Determinants of interregional migration in Spain: Are there differences between foreigners and natives?[1]

Authors and e-mail of the corresponding author:

María Gutiérrez Portilla (maria.gutierrezp@unican.es)

Department: Department of Economics

University: University of Cantabria

Subject area: Population and migratory flows

Abstract: The aim of this paper is to analyze the factors explaining internal mobility of foreigners and natives across the Spanish provinces over the period 2004-2012. For this purpose, an extended gravity model is specified in order to evaluate the determinants of interregional movements between each pair of provinces, this model being estimated for both natives and foreigners. Data on a set of explanatory variables such as per capita income, housing prices, human capital, unemployment rates, climatic conditions or socio-geographic distance is employed. The results of this analysis reveal that the foreign-born population seems to be more influenced, apart from social networks, by economic variables related to the labor market while relatively low housing prices remain important in the case of natives. On the other hand, the role of distance and the value of amenities (climatic conditions) are important factors for both groups of population.

Keywords: Internal migration, foreigners, natives, Spanish provinces

JEL codes: F22, R23, O15, C10

[1] The author is thankful to professors Adolfo Maza, María Hierro and José Villaverde for their helpful advice.

1. Introduction

During the last few decades, there have been important changes in the internal migration flows in Spain. Until the 1970´s there were strong differentials between regions, which led to an increase of the internal migration. Most people moved from rural areas (with fewer job opportunities) to urban areas. In the early eighties, after the economic crisis linked to the first oil price shock, the Spanish economy suffered structural changes and a process of return migrations. Besides, following the expansion of the welfare state, interregional migration flows began to reverse the trend; regions with higher rates of unemployment were receiving migrants coming from more developed regions. Apart from that, location-specific amenities started to exert great influence on migrants' decision to change their place of residency. This resulted in the so-called process of 'inverse migration', that is to say, the traditionally poor and high unemployment regions became net immigration regions, whereas the better-off ones became net outmigration regions (Bover and Velilla, 1999; Hierro and Maza, 2010a). As a consequence, overall net migration rates started to fall considerably.

Since the nineties, Spain has undergone a massive arrival of foreign-born migrants turning from being a country whose people emigrated to Latin America or European countries such as Germany or France into being one of the highest recipients of immigrants across the globe (Bover and Velilla, 1999; Carling, 2007; Carrasco et al., 2008; Reher and Requena, 2009; Hierro and Maza, 2010a; Hierro, 2013). Many papers have shown that this population is more mobile than natives across the Spanish provinces, thus playing an important role on the settlement patterns across the Spanish provinces. This being so, it seems convenient to study how different factors affect native-born and foreigners' mobility across the Spanish provinces.

As can be seen, internal migration patterns are far from being uniform, particularly in model specifications involving migration flows from and to multiple provinces, as it is the case of the present work. The aim of this paper is precisely to assess the factors explaining the internal mobility of the native and foreign-born populations across the Spanish provinces, emphasizing the different nature (economic or non-economic factors such as amenities and quality of life) of the variables determining the internal migration movements. The paper contributes to the existing literature by considering a dimension *ijt* of the panel data (origin, destination and time). Additionally, a variable measuring the socio-geographic distance between pairs of provinces has been computed with the

aim not only to measure the influence of distance between provinces, but also the economic environment of neighboring provinces.

To accomplish this study, we employ annual data from the year 2004 to 2012 coming from the Spanish National Statistical Institute (INE). More specifically, we have extracted the information from the 'Statistic of Residential Variations'. We employ data on 'Internal migration classified by province of origin and province of destination'. This information has been obtained separately for foreigners and natives. By migration, we understand the movement of people from one province to another implying a change of residency. On the basis of this definition, our dependent variable is the gross migration rate (per thousands) between pairs of provinces, that is, the ratio between internal migratory flows (taking into account only interregional flows) and the total population of the province of origin in the previous year. Additional sources from which data has been extracted are the Valencian Institute of Economic Research (IVIE) and the Ministry of Development.

With respect to the empirical framework, this paper offers an econometric specification, based on the popular gravity model, in which a set of explanatory variables are added to the equation with the aim to estimate how these factors affect the gross migration rate from a province *i* to a province *j*.

Considering the level of territorial disaggregation, we opt to use the 50 Spanish provinces (NUTS 3) as it allows to analyze movements across any pair of provinces (included movements across provinces belonging to the same region), something that would not be possible if we took into account the 17 autonomous communities (NUTS 2). It should be noted that the vast majority of studies analyzing migration patterns do it at a national or regional level. We believe the nature of these studies leads to a loss of information because as it is well-known, the greater the level of disaggregation, the more precise the results.

The remainder of the paper is organized as follows. The next section provides a brief literature review on the determinants of internal migration and results of empirical studies focused chiefly on gravity models. Then, a descriptive overview of the internal migration in Spain is carried out. This section is followed by the specification of an augmented gravity model trying to capture factors influencing interregional migration patterns in Spain of natives and foreigners. Finally, the results for both groups of

population are outlined along with a discussion of them. The last section summarizes the findings of this paper.

2. Literature review

The literature addressing the determinants of migration has been prolific in recent decades. A group of studies have delved into the factors shaping internal migration in countries such as Italy or Germany. In this regard, Biagi et al. (2011) examine the differences between long and short distance migration across Italian provinces. They conclude that economic determinants (income and unemployment) mainly affect long distance migration whereas quality of life and amenities are more relevant to explain short distance migration. For his part, Etzo (2011) addresses the study of the determinants of interregional migration flows in Italy, his results revealing that per capita income and the unemployment rate seem to exert a great influence on migrants when it comes to moving to a new region. Additionally, the work published by Schündeln (2014) for the period 1996-2003, addresses the mobility behavior of natives and foreigners for the case of Germany. By using probit and conditional logit regressions, he proves, after taking into account a set of individual characteristics, that immigrants are more likely than natives to internally migrate within Germany.

The core body of research on internal migration flows in Spain has placed special emphasis on its main determinants. That is the case of the work produced by Bentolila and Dolado (1991), who conclude that interregional migration responds significantly to economic variables such as real wage and unemployment differentials although the overall unemployment rate and housing prices differentials also deter migration. Additionally, Antolin and Bover (1997) evaluate the effect of individual characteristics as well as variables such as unemployment, wage and house price differential on the regional migration in Spain. In this vein, another interesting paper is that published by Bover and Arellano (2002) who, using annual Residential Variation Data for the years 1988, 1989, 1990 and 1992, provide evidence on the positive role played by the proportion of employment in the service industry, unemployment, house prices and education on the individual probabilities of intra-regional migration in Spain. Similarly, Juarez (2000), using data for the period 1963 to 1993, shows that interregional mobility responds to regional labor variables (unemployment rates, wages and employment

growth). Likewise, Mulhern and Watson (2009) prove that, between the years 1999 and 2006, differentials in wages and unemployment between provinces are relevant to explain internal migration, although house prices also matter. Another paper by the same authors (2010), supports as well that interregional migration in the 1990s responds to a range of economic variables including wages, unemployment and distance. Maza and Villaverde (2004), employing semiparametric techniques, find that internal migration between 1995 and 2002 is mainly driven by income and climatic condition differentials between origin and destination regions but unemployment and housing prices differentials appear to affect less net migration rates. In a more recent paper, Maza et al. (2013) examine the settlement patterns of foreigners, reaching the conclusion that social networks and again, economic factors, play an important role. As can be seen, most of the research conducted on this topic leads in the same direction.

Some works have studied the differences between foreigners and natives with respect to their internal mobility patterns for the case of Spain. In this regard, Recaño and Roig (2006) carry out an analysis of Spanish internal migration with data corresponding to the years 2003 and 2004. Their results suggest that foreign-born internal migrants tend to be more influenced by the existence of social networks and less sensitive than natives to economic factors and the location of other amenities. Another example is the work produced by Reher and Silvestre (2009) who, using data for 2007 from the National Survey of Immigration, provide evidence on the role played by factors such as educational level, knowledge of the Spanish language, income levels and social networks. Besides, Hierro and Maza (2010a) prove that differences between foreign-born and natives in the way they move within Spain do not have a significantly distinct effect on income convergence. Likewise, these authors provide evidence that foreign-born migration is a significant but weak factor promoting convergence in Spain (Hierro and Maza, 2010b). Additionally, Izquierdo et al. (2014) carry out a recent analysis of migratory flows in Spain during the economic crisis that started in the year 2008. This study provides a detailed comparison between variations of immigration and emigration flows of native and foreign-born population. It also sheds new light on the determinants of migratory flows making special emphasis on the differences between natives and foreigners.

From a methodological point of view, as far as the gravity models are concerned, their use in the field of international trade dates back from the sixties (Tinbergen, 1962) and

continued in force during the early eighties with authors such as Anderson (1979) and Bergstrand (1985). Since then, gravity models have been widely and successfully applied in migration research to integrate the concept of distance between provinces into the empirical models. That is the case of Etzo (2011), who employing data on a regional level, sheds some light on the determinants of interregional migration flows in Italy. The emerging technique of fixed-effect vector decomposition estimator (FEVD) is applied to a gravity model using bilateral flows for the period 1996-2005. For the case of Spain, the paper published by Ródenas (1994) is worth being mentioned. This work, devoted to study interregional migration in Spain using gross bilateral flows, employs a gravity model to prove that during the eighties migration became more responsive to the local composition of the labor force and less sensitive to economic factors such as wages or unemployment rates. Another relevant work is that of Devillanova and García-Fontes (2004). Within the framework of internal migration, a generalized negative binomial regression on gross migration flows between provinces is estimated in order to study internal labor mobility during the period 1978-1992. These authors find that the effect of employment opportunities has increased, migrants being more sensitive to economic conditions after 1984.

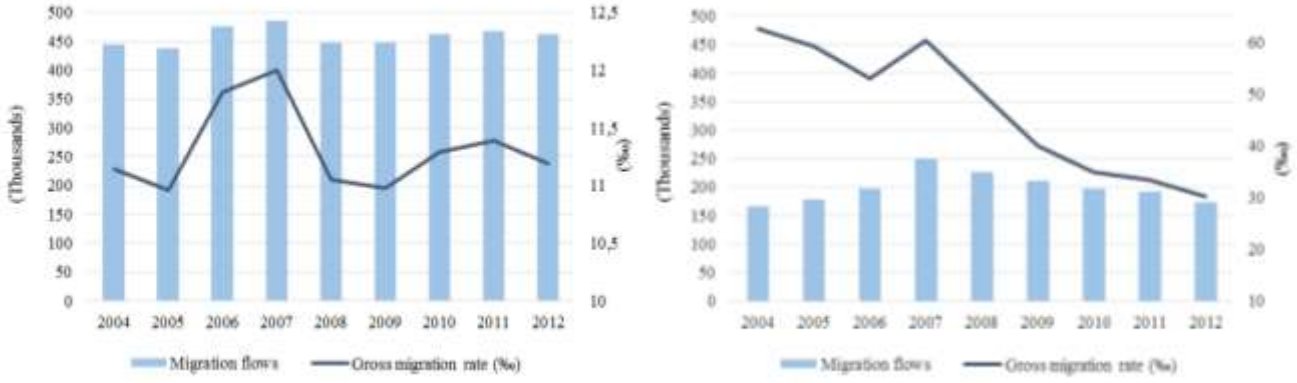
3. Internal migration in Spain: an overview

Following our analysis of interregional migration across Spanish provinces, the aim of this section is to provide a descriptive overview of the patterns of internal migration comparing the native and foreign-born population in Spain during the period 2004-2012.

To begin with, by comparing the evolution over time of internal migration between natives and foreigners (Figure 1), it could be noted that the foreign-born population tends to move more across the Spanish provinces than natives. This population reaches an internal migration rate of 62.45‰ in the year 2004, more than five times the highest rate for the native population in 2007, which was 12‰. Apart from that, Figure 1 also reveals the uneven evolution of interregional migration rates: while in the case of natives, although with ups and downs, it has remained almost unchanged, the interregional migration rate for foreigners has decreased by 51.60% between 2004 and 2012. Finally, the effect of the economic crisis has been more noticeable in the case of

the foreign-born population. Although their internal migration rate shows a declining path since the year 2004 (with a slight upturn in the year 2006), the decreasing trend is more remarkable from the year 2007 onwards. As regards migration flows, they increased until 2007, when they started to go down. In the case of natives, the gross internal migration rate remains much more stable over time although it shows some ups and downs (the lowest value is 11‰ in the years 2005, 2008 and 2009, the highest value (12‰) being reached in 2007).

Figure 1: Evolution of internal migration rate



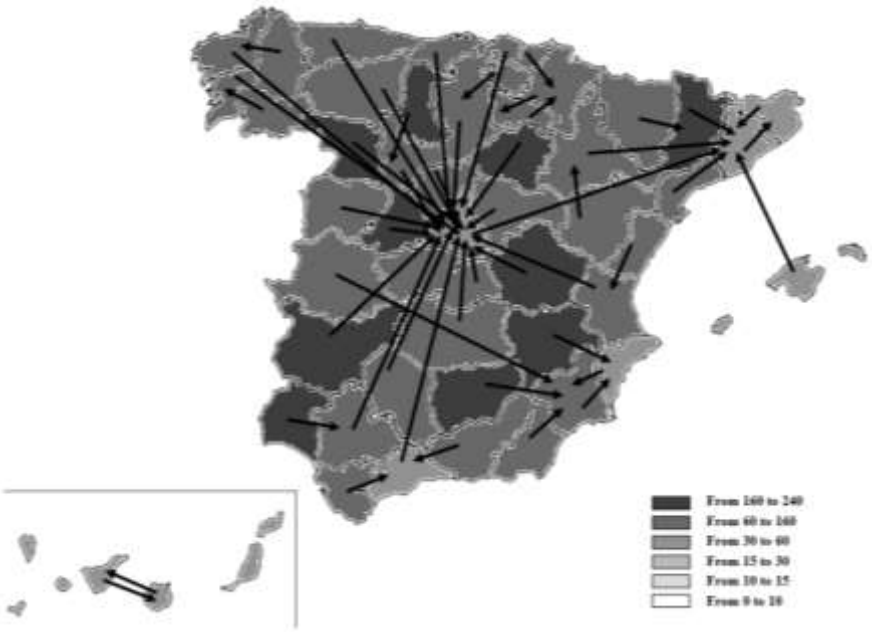
(a) Native-born population *(b) Foreign-born population*

Once the temporal dimension has been covered, it seems pertinent to illustrate where the internal migration flows come from and where they head to, in other words, we want to know more about the dimensions *i* and *j*. To accomplish this aim, Figures 2 and 3 display a pair of provincial maps of Spain (for the initial and final years of the sample) showing the internal gross migration rates for each province. It should be highlighted that the darker the color of a province, the higher the gross internal migration rate of that province. Moreover, the black arrows on the maps indicate the first destination of each of the fifty Spanish provinces, both in the case of natives (a) and foreigners (b).

Thus, a cursory glance at Figures 2 and 3 immediately shows that there exist differences between the foreign-born and native populations. Namely, interregional mobility of the foreign-born population across the Spanish provinces is much higher, as previously seen, regardless of their province of origin. In addition to this, the provinces with the highest migration rates tend to be concentrated around Madrid (in the case of natives in 2004 and foreigners in 2012). For their part, natives in 2012 show the highest internal migration rates in provinces around Madrid and Barcelona, these areas being

characterized by high economic activity. However, no clear pattern can be found in the year 2004 for the foreign-born population. Figure 2b displays precisely that foreigners in the year 2004 used to move more[2] across the Spanish provinces than they recently do (Figure 3b).

(a) Native-born population

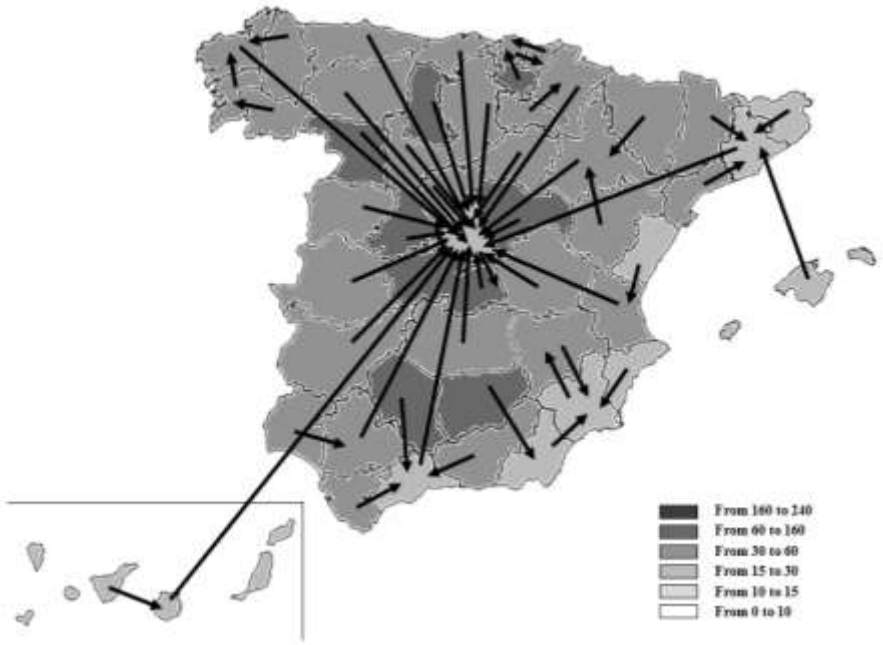


(b) Foreign-born population

Figure 2: Provincial gross migration rates (‰) (Year 2004)

[2] The foreign-born population in provinces such as Zamora or Jaen reached in 2004 internal migration rates of 239‰ and 234‰, respectively.

(a) Native-born population



(b) Foreign-born population

Figure 3: Provincial gross migration rates (‰) (Year 2012)

As regards the provinces of destination, on the whole, it appears that Madrid and to a much lesser extent, Barcelona are the most preferred options, this fact being more evident for the foreign-born population. Additionally, as far as the year 2004 is concerned (Figure 2a, b), natives living in the autonomous communities of Galicia,

Catalonia, Valencia, Andalusia or The Canary Islands seem to move across their own provinces rather than choosing to locate in farther away provinces. Instead, only does this happen in Catalonia and The Canary Islands for the case of foreign-born population. Taking into account the year 2012, a glimpse at Figure 3 (a, b) reveals that the preference for high-income Spanish provinces is even stronger than in the initial year. It can be seen that more native and foreign-born population from Las Palmas migrates to Madrid rather than to Tenerife, which is nearer. This fact could be explained because after the economic crisis, high-income Spanish provinces are thought to provide more employment opportunities (Minondo et al., 2013).

Fully aware of the limited information provided by the descriptive analysis carried out above, a more detailed econometric analysis of dimensions *ijt* is presented hereafter trying to ascertain the factors that are behind the recent settlement patterns of both, foreigners and natives.

4. Empirical model specification

As previously mentioned, we have chosen a gravity model as our empirical framework. In this regard, it seems pertinent to make it clear that the gravity model, in its simplest form, considers migration to depend on the size of the population of the origin and destination country (or province, in our case) and the distance between origin and destination.

The aim of this section is to explain the factors that lead foreigners and natives to move across the Spanish provinces during the period 2004-2012. In particular, we are interested in determining the differences in those factors between both groups of population. To address this issue, we obtained data on the gross migration rates between each pair of provinces, origin and destination[3]. Migrants are presumed to be sensible to changes in economic factors such as the unemployment rate and per capita gross domestic product (GDPpc), while other variables such as the climatic conditions or the level of human capital may also matter when deciding whether or not to migrate to another Spanish province.

[3] The fact of considering interregional migration movements from the province of origin to the province of destination brings consistency to our results as it better reflects the actual situation, in which migrants leave one province in order to move to another one.

The first equation that shall be estimated for natives and foreigners has the following specification:

$$MR^{k}_{ij,t} = \alpha_i + \beta_1 * GDPpc_{ij,t-1} + \beta_2 * UR_{ij,t-1} + \beta_3 * HC_{ij,t-1} + \beta_4 * HP_{ij,t-1} + \beta_5 * AG_{ij,t-1} + \beta_6 * IND_{ij,t-1} + \beta_7 * SER_{ij,t-1} + \beta_8 * CLIM_{ij,t-1} + \beta_9 * DIST_{ij} + \beta_{10} * d_{crisis} + \beta_{11} * d_{Madrid} + \varepsilon_{ij,t} \quad (1)$$

where *k=1,2* (1 for natives and 2 for foreigners) and where *i*, *j* and *t* stand for province of origin, province of destination and year, respectively. Likewise, the dependent variable for our empirical model is MR, which represents the gross migration rate[4] (for natives and foreigners) from a province to another province. As regards the construction of independent variables[5], we have considered relative values of the province of destination with respect to the province of origin, which leads to a better interpretation of the results. Specifically, we have provincial differentials in per capita GDP (GDPpc), in the unemployment rate (UR), in the human capital[6] (HC), in housing prices[7] (HP), in the share of agriculture in GDP (AR), in the share of industry in GDP (IND), in the share of services in GDP (SER) and in the climatic conditions variable[8] (CLIM). The only difference in the equation estimated for foreigners with respect to the one estimated for natives is that an additional variable to account for the network effect is included. In order to solve potential endogeneity problems arising in the estimation between the internal migration rate and the foreign population, we consider the stock of

[4] The gross migration rate (‰) has been constructed as the ratio between the internal migratory flow from province *i* to province *j* and the corresponding population of province *i* for the previous year. We are just considering interregional flows, removing intra-regional ones from our sample as we are only interested in movements across different Spanish regions.

[5] The Consumer Price Index in base year 2011 is used as a deflator.

[6] The human capital is defined as the percentage between the active population with higher education and the total active population.

[7] The housing prices are expressed in euros per square meter.

[8] We have used Boyer and Savageau´s (1985) methodology to measure this variable as it captures the preference for mild temperature rather than extreme climatic conditions. Thus, a score of 1000 points is assigned to each province and subsequently points are subtracted depending on the values of a set of partial indicators: 1. Very hot or very cold months; 2. Seasonal variation in the temperature; 3. Cloudy or overcast days; 4. Days of rain; 5. Extremely high temperatures; 6. Extremely low temperatures.

foreign population of the province of destination in the year 2000 as a proxy to measure social networks ($SN_{j,t-1}$).

In addition, DIST refers to the distance between pairs of provinces expressed in thousand road kilometers[9], d_{crisis} stands for a temporal dummy included to account for the effect of the economic recession, whose outbreak took place in the year 2008[10]. Furthermore, d_{Madrid} denotes a dummy variable to capture all the specific features that make Madrid[11] worth living in, because as seen in Figures 2 and 3, Madrid turns out to be an outlier. Finally, α_i represents provincial fixed effects by province of origin. They have been included in our model to gather in each of them the pairs of provinces origin-destination sharing the same province of origin[12]. They represent specific behaviors which may be due to social, political and institutional factors that are not collected in the specification of the model and may affect interregional migration in Spain. As in any model, the error term encompasses any specification or measurement error associated with the observed interregional flows.

As far as the role of space is concerned, it goes without saying that the distance between the province of origin and the province of destination matters when making the decision to migrate to another Spanish province. However, the fact of including single distances in the model would not properly capture a hypothetical situation in which there exist a province of destination *(a)* which is nearer the province of origin but has a high unemployment rate and another province of destination *(b)* which is farther from the province but has a low unemployment rate. This being so, we think it makes more sense to consider a socio-geographic distance, which takes into account data on distances (expressed in thousand km), combined at the same time with some economic indicator. The basic idea is to weigh the economic environment of a province based on the distance from that province to the rest, so when it comes to the interpretation of the

[9] Data on distances between each pair of provinces was obtained from a website with road distances from Repsol guide: *http://aim-andalucia.com/distancias.html*

[10] This being so, we have considered a value of one from the year 2008 onwards and a value of zero in the previous years to define the dummy variable.

[11] As regards the construction of this dummy variable, it takes a value of one whenever Madrid is the province of destination; otherwise it takes a value of cero.

[12] Our sample being composed by fifty provinces, it goes without saying that we have dropped the corresponding fixed effect for the last province of origin (Zaragoza) to avoid the problem of multicollinearity.

results, the economic logic does not play second fiddle to the role of space. By including the socio-geographic distance in the equation, our specification shall be considered as a gravity model as well but with a compound variable (single distance between pairs of provinces together with the ratio of unemployment rates between pairs of provinces) quantifying the geographic distance from an economic point of view. Not only does it take into account the characteristics of the provinces of origin and destination, but it also considers the situation of the surrounding provinces. In this regard, we assume that migration patterns from and to a particular province are driven by the economic situation of the labor market in the neighboring provinces.

This new variable[13] is defined as follows:

$$SD_{ij,t-1} = \frac{\left(\frac{UR_{j,t-1}}{UR_{i,t-1}}\right) * DIST_{ij}}{\frac{\sum_k \left(\frac{UR_{k,t-1}}{UR_{i,t-1}}\right) * DIST_{ik}}{k}}$$

where *k* denotes neighboring provinces, DIST (as previously explained) refers to the distance[14] between pairs of provinces expressed in thousand road kilometers and UR to the unemployment rate.

Then, the second specification of the model is as follows:

$$\begin{aligned} MR^{k}_{ij,t} = {} & \alpha_i + \gamma_1 * GDPpc_{ij,t-1} + \gamma_2 * HC_{ij,t-1} + \gamma_3 * HP_{ij,t-1} + \gamma_4 * AG_{ij,t-1} \\ & + \gamma_5 * IND_{ij,t-1} + \gamma_6 * SER_{ij,t-1} + \gamma_7 * CLIM_{ij,t-1} + \gamma_8 * SD_{ij,t-1} \\ & + \gamma_9 * d_{crisis} + \gamma_{10} * d_{Madrid} + \varepsilon_{ij,t} \end{aligned} \qquad (2)$$

[13] It is worth mentioning that there are fifty different denominators for this variable, one by province, although in the case of the islands (The Balearics, Las Palmas and Tenerife) the denominator of the SD variable is equal to one as we consider that they have not surrounding area, in other words, each of them borders with itself.

[14] It should be mentioned that the exponent of the distance could be modified depending on the weight given. We opt to use distances themselves. Nevertheless, for the sake of robustness, the same equations have been estimated considering the square of the distance and the distance raised to the 0.5 and it has been proven that while the coefficient remains negative, its absolute value decreases the higher the power of the distance.

Obviously, after including the SD variable in equation (2), the variables UR and DIST have been removed because SD is a compound variable capturing these two other variables. The rest of explanatory variables shown in equation (2) are the same as in equation (1), and again, the foreign population of the province of destination in the year 2000 is included only in the specification for foreigners as a proxy for the network effect.

It is well worth being mentioned that, as can be noted, all the explanatory variables employed in equations (1) and (2), except for the single distance and the dummy variables are lagged one year so that their effects can be taken into account in the migration decision and they have been constructed as the ratio between the specific values of the province of destination *j* and the province of origin *i*. Thus, it enables an easier and more logical interpretation of the results.

Before going any further, some comments on the justification of the variables employed and the directions of their expected effects are mandatory. Firstly, the wage level (proxied by the per capita GDP) and the employment opportunities in the regional labor market are the major determinants of expected earnings in a particular location (Westerlund, 1997; Cameron and Muellbauer, 1998; Maza, 2006; Maza et al., 2013; Izquierdo et al., 2014).

It goes without saying that relative per capita income shapes the destination choices of both foreigners and natives. This idea is backed up by the Neoclassical Theory, which suggests that the bigger the difference between the per capita GDP in the destination and origin provinces, the greater the migratory flows. In the same vein, a higher rate of unemployment in the origin province than in the province of destination should at first, encourage people to emigrate in search of work.

Considering the human capital variable, it seems reasonable to think that migrants would choose a province of destination whose human capital requirement in the labor market is consistent with their own training. So, it could be expected that foreigners, who usually have lower middle training, have a preference for provinces of destination with low levels of human capital so that they can find a job and meet the requirements of the labor market.

The existing literature on migration assumes that housing prices also matter when taking the decision of where to locate. It seems that variations in housing prices

influence regional migration by making people move toward provinces where housing costs are lower (Antolín and Bover, 1997; Bover and Arellano, 2002; Maza and Villaverde, 2004). People probably move in search of cheaper housing and better quality of life (Antolín and Bover, 1997).

Moreover, it seems reasonable to include the shares of agriculture, industry and services in GDP in the equation to assess whether natives and foreign-born migrants are more attracted by one particular sector or, by contrast, whether a higher share of a sector in total GDP in a particular province discourages them from moving to that location.

Another aspect that is bound to affect migration processes is the value of amenities. Accordingly, there are some studies emphasizing the role played by variables measuring the quality of life in the destination location (Maza, 2006; Biagi et al., 2011, Rodriguez-Pose and Ketterer, 2012; Buch et al., 2013). This being so, the climatic conditions variable is part of our model with the aim to capture this effect. Hence, the widespread preference for mild temperature rather than extreme climatic conditions is reflected in the index calculation.

As for the inclusion of the foreign population of the province of destination in the year 2000 in the equations of foreigners, it is expected that the presence of networks in other provinces increases the probability of being willing to migrate to that province. Research on international migration highlights the importance of social networks in sustaining migration flows (Vias, 1998; Moreno and López, 2006; Lewer and Van der Berg, 2008; Maza et al., 2013). It has been proven that thanks to the existence of these networks, migration is more likely to take place as migration costs are reduced. So, these networks clearly influence the settlement of would-be migrants.

As for the inclusion of the distance as an additional explanatory variable, it is evident that being a proxy for the migration costs, it somehow shapes the decision to migrate. As it is well known, migration costs are likely to be correlated with the physical distance between provinces. Thus, it is expected that the bigger the distance between a pair of provinces, the lower the probability of being willing to migrate.

Another variable that is likely to exert influence on the migration decision is the socio-geographic distance. Among the geographic determinants, a greater distance between provinces of origin and destination decreases the predicted number of migrants, as

expected from the gravity model. Nevertheless, concerning our specification of the socio-geographic distance, the interpretation is as follows:

(a) Should the ratio between unemployment rates of the province of destination and the province of origin be greater than unity and should the distance between both provinces be large (both facts entailing a high numerator), and at the same time, should the provinces surrounding the origin province have a high unemployment rate (which leads to a high denominator), then, interregional migration will tend to decrease while the overall variable increases. This reasoning would imply a negative sign in the coefficient associated to the socio-geographic distance.

(b) Similarly, should the ratio between unemployment rates of the province of destination and the province of origin be smaller than unity and should the distance between both provinces be short (both facts entailing a small numerator), and at the same time, should the provinces surrounding the origin province have a low unemployment rate (which leads to a small denominator), then, interregional migration will tend to increase while the overall variable decreases. Thus, this explanation would result in a negative sign in the coefficient associated to the socio-geographic distance.

The approach adopted allows us to compare how the native and foreign-born populations react to economic and non-economic incentives to migrate. In particular, our initial hypothesis that shall be tested, is that foreigners tend to move more due to factors related to the labor market as well as the network effect while natives react more to residential variables such as the housing prices or amenities (climatic conditions).

5. Results

As previously mentioned, in this section we first estimate the equation (1) and then equation (2), both of them for the native and foreign-born population with the aim to determine how the different factors affect interregional migration in Spain. Before doing so, Hausman and Breusch-Pagan tests are carried out to achieve a correct specification of the models for natives and foreign-born (see Table 1).

Table 1: Tests for model specification: natives and foreigners

		Test	*Value*	*Prob.*
Equation (1)				
	Natives	Hausman	505.15	0.000
		Breusch-Pagan	193175.58	0.000
	Foreigners	Hausman	459.97	0.000
		Breusch-Pagan	192293.56	0.000
Equation (2)				
	Natives	Hausman	147.91	0.000
		Breusch-Pagan	179759.77	0.000
	Foreigners	Hausman	504.74	0.000
		Breusch-Pagan	183014.67	0.000

Source: INE, IVIE and Ministry of Development.

To begin with, we perform Hausman test, its result in both models supporting the use of fixed effects estimator rather than a random effects estimator. Secondly, Breusch-Pagan test is computed to test for the presence of heteroscedasticity. The results lead us to reject the null hypothesis of homocedasticity at a significance level of 0.01, again for both models –natives and foreigners-. As a consequence, a Generalized Least Squares (GLS) estimator is preferred to an Ordinary Least Squares estimator. In addition, and in order to tackle potential problems of endogeneity, we initially estimated both equations by the Generalized Method of Moments (GMM) proposed by Arellano and Bond (1991). However, we ruled out this estimation technique as both the second-order serial correlation in residuals and Sargan tests provided really poor results. For this reason, the estimation was finally performed by GLS.

We have estimated the models given by equations (1) and (2) over the period 2004-2012, the information on the lagged independent variables coming from the year 2003. The results are reported in Table 2, where the second and third (fourth and fifth) columns correspond to equation 1 (2), for natives and foreigners, respectively.

Table 2: Determinants of interregional migration of natives and foreigners

	Equation (1)		**Equation (2)**	
Dependent variable: $MR^k_{ij,t}$	**Natives**	**Foreigners**	**Natives**	**Foreigners**
$GDPpc_{ij,t-1}$	0.076*** (0.020)	2.420*** (0.226)	0.048** (0.020)	1.892*** (0.222)
$UR_{ij,t-1}$	-0.001 (0.003)	-0.132*** (0.041)		
$HC_{ij,t-1}$	0.003 (0.005)	-0.208*** (0.070)	0.003 (0.005)	-0.184*** (0.070)
$HP_{ij,t-1}$	-0.069*** (0.011)	1.021*** (0.128)	-0.070*** (0.011)	1.112*** (0.129)
$AG_{ij,t-1}$	-0.002*** (0.000)	0.029*** (0.007)	-0.002*** (0.000)	0.031*** (0.008)
$IND_{ij,t-1}$	-0.044*** (0.007)	1.000*** (0.086)	-0.044*** (0.007)	0.981*** (0.087)
$SER_{ij,t-1}$	-0.176*** (0.035)	6.406*** (0.430)	-0-226*** (0.035)	4.949*** (0.418)
$CLIM_{ij,t-1}$	0.936*** (0.131)	3.442*** (0.606)	0.800*** (0.132)	3.229*** (0.619)
$SN_{j,t-1}$		0.024*** (0.002)		0.021*** (0.002)
$DIST_{ij}$	-0.292*** (0.032)	-2.427*** (0.151)		
$SD_{ij,t-1}$			-0.002*** (0.000)	-0.066*** (0.007)
d_{crisis}	0.003** (0.001)	-0.905*** (0.021)	0.003** (0.001)	-0.893*** (0.021)
d_{Madrid}	2.300*** (0.092)	5.874*** (0.528)	2.399*** (0.935)	7.227*** (0.529)
R^2 adjusted	0.244	0.3465	0.221	0.316

Notes: Standard error in parenthesis;*** Significant at 1%; ** Significant at 5%. Provincial fixed effects have been included in all the estimations. Methodology: GLS. *Source:* INE, IVIE and Ministry of Development.

The most relevant conclusions than can be drawn from this analysis are as follows:

1. Firstly, as expected, per capita income turns out to be a determining factor for both groups of population, natives and foreigners. The four coefficients associated with GDPpc are positive and statistically significant (although the value is logically higher in the case of foreigners), which implies that the higher the relative wage in the destination province, the greater the migrants' willingness to move.
2. As regards the effect of unemployment rates on interregional migration in Spain, the negative and significant coefficient linked to this variable in equation (1) reveals that foreigners look for provinces of destination with a relatively low unemployment rate.
3. The effect of human capital on internal migration in Spain appears not to be significant in the case of natives whereas it does for the foreign-born population in both equations. This reveals that foreign migrants seem to look for provinces with relatively low values of human capital so that they can be inserted into the labor market.
4. As expected, the housing prices variable results statistically significant and carries a negative sign for natives. This is consistent with the idea that they usually tend to react to residential variables more than foreigners, moving to provinces with lower housing prices. The opposite happens in the case of the foreign-born population, the effect of this variable is positive. Figures 2 and 3 of the descriptive analysis show that most foreigners choose Madrid as their province of destination, so this is in line with the fact that they do not necessarily look for lower housing prices. Another possible explanation to this finding is that house prices represent income expectations which arise in the labor market: *"as migration should respond to earnings expectations, there should in this regard be a positive response of net in-migration to higher relative house prices"* (Cameron and Muellbauer, 1998).
5. It is found that the relative shares of agriculture, industry and services in GDP have a significant and negative impact on the gross migration rate among natives while in the case of foreigners, they have a positive effect. It will be necessary to delve further into this result in future research. As can be appreciated, foreigners are more attracted by the services sector.

6. The climatic conditions seem to affect interregional migration in Spain in both groups of population. Thus, not only the value of amenities is relevant in shaping migration patterns for natives, but foreigners are also attracted by mild temperature and optimal weather conditions.
7. Another factor that appears to be behind Spanish interregional migration of the foreign-born population is social networks since the coefficient associated to the variable of foreign population in the model of foreigners turns out to be positive and statistically significant at 1%.
8. Obviously, the coefficient on the single distance variable is significant and positive for natives and foreigners, implying that the higher the distance between provinces, the lower the rates of interregional migration across Spain.
9. As expected, the coefficient on socio-geographic distance is negative and statistically significant for both groups of population, meaning that our variable correctly takes into account the economic environment of a particular province based on the distance from that province to the rest. It should be highlighted that in the case of natives, the negative effect of this compound variable (physical distances and unemployment rates) corrects somehow the non-significance of the unemployment rate in equation (1). Not only that, but also a relevant feature that is worth being mentioned is that the effect of this variable is much higher in the case of foreigners (-0.066 versus -0.002), thus indicating that the relative rate of unemployment from a geographic point of view plays a more important role for foreigners than for natives.
10. Furthermore, the coefficients associated to the dummy variable accounting for the effect of the economic crisis suggest that a period of crisis minimally fosters the internal migration of natives across the Spanish provinces, whereas the crisis seems to discourage interregional movements of the foreign-born population. This finding corresponds exactly to what the temporal evolution in Figure 1 showed.
11. Finally, the dummy variable for Madrid results statistically significant in both equations, its positive impact being much higher for the case of the foreign-born population. This is also consistent with Figures 2 and 3, which reflect that foreigners' preference for Madrid is much more noticeable.

As can be seen, the estimation results hardly change from equation 1 to equation 2, this fact conferring robustness to our results.

6. Conclusions

This paper provides new insights into the factors determining interregional migration flows in Spain for the period 2004-2012. To be precise, it analyzes the effect of different variables (both economic and non-economic) on shaping internal movements and compares how natives and foreigners respond to them.

Our results seem to be consistent with previous research on the determinants of migration arguing that not only labor market conditions matter, but also amenities have a positive impact on shaping the interregional migration in Spain.

After examining the results obtained, it can be inferred that socio-geographic distance succeeds in capturing the role of space or physical distance along with the impact of the economic situation of the labor market in the surrounding provinces. Furthermore, it is shown that the likelihood of mobility increases with lower migration costs and higher relative income differentials in the receiving province. It has also been proven that individuals tend to migrate to provinces with better climatic conditions than in their home provinces. This result highlights that location-specific amenities do influence migrants when choosing their destination.

As a starting point, a brief literature review is presented showing some of the existing studies on the determinants of internal migration (both for Spain and other countries) as well as some papers applying gravity models to the field of internal migration. The next part of the paper offers an overview of the reality of interregional migration in Spain, which reveals that the degree of internal mobility across the Spanish provinces is much higher for the foreign-born population than the native-born and that the major destination for both groups of population is Madrid.

In order to achieve our main goal, two equations of an extended gravity model have been specified and estimated trying to explain the main determinants of the gross migration rate for natives and foreigners. Apart from including economic variables to reflect the labor market situation, additional factors such as the housing prices, the sectoral shares in GDP, the level of human capital, or the climatic conditions have been taken into account. However, the main contribution is the so-called socio-geographic distance variable, aimed at capturing the role of space in the Spanish geography combined with the labor market situation in each province (measured by the rate of unemployment).

Both equations are estimated by GLS to remove potential endogeneity problems that may arise during the estimation process. The results obtained satisfy, to a great extent, our initial hypothesis that while foreigners are more sensitive to economic factors (per capita GDP and unemployment rate), the group of natives responds to quality-of-life variables or the housing prices when it comes to moving to another Spanish province.

Results suggest that variables such as per capita income positively affect both natives and foreigners, implying that the more the relative value in the province of destination with respect to the province of origin, the more likely it is for interregional migration to take place. In the same vein, we find that amenities, measured in our model by the climatic conditions, have a great effect on the interregional mobility for both groups of population. Additionally, it has been proven that the compound variable socio-geographic distance succeeds in capturing the major impact of the combination of both, distance and unemployment rate, for the foreign-born population. Moreover, it has been proven that there exists a network effect in the patterns of migration in the case of foreigners.

Furthermore, our findings also reveal relatively low housing prices is not a determining factor for foreigners to move to another Spanish province. This result supports their preference for Madrid, as a province of destination where housing prices are not low. The dummy variable included in the specifications of both equation results statistically significant (its effect being much higher for foreigners than natives).

References

Anderson, J. E. (1979): "A theoretical foundation for the gravity equation", *American Economic Review*, nº 69 (1), p. 106-116.

Antolin, J. and Bover, O. (1997): "Regional migration in Spain: the effect of personal characteristics and of unemployment, wage and house price differentials using pooled cross-sections", *Oxford Bulletin of Economics and Statistics*, nº 59, p. 215-235.

Bentolila S. and Dolado J. J. (1991): "Mismatch and internal migration in Spain, 1962-1986". Fiorella Padeo Schioppa (ed.) *Mismatch and labour mobility, Cambridge University Press, Cambridge.*

Bergstrand, J. (1985): "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence", *Review of Economics and Statistics,* nº 67 (3), p. 474-481.

Biagi, B., Faggian, A. and McCann, P. (2011): "Long and short distance migration in Italy: the role of economic, social and environmental characteristics", *Spatial Economic Analysis,* nº 6 (1), p. 111-131.

Bover, O. and Arellano, M. (2002): "Learning about migration decisions from the migrants: using complementary datasets to model intra-regional migrations in Spain", *Journal of Population Economics,* nº 15, p. 357-380.

Bover, O. and Velilla, P. (1999): "Migration in Spain: historical background and current trends", Banco de España, documento de trabajo, nº 9909.

Boyer, R., Savageau, D. (1985): Places Rated Almanac. Rand McNally, Chicago.

Buch, T., Hamann, S., Niebuhr, A. and Rossen, A. (2013): "What makes cities attractive? The determinants of urban labour migration in Germany", *Urban Studies,* nº 51, p. 1960-1978.

Cameron, C. and Muellbauer, J. (1998): "The housing market and regional commuting and migration choices", *Scottish Journal of Political Economy,* nº 45 (4), p. 420-446.

Carling, J. (2007): "Unauthorized migration from Africa to Spain", *International Migration Review,* nº 45 (4), p. 3-37.

Carrasco, R., Jimeno, J. F. and Ortega, A. C. (2008): "The effect of immigration on the labor market performance of native-born workers: some evidence for Spain", *Journal of Population Economics,* nº 21, p. 627-648.

Devillanova, C. and García-Fontes, W. (2004): "Migration across Spanish provinces: evidence from the social security records (1978-1992)". *Investigaciones Económicas,* nº 28, p. 461-487.

Etzo, I. (2011): "The determinants of the recent interregional migration flows in Italy: a panel data analysis", *Journal of Regional Science,* nº 51 (5), p. 948-966.

Hierro, M. and Maza, A. (2010a): "Foreign-born internal migrants: are they playing a different role than natives on income convergence in Spain?", *Applied Geography*, nº 30 (4), p. 618-628

Hierro, M. and Maza, A. (2010b): "Per capita income convergence and internal migration in Spain: are foreign-born migrants playing an important role?", *Papers in Regional Science*, nº 89 (1), p. 89-107

Hierro, M. (2013): ''Latin American migration to Spain: main reasons and future perspectives''. *International Migration* (Forthcoming) DOI: 10.1111/imig.12056

Izquierdo, M., Jimeno, J. F. and Lacuesta, A. (2014): ''Los flujos migratorios en España durante la crisis". *Boletín Económico*, Septiembre 2014. Banco de España, p. 53-62.

Juarez, J. P. (2000): ''Analysis of interregional labor migration in Spain using gross flows'', *Journal of Regional Science*, nº 40 (2), p. 377-399.

Lewer, J. and Van der Berg, H. (2008): "A gravity model of immigration", *Economics Letters*, nº 99(1), p. 164-167.

Maza, A. and Villaverde, J. (2004): "Interregional migration in Spain: a semiparametric analysis" *Review of Regional Studies*, nº 34 (2), p. 156-171.

Maza, A. (2006): "Migrations and regional convergence" *Review of Regional Research*, nº 26 (2), p. 191-202.

Maza, A., Villaverde, J. and Hierro, M. (2013): "Explaining the settlement patterns of foreigners in Spain" *Applied Geography*, nº 40, p. 11-20.

Minondo, A., Requena, F. and Serrano, G. (2013): "Movimientos migratorios en España antes y después de 2008" *Papeles de Economía Española*, nº 138, p. 80-97.

Moreno, I. and López, G. (2006): "Los determinantes de la inmigración internacional en España", *Investigaciones Regionales*, nº 9 , p. 23-46.

Mulhern, A. and Watson, J. (2009): "Spanish internal migration: is there anything new to say?", *Spatial Economic Analysis*, nº 4 (1), p. 103-120.

Mulhern, A. and Watson, J. (2010): "Spanish inter-regional migration: an enigma resolved", *Applied Economics Letters*, nº 17, p. 1355-1359.

Recaño, J. and Roig, M. (2006): "The internal migration of foreigners in Spain", *Centro de Estudios Demográficos* (presented in the European Population Conference, Liverpool).

Reher, D. and Requena, M. (2009): "The national immigrant survey of Spain: a new data source for migration studies in Europe", *Demographic Research,* nº 20 (2), p. 253-278.

Reher, D. and Silvestre, J. (2009): "Internal migration patterns of foreign-born immigrants in a country of recent mass immigration: evidence from new micro data for Spain", *International Migration Review*, nº 43 (4), p. 815-849.

Ródenas, C. (1994): "Migraciones interregionales en España (1960-1989): cambios y barreras", *Revista de Economía Aplicada,* nº 2, p. 5-36.

Rodríguez-Pose, A. and Ketterer, T. D. (2012): "Do local amenities affect the appeal of regions in Europe for migrants?", *Journal of Regional Science*, nº 52 (4), p. 535-561.

Schündeln, M. (2014): "Are immigrants more mobile than natives? Evidence from Germany", *Journal of Regional Science*, nº 54 (1), p. 70-95.

Tinbergen, J. (1962): Shaping the World Economy, Twentieth Century Fund. New York.

Vias, A. C. (1998): "Detecting differences in interstate migration streams using an enhanced gravity model", *Southeastern Geographer*, nº 1, p. 22-40.

Westerlund, O. (1997): "Employment opportunities, wages and interregional migration in Sweden 1970-1989", *Journal of Regional Science*, nº 37 (1), p. 55-73.