Spanish Internal Migration.

Is there anything new to say?

Alan Mulhern[1] and John Watson[2]

June 2009

[1] School of Economics, Kingston University, Penrhyn Road, Kingston upon Thames KT1 2EE. UK. Email: alan.mulhern@kingston.ac.uk.
[2] Aplicaciones Estadisticas, Madrid. Spain. Email: jgwatson@gmail.com.

Spanish Internal Migration.
Is there anything new to say?

Abstract

Spanish internal migration has long been resistant to traditional economic explanations. However this paper examines the data from 1999-2006 after considerable changes in the Spanish economy. Moreover it examines migration at the disaggregated level of Spanish provinces rather than regions, the usual unit of measurement. Using a spatial error model as well as a spatial autoregression model it finds the differentials in wages and unemployment between provinces to be significant explanatory variables. Housing prices are also important in accounting for the dynamics of internal migration.

Keywords: Internal Migration; Spain; Spatial Error Model.

JEL Classification: R23, O15, J61.

1. Introduction

Internal migration has traditionally been explained by the differences in wages and unemployment that exist between regions or sectors. These variables can be viewed as the core economic explanatory factors of such migration and constitute the Harris-Todaro (H-T) model. Additional variables that have been added are expected to have some influence on migration – e.g. the level of education of the migrant, differentials in housing prices and infrastructure, network effects and a range of other variables. These might be thought of as non-core variables and constitute, together with the core variables, an extended H-T model. Such extended H-T models have been recently successfully applied to regional migration in countries such as Russia (Andrienko and Guriev 2004) and Poland (Ghatak et al. 2008), where in both cases the core and a range of non-core variables were found to be significant.

However none of the considerable research on Spanish internal migration finds clear significance in even core variables. Major problems with these variables haunt every paper- as is shown in the Literature Review below. Jimeno and Bentolila (1997 p 33), for example, conclude that 'inter-regional migration flows and regional labour participation decisions are scarcely responsive to regional real wages and employment and that the response of Spanish migration to demand shocks has been low compared to that of the US and the EU'. Lindley et al. (2002 p56) even claim that 'Spain is effectively an economy with almost no migration' and that there is moreover 'evidence, consistent with previous work, that migration is actually negatively related to the regional unemployment rate'.

The reasons for this comparatively low level of internal migration and its poor response to traditional explanatory variables are well known. Firstly for many decades national unemployment in Spain has been at very high levels. Thus no matter what the differentials in employment between regions the chances of finding work in other regions has been low. Very high national unemployment discourages internal migration (Bover and Velilla 2002) and instead promotes migration to other countries.

Secondly labour markets in Spain have been characterised as rigid by many observers (Antolin and Bover 1997; Bande et al. 2007; Bande and Karanassou (forthcoming); Bentolila 1997, 2001; Devillanova and Garcia Fontes 1998; Fonseca 2003; Jimeno-Serrano and Bentolila 1997; Maza and Moral-Arce 2006). Wages have often been determined centrally by strong trade unions and marked segmentation of the labour market has existed for many years (Gil Martin 2004). In addition the range of benefits available to the unemployed in Spain has been another factor militating against the incentive to migrate (Bover and Velilla 2002). There exist a host of other factors in Spanish labour markets (Bentolila and Dolado 1990; Gil Martín 2004) that have deterred migration but which have been increasingly addressed by government policy through the 1990s. These include mismatching of demand and supply of jobs as well as a lack of information concerning job availability. Gil Martín specifies in detail the mechanisms of labour market rigidity explaining the low elasticity of employment in relation to economic growth, for example, high hiring and firing costs as well as segmentation of the labour market by which insiders bid up their wages and leave employers to make the jobs of outsiders more precarious and lower paid.

Thirdly, Antolin and Bover (1997 p2) believe, in a view that has had considerable acceptance in the Spanish literature, that internal migration is not explicable in terms of traditional variables and that ‘high regional unemployment does not trigger any more migrations to more prosperous regions’. They suggest greater attention needs to be paid the personal characteristics of the unemployed (educational and other) in order to explain the conundrum.

The above characteristics of the Spanish labour markets and the phenomena of regional immobility have been consistently outlined in the literature over decades. They account for the low absolute levels of Spanish internal migration as well as its lack of response to wage and unemployment differentials between areas.

The purpose of this paper is to explore the question of Spanish internal migration using contemporary data. This is important for several reasons. Firstly, substantial economic change has occurred in the Spanish economy – among them greater flexibility in labour markets and a fall in the rate of unemployment. Such changes should affect internal migration and its response to differentials wage and unemployment gaps.

Secondly, from the point of view of the researcher, data is available on a wider range of variables giving new opportunities – e.g. data at provincial level which is far more detailed than that previously available at regional level. Quite simply research across the 17 *regions* of Spain does not describe the detailed dynamics of internal migration. However data at the *provincial* level (52 provinces) is picking up far greater internal migration since the majority of internal migration is within regions rather than across them. Also an extended specification of the traditional H-T model is now possible since data is more available in such areas as housing prices, a crucial variable for understanding internal migration in Spain as well as other countries. In addition the researcher has available more sophisticated models that can test the data more rigorously. For example Maza and Villaverde (2004) use nonparametric and semiparametric estimation techniques on relatively recent regional data, while this current paper takes account of spatial heterogeneity by using a spatial error and autoregressive model.

The subject of our investigation is therefore migration across the 52 provinces of Spain in the years 1999-2006 - see Figure 1 for a map of the provinces, the spatial units of our investigation. We are using only the data for internal migration of Spanish residents not those of recent immigrants, who are naturally moving around the provinces and regions to a far greater extent. The internal migration of recent immigrants therefore is not the subject of our study since data is only very recently available since 2003 for this group. Moreover Hierro (2007) has indicated that their impact upon the migration tendencies of Spanish residents has been negligible. We have therefore not chosen to study the totality of Spanish internal migration but only that portion of it done by Spanish residents. We wish to establish that this is indeed responsive to the traditional core variables of the H-T model, i.e. wage and unemployment differences, in this case between Spanish provinces. A descriptive account of Spanish migration between regions is given in the appendix. The motivations of the paper are then: the needs of economic and social research in times of rapid change; the opportunities arising from new data at far more disaggregated levels; data availability allowing the use of an extended H-T specification (i.e. new variables); more sophisticated modelling techniques; and finally intellectual curiosity

– given these changes, data availability and better modelling, is there then anything new to say on Spanish internal migration?

Section 2 provides the literature review. Section 3 introduces the theoretical model. Section 4 details the empirical specification and results. Section 5 reflects upon the results while section 6 briefly concludes.

2. Literature Review

There is of course a large theoretical and empirical literature on international migration, for example Borjas (1999), Greenwood (1997), Ghatak et al. (1996), Hatton and Williamson (1998), Lucas (1997). Comparable variables driving such migration are supposed to underlie the dynamics of internal migration within a country. However Spanish internal migration specifically has been characterized as low by international standards and as unresponsive to traditional explanatory variables such as wages and unemployment (Antolin and Bover 1997, Bentolila 1997 and 2001, Bentolila and Jimeno Serano 1998, Fonseca 2003, Juarez 2000, Lindley et al. 2002).

Econometric investigation into the explanatory factors causing internal migration in Spain have found variables that have been wrongly signed, unexpectedly insignificant or with low elasticity. For Antolin and Bover (1997) the signs on regional unemployment differentials and wages variables are perverse. Bentolila (1997) points out that migration among Spanish regions fell significantly since the 1970s in spite of large and widening regional unemployment rate differentials. Bentolila and Dolado (1990) find that a region's relative wage and relative unemployment differentials do cause some, albeit small, net migration. Relative employment growth, however, was not a significant variable. Lindley et al. (2002) also find the regional unemployment variable to be perverse while Bover and Velilla (2002) claim that high regional unemployment does not trigger migrations to more prosperous regions. In their general historical survey the latter point out...

> since the mid-1980s we are witnessing in Spain what may seem a migration puzzle: despite persistent unemployment differentials, high unemployment regions are not any more net out-migration regions while rich and low unemployment ones are no longer net immigration regions.

An interesting recent paper, Maza and Villaverde (2004), examines the period 1995-2000 using new techniques and an extended H-T specification. Regional migration is the dependent variable while explanatory variables include differentials in per capita income (as a proxy for wages), unemployment, human capital and housing prices. Many of these variables prove of importance in explaining some part of inter-regional migration but most importantly Maza and Villaverde note its marked inertia as well as finding that unemployment rates are relatively unimportant determinants.

They conclude

> This appears to indicate that along with the traditional economic factors there are other determining factors of migration that are non-economic in nature and whose influence is difficult to quantify.

Antolin and Bover (1997), finding traditional economic explanatory variables such as unemployment and wages problematic (and sometimes wrongly signed), tested for personal characteristics of migrants as the key to the puzzle. The regional unemployment differential was found to have a different effect on individuals of differing circumstances. Higher education levels were found to promote migration responses, while having children or living with relatives reduced it. Very importantly they found that while the non-registered unemployed do respond, the employed scarcely respond (unless highly educated) while the least responsive are the registered unemployed. They conclude that the Spanish labour market (registered unemployed) is not responding to personal or regional unemployment. Following this work Bover and Velilla (1999) agree that the registered unemployed, living in regions with high unemployment, rarely migrate, probably reflecting the level and availability of unemployment benefits. The more educated, however, are more likely to migrate. Moreover, in their study, wages did not prove at all to be a significant explanatory variable for inter-regional migration. Other work (Ahn et al.1999) indicates the same problem with unemployment benefits.

Juarez (2000) is an example of how more flexible measurement of variables, for example gross instead of net migration flows between regions or the rate of change in relative wages instead of the differentials in regional wages, has been required to produce more encouraging results. Given the manifest difficulties with the *inter-regional* data, some research has focused on *intra-regional* migration. Bover and Arellano (2002) found that a series of economic determinants were indeed significant in explaining such migration within a region. These included unemployment, housing prices, the education level of the migrant, and employment in the service industry (reflecting Spain's tourist industry). However wages were not found to be a significant explanatory variable. Devillanova and Garcia Fontes (1998) testing for migration between provinces (a sub-regional category) found that pre-1986 migration was unresponsive to economic incentives (e.g. unemployment rate and employment growth differentials) but from 1986-1992 there was some response. However the sign on wages was counter-intuitive. Apart from the generally agreed evidence that more highly qualified workers tended to migrate more than the less qualified in response to economic incentives (Bover and Arellano 2002; Garcia. et alia 1999; Mauro and Spolimbergo 1999) the literature has revealed a problematic research field.

3. General Theoretical Model

Our general theoretical model consists of two provinces: a home and a target province. The former will be denoted with the superscript h and the latter with the superscript t. Each province produces a homogeneous output, Q, by means of capital and labour. The production function takes the form:

$$Q^j = F^j(a,k,l) \quad with\ j = h,t \tag{2.1}$$

where a is total factor productivity, k is capital and l is labour. Production functions are assumed to be increasing, concave and homogeneous of degree one. They have positive marginal products of capital and labour and in addition:

$$F^j{}_{la} > 0,\ F^j{}_{lk} > 0,\ F^j{}_{ll} < 0 \tag{2.2}$$

For the purposes of simplicity in the model we assume that the workforce in each province, N^h in home province and N^t in target province, is fixed. However, population can migrate to the other province (M), thus at any point in time s the available workforce is N^h - M(s) in home province and N^t + M(s) in target province.

The marginal product of labour, i.e. F^j_l, sets real wages (hence referred to as wages), which for full employment in target province are $F^t_l(a^t,k^t, N^t + M)$, and in home province $F^h_l(a^h,K^h, N^h - M)$. However, due to inflexibility of wages, actual wages may be higher, thus explaining unemployment. Real wages thus consist of two components: full employment wages and an excess wage W^j_E. Therefore,

$$F^h{}_l(a^h, k^h, L^{h*}) = F^h{}_l(a^h, k^h, N^h - M) + W^h{}_E$$
$$F^t{}_l(a^t, k^t, L^{t*}) = F^t{}_l(a^t, k^t, N^t + M) + W^t{}_E \tag{2.3}$$

where L^{h*} and L^{t*} are the actual levels of employment in each province, which is endogenously established within the model. Thus if $W^t_E > 0$, then $L^{t*} < N^t + M$ and if $W^h_E > 0$, then $L^{h*} < N^h - M$. Unemployment in each province depends on the level of wage inflexibility, the more inflexible wages are, i.e., the higher the excess wage, the higher the level of unemployment. Moreover, if wages are fully flexible, i.e. $W^j_E = 0$, then full employment is achieved.

Let R^j be a vector of k amenities in province j. Formally,

$$R^j = \{r^j{}_1, r^j{}_2, ... r^j{}_k\} \quad with\ j = h,t. \tag{2.4}$$

Each provincial amenity represents any characteristic having a positive marginal utility in individuals' utility function, such as public services, weather, the negative of price of housing, etc.

Following the H-T model of rural urban migration, wage and unemployment gaps are the main determinants of migration. In addition, provincial amenities are included as determinants of migration. The derivative of migration with respect to time, $\dot{M}$ is given by

$$\dot{M} = \alpha\left[E^t W^t + \sum_k r^t{}_i - E^h W^h - \sum_k r^h{}_i - C\right] \tag{2.5}$$

where E^j is the probability of finding employment in province j, W^j is actual wages in province j, C is the direct cost of migration and α is an adjustment constant.

The probability of finding employment is:

$$E^h = \frac{L^{h*}}{N^h - M}, \quad E^{t*} = \frac{L^{t*}}{N^t + M} \tag{2.6}$$

Therefore, we can write the steady state of (2.5) as

$$E^t \cdot W^t - E^h \cdot W^h = C + (\sum_k r^h{}_i - \sum_k r^t{}_i)$$
$$with \quad E^j = E^j(a^j, k^j, w^j{}_E, M) \quad and \quad W^j = W^j(a^j, k^j, w^j{}_E, M) \qquad (2.7)$$

Both employment and wages are functions of total factor productivity, capital and excess wage in the correspondent province and the stock of migrants M.

We can obtain the derivative of M with respect to any other variable in the same equation by implicitly deriving (2.7). For example, the derivative of M with respect to the costs of migration is:

$$\frac{dM}{dC} = \frac{1}{W^t \frac{\partial E^t}{\partial M} + E^t \frac{\partial W^t}{\partial M} - W^h \frac{\partial E^h}{\partial M} - E^h \frac{\partial E^h}{\partial M}} \qquad (2.8)$$

From (2.2), (2.3) and (2.6) it follows that both the derivatives of employment and wages in target province with respect to migration are negative, while both derivatives of employment and wages in home province are positive. Thus, the numerator in (2.8) is negative.

Similarly, with respect to provincial amenities we have:

$$\frac{dM}{dr^h{}_i} = \frac{1}{W^t \frac{\partial E^t}{\partial M} + E^t \frac{\partial W^t}{\partial M} - W^h \frac{\partial E^h}{\partial M} - E^h \frac{\partial E^h}{\partial M}} \quad \forall i = 1,2,..k \qquad (2.9)$$

$$\frac{dM}{dr^t{}_i} = \frac{-1}{W^t \frac{\partial E^t}{\partial M} + E^t \frac{\partial W^t}{\partial M} - W^h \frac{\partial E^h}{\partial M} - E^h \frac{\partial E^h}{\partial M}} \quad \forall i = 1,2,..k \qquad (2.10)$$

Since the result of (2.9) is less than zero and in (2.10) greater than zero, it follows that any marginal increase in any amenity in the target province will increase the stock of migrants M, while any increase in any amenity in the home province will deter migrants.

However, the relationship of migration with respect to excess wages is not as clear as in the previous examples, as shown in the following equations:

$$\frac{\partial M}{\partial E^t} = \frac{-W^t \dfrac{\partial E^t}{\partial W^t{}_E} - E^t \dfrac{\partial W^t}{\partial W^t{}_E}}{W^t \dfrac{\partial E^t}{\partial M} + E^t \dfrac{\partial W^t}{\partial M} - W^h \dfrac{\partial E^h}{\partial M} - E^h \dfrac{\partial E^h}{\partial M}} \tag{2.11}$$

$$\frac{\partial M}{\partial E^t} = \frac{W^h \dfrac{\partial E^h}{\partial W^h{}_E} + E^h \dfrac{\partial W^h}{\partial W^h{}_E}}{W^t \dfrac{\partial E^t}{\partial M} + E^t \dfrac{\partial W^t}{\partial M} - W^h \dfrac{\partial E^h}{\partial M} - E^h \dfrac{\partial E^h}{\partial M}} \tag{2.12}$$

Because $\frac{\partial E^j}{\partial W^j{}_E}$<0 and $\frac{\partial W^j}{\partial W^j{}_E}$>0 for j=h,t, the sign of the derivatives in (2.11) and (2.12) is not clear. If the effect of employment is greater than the wage effect, then an increase of excess wages in target (home) province will increase (reduce) migration. The opposite, however, will occur if the effect of wages is greater than the effect of employment. Therefore any increase in wages due to capital or total factor productivity increase in any province will have a straightforward effect of attracting migrants to that province. However when the effect of an increase in wages pushes them above the marginal product of labour the effect is not so clear. It will attract migrants if the wage effect is greater than the unemployment effect, and will deter migration otherwise.

4. Econometric Estimation

Traditional gravity models of migration flows do not take into account the spatial relationship between regions. OLS models do not take into account spatial dependence in the data. LeSage and Pace (2005) developed a methodology for expanding traditional models in order to include either spatial autocorrelation or spatial models. This allows a better analysis of migration flows across a country. We follow their methodology, but instead of using one year, we expand it to include a time series of migration flows.

Let M be a nxn matrix of migration flows for n, where the element Mht denotes migration from province h to province t. Migration will be measured as the logarithm of migrants from one province to another divided by origin destination. We can obtain two n^2 vectors of dependent variables with one which is origin-centric and another which is destination-centric. The former, Mh can be constructed by Mh=vec(M) and the later, Mt, by Mt=vec(M'). This is, in the first case, the n first elements of vector Mh and contains migration flows from province 1 to all other provinces, while the last n elements denote flows from province n to all other provinces. Equivalently, the first n elements of Mt contain migration flows from all provinces to province 1 and the last n contains flows from all provinces to province 1. For the rest of this section we use the origin-centric vector M=Mh as the dependent variable to be explained.

Let W be an nxn spatial weight matrix which takes into account spatial contiguity. That is, if province h is a neighbor of province t, then the element Wht is greater than zero, otherwise it is zero. This matrix contains zeros in its diagonals as we avoid a province from being a neighbour of itself. We use 1 for the element to denote contiguity, but W is row standardized.

Similarly to matrix M, we can obtain two different matrixes for the weight matrix, an origin based matrix Wh which we define as $Wh = In \otimes W$; and a destination based matrix $Wt = W \otimes In$.

Following the previous definition we can present the model which we aim to investigate:

$$m = \alpha + \beta x_t + \beta x_h + D + u$$

$$u = (\rho_1 Wh + \rho_2 Wt)\varepsilon + e \qquad (4.1)$$

Equation (4.1) is a spatial error (SE) specification which takes into account geographical interaction of errors through the error term. Xt denotes explanatory variables for destination provinces and Xh explanatory variables for origin provinces. D denotes distance. The first element of the error term includes this spatial correlation through two terms. The Wh lag is the origin spatial relation. It shows correlation between variables sharing the neighboring origin province. This seems intuitive as a shock in any origin province will affect neighboring provinces as well. The Wt lag shows a similar relationship but for neighboring destination provinces. Finally, the second element e is an $N \sim (0, \sigma)$ error term.

Three problems arise from estimation of equation (4.1). Firstly, it is common in migration flow matrices that elements in diagonals, i.e. intra-provincial migration, show large numbers in comparison to those off the diagonal. This implies long tails which make difficult estimation using traditional techniques. Secondly it is necessary to compute log-determinants for this particular case as we need to deal with two different weight matrices. Thirdly, computation limitations may be problematic as we are dealing with very large matrices. In order to solve these problems we follow LeSage and Pace (2005). For the first problem we will use a Bayesian approach, while for the other two we will make use of properties from matrices Wt and Wh which will simplify the calculation log-determinants.

LeSage (1997) introduces a Bayesian approach for estimation of spatial regression models which is based on Markov Chain Monte Carlo methods (MCMC) from Gelfand and Smith (1990). This approach is particularly interesting for this estimation as it can handle dependent variable vectors showing fat-tails. It relaxes specification (4.1) allowing for non-constant variance of error terms. Thus error term e in (4.1) is

$$e \sim N(0, \sigma diag(V))$$

$$V = v1, v2 \dots vn^2 \qquad (4.2)$$

Traditional techniques are inappropriate for this estimation as there are not enough degrees of freedom to estimate n^2 variables. Therefore, LeSage and Pace (2005) rely on an approach from Geweke (1994) which uses a χ^2(r)/r prior on vi terms. Thus we can estimate the additional vi parameters through a single parameter r.

The conditional distributions for the explanatory variables β take the form of a normal multivariate distribution, while the conditional distribution for σ is $\chi^2(n^2)$ (LeSage 2004). The conditional distributions for the ρ1 and ρ2 parameters are

$$P(\rho 1|\beta, \sigma, V, \rho 2) = \log(In^2 - \rho 1Wh - \rho 2Wt)\frac{n^2-2k}{2}\frac{\log(e'V^{-1}e)}{n^2-k} \qquad (4.3)$$

LeSage and Pace (2005) propose, following a Metropolis Hasting algorithm, to sample this distribution. The details on the algorithm in order to obtain posterior distribution estimates for all parameters are described in LeSage (1999).

We rely on the Matlab routines for Bayesian estimation developed by LeSage (2004) in order to pursue this estimation. We use a restricted version of (4.1) in which we use symmetrical origin destination explanatory variables β, i.e. βh=-βt, thus assuming perfect information on provinces by migrants.

For provincial partition we use Spanish provinces. There are 52 provinces in Spain, but wage data is unavailable for some provinces in the Basque Country, Ceuta and Melilla. Therefore our study is constricted to 47 provinces (since the above 3 areas have 5 provinces between them). The period of study is from 1999 to 2006. We have 13246 observations. It is known that migration can also affect explanatory variables such as unemployment, wages and housing prices. Increasing population in a province can increase unemployment and housing prices as well as decrease wages in the short-run. Therefore, to avoid a causality problem, we use lagged explanatory variables. The following Table 1 shows the variables used while variable definition and sources are given in Table 2.

Insert Table 1

Insert Table 2

Results of the regressions are shown in the following table:

Insert Table 3

Housing prices, wages and unemployment show the expected signs and are significant. Distance has also the expected sign and is significant. Infrastructure has the expected sign, but however is not significant. The R-Squared, 0.62 has an acceptable value. The coefficients show strong elasticity of inter-provincial migration in response to wage and housing price differentials. Differentials in housing prices, along with wage and unemployment gaps, are key to understanding internal migration. Unemployment differentials between provinces are significant although the elasticity is low. The results show that key variables can explain internal migration flows in Spain.

This is a heteroskedastic model and every region is allowed a different distribution of errors (common models are homoskedastic). This takes the standard deviation of errors to be a random variable which is distributed as a chi squared with R as a parameter. The lower section of Table 3 gives various pieces of information. The *mean of sige* and *sige epe/(n-k)* indicate the methodology i.e. the use of a Bayesian model used by LeSage. The *r value* takes the chi distribution which has a mean of 13.90 which is good because it means we do indeed have a heteroskedastic model - if we didn't use such a model the results would be biased. Likewise, Pace et alia (2000) clarify our methodological source. The *fixed effects* show a good range and a distribution of positive and negative signs. It is not possible to give their significance since we are taking averages in order not to lose degrees of freedom.

The above model is a contiguity matrix that takes 1 where two provinces are next to one another. In order to test the robustness of these results we decided to test the data using a distance matrix (non fixed effects) that takes the inverse of the matrix (the higher the number the closer the regions are). This extra testing is then a Bayesian spatial autoregressive model for the Spanish provinces 1998-2006. Right hand variables are again lagged. Results are given in Table 4.

Insert Table 4.

All variables are significant and right signed. Capital was not included since in the first model (Table 3) it proved insignificant and in this model was not useful. The R squared is acceptable. Rho is the coefficient of the spatial average of the dependent variable and should be near to 1. In the table it is 0.86 and confirms that it is a spatial model with a geographical relationship between regions and migration - closer regions behave in a similar way.

5. Reflections and Observations on Results

The results showed wage and unemployment differences to be significant explanatory variables of internal migration in Spain. Distance proved significant also. Housing prices, as expected, proved to be a significant variable. This makes intuitive sense since wages and even unemployment signals have to be modified by the cost of housing. Infrastructure failed to make significance. The core variables of wage and unemployment differentials then prove significant and the non-core variables such as housing prices and infrastructure prove to be very useful but do not always have the desired significance level (in this case the infrastructure variable). The difficulty of finding good proxies for some of this data (e.g. infrastructure) may partially explain this.

We can conclude however that Spanish internal migration is determined by a combination of economic explanatory variables. As we deduce from the Literature Review these results are rather different from those in the previous research literature. Exactly why this is the case is difficult to answer. However let us mention a few summary points. A combination of factors has probably influenced this matter. Firstly most previous research covered earlier time periods when indeed the Spanish labour force was less motivated to migrate. Changes in labour laws in the 1990s have created more flexibility in labour markets and this has encouraged, we believe, greater internal migration. Secondly we used a Spatial Error Model on a contemporary data series at the most disaggregated level possible for country-wide data – provinces. Previous research[3] has generally used OLS models which have not taken into account spatial heterogeneity. Much of the previous research was also at regional level where it was inherently more difficult to discover the migration dynamics, since most internal migration was within provinces rather than between them. Thirdly our extended specification was important since without the inclusion of the additional variables (especially housing prices, not previously available) then core variables showed disappointing results – much as previous investigation had shown. It is not possible to run our model on earlier time series since the data required for an extended specification is not available in these earlier periods. We have run our model from the

[3] Maza and Villaverde (2004) come closest to our results and are using sophisticated techniques on a recent data series. They have a similar variable specification to ours and are testing at the more aggregated level of regions as opposed to our paper which deals with provinces. However they believe that their economic variables do not truly explain inter-regional migration and that there must be non-economic determinants.

earliest to the latest dates possible depending on data availability. Fourthly it is often mentioned that international migration has increased enormously into Spain in the last 15 years and must be an influence on internal migration. However its impact on the Spanish native population's propensity to migrate - the subject of our investigation - has according to Hierro (2007) been negligible. Reliable data at provincial level right across Spain is only recently available (2003 onwards) for internal migration of those without Spanish nationality or limited residency status. In conclusion objective and subjective conditions have changed: objective in the sense of real changes in the Spanish economy; subjective in the sense of new possibilities for the researcher in terms of techniques and data. It has not proved possible to be more precise that this and indeed, this has not been central to this paper which has chosen a more modest though still important question: does internal migration by Spanish residents in modern times now conform to economic expectations? We have answered in the affirmative.

6. Conclusions

Internal migration in Spain, using inter-provincial data can now be demonstrated to follow traditional economic reasoning of the Harris-Todaro variety with respect to gaps in wages and unemployment. This migration has proved to be also very responsive to housing price differentials as well as distance factors. The authors feel that a combination of changes in objective conditions in the Spanish economy (greater labour market flexibility for example) and new possibilities for the researcher (contemporary provincial data especially supporting an extended specification, as well as new modelling) underlies the emergence of a picture of Spanish internal migration that is at last following economic expectations.

References

Ahn, N., De la Rica, S. and Ugidos, A., (1999) Willingness to move for work and unemployment duration in Spain. *Economica*, 66, pp335-357.

Andrienko,Y. and Guriev, S. (2004) Determinants of Inter-Regional Mobility in Russia. *Economics of Transition*, 12. pp.1-27.

Antolin, P. and Bover,O. (1997) Regional Migration in Spain: The Effect of Personal Characteristics and of Unemployment, Wage and of Housing Price Differentials Using Pooled Cross-Sections. *Oxford Bulletin of Economics and Statistics, 59.2.*

Bande, R., Fernández, M. and Montuenga, V.M. (2007) Regional Disparities in the Unemployment Rate: The Role of the Wage-Setting Mechanism in Spain, 1987-92. *Regional Studies*, 41(2), pp. 235 – 251.

Bande, R. and Karanassou, M. (forthcoming) Labour Market Flexibility and Regional Unemployment Rate Dynamics: Spain 1980–1995. *Papers in Regional Science*.

Bentolila, S. (1997) Sticky Labor in Spanish Regions. *European Economic Review,* 41(3-5) pp.591-598.

Bentolila, S. (2001) Las migraciones interiores en España. *Documentos de Trabajo,* 2001-07. Fundación de Estudios de Economía Aplicada. Madrid.

Bentolila, S. and Dolado. J.J. (1990) Mismatch and Internal Migration in Spain. In F.Padoa-Schioppa, ed, *Mismatch and Labor Mobility*, Cambridge University Press.

Bentolila, S. and Jimeno Serrano, J.F. (1998) Regional Unemployment Persistence (Spain *1976-94*). *Labour Economics*, 5(1) pp.25-51.

Borjas, G.J. (1999) The economic analysis of immigrations. In AC. Ashenfelter and D. Card, (eds), Handbook of Labor Economics, vol. 3, pp.1697-1760. Elsevier, Amsterdam.

Bover, O. and Velilla, P. (1999) Migrations in Spain: Historical Background and Current Trends, *Documento de Trabajo* no.9909, Banco de España.

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

Bover, O. and Arellano, M. (2002) Learning about Migration Decisions from the Migrants: Using Complementary Data Sets to Model Intra-regional Migrations in Spain. *Journal of Population Economics*, 15: pp.357-80.

Devillanova, C. and Garcia Fontes, W (1998) Migration across Spanish Provinces: Evidence from the Social Security Records. 1979-1992. *Investigaciones Económicas*: 28 **(**3): pp.461-487, Fundación SEPI.

Fonseca, R. (2003) On the Interaction between Unemployment and Inter-regional Mobility. CSEF Working Papers 105. Centre for Studies in Economics and Finance (CSEF), University of Salerno, Italy.

García, J., Gómez, J., Muñoz Sánchez, E., Solana Ibañez, J. (1999) Modelos Migratorios: Teoría del Capital Humano, *Jornadas de la Asociación de Economía de la Educación*. 10: pp**.**121-123.

Gelfand, A.E. and Smith A.F.M. (1990) Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association,* 85: pp.398-409

Geweke, J. (1994) Bayesian Comparison of Econometric Models. *Econometrica,* 57, pp.1317-1340.

Ghatak, S., Levine, P. and Wheatley Price, S. (1996) Migration Theories and Evidence: An Assessment. *Journal of Economic Surveys,* 10:1, pp.59–98.

Ghatak, S., Mulhern, A. and Watson, J. (2008) Inter-Regional Migration in Transition Economies: The Case of Poland. *Review of Development Economics,* 12(1), pp,209-222.

Gil Martin, S. (2004) An Overview of Spanish Labour Markets Reforms, 1985-2002. *Unidad de Políticas Comparadas*. Working Paper 17-02. Consejo Superior de Investigaciones Científicas. Madrid.

Greenwood, M.J. (1997) Internal migration in developed countries. In Rosenzweig, M. and Stark, O. (eds.), Handbook of Population and Family Economics 1b. Ch 12 pp. 647-720. Elsevier, Amsterdam.

Harris, J. and Todaro, M. (1970) Migration, Unemployment and Development: A Two-Sector Analysis. *American Economic Review*, 60: pp.126-142.

Hatton, T. and Williamson, J. (1998) *The Age of Mass Migration*. Oxford University Press.

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

Instituto Nacional de Estadística: Madrid. http://www.ine.es. Accessed May 18th 2008.

Jimeno-Serrano, J.F. and Bentolila, S. (1997) Regional Unemployment Persistence (Spain 1976-94). *Labour Economics*, 4, pp.68-75.

Juarez, J.P. (2000) Analysis of Interregional Labor Migration in Spain Using Gross Flows. *Journal of Regional Science*, 40 **(**2), pp.377-399.

LeSage, J.P. (1997) Bayesian estimation of Spatial Autoregressive Models. *International Regional Science Review*, 20, pp.113-129.

LeSage, J.P. (1999) The Theory and Practice of Spatial Econometrics. University of Toledo: http://www.spatial-econometrics.com. Accessed 06.11.2008.

LeSage, J.P. (2004) Spatial Regression Models in Numerical Issues. In *Statistical Computing for the Social Scientist.* Micah Altman, Jeff Gill, Michael P. McDonald . John Wiley and Sons.

LeSage, J.P. and Pace, R.K. (2005) *Spatial Econometric Modeling of Origin-Destination flows*. http://www4.fe.uc.pt/spatial/doc/lecture7.pdf. Accessed 06.11.2008.

Lindley, J. U., Wright, P. (2002) Regional mobility and unemployment transitions in the UK and Spain. *Leverhulme Centre for Research on Globalisation and Economic Policy.* University of Nottingham. Oct 4th.

Lucas, R. (1997) Internal migration in developing countries. In Rosenzweig, M.R. and Stark, O. (eds.), *Handbook of Population and Family Economics 1b*. Elsevier, Amsterdam.

Mauro, P. and Spolimbergo, A. (1999) How do Skilled and the Unskilled Respond to Regional Shocks? *IMF Staff papers*. 46(1).

Maza, A. and Moral-Arce, I. (2006) An Analysis of Wage Flexibility: Evidence from the Spanish Regions. *Annals of Regional Science*, 40(3), pp. 621-637.

Maza, A. and Villaverde, J. (2004) Interregional migration in Spain: a semiparametric analysis. *The Review of Regional Studies*, 34(2), pp.37-52.

Pace, R. K., Barry, R., Gilley, O.W. and Sirmans, C.F. (2000) A Method for Spatial-temporal Forecasting with an Application to Real Estate Prices. *International Journal of Forecasting*, 16(2), pp. 229-246.

Appendix: descriptive features of Spanish internal migration.

The subject of investigation has been inter-provincial migration. However given the fact that there are 52 provinces a descriptive account of this data is beyond the scope of this paper. What we can note is that the volume of this inter-provincial migration within regions has been growing - actually by 36 % from 1998-2006. Fig 2 shows the trend. What we also know is that it has been responding to economic variables such as wage and unemployment gaps as well as housing prices which have significant impact on this migration.

Insert Figure 2

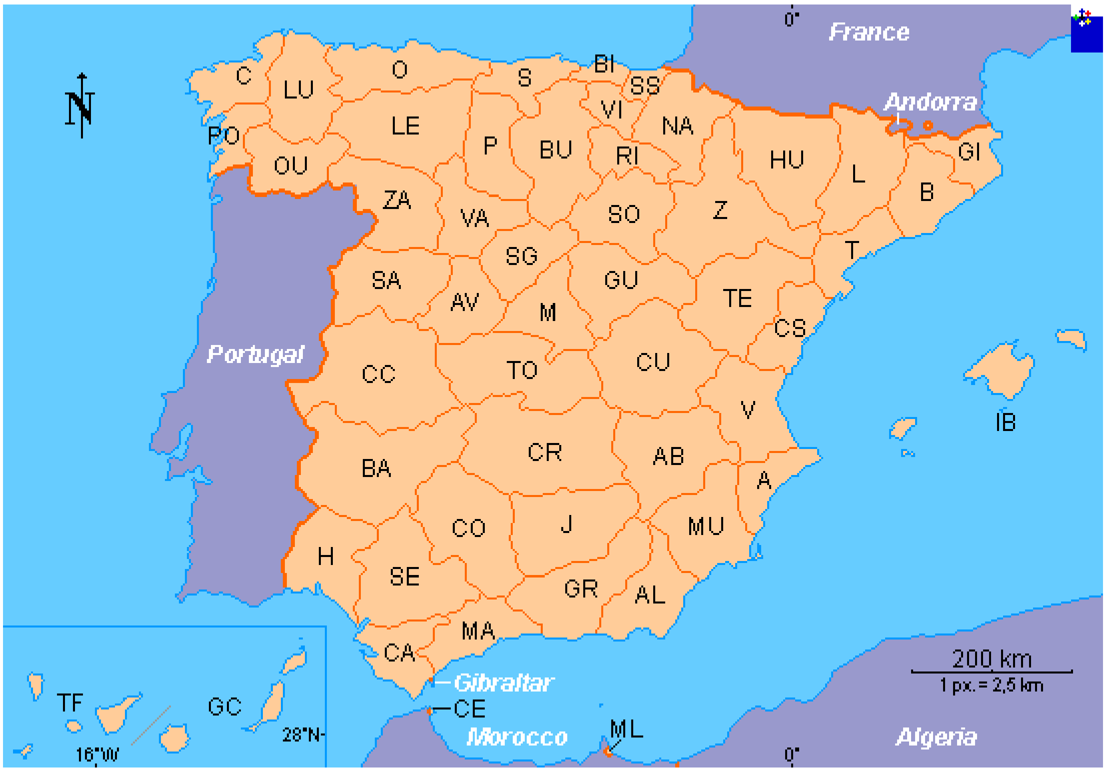
Inter-regional migration is of course in far smaller numbers than intra-regional (e.g. provincial) migration. Yet the outline picture of migration is easier to make out in the regions where one is not swamped with detail. The determinants of migration dynamics between regions are basically the same as within regions with flows responding to gaps in employment and wages as well as housing prices. Table 5 shows the 17 Spanish regions.

Insert Table 5

It shows the balance of migration (in-migration and out-migration) for each region from 1995-2006. Some regions are experiencing increased figures (both positive and negative) of net migration over this period. Two of the most powerful regions, Madrid and Catalunya, have been experiencing increasing levels of net out-migration, surprising at first glance since they are responsible for much of Spain's industry and employment. However the secret of this negative flow lies in the exceptional rise in housing prices (and consequently land prices, office rentals as well as accompanying diseconomies of scale) in these regions. Surrounding regions, with lower house prices but yet with expanding economic opportunities have captured this outflow, for example Castille y Leon and Valencia.

Figures 3 and 4 show select regions that have experienced this increased amplitude but which have also experienced changes from a negative to a positive balance in migration – Andalucía is a case in point reflecting the steady success of its regional economy. Other regions such as Melilla, Ceuta, Rioja and Aragon are further from the centre of economic power and have not experienced strong migration dynamics since they generate just sufficient economic activity to hold their population from increased outflow to more prosperous regions. It seems that the more powerful economic regions such as Madrid and Castille y Leon on the one hand, and Catalunya and Valencia on the other, generate greater migration movement – both inward and outward. Thus the centres of economic growth generate larger migration dynamics.

Figure 1: Map of Spanish provinces



Legend key:

A Alacant / Alicante
AB Albacete
AL Almería
AV Ávila
B Barcelona
BA Badajoz
BI Bizkaia / Vizcaya
BU Burgos
C A Coruña / La Coruña
CA Cádiz
CC Cáceres
CE Ceuta
CO Córdoba
CR Ciudad Real
CS Castelló / Castellón
CU Cuenca
GC Las Palmas de Gran Canaria
GI Girona
GR Granada
GU Guadalajara
H Huelva
HU Huesca
IB Balears / Baleares
J Jaén
L Lleida
LE León
RI La Rioja
LU Lugo
M Madrid
MA Málaga
ML Melilla
MU Murcia
NA Nafarroa / Navarra
O Asturias
OU Ourense
P Palencia
PO Pontevedra
S Cantabria
SA Salamanca
SE Sevilla
SG Segovia
SO Soria
SS Gipuzkoa / Guipúzcoa
T Tarragona
TE Teruel
TF Santa Cruz de Tenerife
TO Toledo
V València / Valencia
VA Valladolid
VI Araba / Álava
Z Zaragoza
ZA Zamora

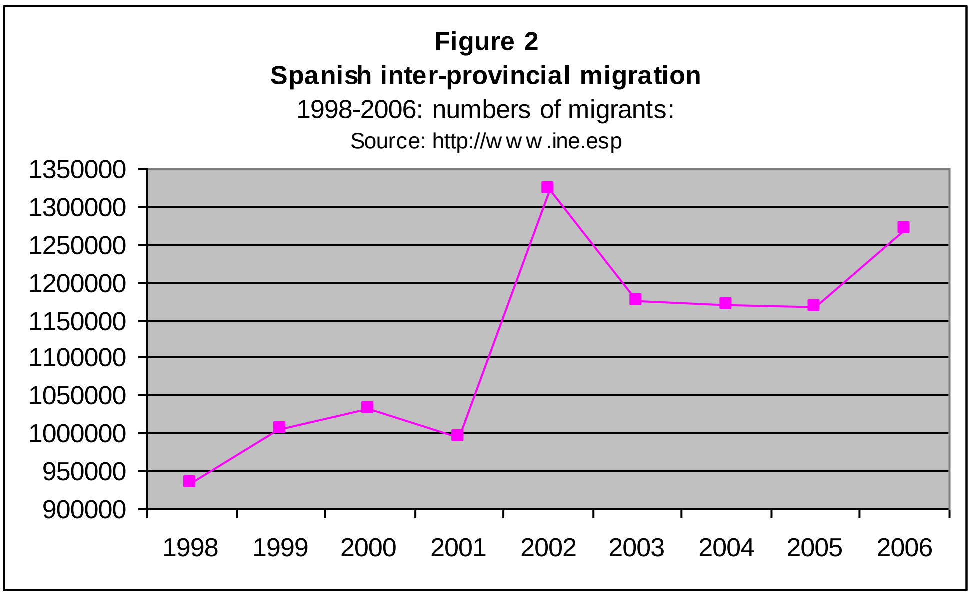


Figure 2
Spanish inter-provincial migration
1998-2006: numbers of migrants:
Source: http://w w w .ine.esp
1350000
1300000
1250000
1200000
1150000
1100000
1050000
1000000
950000
900000
1998
1999
2000
2001
2002
2003
2004
2005
2006

Figure 3
Net Spanish Inter-Regional Migration
1995-2006: select regions. source: http://w w w .ine.esp

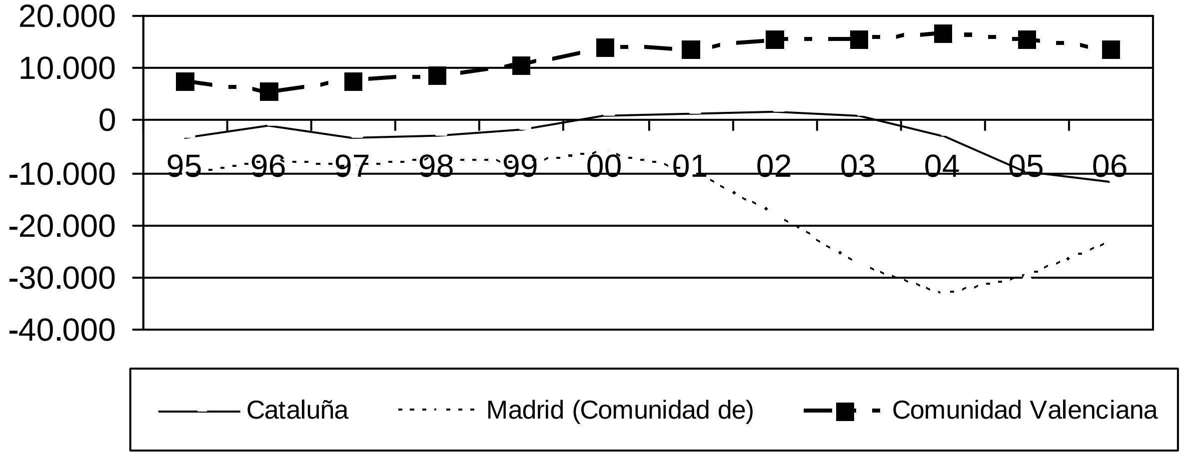


Figure 4
Net Spanish Inter-Regional Migration
1995-2006. Select regions. Source: http://w w w .ine.esp

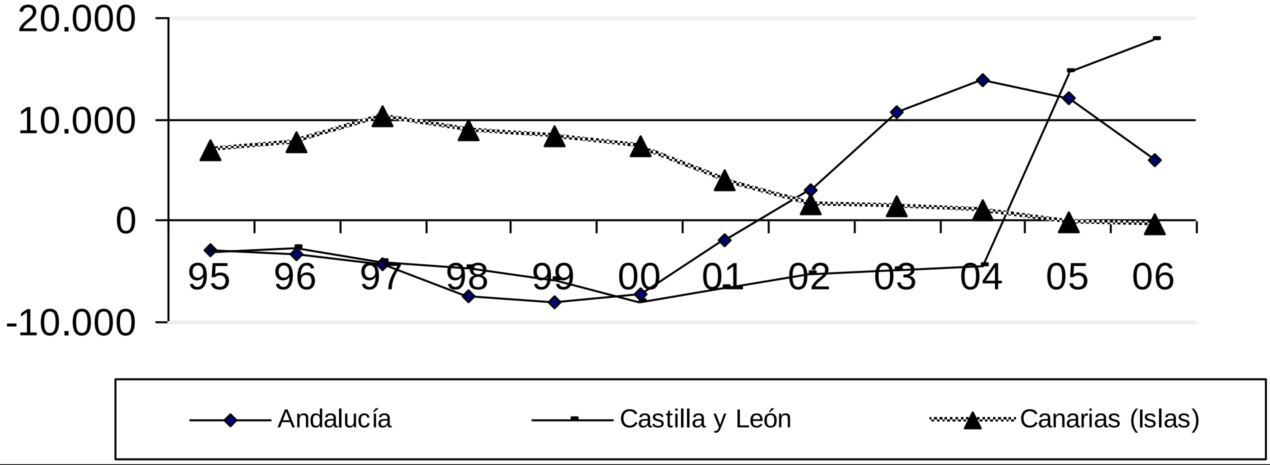


Table 1
Variables

DEPENDENT:	Logarithm of migrants from one province to another divided by origin destination.
LPH:	Logarithm of the difference of housing prices in destination and origin province (lagged 1 year).
LW:	Logarithm of the difference of real wages in destination and origin province (lagged 1 year).
LK:	Logarithm of the difference of per capita capital in infrastructure in destination and origin province (lagged 1 year).
LD	Logarithm of distance.
LU	Logarithm of the difference of unemployment in destination and origin province (lagged 1 year).

Table 2
Variable Definition and Sources

Migration	From home to target province of natives as a proportion of population of home province. Official statistics of those with residency qualifications of at least two years – therefore excluding recent migrants. Source: INE = Instituto Nacional de Estatistica. www.ine.es
Wages	Average wage per month deflated by Consumer Price Index (not including housing prices). Source: Agencia Tributaria. Madrid. www.agenciatributaria.es
Unemployment	Percentage of unemployed over active population. Source: Encuesta de Población Activa. INE. www.ine.es
Price of housing	Average housing prices by mtrs2. Deflated by CPI. Source: Ministerio de Vivienda. www.mviv.es
Distance	KM of road distance between capitals. Canarias and Balears show aerial distance.
Infrastructure	Capital Stock in urban infrastructure, roads, trains and airports divided by area. Source: BDMORES regional database.

Table 3 RESULTS			
Migration in Spanish Provinces 1998-2006			
Spatial error model - fixed effects			
	BETA	T-STATS	
Intercept	2.93	76.68	
LPH	-0.24	-24.37	
LW	0.43	21.29	
LK	0.00	1.15	
LD	-0.21	-74.35	
LU	-0.04	6.01	
OBS	10349		
R-SQUARED	0.62		
F-TEST	0		
mean of sige draws = 0.0062			
sige, epe/(n-k) = 0.0457			
mean r-value = 13.9083			
std r-value = 1.2832			
Pace and Barry, 1999 MC lndet approximation used			
order for MC appr = 50			
iter for MC appr = 30			
FIXED EFFECTS			
02	0.27202435		-0.25963688
03	0.51371637		0.46373207
04	0.57153161		0.10058003
05	-0.04295552		0.21806314
06	0.45839484		0.61446375
07	-0.38917815		0.02985492
08	0.22407078		0.38683417
09	0.27474543		0.11844952
10	-0.02014151		0.07089871
11	0.47733744		0.1860359
12	0.25428777		0.47580404
13	0.04695358		0.24452997
27	0.08343655		-0.03864367
28	0.01111406		0.63784122
29	-0.0978703		-0.27003984
30	0.22554259		0.62242675
32	-0.30176425		0.29595972
33	0.31782615		0.6435223
34	0.31721033		0.24443693
35	-0.05427689		-0.05361513
36	0.01665665		0.15616394
37	0.13444149		0.28348268

Table 4: Results

Migration in Spanish Provinces 1998-2006

Bayesian spatial autoregression model - no fixed effects

Heteroscedastic version with r-value estimate

R-squared	0.6843
Rbar-squared	0.6842
mean of sige draws	0.0044
sige, epe/(n-k)	0.0637
mean r-value	120,891
std r-value	31,411
Nobs, Nvars	14812, 6
ndraws,nomit	10, 1
total time in secs	112,700
time for lndet	93,590
time for sampling	13,810

Pace and Barry, 1999 MC lndet approximation used

order for MC appr	50
iter for MC appr	30
min and max rho	-1.0000, 1.0000

****

Posterior Estimates

Variable	Coefficient	t stat
C	0.60	9.88
LW	0.05	8.20
LU	-0.01	-6.03
LPH	-0.02	-4.38
LD	-0.02	-4.17
rho	0.86	37.73

Table 5

Table 5:

Net Regional Spanish Migration: 1995-2006.

Region	1995	1,996	1997	1998	1999	2000	2001	2002	2003	2004	2005	20
Andalucía	-2,964	-3,317	-4,317	-7,506	-7,995	-7,227	-1,932	3,062	10,670	13,909	12,175	6
Aragón	290	-347	-280	-486	-633	-820	-4	-1,218	-597	461	239	
Asturias	-1,040	-430	-923	-2,305	-2,094	-2,185	-2,127	-2,337	-839	-820	-696	
Baleares	4,749	3,953	5,845	9,979	10,931	11,127	8,376	5,542	-978	-1,335	1,360	3
Canarias	6,882	7,676	10,351	8,945	8,271	7,355	3,952	1,646	1,501	980	-127	
Cantabria	766	325	735	1,275	1,626	1,829	1,566	1,513	2,379	1,658	1,164	1
Castilla -laM	5,417	1,828	2,729	2,210	3,054	1,775	3,420	8,977	12,410	12,232	-4,155	-2
Castille Leó	-3,068	-2,786	-4,172	-4,608	-5,808	-7,971	-6,635	-5,234	-5,017	-4,457	14,618	17
Cataluña	-3,468	-1,136	-3,286	-2,896	-1,737	920	1,458	1,549	1,017	-2,988	-9,776	-11
Valencia	7,440	5,634	7,389	8,364	10,489	13,720	13,434	15,370	15,502	16,600	15,362	13
Extremadura	-447	4	-550	-2,028	-2,242	-3,202	-3,375	-3,009	-1,285	-390	-1,368	
Galicia	-399	-1,676	-2,138	-3,574	-5,138	-5,653	-4,229	-3,652	-3,627	-1,877	93	-1
Madrid	-10,671	-8,084	-9,183	-7,424	-8,420	-6,341	-9,748	-18,108	-27,856	-33,331	-29,994	-23
Murcia	820	1,232	1,390	1,933	1,865	701	-620	833	191	3,558	5,395	2
Navarra	1,300	651	741	1,259	1,391	1,115	458	136	-15	126	323	
País Vasco	-5,746	-4,310	-4,290	-4,591	-5,139	-4,954	-3,866	-3,549	-2,922	-3,590	-3,707	-4
Rioja (La)	181	304	377	573	1,058	1,203	814	248	977	1,003	42	
Ceuta	-395	-204	-202	419	337	-530	-319	-814	-796	-1,024	-577	
Melilla	353	683	-216	461	184	-862	-623	-955	-715	-715	-371	

source: Estadística de Variaciones Residenciales. INE