

INTERNAL MIGRATION IN SPAIN: DEALING WITH MULTILATERAL RESISTANCE AND NONLINEARITIES

Adolfo Maza, María Gutiérrez-Portilla, María Hierro and José Villaverde
University of Cantabria

ABSTRACT

This paper investigates the determinants of internal mobility across Spanish provinces of both foreigners and natives over the decade 2004-14. Building on an extended gravity model, our econometric strategy controls for multilateral resistance to migration by including different fixed effects structures. Additionally, the paper allows for some nonlinearities in the key economic determinants of migration, wages and unemployment. The main finding is that the impact of economic factors on internal migration is higher for foreigners than for natives; furthermore, the effect of these factors on internal migration is clearly nonlinear for the group of natives, while this only happens for foreigners when dyadic fixed effects of origin-destination are considered. Finally, the paper shows that the nature of the amenities with the greatest impact on internal movements differs between the two groups: foreigners look for social services and cultural amenities, whereas natives are more attracted by good climate conditions.

KEYWORDS: Internal migration, foreigners, natives, amenities, threshold, Spanish provinces

JEL CODES: R23, O15, C23, C24

INTRODUCTION

Identifying the determinants of labour mobility has attracted considerable attention in the field of regional economics for decades, among other reasons because of the impact of migration on convergence. Ravenstein's ground-breaking work 'The Laws of Migration' (1885) formulated some of the hypotheses on which most of migration research has been based, among which those highlighting the role played by economic factors as major causes of migration. The neoclassical migration theory (Lewis, 1954), which regards internal labour migration as an integral part of economic development, focused on wage differentials as the main reason for migration. Subsequently, Todaro (1969) extended this theory to account for the fact that migrants take the decision to move by choosing the labour market location that maximises their expected rather than current earnings.

Later on, two main distinct approaches, with opposite conclusions regarding territorial cohesion, appeared in the debate. On the one hand, the so-called 'disequilibrium' model (Muth, 1971; Greenwood, 1975, 1985), which postulates that migration is simply a response to economic incentives (economic differentials) in labour markets, restoring eventually the equilibrium between them; among the most recent papers in this group, Etzo (2011) and Détang-Dessendre et al. (2016) stand out. On the other hand, an alternative approach, known as 'equilibrium' model of migration, emerged from the works of Graves (1976, 1980, 1983) and Knapp and Graves (1989). Unlike the previous one, it postulates that differences in economic variables are partially compensated by non-economic factors and, then, keep in time; in other words, it underscores the role of the so-called amenities or local attributes (quality of life factors, such as climate) in shaping migrant's decisions, as they are able to increase their utility. The basic idea of this approach is that "*migration takes place as a result of changes in demand for location-fixed amenities*" (Graves, 1980: 227).

Although opposite in nature, both approaches are still relevant and have become complementary over the years. Individuals seeking to migrate maximise their utility by taking into account a set of economic and non-economic factors in potential alternative destinations. In any case, as pointed out by Biagi et al. (2011), there seems to be a consensus in the literature devoted to the European case that, when it comes to internal migration, economic motivations are dominant while amenities tend to play a minor role (see, for instance, Etzo, 2011; Piras, 2012; Détang-Dessendre et al., 2016). In the United States, on the contrary, empirical evidence conveys the message that internal migration is more amenity-driven (see, for example, Deller et al., 2001; Partridge and Rickman, 2003; Rappaport, 2007; Partridge, 2010).

Against this backdrop, this paper uses the Spanish case as a sort of laboratory to analyse internal migration. There are two main reasons in support of this choice. First, the literature on internal migration in Spain is not very conclusive. Some papers emphasise the relevance of economic drivers (García-Ferrer, 1980; Santillana, 1981; Bentolila and Dolado, 1991; Antolín and Bover, 1997; Devillanova and García-Fontes, 2004; Mulhern and Watson, 2009, 2010) while others point to the increasing role of amenities (Ródenas, 1994; De la Fuente, 1999; Lago and Aguayo, 2004; Maza and Villaverde, 2004). Second, the pattern and composition of internal migration in Spain has changed remarkably over the last decade; this is the result of the massive arrival of immigrants in the early 2000s (Reher and Requena, 2009; Reher and Silvestre, 2009; Amuedo-Dorantes and De la Rica, 2010; Hierro, 2016) and their greater mobility (Recaño, 2002; Recaño and Roig, 2006; Reher and Silvestre, 2009; Hierro and Maza, 2010).

Accordingly, this paper has two aims: on the one hand, to assess whether economic or quality of life factors are more relevant when shaping internal migration in Spain and, on the other, to unveil whether there are significant differences between natives and foreigners' internal migration patterns. To do so, it carries out the analysis separately for both groups of population; specifically, the hypothesis to be tested is that natives' preferences are more amenity-based/less economic-oriented than those of foreigners. To the best of our knowledge, there are just a few papers for other countries,¹ and only a recent study for Spain (Gutiérrez-Portilla et al., 2018),² carrying out a comparative assessment of internal migration motives between foreigners and natives.

Apart from the above contribution, this study also differs from others in two methodological respects. First, based on the well-known Random Utility Maximisation (RUM) model for migration, it applies an extended gravity model where, to account for the multilateral resistance to migration (i.e. the influence of alternative destinations on bilateral migration rates), different structures of fixed effects (monadic and dyadic) are used. Second, the paper tries to provide new insights into the existence of nonlinearities in the main economic variables explaining migration; this is done by resorting to panel threshold techniques where the threshold level is endogenously determined (Hansen, 1999). Although there is some evidence confirming the existence of nonlinear effects, most of these previous papers tend to select arbitrary or exogenous cut-off values (see, for instance, Burda et al., 1998; Juárez, 2000; Andrienko and Guriev, 2004; or Congregado et al., 2011) instead of endogenous ones.³ As far as we know, this is the first paper in which, as well as controlling for multilateral resistance, the threshold level for variables affecting migration is endogenously determined.

To accomplish these goals, bilateral origin-destination gross migration flows among Spanish provinces are used. Spanish provinces (50) correspond to the third level of the Nomenclature of Territorial Units for Statistics (NUTS-3) (see Figure 1). The sample period spans, for reasons of data availability, from 2004 to 2014.

INSERT FIGURE 1 AROUND HERE

The remainder of the paper is structured as follows. Section 2 provides a brief literature review on the determinants of internal migration in Spain. Section 3 describes the empirical strategy. Section 4 presents the model specification and data. Section 5 discusses the results. Finally, the main conclusions and implications for policy are outlined in Section 6.

LITERATURE REVIEW ON THE DETERMINANTS OF INTERNAL MIGRATION: THE SPANISH CASE

As pointed out in the Introduction, papers on migration determinants can be classified into two categories. On the one hand, those mainly highlighting the importance of economic factors (such as unemployment rates, wages and employment growth) and consequently, in line with the disequilibrium model of migration. On the other, those studies stressing the relevant role played by amenities, thereby giving support to the equilibrium model.

In the first group, García-Ferrer (1980) and Santillana (1981) can be considered the pioneering papers proving that migration across Spanish provinces is mainly driven by economic factors. Then, Bentolila and Dolado (1991), Juárez (2000) and Devillanova and García-Fontes (2004) confirm this conclusion, although the former raises some doubts about the intensity of these factors when explaining migratory flows. For their part, Mulhern and Watson (2009, 2010) prove that, apart from traditional economic variables, housing prices and distance (which can be considered a proxy for transportation costs) also influence the decision to migrate. By adopting a more microeconomic approach than previous papers, Antolín and Bover (1997) and Bover and Arellano (2002) demonstrate that personal characteristics not only have an important direct effect on migration but also an indirect one through their interaction with economic variables; specifically, these papers indicate that the higher the education level, the greater the probability of migrating.

On the other hand, there is also a bunch of papers highlighting the influence of non-economic factors. Ródenas (1994) and De la Fuente (1999), for example, detect signs of the relevant role played by amenities, along with a strong reduction in the incentive to migrate as a response to

labour factors. Other studies emphasizing the importance of quality of life variables are the ones developed by Maza and Villaverde (2004, 2008). The former, by employing semiparametric techniques, finds that Spanish internal migration is mainly driven not only by income but also by climate condition differentials between origin and destination regions. Likewise, unemployment and housing prices appear to affect net migration rates, although to a lesser extent. In the subsequent study the same authors, by using spatial econometric techniques, confirm previous results and highlight the importance of variables such as human capital and population density. Finally, the paper by Lago and Aguayo (2004) also points to the importance of variables measuring quality of life and climatology.

While previous papers do not differentiate between natives and foreigners, it is important to highlight that there exists another group of studies trying to assess, specifically, factors behind the internal migration of foreign population. In this regard, Recaño and Roig (2006), for the years 2003 and 2004, prove that foreign internal migrants tend to be influenced by the existence of social networks and distance between provinces, with economic factors still not having a great impact. Additionally, Reher and Silvestre (2009), using more recent data (year 2007) taken from the National Survey of Immigration and estimating logit and binomial regression models, highlight the importance of factors such as educational degree attainment, knowledge of the Spanish language, income levels and social networks on the propensity to move of foreigners. Likewise, covering a longer period of time (2000-09), Maza et al. (2013) examine the settlement patterns of foreigners across the Spanish provinces. They estimate a model for both total foreign population and four broad areas of origin separately (EU-15, New Member States, South America and Africa). Their results reinforce that social networks and economic factors play a significant role on migration decisions, although there are also some key differences in intensity among these groups. Finally, a recent paper by Gutiérrez-Portilla et al. (2018) provides new insights into the factors that shape internal migration flows of foreigners (and natives) in Spain both before and after the outbreak of the crisis. Their results demonstrate that the role played by labour factors is especially important for foreigners.

EMPIRICAL STRATEGY

As mentioned in the Introduction, this paper combines a gravity equation derived from the RUM model for migration (so it deals with multilateral resistance by using combinations of fixed effects), with threshold regression techniques to test for the existence of nonlinearities.

This section is aimed at presenting the empirical procedure that we will follow to identify endogenous thresholds. We think that this feature cannot be overlooked, as some papers have already reported the existence of thresholds, or nonlinear behaviours, in labour market variables. To address this issue, however, they have mainly used arbitrary or exogenous thresholds (Burda et al., 1998; Andrienko and Guriev, 2004; Juárez, 2000; Congregado et al., 2011), while here we draw upon endogenous methods. Specifically, we follow Hansen (1999), which allows us to test for the existence of one or multiple thresholds in a specific explanatory variable and obtain an endogenous estimation of the threshold parameter(s).

Let us consider a three-dimensional (as in our case study) equation given as follows:

$$Y_{jk,t} = \theta \mathbf{X}_{jk,t} + \varepsilon_{jk,t} \quad (1)$$

where $Y_{jk,t}$ is the dependent variable, $\mathbf{X}_{jk,t}$ denotes a vector encompassing the regressors, and $\varepsilon_{jk,t}$ is the error-term.

The general model proposed by Hansen (1999) to identify nonlinearities is as follows:

$$Y_{jk,t} = \begin{cases} \rho_1 Q_{jk,t} + \theta_1 \mathbf{X}_{jk,t} + \varepsilon_{jk,t} , & Q_{jk,t} \leq \gamma \\ \rho_2 Q_{jk,t} + \theta_2 \mathbf{X}_{jk,t} + \varepsilon_{jk,t} , & Q_{jk,t} > \gamma \end{cases} \quad (2)$$

where $Q_{jk,t}$ represents the threshold variable (extracted from $\mathbf{X}_{jk,t}$) and γ is the threshold parameter. It can be assumed that, for a given value of the threshold parameter linked to the threshold variable, $\hat{\rho}_1(\gamma)$, $\hat{\theta}_1(\gamma)$, $\hat{\rho}_2(\gamma)$ and $\hat{\theta}_2(\gamma)$ denote the corresponding estimates of the slope coefficients. Then, the most straightforward computing method to estimate the threshold parameter is through concentration. Consequently, the sum of squared errors $S(\gamma)$ conditioned to a value of γ can be expressed as follows:

$$S(\gamma) = \sum_{j=1}^N \sum_{k=1}^N \sum_{t=1}^T \hat{\varepsilon}_{jk,t}^2(\gamma) \quad (3)$$

Therefore, the level of the threshold that minimises $S(\gamma)$ is the consistent estimate of the threshold:

$$\hat{\gamma} = \underset{\gamma}{ArgMin} S(\gamma) \quad (4)$$

To avoid that during the minimization process a threshold $\hat{\gamma}$ sorts too few observations into each regime, it is convenient to restrict the search in (4) to values of γ so that a minimal percentage of the observations lie on each side of the threshold (Hansen, 1999). Therefore, we use a grid search over the potential values of the threshold variable with a 5% trimming.

Apart from that, we should determine the number of thresholds to specify the correct model. So, the first step consists of testing whether the threshold effect is statistically significant in the one-single-threshold model. In this case, the null hypothesis ($H_0: \rho_1 = \rho_2; \theta_1 = \theta_2$), which corresponds to a linear model, could be tested using a standard test. If S_0 denotes the sum of squares of the linear model, the approximate likelihood ratio test of H_0 can be specified as:

$$F_1 = \frac{S_0 - S(\hat{\gamma})}{\hat{\sigma}^2} \quad (5)$$

being $\hat{\sigma}^2$ a convergent estimate of σ^2 . The main problem here is that the threshold parameter γ is not identified under the null hypothesis of no threshold effect and the asymptotic distribution of F_1 is not standard. To overcome this issue, Hansen (1999) proposes to use bootstrap simulations to compute the p-value of the distribution in the context of panel models and test for the nonlinearity hypothesis. Then, if the p-value associated to F_1 did not reject the linear hypothesis, obviously there would be no thresholds. If, on the contrary, the p-value led to rejecting the linear hypothesis, there would be, at least, one threshold.

This being the case, the hypothesis of one threshold against the alternative of two thresholds should be tested. Once the corresponding changes in equation (2) are made, the likelihood ratio statistic is given by:

$$F_2 = \frac{S(\hat{\gamma}) - S(\hat{\gamma}_1, \hat{\gamma}_2)}{\hat{\sigma}^2} \quad (6)$$

where obviously $S(\hat{\gamma}_1, \hat{\gamma}_2)$ denotes the corresponding residual sum of squares and $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are the threshold estimates of a double-threshold model. Likewise, if the bootstrap p-value associated to F_2 led to rejecting the null hypothesis of one threshold, we then should discriminate between two and three thresholds and so on. Conversely, if the p-value is such that we cannot reject the null of one threshold, the correct specification is the model with one threshold.

MODEL SPECIFICATION AND DATA

Once the methodology to identify endogenous thresholds has been explained, this section presents the model used to estimate the main factors behind internal movements across Spanish provinces of foreigners and natives, as well as data description. The baseline specification of our migration model is as follows:

$$MIGR_{jk,t} = \theta X_{kj,t-1} + \varepsilon_{jk,t} \quad (7)$$

where the dependent variable, $MIGR_{jk,t}$, refers to the gross migration rate between pairs of provinces j and k at time t , and $X_{kj,t-1}$ is a vector of independent variables that is defined in relative terms (ratio destination/origin) -as migrants compare the situation of these variables in the potential destination (k) to that in their origin province (j)-. Independent variables are also lagged one year (apart from those variables that are time-fixed) to avoid endogeneity problems and to reflect that they normally affect migration in the following period.⁴

Regarding the dependent variable, it refers to annual internal migration by province of origin and destination, for both foreigners⁵ and natives, between 2004 and 2014. Internal migration data come from the Statistic of Residential Variations (EVR) provided by the Spanish National Statistics Institute (INE). This archive consists of official changes of residence between municipalities registered in the Civil Register, and it is considered the most reliable database on internal migration in Spain (Martí and Ródenas, 2004). The main advantage of this database (over Census data) is that it allows us to construct a three-dimensional panel (jkt) to properly control for the multilateral resistance to migration by using different structures of fixed effects (see below).

As for the deterministic factors, here $X_{kj,t-1} = [WAGE_{kj,t-1}, UR_{kj,t-1}, HOUS_{kj,t-1}, AGR_{kj,t-1}, CONST_{kj,t-1}, SER_{kj,t-1}, DIST_{jk}, CLIM_{kj}, CFI_{kj}, SSI_{kj}]$. Table 1 describes these variables, their definition, and data source. Although the selection of variables is based on literature on this topic, it is admittedly ad hoc so that giving reasons supporting it seems to be pertinent. Regarding economic determinants, it has been proved in the literature that migrants are attracted by labour-market opportunities that maximise their expected wages, choosing areas with higher wages and lower unemployment rates (Mulhern and Watson, 2009, 2010; Etzo, 2011; Maza et al., 2013). Consequently, we include wages ($WAGE$) and unemployment rates (UR).

Additionally, amenities or quality of life factors are also driving forces of internal migration (Knapp and Graves, 1989; Rappaport, 2007; Partridge, 2010; Rodríguez-Pose and Ketterer, 2012; Krivokapic-Skoko and Collins, 2016). We use three different variables to measure amenities. On the one hand, as suggested by Rappaport (2007), people tend to move toward places with nice weather. So we use a climatic conditions index ($CLIM$) which penalises extreme weather conditions in favour of mild climate. On the other, based on the quality of the life index computed in Royuela et al. (2003), two variables measuring human-made amenities are taken into account: the cultural facilities index (CFI) and the social services index (SSI).

Additionally, the accessibility of each destination is usually proxied by the distance from origin to destination, so migrants show less willingness to move as distance (*DIST*) increases (evidence for Spain can be found in Ródenas, 1994; Recaño and Roig, 2006; Mulhern and Watson, 2009, 2010). Furthermore, it is also well accepted that housing prices differentials (*HOUS*) are a major determinant of migration decisions, making people move toward places where these prices are lower (Jackman and Savouri, 1992a,b). Finally, we also control for the sectoral structure of the economy by including the shares of agriculture (*AGR*), construction (*CONST*) and service (*SER*) sectors of GDP; we leave the industry sector out of the model to avoid multicollinearity problems. The inclusion of these industry-mix variables allows us to detect the role of each sector in fostering or discouraging migration.

INSERT TABLE 1 AROUND HERE

Now, we want to test for nonlinearities in the two key economic variables affecting migration: wages and unemployment rates. To consider both variables at the same time, we built the so-called expected wage variable (*ExpWAGE*) -defined as wages multiplied by employment probabilities (1-unemployment rates)- to apply the threshold methodology explained before. To do so, Equations (8) and (9) are used.

$$MIGR_{jk,t} = \rho_1 ExpWAGE_{kj,t-1} + \theta_2 HOUS_{kj,t-1} + \theta_3 AGR_{kj,t-1} + \theta_4 CONST_{kj,t-1} + \theta_5 SER_{kj,t-1} + \theta_6 DIST_{jk} + \theta_7 CLIM_{kj} + \theta_8 CFI_{kj} + \theta_9 SSI_{kj} + \varepsilon_{1,jkt} \quad \text{if } ExpWAGE_{kj,t-1} \leq \gamma \quad (8a)$$

$$MIGR_{jk,t} = \rho_2 ExpWAGE_{kj,t-1} + \theta_2 HOUS_{kj,t-1} + \theta_3 AGR_{kj,t-1} + \theta_4 CONST_{kj,t-1} + \theta_5 SER_{kj,t-1} + \theta_6 DIST_{jk} + \theta_7 CLIM_{kj} + \theta_8 CFI_{kj} + \theta_9 SSI_{kj} + \varepsilon_{2,jkt} \quad \text{if } ExpWAGE_{kj,t-1} > \gamma \quad (8b)$$

$$MIGR_{jk,t} = \rho_1 ExpWAGE_{kj,t-1} + \theta_2 HOUS_{kj,t-1} + \theta_3 AGR_{kj,t-1} + \theta_4 CONST_{kj,t-1} + \theta_5 SER_{kj,t-1} + \varepsilon_{1,jkt} \quad \text{if } ExpWAGE_{kj,t-1} \leq \gamma \quad (9a)$$

$$MIGR_{jk,t} = \rho_2 ExpWAGE_{kj,t-1} + \theta_2 HOUS_{kj,t-1} + \theta_3 AGR_{kj,t-1} + \theta_4 CONST_{kj,t-1} + \theta_5 SER_{kj,t-1} + \varepsilon_{2,jkt} \quad \text{if } ExpWAGE_{kj,t-1} > \gamma \quad (9b)$$

The difference between them is that while Equation (8) encompasses all the variables previously defined, Equation (9) does not include the time-invariant explanatory variables. We proceed in this way as different structures of fixed effects will then be incorporated, and thus, we seek to be consistent with the specification.⁶

Once threshold estimates are obtained, we have to create interaction variables with the economic determinants depending on whether the value of the expected wage of each dyad origin-destination lies below or above this estimated threshold. To do so, firstly two dummy

variables ($d \leq \hat{y}$ and $d > \hat{y}$) are defined, which encompass the dyads for which the value of the expected wage is lower or higher than the estimated threshold in each case (Equations 8 and 9); then, we create the interaction variables as the product between these two dummies and the key economic determinants of our model (wages and unemployment). The resulting variables allow us to capture the magnitude of the change in the behaviour (of both natives and foreigners depending on the case) with respect to these factors when migrating internally across Spanish provinces. In other words, the resulting interaction variables, for wages, $WAGE_{kj,t-1} * (d \leq \hat{y})$ and $WAGE_{kj,t-1} * (d > \hat{y})$ and unemployment, $UR_{kj,t-1} * (d \leq \hat{y})$ and $UR_{kj,t-1} * (d > \hat{y})$, will identify potential nonlinearities once the value of the threshold in the expected wage variable has been endogenously determined.

After that, as pointed out before, in order to account for the multilateral resistance to migration, different combinations of fixed effects will be included in the final model. This phenomenon, overlooked in most migration studies, is defined as the confounding influence that the attractiveness of alternative destinations exerts on the bilateral migration rate. Its importance has been highlighted by Bertoli and Fernández-Huertas Moraga (2013), who show that multilateral resistance to migration is properly captured with the inclusion of different structures of fixed effects, giving rise to consistent estimates.⁷ More specifically, as the data employed has three dimensions (origin j , destination k , and moment in time t), four different structures of fixed effects have been considered (models (i) to (iv)). Specifically, for models (i), (iii) and (iv), we draw on the results regarding thresholds obtained in Equation (8), and for model (ii), in Equation (9).⁸

(i) Model with monadic fixed effects of origin, destination and time:

$$\begin{aligned} MIGR_{jk,t} = & \beta_1 WAGE_{kj,t-1} * (d \leq \hat{y}) + \beta_2 WAGE_{kj,t-1} * (d > \hat{y}) + \beta_3 UR_{kj,t-1} * (d \leq \hat{y}) + \\ & \beta_4 UR_{kj,t-1} * (d > \hat{y}) + \beta_5 HOUS_{kj,t-1} + \beta_6 AGR_{kj,t-1} + \beta_7 CONST_{kj,t-1} + \\ & \beta_8 SER_{kj,t-1} + \beta_9 DIST_{jk} + \beta_{10} CLIM_{kj} + \beta_{11} CFI_{kj} + \beta_{12} SSI_{kj} + \alpha_j + \alpha_k + \alpha_t + \varepsilon_{jkt} \end{aligned} \quad (10)$$

where α_j and α_k represent provincial fixed effects of origin and destination, respectively. As shown in the literature (Mayda, 2010; Beine and Parsons, 2015; Royuela and Ordóñez, 2016), they allow us to control for specific effects of each origin and destination that are not taken into account by deterministic components of utility, either time-invariant origin push (destination pull) factors or time-invariant origin (destination) related costs variables. For instance, these dummies control for specific migration policies (McKenzie et al., 2014; Beine et al., 2016).

Besides, temporal fixed effects, denoted by α_t are also included to capture common shocks to all provinces considered in the sample in each year.

(ii) Model with dyadic fixed effects of origin-destination and monadic fixed effects of time:

$$\begin{aligned} MIGR_{jk,t} = & \beta_1 WAGE_{kj,t-1} * (d \leq \hat{\gamma}) + \beta_2 WAGE_{kj,t-1} * (d > \hat{\gamma}) + \beta_3 UR_{kj,t-1} * (d \leq \hat{\gamma}) + \\ & \beta_4 UR_{kj,t-1} * (d > \hat{\gamma}) + \beta_5 HOUS_{kj,t-1} + \beta_6 AGR_{kj,t-1} + \beta_7 CONST_{kj,t-1} + \\ & \beta_8 SER_{kj,t-1} + \alpha_{jk} + \alpha_t + \varepsilon_{jkt} \end{aligned} \quad (11)$$

where, apart from temporal fixed effects, α_{jk} denotes fixed effects for each combination of origin and destination provinces; they are introduced to control for time-invariant features common to each pair of provinces (Mayda, 2010; Ortega and Peri, 2013). Some of the factors that are captured with this dyadic structure include migration networks between pairs of provinces and bilateral migration costs such as distance or culture proximity (Beine et al., 2016). As can be appreciated, in this case we leave out of the equation the variables $DIST_{jk}$, $CLIM_{kj}$, CFI_{kj} and SSI_{kj} since they are constant within province pairs and, consequently, would be perfectly collinear with this structure of fixed effects.

(iii) Model with dyadic fixed effects of origin-time and monadic fixed effects of destination:

$$\begin{aligned} MIGR_{jk,t} = & \beta_1 WAGE_{kj,t-1} * (d \leq \hat{\gamma}) + \beta_2 WAGE_{kj,t-1} * (d > \hat{\gamma}) + \beta_3 UR_{kj,t-1} * (d \leq \hat{\gamma}) + \\ & \beta_4 UR_{kj,t-1} * (d > \hat{\gamma}) + \beta_5 HOUS_{kj,t-1} + \beta_6 AGR_{kj,t-1} + \beta_7 CONST_{kj,t-1} + \\ & \beta_8 SER_{kj,t-1} + \beta_9 DIST_{jk} + \beta_{10} CLIM_{kj} + \beta_{11} CFI_{kj} + \beta_{12} SSI_{kj} + \alpha_{jt} + \alpha_k + \varepsilon_{jkt} \end{aligned} \quad (12)$$

where α_{jt} represents dyadic fixed effects that are specific to each province of origin and each year. They allow us to capture all the push determinants of migration together with the multilateral resistance derived from heterogeneity in migration preferences by origin (Royuela and Ordóñez, 2016). Provincial fixed effects by province of destination (α_k) have also been considered in this Equation.

(iv) Model with dyadic fixed effects of destination-time and monadic fixed effects by origin:

$$\begin{aligned} MIGR_{jk,t} = & \beta_1 WAGE_{kj,t-1} * (d \leq \hat{\gamma}) + \beta_2 WAGE_{kj,t-1} * (d > \hat{\gamma}) + \beta_3 UR_{kj,t-1} * (d \leq \hat{\gamma}) + \\ & \beta_4 UR_{kj,t-1} * (d > \hat{\gamma}) + \beta_5 HOUS_{kj,t-1} + \beta_6 AGR_{kj,t-1} + \beta_7 CONST_{kj,t-1} + \\ & \beta_8 SER_{kj,t-1} + \beta_9 DIST_{jk} + \beta_{10} CLIM_{kj} + \beta_{11} CFI_{kj} + \beta_{12} SSI_{kj} + \alpha_{kt} + \alpha_j + \varepsilon_{jkt} \end{aligned} \quad (13)$$

where apart from provincial fixed effects by province of origin (α_j), we also include dyadic fixed effects that are specific to each province of destination and year (α_{kt}). This structure of fixed effects enables us to control for the pull determinants of migration and the multilateral

resistance derived from heterogeneity in the expectations about each potential province of destination (Royuela and Ordóñez, 2016).⁹

RESULTS

The aim of this section is to discuss the results. First, those corresponding to the threshold parameter estimation; second, the results of the gravity model estimation including the thresholds previously obtained.

Table 2 presents the results of the tests to determine the number of thresholds (tests statistics F_1 and F_2 and their bootstrap p-values (Hansen, 1999)) for natives and foreigners and for Equations (8) and (9). As can be seen, the p-values associated to F_1 lead us to reject the null hypothesis of no threshold at the 5% significance level; with respect to the test for double threshold, the null hypothesis of one threshold cannot be rejected. This evidence confirms that we should specify and estimate the one-single-threshold model. As regards the estimated thresholds, Table 2 shows that, when all variables are included (Equation 8), the estimated thresholds in the expected wage variable for natives and foreigners are 1.473 and 1.594, respectively; in Equation (9), the estimated threshold coincides for natives and foreigners, 1.474.

INSERT TABLE 2 AROUND HERE

Table 3 presents the results of estimating Equations (10) to (13) for natives and foreigners. The estimation is performed by Generalised Least Squares (GLS) as the Breusch-Pagan test points to the presence of heteroscedasticity.

INSERT TABLE 3 AROUND HERE

To begin with, as regards the economic determinants of internal migration, it can be seen that the relative wage between the provinces of origin and destination shapes internal migration flows. The two associated interaction variables result positive and statistically significant in all equations for both natives and foreigners. As for natives, the interaction variable above the threshold always shows a higher coefficient than below it. This nonlinear effect is more evident, in any case, when controlling for origin, destination and time fixed effects (Equation 10), and for destination's specific circumstances (Equation 13). When comparing the two groups of migrants, foreigners seem to be more affected than natives by wage differentials between provinces of origin and destination, a result that is in line with the evidence found by Gutiérrez-

Portilla et al. (2018). However, only Equation 11 (which controls for time-invariant features common to each pair of provinces) does reflect a nonlinear behaviour on the wage variable.¹⁰

As for the unemployment rate, some relevant results arise: firstly, in the case of natives, when above the endogenously estimated threshold, a higher unemployment rate in the province of destination than in that of origin seems to discourage them from migrating, except when controlling for specific factors in origin (Equation 12). On the other hand, below the threshold -that is, when there exist small economic differentials between origin and destination- only the inclusion of features common to each pair of provinces gives rise to a negative and statistically effect of unemployment (Equation 11); when other combinations of fixed effects are considered, a rather small, although positive effect is found. Consequently, we can conclude that the effect of unemployment in the group of natives is nonlinear (Juárez, 2000, Clemente et al., 2016). In the case of foreigners, the deterrent effect of higher unemployment rates in the province of destination with respect to that of origin on internal migration is greater than the effect found for natives, although this does not happen, however, when dyadic fixed effects of destination-time are included (Equation 13). For foreigners, as in the case of wages, the parameters below and above the threshold only statistically differ from each other when dyadic fixed effects of origin-destination are taken into account. Thus, the nonlinear behaviour of the unemployment for foreigners is not as clear as for natives.

The housing price variable also discloses interesting findings: it shows a negative (positive) and statistically significant coefficient in the case of natives (foreigners), regardless of the structure of fixed effects considered, with the only exception of Equation (13) for foreigners (where is non-significant). This implies that natives look for lower housing prices when moving internally across Spanish provinces. This preference for cheaper housing has been reported by previous literature (Berger and Blomquist, 1992; Bover and Arellano, 2002; Maza and Villaverde, 2004; Gutiérrez-Portilla et al., 2018). Regarding foreign population, two tentative explanations could be behind the positive effect: first, as pointed out by Rappaport (2007: 377) *“individuals are willing to endure greater crowdedness and the associated higher price of housing in order to directly enjoy higher quality of life and to indirectly enjoy higher productivity via the higher wage it affords”*; second, it could be due to the fact that foreign population is more likely to rent rather than buy a house.

Concerning sectoral structure, the results suggest that, in most cases, a relatively higher share of the agricultural (service) sector in destination with respect to origin hinders (fosters) migration, both for foreigners and natives. This is a rather expected finding due to the relatively

greater employment opportunities offered by the service sector when compared with the agriculture one (Bover and Arellano, 2002; Paluzie et al., 2009). Only in Equation (11), when controlling for specific characteristics for each pair of origin-destination provinces, agriculture (services) shows a positive (negative) and statistically significant coefficient for foreigners (natives). As for construction, it exhibits huge differences between foreigners and natives (Fromentin, 2016); more specifically, it slightly discourages migration for natives whereas it reveals a positive impact on the migration process for foreigners.

With regard to the distance variable, which represents the main costs associated to migration decisions, it has the expected negative sign, in line with previous evidence (Ródenas, 1994; Recaño and Roig, 2006; Mulhern and Watson, 2009, 2010). The dissuasive effect of distance is greater for foreigners than natives, a result in accordance with Gutiérrez-Portilla et al. (2018); this difference becomes especially high after controlling for specific effects of each origin and destination (Equation 10).

Finally, as for amenities or quality of life factors, several features are worth being highlighted. First, more pleasant climatic conditions in the province of destination than in origin seem to be a pull factor for internal migration among native population. This conclusion is in accordance with those obtained by Maza and Villaverde (2004), Faggian and Royuela (2010) and Gutiérrez-Portilla et al. (2018). More precisely, the effect is higher when specific characteristics of the origin and destination provinces are captured individually (Equation 10). Second, foreigners are affected by cultural facilities and social services to a higher degree than natives; the coefficients associated to these variables show positive and statistically significant values in both cases. As can be appreciated, the sign and value of these coefficients as well as their level of significance are quite similar in all the specifications, irrespective of the combination of fixed effects used. This provides evidence on the fact that the multilateral resistance to migration does not bias the effect of quality of life factors on internal migration. This also happened with the distance variable, which effect is robust to variations in the model specification.

To sum up, the main conclusion is that the hypothesis laid out in the Introduction that natives' preferences are more amenity-based/less economic-oriented than foreigners' is partly true. On the one hand, the impact of economic determinants (wages and unemployment) is higher for foreigners than for natives, a nonlinear effect of these two variables being more frequently detected for the latter. This effect does only occur within the group of foreigners when time-invariant features common to each pair of provinces are taken into account. On the other hand, the nature of the most influential amenities for each group of population differs: while natives

tend to seek for provinces with good climate, foreigners tend to be attracted by provinces with more cultural and social amenities.

CONCLUSIONS

This paper examines the factors driving internal mobility of foreigners and natives across Spanish provinces over the period 2004-14. Specifically, it focuses on two types of factors, economic and non-economic (amenities) factors. To accomplish this aim, a gravity model with four different combinations of fixed effects is estimated. By proceeding in this way, we obtain robust estimations of the parameters, which are free from the bias induced by multilateral resistance to migration.

Another further feature of our analysis is the consideration of nonlinearities in the economic determinants of the proposed model, which allows us to verify whether there exists a nonlinear effect of wage and/or unemployment rate differentials on internal movements of both foreigners and natives. To do so, the Hansen (1999)'s endogenous method of threshold selection is applied and, once the thresholds are found, interaction variables (referring to wage and unemployment rates) are created.

The findings of the paper indicate that, in line with the evidence found in Germany (Schündeln, 2014), foreigners show more responsiveness to economic determinants than natives when moving across Spanish provinces. As for their nonlinear effect on internal migration, there exist differences between foreigners and natives. With respect to native population, the impact of a higher wage in the province of destination with respect to the province of origin is greater above than below the endogenously determined threshold, regardless of the structure of fixed effects considered; with respect to unemployment, there also seems to exist a nonlinear behaviour, although less marked than that found in wages. In the case of foreigners, a nonlinear effect of both variables is just found when controlling for dyadic origin-destination fixed effects.

Results concerning the rest of variables reveal that distance between the province of origin and destination arises as a deterrent factor to migration, and that amenities emerge as determining factors behind the decision to migrate for both groups of population. The nature of the amenities playing a higher role is different between both groups, though; foreigners look for social services and cultural amenities, while pleasant climate conditions are more appealing to natives.

In conclusion, the evidence shown in this paper points to the fact that the disequilibrium and the equilibrium models of migration cannot be totally separated in order to explain internal

migration of foreigners and natives in Spain. Although specific for this country, these results are somehow related to those found by Biagi et al. (2011) for internal migration in Italy, and Rodríguez-Pose and Ketterer (2012) for migration across European regions, as they suggest that both approaches (disequilibrium and equilibrium) must be considered complementary rather than alternative to get a good knowledge of the migration phenomenon.

Some recommendations for policy-making can be drawn from this study. First, according to what we have just stated it is obvious that the role played by migratory movements in the search of internal cohesion and reduction of provincial disparities in Spain is not instrumental; economic determinants, although important, are not the only factor at play. Therefore, additional policies trying to foster economic growth of the poorest provinces are necessary.

Second, as the distinction between natives and foreigners revealed that the latter show more response to economic determinants than natives do, some policies could be welcome when it comes to increasing the intensity of ‘equilibrium’ migration. As foreigners tend to be at a disadvantage compared with natives when looking for vacancies in other provinces (Corkill, 2001), policies aimed at improving information channels would be desirable if feasible. In the same vein, policies trying to ease the match between vacancies and labour skills, a problem that once again is more acute for foreigners, would be pertinent. We are referring to, for instance, training and active labour market policies to alleviate job-skill mismatches (Gutiérrez-Posada et al., 2017).

Third, our results uncovered the fact that, especially for natives, the relationship between wage disparities and migratory flows is not linear. It is more intense the higher the difference between wages in the origin and destination provinces. That being so, one could expect that a more flexible wage formation system with a less centralised bargaining process in Spain might contribute, by enlarging those differences, to trigger, as a side effect, internal mobility.

Finally, we want to stress some limitations of this paper. Among them, it is important to note that the study cannot be carried out at a more spatially disaggregated level (because of data availability problems), so intra-provincial movements of foreign and native populations are not considered. Additionally, the so-called Modifiable Areal Unit Problem (MAUP) also comes into action. When possible, the use of analytical areas (Rubiera-Morollón and Viñuela, 2013; Viñuela et al., 2014), internally more homogeneous than provinces, would overcome these drawbacks.

Another weakness of the paper has to do with the lack of data on migratory flows between pairs of provinces by the age of migrants, which prevents us from incorporating this variable into the analysis. In any case, as data for flows of native and foreign populations from each province to the remaining ones by age ranges do exist, we have tried to somewhat overcome this limitation by computing an indicator of the average age at which natives and foreigners leave each province. This indicator reveals that, on average, natives migrating across Spanish provinces are slightly older than foreigners (38 versus 36 years old). Taking each province separately, differences are not either quite remarkable. For these reasons, we do not believe the previous conclusions would substantially change if the variable ‘age’ were included in the model.

¹ International evidence on the determinants of internal migration comparing natives and foreigners is rather scarce: Schündeln (2014) points to higher responsiveness to labour market differentials in the immigrant than in the native population in Germany. Lamonica and Zagaglia (2013) conclude that economic and demographic factors affect Italians and foreigners differently: first, the demographic situation of the sending regions is proved to be a push factor only for the mobility of the Italian population; and second, while the economic conditions of the sending regions have a much greater impact on the Italians, foreigners are more influenced by better economic conditions in the destination regions. Although dealing with a different topic, Auer et al. (2017) find that migration-related factors such as weaker work values or lower quality of informal networks help to explain the overall disadvantage in unemployment duration of immigrants with respect to Swiss nationals.

² Although dealing with population growth in Spain rather than migration, Gutiérrez-Posada et al. (2017) is another paper deserving attention. It adopts a geographically weighted regression approach to confirm that important factors such as distance (either to the big cities or to the coast) and size (initial population) can have different effects on population growth across both space and time.

³ The only exceptions are the papers by Clemente et al. (2016) and Gutiérrez-Portilla et al. (2018).

⁴ In this regard, as indicated by Beine et al. (2016: 504), “*controlling for multilateral resistance to migration can make instrumentation unnecessary as long as endogeneity problem is not due to reverse causality, or as long as the resistance terms capture a big part of the omitted factors*”.

⁵ Foreigners are defined by nationality.

⁶ To be more precise, the specification of these two Equations (required to endogenously estimate the threshold) must be such that, in the subsequent stage, when four different combinations of fixed effects are included, there are no variables that might cause multicollinearity problems.

⁷ Other papers supporting the use of different structures of fixed effects are Beine et al. (2011), Ortega and Peri (2013), McKenzie et al. (2014) and Beine and Parsons (2015).

⁸ This is so because model (ii) incorporates dyadic fixed effects of origin-destination to control for time-invariant characteristics of each pair of provinces, so it would not be possible to use the results of thresholds obtained in Equation (8) as it encompasses all the variables (including the time-invariant ones) and consequently, multicollinearity problems would arise in the estimation.

⁹ These dummies can also capture migration policies, which are difficult to be considered otherwise (Beine and Parsons, 2015).

¹⁰ Wald tests for equality of parameters suggest that, in the case of foreigners, the coefficients of the interaction variables below and above the threshold do not statistically differ from each other in Equations (10), (12) and (13).

REFERENCES

- Amuedo-Dorantes, C. and De la Rica, S.
2010 “Immigrants’ responsiveness to labor market conditions and their impact of regional employment disparities: evidence from Spain”, *SERIEs*, 1(4): 387-407.
- Andrienko, Y. and Guriev, S.
2004 “Determinants of interregional mobility in Russia”. *Economics of Transition*, 12(1): 1-27.
- Antolín, P. 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*, 59(2): 215-235.
- Auer, D., Bonoli, G. and Fossati, J.
2017 “Why do immigrants have longer periods of unemployment? Swiss evidence”, *International Migration*, 55(1): 157-174.
- Beine M, Bertoli S and Fernández-Huertas Moraga, J.
2016 “A practitioners’ guide to gravity models of international migration”, *The World Economy*, 39(4): 496-512.
- Beine, M., Docquier, F. and Ozden, Ç.
2011 “Diasporas”, *Journal of Development Economics*, 95(1): 30-41.
- Beine, M., and Parsons, C.
2015 “Climatic factors as determinants of international migration”, *Scandinavian Journal of Economics*, 117(2): 723-767.
- Bentolila, S. and Dolado J.J.
1991 “Mismatch and internal migration in Spain, 1962-1986”. In Fiorella Padeo Schioppa (Ed.) *Mismatch and labor mobility*, Cambridge University Press, Cambridge.
- Berger, M.C. and Blomquist, G.C.
1992 “Mobility and destination in migration decisions: the roles of earnings, quality of life and housing prices”, *Journal of Housing Economics*, 2(1): 37-59.
- Bertoli, S. and Fernández-Huertas Moraga, J.
2013 “Multilateral resistance to migration”, *Journal of Development Economics*, 102: 79-100.
- 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*, 6(1): 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*, 15(2): 357-380.

Boyer, R. and Savageau, D.

1985 *Places Rated Almanac*. Rand McNally, Chicago.

Brakman, S., Garretsen, H. and Van Marrewijk, C.

2009 “Economic geography within and between European nations: the role of market potential and density across space and time”, *Journal of Regional Science*, 49(4): 777-800.

Bruna, F., López-Rodríguez, J. and Faiña, A.

2016 “Market potential, spatial dependences and spillovers in European regions”, *Regional Studies*, 50(9): 1551-1563.

Burda, M.C., Härdle, W., Müller, M. and Werwatz, A.

1998 “Semiparametric analysis of German East-West migration intentions: facts and theory”, *Journal of Applied Economics*, 13(5): 525-541.

Clemente, J., Larramona, G. and Olmos, L.

2016 “Interregional migration and thresholds: evidence from Spain”, *Spatial Economic Analysis*, 11(3): 276-293.

Congregado, E., Golpe, A.A. and Van Stel, A.

2011 “Exploring the big jump in the Spanish unemployment rate: evidence on an ‘added-worker’ effect”, *Economic Modelling*, 28(3): 1099-1105.

Corkill, D.

2001 “Economic migrants and the labour market in Spain and Portugal”, *Ethnic and Racial Studies*, 24(5): 828-844.

De la Fuente, A.

1999 “La dinámica territorial de la población española: un panorama y algunos resultados provisionales”, *Revista de Economía Aplicada*, 20(7): 53-108.

Deller, S.C., Tsung-Hsiu, T., Marcouiller, D.W. and English, D.B.K.

2001 “The role of amenities and quality of life in rural economic growth”, *American Journal of Agricultural Economics*, 83: 352-365.

Détang-Dessendre, C., Partridge, M.D. and Piguet, V.

2016 “Local labor market flexibility in perceived low migration country: the case of French labor markets”, *Regional Science and Urban Economics*, 58: 89-103.

Devillanova, C. and García-Fontes, W.

- 2004 “Migration across Spanish provinces: evidence from the social security records (1978-1992)”, *Investigaciones Económicas*, 28: 461-487.
- Etzo, I.
- 2011 “The determinants of the recent interregional migration flows in Italy: a panel data analysis”, *Journal of Regional Science*, 51(5): 948-966.
- Faggian, A. and Royuela, V.
- 2010 “Migration flows and quality of life in a metropolitan area: the case of Barcelona-Spain”, *Applied Research in Quality of Life*, 5(3): 241-259.
- Fromentin, V.
- 2016 “The global economic crisis and migrant workers: the case of the construction sector in Europe”, *International Economic Journal*, 30(1): 147-163.
- García-Ferrer, A.
- 1980 “Interactions between internal migration, employment growth, and regional income differences in Spain”, *Journal Development Economics*, 7(2): 211-229.
- Graves, P.E.
- 1976 “A reexamination of migration, economic opportunity, and the quality of life”, *Journal of Regional Science*, 16: 107-112.
- 1980 “Migration and climate,” *Journal of Regional Science*, 20: 227-237.
- 1983 “Migration with a composite amenity: the role of rents,” *Journal of Regional Science*, 23: 541-546.
- Greenwood, M.J.
- 1975 “Research on internal migration in the United States: a survey”. *Journal of Economic Literature*, 13(2): 397-433.
- 1985 “Human migration: theory, models, and empirical studies”, *Journal of Regional Science*, 25(4): 521-544.
- Gutiérrez-Portilla, M., Maza, A. and Hierro, M.
- 2018 “Foreigners versus natives in Spain: different migration patterns? Any changes in the aftermath of the crisis”, *The Annals of Regional Science*. DOI: 10.1007/s00168-018-0862-9.
- Gutiérrez-Posada, D., Rubiera-Morollón, F. and Viñuela, A.
- 2017 “Heterogeneity in the determinants of population growth at the local level: analysis of the Spanish case with a GWR approach”, *International Regional Science Review*, 40(3): 211-240.
- Hansen, B.

- 1999 "Threshold effects in non-dynamic panels: estimation, testing and inference", *Journal of Econometrics*, 93(2): 345-368.
- Hierro, M.
- 2016 "Latin American migration to Spain: main reasons and future perspectives", *International Migration*, 54(1): 64-83.
- Hierro, M. and Maza, A.
- 2010 "Per capita income convergence and internal migration in Spain: are foreign-born migrants playing an important role?", *Papers in Regional Science*, 89(1): 89-107.
- Jackman, R. and Savouri, S.
- 1992a "Regional migration versus regional commuting: the identification of housing and employment flows", *Scottish Journal of Political Economy*, 39(3): 272-287.
- 1992b "Regional migration in Britain: an analysis of gross flows using NHS Central Register data", *Economic Journal*, 102: 1433-1450.
- Juárez, J.P.
- 2000 "Analysis of interregional labor migration in Spain using gross flows", *Journal of Regional Science*, 40(2): 377-399.
- Knapp, T.A. and Graves P.E.
- 1989 "On the role of amenities in models of migration and regional development", *Journal of Regional Science*, 29(1): 71-87.
- Krivokapic-Skoko, B. and Collins, J.
- 2016 "Looking for rural idyll 'down under': international immigrants in rural Australia", *International Migration*, 54(1): 167-179.
- Lago, C. and Aguayo, E.
- 2004 "Migraciones interiores en España: un modelo econométrico regional 1980-2000", *Regional and Sectoral Studies*, 4(1): 121-138.
- Lamonica, G.R. and Zagaglia B.
- 2013 "The determinants of internal mobility in Italy, 1995-2006: a comparison of Italians and resident foreigners," *Demographic Research*, 29(16): 407-440.
- Lewis, A.
- 1954 "Economic development with unlimited supply of labour", *The Manchester School of Economic and Social Studies*, 22(2): 139-191.
- Martí, M. and Ródenas, C.
- 2004 "Migrantes y migraciones: de nuevo la divergencia en las fuentes estadísticas", *Estadística Española*, 46(156): 293-321.

Mayda, A.M.

2010 “International migration: a panel data analysis of the determinants of bilateral flows”, *Journal of Population Economics*, 23: 1249-1274.

Maza, A. and Villaverde, J.

2004 “Interregional migration in Spain: a semiparametric analysis”, *Review of Regional Studies*, 34(2): 156-171.

2008 “A note on the need to account for spatial dependence: a case of migratory flows in Spain”, *Review of Regional Studies*, 38(1): 105-111.

Maza, A., Villaverde, J. and Hierro, M.

2013 “Explaining the settlement patterns of foreigners in Spain”, *Applied Geography*, 40: 11-20.

McKenzie, D.J., Theoharides, C. and Yang, D.

2014 “Distortions in the international migrant labor market: evidence from Filipino migration and wage responses to destination country economic shocks”, *American Economic Journal: Applied Economics*, 6(2): 49-75.

Mulhern, A. and Watson, J.

2009 “Spanish internal migration: is there anything new to say?”, *Spatial Economic Analysis*, 4(1): 103-120.

2010 “Spanish inter-regional migration: an enigma resolved”, *Applied Economics Letters*, 17(14): 1355-1359.

Muth, R.

1971 “Migration: chicken or egg?”, *Southern Economic Journal*, 37(3): 295-306.

Ortega, F. and Peri, G.

2013 “The effect of income and immigration policies on international migration”, *Migration Studies*, 1: 47-74.

Paluzie, E., Pons, J., Silvestre, J. and Tirado, D.

2009 “Migrants and market potential in Spain over the twentieth century: a test of the new economic geography”, *Spanish Economic Review*, 11(4): 243-265.

Partridge, M.D.

2010 “The dueling models: NEG vs. amenity migration in explaining US engines of growth”, *Papers in Regional Science*, 89(3): 513-536.

Partridge, M.D. and Rickman, D.S.

2003 “The waxing and waning of US regional economies: the chicken-egg of jobs versus people”, *Journal of Urban Economics*, 53(1): 76-97.

Piras, R.

2012 “Internal migration across Italian regions: Macroeconomic determinants and accommodating potential for a dualistic economy”, *The Manchester School*, 80(4): 499-524.

Rappaport, J.

2007 “Moving to nice weather”, *Regional Science and Urban Economics*, 37(3): 375-398.

Ravenstein, E.G.

1885 “The laws of migration”, *Journal of the Statistical Society of London*, 48(2): 167-235.

Recaño, J.

2002 “La movilidad geográfica de la población extranjera en España: un fenómeno emergente”, *Cuadernos de Geografía*, 72: 135-156.

Recaño, J. and Roig, M.

2006 “The internal migration of foreigners in Spain”, Centro de Estudios Demográficos (presented in the European Population Conference, Liverpool).

Redding, S.J. and Venables, A.J.

2004 “Economic geography and international inequality”, *Journal of International Economics*, 62: 53-82.

Reher, D. and Requena, M.

2009 “The national immigrant survey of Spain: a new data source for migration studies in Europe”, *Demographic Research*, 20(12): 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 microdata for Spain”, *International Migration Review*, 43(4): 815-849.

Ródenas, C.

1994 “Migraciones interregionales en España (1960-1989): cambios y barreras”, *Revista de Economía Aplicada*, 2: 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*, 52(4): 535-561.

Royuela, V. and Ordóñez, J.

2016 “Internal migration in a developing country: a panel data analysis of Ecuador (1982-2010)”, *Papers in Regional Science*, forthcoming.

Royuela, V., Suriñach, J., and Reyes, M.

- 2003 “Measuring quality of life in small areas over different periods of time. Analysis of the province of Barcelona” *Social Indicators Research*, 64(1): 51-74.
- Rubiera-Morollón, F., and Viñuela, A.,
- 2013 “From local units to economic regions in Spain: where the agglomerations economies are meaningful”. In Fernández-Vázquez, E. and Rubiera-Morollón, F. (Eds.) *Defining the spatial scale in modern regional analysis: new challenges from data at local level* (pp. 23-44). Berlín, Germany: Springer.
- Santillana, I.
- 1981 “Los determinantes económicos de las migraciones internas en España, 1960-1973”, *Cuadernos de Economía*, 25: 381-407.
- Schündeln, M.
- 2014 “Are immigrants more mobile than natives? Evidence from Germany”, *Journal of Regional Science*, 54(1): 70-95.
- Todaro, M.P.
- 1969 “A model of labor migration and urban unemployment in less developed countries”, *The American Economic Review*, 59: 138-148.
- Viñuela, A., Rubiera-Morollón, F., and Fernández-Vázquez, E.
- 2014 “Applying economic-based analytical regions: a study of the spatial distribution of employment in Spain”, *The Annals of Regional Science*, 52(1): 87-102.

TABLES AND FIGURES

TABLE 1: Variables definitions and sources

Variable name	Definition	Source
$MIGR_{jk,t}$	Gross migration rate (‰) between pairs of provinces Internal migratory flows (considering only interprovincial flows) divided by the corresponding population, foreign or native, in each case, of the province of origin	Statistic of Residential Variations, Spanish National Statistics Institute (INE)
$WAGE_{kj,t-1}$	Wage (proxied by per capita income) ^a	INE
$UR_{kj,t-1}$	Unemployment rate	INE
$HOUS_{kj,t-1}$	Housing prices, expressed in euros per square meter	Ministry of Development
$AGR_{kj,t-1}$	Share of the agriculture sector over GDP	INE
$CONST_{kj,t-1}$	Share of the construction sector over GDP	INE
$SER_{kj,t-1}$	Share of the service sector over GDP	INE
$DIST_{jk}$	Distance between pairs of capitals, expressed in thousand road kilometers	Repsol guide: http://aim-andalucia.com/distancias.html
$CLIM_{kj}$	Climate Conditions Index ^b 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 ^c	INE
CFI_{kj}	Cultural Facilities Index ^d 0.3*libraries per 1000 inhabitants + 0.4*museums per 1000 old age inhabitants + 0.3*cinemas per 1000 inhabitants	http://directoriobibliotecas.mcu.es http://directoriomuseos.mcu.es http://www.filmaffinity.com/es/theaters.php
SSI_{kj}	Social Services Index ^d 0.5*number of old age residences per 1000 old age inhabitants + 0.5*number of old age open-day residences per 1000 old age inhabitants	http://envejecimiento.csic.es/recursos/residencias/por_provincia.html http://envejecimiento.csic.es/recursos/centrosdia/index.htm

Notes: The explanatory variables, except for distance, are defined as the value of the province of destination k divided by the value of the province of origin j . ^a Following Redding and Venables (2004), Brakman et al. (2009) and Bruna et al. (2016) we take per capita income as a proxy for wages. ^b We apply Boyer and Savageau's (1985) methodology to define this variable. ^c The mean of the monthly data for the longest period available has been considered whenever there is no data for the whole period. ^d We follow the work by Royuela et al. (2003), in which the authors compute a composite quality of life index for the area of Barcelona, to create these variables.

TABLE 2: Tests for threshold effects and threshold estimates: $ExpWAGE_{kj,t-1}$

	Equation 8		Equation 9	
	Natives	Foreigners	Natives	Foreigners
Test for single threshold				
F_1	1616.731	1631.049	2230.530	2205.517
P -value	0.000	0.000	0.000	0.000
Test for double threshold				
F_2	410.657	572.886	453.312	658.816
P -value	1.000	1.000	1.000	1.000
Threshold estimates				
$\hat{\gamma}$	1.473	1.594	1.474	1.474
95% confidence interval	[1.472,1.474]	[1.593,1.616]	[1.472,1.474]	[1.472,1.474]
Residual sum of squares	10579.442	223319.306	11550.180	245615.052

Notes: p-values are computed from 50 simulations. F_1 denotes the Fisher type statistic associated to the test of the null of no threshold. F_2 corresponds to the test of one threshold against two thresholds. The confidence interval for the threshold parameter corresponds to the no rejection region of confidence level 95% associated to the likelihood ratio statistic for test on the values of the threshold parameter (Hansen, 1999). *Source:* own elaboration.

TABLE 3: Results of Equations (10) to (13)

	Equation 10		Equation 11		Equation 12		Equation 13	
	Natives	Foreigners	Natives	Foreigners	Natives	Foreigners	Natives	Foreigners
$WAGE_{kj,t-1} * (d \leq \hat{\nu})$	0.227*** (0.012)	2.215*** (0.105)	0.029*** (0.033)	2.290*** (0.071)	0.437*** (0.016)	2.410*** (0.110)	0.380*** (0.015)	1.591*** (0.087)
$WAGE_{kj,t-1} * (d > \hat{\nu})$	0.293*** (0.011)	2.264*** (0.095)	0.033*** (0.003)	2.355*** (0.070)	0.475*** (0.014)	2.411*** (0.100)	0.435*** (0.013)	1.604*** (0.096)
$UR_{kj,t-1} * (d \leq \hat{\nu})$	0.021*** (0.001)	-0.102*** (0.015)	-0.001** (0.0005)	-0.123*** (0.010)	0.048*** (0.002)	-0.194*** (0.018)	0.022*** (0.002)	0.059*** (0.013)
$UR_{kj,t-1} * (d > \hat{\nu})$	-0.018* (0.011)	-0.225*** (0.079)	-0.005*** (0.002)	-0.348*** (0.069)	0.041*** (0.012)	-0.222*** (0.076)	-0.006* (0.004)	0.167** (0.066)
$HOUS_{kj,t-1}$	-0.029*** (0.005)	0.096** (0.047)	-0.012*** (0.001)	0.486*** (0.038)	-0.039*** (0.006)	0.144*** (0.049)	-0.044*** (0.006)	0.008 (0.040)
$AGR_{kj,t-1}$	-0.001* (0.000)	-0.001* (0.000)	-0.001*** (0.000)	0.004*** (0.001)	-0.0002*** (0.000)	-0.002*** (0.0004)	-0.0002*** (0.000)	-0.001* (0.000)
$CONST_{kj,t-1}$	-0.019*** (0.004)	0.238*** (0.036)	-0.009*** (0.001)	0.198*** (0.023)	-0.031*** (0.005)	0.238*** (0.041)	-0.015*** (0.005)	0.041 (0.038)
$SER_{kj,t-1}$	0.139*** (0.018)	1.561*** (0.158)	-0.043*** (0.005)	1.933*** (0.105)	0.251*** (0.025)	0.950*** (0.179)	0.267*** (0.024)	1.684*** (0.159)
$DIST_{jk}$	-0.644*** (0.004)	-2.137*** (0.027)	-	-	-0.651*** (0.004)	-1.971*** (0.027)	-0.652*** (0.004)	-1.984*** (0.025)
$CLIM_{kj}$	0.185*** (0.057)	-0.849* (0.483)	-	-	0.123** (0.059)	-0.678 (0.455)	0.084 (0.057)	-0.643* (0.344)
CFI_{kj}	0.051*** (0.003)	0.168*** (0.022)	-	-	0.044*** (0.003)	0.200*** (0.019)	0.051*** (0.002)	0.205*** (0.015)
SSI_{kj}	0.014*** (0.002)	0.203*** (0.023)	-	-	0.021*** (0.002)	0.212*** (0.019)	0.020*** (0.002)	0.180*** (0.016)
Fixed effects	Origin, destination and time		Origin-destination and time		Origin-time and destination		Destination-time and origin	
Observations/Groups	26950/ 2450		26950/ 2450		26950/ 2450		26950/ 2450	
Adjusted R2	0.357	0.398	0.977	0.821	0.357	0.409	0.360	0.441

Notes: Standard error in parenthesis. Significance: *** 1%; ** 5%; * 10%. Methodology: GLS. Source: INE and Ministry of Development.

FIGURE 1: Provincial map of Spain.



Note: Regions are represented by bold lines.