offers the possibility of mapping the mobility information contained in a spatial transition matrix into a number of metrics that summarize the degree of mobility in the distribution taking, additionally, provinces' spatial location into account. Although spatial mobility is not exactly what we want to measure, it is quite close, as far as a contagion

[13] Although discretization of the distribution involves an unavoidable degree of arbitrariness, the advantage of a discrete approach over a continuous one is that it allows us to define specific contagion metrics.

process implies spatially conditioned mobility of provinces towards higher or lower immigration levels. According to this latter, the difference is that instead of spatial mobility overall, a contagion process pertains to upward and downward spatial mobility.

Taking the above considerations into account, we propose two measures of positive and negative contagion, respectively. These measures have been built on the foundation of the mobility measure proposed by Maza et al. (2010)[14] for measuring mobility in income distributions, but properly adapted to capture spatial and upward/downward mobility considerations. For a given lag $k = 2$ the *positive contagion index* for the period $(t, t+s)$, hereafter denoted as C^{+}, is defined as:

$$C_2^{+}(t, t+s) = \sum_i \sum_{j>i} p_{i|2}(t)\, p_{ij|2}\,(t, t+s)\, \frac{d_{ij|2}(t)}{c_{i|2}(t)} \tag{8}$$

where $p_{i|2}(t)$ is the initial proportion of regions in *i*'s migration state for a spatial lag $k = 2$ at time t; $p_{ij|2}(t, t+s)$ denotes, as we have already indicated, the probability of moving from a migration state i to another j between t and $t+s$ for a given spatial lag $k = 2$; $d_{ij|2}(t)$ is, for a given spatial lag $k = 2$, a distance measure between migration states i and j at time t, defined as $d_{ij|2}(t) = \left|\bar{y}_{j|2}(t) - \bar{y}_{i|2}(t)\right|$, namely the absolute difference between the average per capita international migration of the states under consideration; and, finally, $c_{i|2}(t)$ is introduced into the expression in order to normalize the contagion index, and defined as the largest value in the i's row of $D_2(t)$ (a distance matrix with generic elements $d_{ij|2}(t)$).

Some points must be pointed out regarding this measure. Firstly, as a basic premise for positive contagion is that provinces are surrounded by above-average immigration provinces, this measure is defined for $k = 2$. Secondly, summation is defined over states $j > i$ in order to only include probabilities in the upper (off-diagonal) triangle part of the transition matrix, and therefore it only takes in upward movements. Finally, the index is perfectly bounded between 0 and 1; particularly, the closer its value to 1, the higher the positive contagion effect is.

As can be noted, the metric defined above measures positive spatial contagion by way of considering the influence of: (1) neighbouring provinces' international migration, (2)

[14] For more details about advantages coming from this measure over other measures proposed in this strand of the literature, see Maza et al. (2010).

upward transitions between migration states, (3) the size of migration states and (4) how far each other migration states are.

Likewise, we define the *negative contagion index* $C_k^-(t, t+s)$ for a given lag $k = 1$ as:

$$C_1^-(t, t+s) = \sum_i \sum_{j<i} p_{i|1}(t) p_{ij|1}(t, t+s) \frac{d_{ij|1}(t)}{c_{i|1}(t)} \quad (9)$$

This measure differs in three respects from that in equation (8): Firstly, summation is defined over states $j < i$ in order to only include probabilities in the lower (off-diagonal) triangle part of the transition matrix, and therefore it only considers downward movements. Secondly, as for the case of negative contagion provinces must be surrounded by below-average immigration provinces, so that this measure is defined for $k = 1$. Finally, although bounds being the same, the closer its value to 1, the higher the negative contagion effect is.

Before proceeding to the empirical results, some remarks must be made because, as it seems obvious, metrics defined above critically depend on the definition of migration states and the transition period length. Regarding the first point, we have defined seven exhaustive and mutually exclusive per capita migration states: [0, 50), [50, 75), [75, 90), [90, 110), [110, 125), [125, 150), [150, $+\infty$). As for the transition period length, we opted for estimating a five-year transition probability matrix.

Table 7. Spatial transition matrices $P_k(t, t+5)$ and positive and negative contagion indexes (1998-2009)

Spatial lag $\boldsymbol{k}$	$\boldsymbol{i \setminus j}$	**[0, 50]**	**(50, 75]**	**(75, 90]**	**(90, 110]**	**(110, 125]**	**(125, 150]**	**(150, ∞)**	$\boldsymbol{p_{i\|1}(t)}$	$\boldsymbol{C_1^-(t, t+s)}$
	[0, 50]	0.734	0.216	0.029	0.022	0.000	0.000	0.000	0.519	
	(50, 75]	0.328	0.293	0.207	0.103	0.052	0.017	0.000	0.216	
	(75, 90]	0.077	0.308	0.154	0.154	0.077	0.154	0.077	0.049	
1	**(90, 110]**	0.133	0.133	0.000	0.133	0.333	0.200	0.067	0.056	**0.031**
	(110, 125]	0.000	0.167	0.000	0.000	0.000	0.167	0.667	0.022	
	(125, 150]	0.000	0.143	0.000	0.286	0.429	0.143	0.000	0.026	
	(125, ∞)	0.000	0.000	0.000	0.000	0.000	0.067	0.933	0.112	

Spatial lag $\boldsymbol{k}$	$\boldsymbol{i \setminus j}$	**[0, 50]**	**(50, 75]**	**(75, 90]**	**(90, 110]**	**(110, 125]**	**(125, 150]**	**(150, ∞)**	$\boldsymbol{p_{i\|2}(t)}$	$\boldsymbol{C_2^+(t, t+s)}$
	[0, 50]	0.167	0.333	0.167	0.333	0.000	0.000	0.000	0.045	
	(50, 75]	0.136	0.318	0.182	0.273	0.091	0.000	0.000	0.167	
	(75, 90]	0.000	0.000	0.333	0.444	0.111	0.000	0.111	0.068	
2	**(90, 110]**	0.000	0.000	0.000	0.077	0.769	0.077	0.077	0.098	**0.148**
	(110, 125]	0.000	0.000	0.000	0.000	0.400	0.500	0.100	0.076	
	(125, 150]	0.000	0.000	0.000	0.000	0.000	0.214	0.786	0.106	
	(125, ∞)	0.000	0.000	0.000	0.000	0.000	0.138	0.862	0.439	

Source: Own elaboration from INE data.

Table 7 displays the two spatial transition matrices for the period 1998-2009. A quick glance to the first one shows that while negative contagion among some provinces surrounded by below-average immigration provinces exists (see, for instance cells [2, 1], [3, 2], [6, 5]), it does not seem very high. By contrast, results for provinces surrounded by above-average immigration provinces point clearly to the existence of high positive contagion; looking at the second spatial matrix, we see that upward spatial transition probabilities are considerably higher than downward spatial transition probabilities as a whole.

Some caution is needed, however, when interpreting these results. As indicated in this Section, precise conclusions on the scale of contagion degree calls for further information on the size of migration states and the distance between them. Accordingly, and based on previous estimations, Table 7 also presents the value obtained for contagion indexes C_1^- and C_2^+. For a correct interpretation, it is worth clarifying that a situation of maximum negative (positive) contagion, given by $C_1^- = 1$ ($C_2^+ = 1$), which arises when all transitions are downward (upward) movements towards the more distanced migration state, is almost impossible to occur. Thus, for a correct interpretation of the results we have carried out several simulations to distinguish between high, medium and low negative (positive) contagion degrees. On the basis of these simulations and to ease interpretation, we adopt the following criteria: we label a situation as "high negative contagion degree" if C_1^- is over 0.064 (obtained when 60% of the provinces move downwards to a contiguous state); "medium mobility degree" if C_1^- is between 0.064 and 0.033 (between 60% and 30% of the provinces move downwards to a contiguous state); and "low mobility degree" if C_1^- is below. As for positive contagion: we label a situation as "high positive contagion degree" if C_2^+ is over 0.093 (obtained when 60% of the provinces move upwards to a contiguous state); "medium mobility degree" if C_2^+ is between 0.093 and 0.047 (between 60% and 30% of the provinces move upwards to a contiguous state); and "low mobility degree" if C_2^+ is below. In view of that, the value obtained for C_1^-, of 0.031, indicates a low negative contagion within the distribution. In addition, the value obtained for C_2^+, of 0.148, reveals, as it was expected, a high positive contagion within the distribution.

5 Conclusions

In this paper we explored the international migration distribution across the Spanish provinces over the 1998-2009 period. In order to conduct the study, after a descriptive analysis we examined some relevant characteristics of the distribution. Thus, we were interested in computing inequality levels, the results indicating a significant reduction of provincial disparities, especially in the early and late 2000s. In addition, we analyzed the external shape of the distribution by means of a non-parametric analysis, revealing some distribution clustering around low and high values. We were also interested in examining polarization degree by means of the polarization measure proposed by Esteban et al. (2007), the results revealing that, while substantial decrease in polarization levels has occurred, groups have become more identified. Finally, we paid attention to spatial association in the distribution, obtaining that the geographical context has played an increasing role in the distribution.

In the second part of the study we tried to gain new insights into the existence of geographical contagion effects in the distribution. In order to do it we proposed, using a spatial Markov chain approach, two new measures of positive and negative contagion. The results obtained do reveal that some contagion force prominently underlies the international migration distribution in Spain and, furthermore, that positive contagion among provinces surrounded by high immigration provinces is the most significant one. This result has profound social, economic and political implications. Clearly, positive contagion can receive either a positive or negative interpretation, depending upon immigration is perceived as having positive demographic and economic benefits for a territory or as being an economic and cultural threat to the native population. On the other hand, it is also clear that as positive contagion progress further towards geographically close territories, interdependencies among neighboring provinces become increasingly higher and complex. This being so, even though national authorities exercise the main competences in immigration issues (i.e. migratory flows control, citizenship, legal status), higher cooperation in the field of social integration to combat social exclusion of immigrants in society (with initiatives pertaining to employment, education, health care, housing, social services or civic participation) would be highly recommended as regional and local authorities, responsible for the policy measures involved in the sphere of integration, might benefit sharing experience, knowledge and best practices. In such a way, assuming that sub-national authorities being guided regardless of their political color by the same principles and goals in this realm, it seems clear that a common workspace for integration oriented-discussion would contribute to increase the capacity for response to multiple challenges posed by immigration.

References

Anselin L (1988). *Spatial econometrics: Methods and models*. Kluwer: Dordrecht.

Anselin L (1992). *SpaceStat tutorial: A workbook for using SpaceStat in the analysis of spatial data*. Illinois, University of Illinois.

Anselin L (1995). "Local Indicators of Spatial Association-LISA", *Geographical Analysis* 27, 93–115.

Arango, J (2003). "Europa, ¿una sociedad multicultural en el siglo XXI?", *Papeles de Economía Española* 98: 2–15.

Arango J, Finotelli C (2009). "Past and future challenges of a Southern European migration regime: The Spanish case", *IDEA Working Papers* nº 8.

Arbia G, Paelinck JHP (2003). "Spatial econometric modeling of regional convergence in continuous time", *International Regional Science Review* 26, 342–362.

Bosker M (2009). "The spatial evolution of regional GDP disparities in the 'old' and the 'new' Europe", *Papers in Regional Science* 88(1), 3–27.

Bover O, Velilla P, (2002). "Migrations in Spain: Historical Background and Current Trends", in K. Zimmermann (ed.), European Migration: What Do We Know? CEPR and Oxford University Press.

Bruquetas M, Garcés B, Morén R, Penninx R, Ruiz E (2008). "Immigration and integration policymaking in Spain", IMISCOE Working Paper No. 21.

Carling J (2007). "Unauthorized migration from Africa to Spain", *International Migration* 45(4), 3-37.

Dall'erba S (2005). "Productivity convergence and spatial dependence among Spanish regions", *Journal of Geographical Systems* 7, 207–227.

Delgado L (2002). "Immigration in Europe: realities and policies", Unidad de Políticas comparadas (CSIC), Working Paper 02-18, University Rey Juan Carlos (Spain).

Dietz B (2000). "German and Jewish migration from the former Soviet Union to Germany: Background, trend and implications", *Journal of Ethnic and Migration Studies* 26(4), 635–652.

Durlauf SN, Johnson PA,Temple JRW (2005). "Growth econometrics", In: Aghion P and Durlauf SN (Eds.), *Handbook of Economic Growth*. Vol. 1A. Amsterdam: North Holland.

Esteban JM, Gradín C, Ray D (2007). "An extension of a measure of polarization, with an application to the income distributions of five OECD countries", *Journal of Economic Inequality* 5, 1–19.

Fassmann H, Münz R (1992). "Patterns and trends of international migration in Western Europe", *Population and Development Review* 18(3), 457–480.

Fernández C, Ortega C (2008). "Labor market assimilation of immigrants in Spain: employment at the expense of bad job-matches?", *Spanish Economic Review* 10(2): 83–107.

Fingleton B (2001). "Theoretical economic geography and spatial econometrics: dynamic perspectives", *Journal of Economic Geography* 1, 201–225.

Fingleton B, López-Bazo E (2006). "Empirical growth models with spatial effects", *Papers in Regional Science* 85(2), 177–198.

Florax RJ, Rey S (1995). "The impact of misspecified spatial structure in linear regression models", In: Anselin L and Florax RJ (Eds.) *New directions in spatial econometrics*, pp. 21-74. New York, Springer-Verlag.

Geary RC (1954). "The contiguity ratio and statistical mapping", *The Incorporated Statistician* 5(3), 115–145.

Hierro M (2007). "The effect of foreign-born residents on migratory patterns of natives in Spain", *Economics Bulletin* 10(3), 1–6.

Hierro M (2011). "Latin American Migration to Spain: Main reasons and future perspectives", *International Migration* (in press).

Hierro M, Maza A (2010a). "Per capita income convergence and internal migration in Spain: Are foreign-born migrants playing an important role?", *Papers in Regional Science* 89, 89–107.

Hierro M, Maza A (2010b). "Foreign-born internal migrants: Are they playing a different role than natives on income convergence in Spain?", *Applied Geography* 30(4), 618-628.

Izquierdo A, Carrasco C (2005). "Flujos y tendencias y signos de instalación de los extranjeros en España", *Papeles de Economía Española* 104, 92–122.

Izquierdo M, Lacuesta A, Vegas R (2009). "Assimilation of immigrants in Spain: A longitudinal analysis", *Labour Economics* 16, 669–678.

King R, Fielding AJ, Black R (1997). "The international migration turnaround in Southern Europe". In: King R, Black R (eds.). *Southern Europe and the New Immigrations*. Brighton: Sussex Academic Press, pp. 1–25.

Le Gallo J (2004). "Space-time analysis of GDP disparities among European regions: A Markov chains approach", *International Regional Science Review* 27(2), 138–163.

LeSage JP, Fischer MM, Scherngell T (2007). "Knowledge spillovers across Europe: Evidence from a Poisson spatial interaction model with spatial effects", *Papers in Regional Science* 86(3), 393–421.

Maza A, Hierro M, Villaverde J (2010). "Measuring intra-distribution income dynamics: An application to the European Regions", *Annals of Regional Science* 45, 313–329.

Maza A, Villaverde J (2009). "Provincial wages in Spain: Convergence and flexibility", *Urban Studies* 46 (9), 1969–1993.

Messner S, Anselin L (2004). "Spatial Analyses of Homicide with Areal Data", In: M. Goodchild and D. Janelle (Eds.) *Spatially Integrated Social Science*, pp. 127–44. New York: Oxford University Press.

Moran P (1948) "The interpretation of statistical maps", *Journal of the Royal Statistical Society B* 10, 243–251.

Pumares P, García-Coll A, Asensio A (2006). *La movilidad laboral y geográfica de la población extranjera en España.* Madrid: Subdirección General de Información Administrativa y Publicaciones, Observatorio Permanente de la inmigración, Ministerio de Trabajo y Asuntos Sociales.

Recaño J (2002). "La movilidad geográfica de la población extranjera en España: un fenómeno emergente", *Cuadernos de Geografía* 72, 135–156.

Recaño J (2004). "Migraciones internas y distribución espacial de la población española". En Leal, J. (ed.) *Informe sobre la situación demográfica en España.* Madrid: Fundación Abril Martorell, pp. 187–228.

Recaño J, Domingo A (2006). "Evolución de la distribución territorial y la movilidad geográfica de la población extranjera en España". En: Aja, E. y Arango, J.: *20 años de inmigración en España*, Fundación CIDOB-Fundación Ortega y Gasset, Barcelona, pp. 303–337.

Rey SJ (2001). "Spatial empirics for economic growth and convergence", *Geographical Analysis* 33, 195-214.

Rey SJ, Janikas MV (2005). "Regional convergence, inequality, and space", *Journal of Economic Geography* 5(2), 155-176.

Salgado-Ugarte I, Shimizu M, Taniuchi T (1997). "Nonparametric assessment of multimodality for univariate data", *Stata Technical Bulletin* 27, 5–19.

Silverman BW (1986). *Density estimation for statistics and data analysis*. Chapman and Hall: London.