CENTRUM Católica's Working Paper Series

No. 2014-05-0007 / May 2014

Price dynamics in the Spanish housing market between 1995 and 2008. Evidence from a panel of provinces

Ramiro Gil-Serrate

CENTRUM Católica – Pontificia Universidad Católica del Perú

Price dynamics in the Spanish housing market between 1995 and 2008. Evidence from a panel of provinces

Ramiro Gil-Serrate*

ABSTRACT:

In this paper we follow the specific literature in order to obtain a theoretical framework for the analysis of the dynamics of house prices. From this framework results a long run relationship between the house price variable and its fundamentals. This relationship is estimated using a static and a dynamic panel for the 50 Spanish provinces and the period 1995-2008. Previous to the estimation a detailed panel unit root analysis is done. The results obtained from the estimation are according to the theory and present clear evidence of serial correlation in house prices and of income elasticity of 0.3. However, the results also suggest the existence of additional information that has not been considered in the empirical analysis. This is the existence of a spatial pattern in the data for which we provide clear evidence. Consequently, cross section dependence has to be explicitly taken into account in subsequent analysis of Spanish house prices.

Keywords: Panel cointegration, housing prices, adjustment dynamics, Spanish housing market

JEL classification: C23; R21; R31

1. Introduction

Housing plays a key role in household budget decisions due to its considerable price. Moreover, when buying a house, households are not only demanding housing services but also looking for a low-risk asset to invest in. As a result of these features, the housing market has significant implications in the behaviour of the whole economy, as we have had the opportunity to witness in the last two years. The interest of the economists in the role of housing in the economy is not something new[1], however, in the last years there has been a surge of interest in the behaviour of the housing market internationally, and particularly of house prices, Girouard et al. (2006). This is mainly because the majority of the OECD countries experienced a significant increase in house prices in the last decade (see figure 1). One of the countries that experienced higher rise in house prices during this period is Spain. In consequence, this case is an interesting one to study when the objective is to identify the main determinants of the above-mentioned increase.

At the same time that this renewed interest in house price dynamics has emerged, there has been a considerable development of spatial econometrics when working with panels of economic data, that is to say, econometrics focused on taking explicitly into account possible spatial interactions and interdependencies among units in such economic data. This is particularly interesting when the common observed factors considered in the panel do not capture adequately these possible spatial interactions.

The aim of this paper is to study and understand house price dynamics in Spain applying the common, and widely accepted, economic theory and econometrics tools, but also, when applying these econometric tools, determine the possible existence of a spatial structure in the panel data. In case this spatial structure exists should be explicitly taken into account in favour of accuracy in the results and, in order to do this properly, a spatio-temporal model has to be estimated. The latter task falls beyond the objective of this paper due to space and time constraints, mainly due to the fact that the specific spatial econometric tests are not implemented in the usual econometrics packages[2] so have to be applied *manually*. Consequently, this paper should be considered as the first step of two: 1st. To apply the customary econometric tools to the study of house price dynamics in Spain and identify, in case it exists, a spatial pattern in the data. 2nd. To explicitly take into account in the

[1] Some seminal works date back to the 60s, see for example Muth (1960) and Reid (1962).
[2] Let's say Stata or EViews in their latest versions.

econometric analysis such spatial pattern. The second step would be developed in a subsequent paper.

In order to carry out the first step, we theoretically obtain from the definition of the after-tax cost of homeownership an equilibrium relationship between housing prices and fundamentals, which may be tested empirically. These fundamentals are household income, the interest rate and demographic factors that influence the rate of household formation[3]. Once we have the theoretical foundations, we proceed to do the econometric analysis to validate this relationship between house prices and fundamentals[4]. To do this, we use the most commonly tests for unit roots in panels and we estimate our general model in a static and in a dynamic framework[5]. The main results confirm the theoretical relationship, with income as the main driver of house price behaviour in Spain during the period analyzed. However, the results also suggest the existence of some additional information that has not been taken into account in the analysis, as could be the cross-section dependence. So, next, we analyze the possible existence of a spatial structure in the data in two different ways, first, intuitively through simple correlation coefficients between the different spatial units and, second, through the test of Moran, one of the best-known spatial independence tests in the literature of spatial econometrics. In both cases the results suggest the existence of a spatial structure in the data.

An important caveat has to be done. To try to capture adequately as much information as possible in the behaviour of Spanish housing prices, we consider data at the available lower spatial unit, which is the province level[6]. Unfortunately, the period for which these data exist is not very long, 14 years from 1995 to 2008 and does not cover a full price cycle in the Spanish case (see figure 2), so we have to say that this might bias our results.

[3] In the specific literature there are other determinants considered in addition to these three fundamentals, such are labour market conditions, construction costs and building regulations. However, we are not able to take into account these variables due to data limitation. On the other hand, size of the population and income growth are usual proxies for some determinants such as information costs or expectations, see Capozza et al. (2004).

[4] As Case and Shiller (2003) point out and subsequent literature on the topic reminds, in case there is not such long-run relationship, house prices are assumed to behave erratically and the possibility of bubbles increases. In case the relationship exists, house prices are just above or below their long-term trend.

[5] We cannot test for cointegration since appropriate panel cointegration tests, as for example the Pesaran CIPS's test, see Pesaran (2006b), are not implemented in the usual econometrics packages. We leave this task for the above-mentioned second step.

[6] The Spanish province is an administrative unit which is below the region and above the city. There are 50 Spanish provinces.

Previous literature which has recently studied house price dynamics in a regional or local context includes, among others, are Malpezzi (1999) and Capozza et al. (2004) for the US and Meen (1999) and Cameron et al. (2006) for the UK. Nagahata et al. (2004) and Holly et al. (2006) require special attention because they take into account explicitly spatial aspects as we intend to do. Nagahata et al. (2004) found a long-run equilibrium relationship between house prices in 47 Japanese prefectures and its main determinants, which are income, interest rate and expectations. Holly et al. (2006) found a cointegrating relationship between house prices and income and identify a small role for real interest rates in the case of the US states. With regard to specific literature on the Spanish case, Martínez and Maza (2003) analyzed house price determinants at national level[7] and Larraz and Alfaro (2008) studied the dynamics of regional housing prices trying to identify existing asymmetries between regions[8].

The structure of the paper is as follows. Section 2 presents the basic economic theory underlying house price determination. Section 3 describes the main characteristics of the Spanish housing market during the period analyzed. Section 4 presents the estimation methodology and its main results for the relationship between house prices in Spanish provinces during the period 1995-2008 and their fundamentals. Section 5 Analyses the existence of spatial structure in the data. Finally, Section 6 provides some concluding remarks and guide for future work.

[7] For the period 1978-2002 they found that income and nominal interest rates are clear explanatory factors of Spanish house price behaviour which, by the end of the period, were above their long-term equilibrium level but by an amount not different to other times in the past.

[8] However, the type of analysis and tools used in this paper raise some doubts.

2. Theoretical foundations

In the specific literature most of the recent theoretical work analyzing the behaviour of real house prices takes as a starting point the standard definition of the real housing user cost of capital. This definition results from the customary problem of utility maximization in an inter-temporal model of consumption, with two different goods: a composite consumption good (*C*) and housing services. As Cameron et al. (2006) and Holly et al. (2006) point out, this model is widely considered, see Dougherty and Van Order (1982) or Meen (1990 and 1999)[9], inter alia. It is usually assumed for simplicity that the flow of housing services is directly proportional to the housing stock (*H*), so that the latter can be directly considered as an argument in an additive utility function[10]:

$$\int_0^\infty [\,u(C) + v(H)\,]\, e^{-\lambda t}\, dt, \text{ with } 0 < \lambda < 1 \qquad (1)$$

Where λ is the inter-temporal discount rate.

The household carry out a dynamic optimisation of (1) subject to the budget and technical constraints (2, 3 and 4, respectively):

$$P\,N + S + C = (1 - \tau)\,(Y + i\,A) \qquad (2)$$

$$\overset{*}{H} = N - \delta\,H \qquad (3)$$

$$\overset{*}{A} = S - \pi\,A \qquad (4)$$

Where the price of the composite consumption good is scaled to unity, *P* is the real price of housing; *N* is new purchases of housing; *S* is real savings; τ is the marginal household tax rate; *Y* is real household income; *i* is the nominal interest rate; *A* is real non-housing assets and δ and π represent the physical depreciation rate on the housing stock and the general rate of inflation (hence the depreciation rate on non-housing assets), respectively. * denotes the time derivative of a variable.

[9] Although is not the only one, house prices are also derived as a reduced form from separate housing demand and supply equations, see for example, McAvinchey and Maclennan (1982) or Hendry (1984).
[10] In the interest of clarity, explicit time's notation is omitted from the equations.

From the first order conditions of the maximization problem can be obtained the marginal rate of substitution between housing and the composite consumption good which is given by:

$$\frac{v'(H)}{u'(C)} = P\,[(1-\tau)\,i - \pi + \delta - \overset{*}{P} / P] \tag{5}$$

Where v' (H) and u' (C) are the marginal utilities of housing services and consumption, respectively. This reflects the real housing user cost of capital[11].

Usually (5) is rewritten to include expectations as in (6):

$$\frac{v'(H)}{u'(C)} = P\,[(1-\tau)\,i - \pi^e + \delta - (\overset{*}{P} / P)^e] \tag{6}$$

Where $\pi^e + (\overset{*}{P} / P)^e$ is the expected nominal gain on housing[12].

The equation in (6) can also be derived in a different way as, for example, Holly et al. (2006) shows. Instead of considering the housing market as a market for the services of the housing stock it can be considered as a market for housing as an asset[13]. In this second approach, the real return to housing results from any expected capital gains on housing minus depreciation[14] plus the real rental price of housing services in each time period, *R*. From the arbitrage condition this real return of housing equals the post-tax return on alternative assets, usually measured as the return on an interest bearing financial asset[15], resulting the following equation:

$$\pi + (\overset{*}{P} / P) - \delta + R / P = (1-\tau)\,i \tag{7}$$

Re-arranging (7) we have a very similar equation to (5):

$$R = P\,[(1-\tau)\,i - \pi + \delta - \overset{*}{P} / P] \tag{8}$$

---

[11] Specifically the real housing user cost of capital expressed as a proportion of the price of the house is $[(1-\tau)\,i - \pi + \delta - \overset{*}{P} / P]$.

[12] In an empirical analysis, usually, only the general rate of inflation is considered, arguing that "over long horizons housing prices have appreciated at rates remarkably close to the rate of inflation" (Cappoza et al. (2004, p.14).

[13] In fact the housing market is both a market for housing services and for housing as an asset.

[14] Also can be considered additional costs such as maintenance expenditures and property taxes.

[15] This return is usually taken to be the interest rate on a risk-free asset or alternatively as the mortgage interest rate.

And from (8) can easily be obtained a real house price equation:

$$P = R / [(1-\tau)\, i - \pi + \delta - \overset{*}{P} / P] \qquad (9)$$

When analyzing the relationship in (9), should be point out, as noted by Poterba (1984), among others, that a rise in inflation increase the real price of houses, since, coherently, the housing stock is considered fixed in the short run. Consistently, a move from a scenario of moderate inflation to a scenario of low inflation[16] has to be associated with a fall in real house prices. The same argument can be applied to an improvement and a worsening in the tax treatment of housing[17].

Yet in the literature equation (9) is hardly tested. As Meen (1999) points out, this is due mainly to two reasons. First, because there is considerable evidence against efficiency in housing markets (that is to say that prices reflect all information available)[18]. The evidence is in favour of a fairly slowly process of adjustment towards the new equilibrium generated by any new information in the market[19]. Second, because there is a problem of suitable data on real rental price, *R*. This variable is unobservable and does not correspond to any published data, as for example, imputed rental incomes in the official accounts of a country or region. The general practice is to replace *R* by its determinants. These are derived from inverting the standard demand equation of the market for housing services. As a result, such determinants are: existing stock of housing, income[20] and a set of factors which affect household formation that can be summarized as demographic factors (*D*).

Now, taking logarithms in (9) we have:

$$\ln P = \ln R - \ln [(1-\tau)\, i - \pi + \delta - \overset{*}{P} / P] \qquad (10)$$

However, the second term on the right hand side can, in practice, be considered as a real interest rate (*r*), so logarithms cannot be taken here since this variable can take negative

[16] Which is expected to persist as, for example, under the current global economic downturn.
[17] There are some preliminary results on this topic, for the Spanish housing market, in the author's GY458 paper "Does tax treatment of owning affect housing prices in Spain? An analysis for the period 1996-2007".
[18] See, for example, Cho (1996).
[19] From this it follows that the literature has paid special attention to model and analyze the process of dynamic adjustment. Among the most recent papers, see as an example, Capozza et al. (2004), Jacobsen and Naug (2005), McCarthy and Peach (2004), Cameron et al. (2006) or Holly et al. (2006).
[20] However there are some other theoretical approaches, as for example, deriving jointly the demand for housing services and consumer goods so that the demand for housing is conditioned on consumption rather than income, see Pain and Westaway (1997).

values[21]. The usual solution to this problem has been to consider a semi-logarithmic relationship between real house prices and real interest rates.

All the previous caveats lead us to the following realtionship, which is the basis for our empirical analysis:

$$\ln P = f(\ln Y, \ln H, \ln D, r) \qquad (11)$$

However, before we proceed to carry out the empirical analysis and validation of the latter relationship it is useful to present the main features of the market on which the mentioned relationship is going to be analyzed, paying special attention to the evolution of the variables involved in (11).

3. Main features of the Spanish housing market from 1995 to 2008[22]

Spain is one of the developed countries where the increase in housing prices was more marked from 1995 to 2008 (see figure 1), specifically, is among the three ones which experienced the highest average annual real growth in house prices in the group of the twenty main developed countries. The increase in housing prices was also very marked in Ireland and the United Kingdom, while in other countries, like Japan, Germany and Switzerland was much less intense. During the period of study, Spanish nominal house prices rose by more than 200 percent and real house prices by more than 100 percent[23](see figure 2).

At the same time, the Spanish economy experienced a considerable period of real growth, with an annual average real GDP per capita growth of more than 2.5 percent[24] while the Spanish population increased in more than 6 million people (around 15 percent of the initial population) from which the majority were working immigrants[25]. Also, the entrance of Spain

---

[21] As, for example, happened in Spain, at a national level, during year 2005 (see figure 3).

[22] Some of the ideas and references cited in this section follow section 1 of the author's GY458 paper "Does tax treatment of owning affect housing prices in Spain? An analysis for the period 1996-2007".

[23] 203.2 percent and 105.3 percent, respectively, according to the Spanish Department of Housing data (www.mviv.es), and 174.2 percent and 86.2 percent, respectively, according to Sociedad de Tasación S.A data (web.st-tasacion.es).

[24] According to Instituto Nacional de Estadística, INE (www.ine.es).

[25] According to INE, Between 1998 and 2007, the migratory balance increased twelvefold (and ninefold between 1998 and 2008, in 2008 the immigration flow decreased to the level of 2005 due to the global economic downturn). Entrance policy tightened in the last years of the period but immigration flows remained considerable with arrivals from East Europe taking over from Latin America.

in the Euro area in 1999 resulted in an accentuated fall in nominal and real interest rates, the latter were even negative at a certain point during the period considered (see figure 3).

According to our theoretical model, these economic and demographic shocks are behind the marked increase showed by Spanish housing prices. That is to say, theoretically, a rise in housing prices is basically explained through an increase in disposable household income, a growth of adult population and a reduction in interest rates. What's more, during the period considered some other relevant sociological and institutional factors occurred in the Spanish scenario that boosted the effect of such fundamentals on housing prices. On the one hand, the baby-boomers massively joined the labour force causing an additional rise in the percentage of working adults[26], who are more likely to buy a home, and the traditional family unit came to an end, which resulted in young adults leaving home earlier contributing to the rise in the number of Spanish households and the reduction in their average size[27]. On the other hand, the loosening of credit issuance conditions, which facilitated the access to credit, due to intense competition among credit institutions. Examples of this are the offer of partially deferred loans, which considerably reduce the effort during the first years of repayment, or the extension of legal term for mortgages[28]. In addition, unlike United States or Germany, in Spain most issued mortgages were variable rate, consequently, continually high inflation meant close to zero or even negative real interest rates.

Other factors also played a role, such as: a significant increase in foreign investment in property during the period analyzed[29], an inflow of funds into property from the stock market[30], the concentration of demand mainly on urban areas where land was scarce[31] and the

[26] In 1995, approximately 46 percent of the population was between 25 and 59 years old while in 2008 this percentage had risen to 53 percent of the population (INE). The current average retirement age in Spain is 59 years old (INE).
[27] Between 1995 and 2008 the increase in the number of households, approximately 33 percent, was much higher than that of the population, approximately 15 percent (INE). As a consequence the average size of the Spanish household decreased from 3.2 to 2.8 (INE).
[28] The average mortgage moved from around 15 years in the mid-1990s to 26 years and 9 months in 2008, Colegio de registradores de la propiedad, bienes muebles y mercantiles de España (2009). Furthermore, some credit institutions offered a new type of mortgage loan with a 50 years legal term for less than 35 years old customers.
[29] Foreign property inflows were EUR 2.93bn in 1999, 7.05bn in 2003, 5.35bn in 2007 and 5.44 in 2008 (Banco de España, balance of payments). Information available does not differentiate between commercial and residential property, however, these figures clearly reflect and increasing number of foreigners owning a holiday place in Spain during this period favoured by the expansion of low-cost airlines.
[30] The IBEX benchmark index halved between mid 2000 and autumn 2002. The percentage of real estate on total wealth of Spanish households increased from 78.7 in 2002 to 80 in 2005 (Banco de España, Encuesta financiera de las familias 2002 y 2005).

limited size of the rental market. The Spanish housing market is also striking because of the high rate of ownership (see figure 4)[32]. The latter factor is partially a result of owner's unwillingness to offer vacant propertiesin the market due to low legal support for landlords in comparison with other European countries[33]. However, also is due to the Spanish housing policy which favours ownership through subsidies to the purchase of homes by low-income households, the deductibility of mortgage interest and principal payment from personal income tax and the exemption from capital gains taxation for owners of principal owner-occupied properties[34]. In addition, owners of principal owner-occupied properties are exempt from paying taxes on imputed rents since 1999's reform.

In conclusion, the Spanish housing market during the last 15 years is a very interesting case in order to test empirically the relationship between house prices and their fundamentals according to the theory. This is because not only prices but also these fundamentals have experienced a considerable dynamism during the period analyzed, that is to say variability, what increases the amount of information in the data and consequently the quality of the estimations.

[31] Hilber and Mayer (2009) show that locations with less undeveloped land have more inelastic supply. Nevertheless, the rise in Spanish house price was mainly due to demand-side influences, that is to say the insufficient supply was not a main factor, unlike the case in United Kingdom (European Central Bank, 2003).

[32] According to the data this rate was 85 in 2002, the higher among all European countries, the closest cases to Spain were Greece, 80 and Ireland 78, while the lowest rates were in Switzerland, 30 Germany, 39 and The Netherlands and Sweden, 53. For a thorough explanation of the variation of homeownership across European countries see Hilber (2007). According to this author, who exploits the European Community Household Panel: 1994-2001, the substantial differences in homeownership rates reflect a number of influencing factors including, among others, the accommodation type, the provision of public housing and the tax treatment. At the same time, an increasing trend for ownership as opposed to renting among the European countries is clearly observed for the period covered by the European Community Household Panel. Particularly, in Spain the reported ownership rate was 78.8 percent in 1994 and 83.4 percent in 1998.

[33] According to estimates of the Spanish Department of Housing the number of vacant properties at the end of 2008 was, approximately, 2.5 million.

[34] The exemption is applicable only if capital gains are reinvested on a principal dwelling within two years.

4. Estimation methodology and main results

According to (11), which for convenience is restated here:

$$\ln P = f(\ln Y, \ln H, \ln D, r) \qquad (11)$$

we construct a panel of annual data[35] for the 50 Spanish provinces[36] and the period 1995-2008. Housing prices (P_{it})[37] are from the Spanish Department of Housing, where provinces are the smaller spatial unit for which data are published[38][39]. Provincial GDP (Y_{it}) and the population figures (D_{it}): total population and working adult population by province[40] are obtained from the national statistical office (INE – Instituto Nacional de Estadística). The stock of houses (H_{it}) is an estimate and is, also, from the Spanish Department of Housing. Finally, the real interest rate (r_{it}) is work out from the mortgage interest rate and the provincial CPI (2001 = 1). The mortgage interest rate is the average rate offered by the Spanish credit institutions for new mortgages according to the Spanish Central Bank (Banco de España). The CPI of each province is obtained from INE. Table 1 provides summary statistics on the data series.

In order to carry out the empirical analysis, prices are deflated by the CPI and (natural) logarithms are taken (p_{it}). Provincial GDP is also deflated by the CPI and provincial population and (natural) logarithms are taken (y_{it}). Working adult population by province is divided by total population and (natural) logarithms are taken (d_{it}). Lastly, the stock of houses is also divided by total population and (natural) logarithms are taken (h_{it}). As a result, the long-run relation stated by the theory, showed in (11), can be re-written in the following log-linear form:

$$p_{it} = \alpha_i + \beta_i' x_{it} + u_{it}; \quad \text{with } u_{it} \sim I.I.D.\ (0, \sigma_u^2) \text{ and } i = 1, 2, ..., N;\ t = 1, 2, ..., T \qquad (12)$$

[35] When only quarterly data are available, we use annual simple averages of the four quarters.
[36] The autonomous cities of Ceuta and Melilla, located on the Moroccan coast, are excluded.
[37] Average price per square metre.
[38] The Spanish Department of Housing also publishes data at regional (17 regions) and national level.
[39] As pointed out above, we are aware of an additional data source, that of private company Sociedad de Tasación S.A. This company publishes biannual data at urban level (main cities in each province) since December 1985. However, after a detailed analysis, we decided to take the data from the Spanish Department of Housing because we considered official data more trustable. One possible extension of this paper would be to check its results with the data base from Sociedad de Tasación S.A.
[40] As indicated in the previous section, the main demographic changes experienced in Spain during the period analyzed are the considerable growth of immigration, which is included in the total population figure, and the relevant increase in the percentage of working adult population.

where α_i is a term which represents the specific unobservable factors of each province, *βi*' and x_{it} are matrices 1 * J and J * 1 which represent, respectively, the coefficients and the *j* explanatory variables. The latter are comprised of the control variables: y_{it}, h_{it}, d_{it} and r_{it}; and, finally, u_{it} is the error term. With regard to the coefficients of the explanatory variables, in accordance with the theory, the income and demographic factors coefficients are expected to be positive and the stock and real interest coefficients are expected to be negative.

4.1. Panel unit root tests and panel cointegration

In the usual process of estimation, the first stage is to test for cointegration in the relationship stated by the theory[41]. To test for cointegration in a panel requires specific methods[42]. Particularly, since we are interested in taking into account the possible spatial structure, it is convenient to test whether the price and fundamentals display significant cross section dependence before checking whether the variables are I (1). If cross section dependence exists, should be applied second generation tests of integration, such as the CIPS one proposed by Pesaran (2006b), which are not implemented in the usual econometric packages. What's more, to properly determine the possible existence of cross section dependence should be computed a test of error cross dependence developed by Pesaran (2006a) that is applicable to short T and large N panels, according to Holly et al. (2006). Again this test is not implemented in econometric packages.

As a suboptimal solution we consider the panel unit root tests developed in Levin Lin and Chu (2002), Im, Pesaran and Shin (1995) and Hadri (2000) known as: Levin, Lin & Chu "t*", Im, Pesaran and Shin "W" and Hadri "Z", respectively, and which are usually implemented in econometric packages[43]. The tests results for our variables, summarized in table 2, show that for p_{it}, y_{it}, and d_{it} the unit root hypothesis cannot be rejected if the trended nature of these variables is taken into account[44]. Moreover, it seems that the tests are not able to reject this hypothesis for the first difference of the variables, in the majority of the cases, suggesting that these variables are I (2). On the contrary, although not clearly, the results point out that h_{it}

[41] In order to validate that such relationship is a long-run equilibrium relationship.

[42] In Breitung and Pesaran (2008) there is a review of the panel counter part of the classical literature on cointegration techniques developed by Engle and Granger (1987), Johansen (1995) and Phillips (1991) among others.

[43] We consider that these are the most adequate test among the different panel unit root tests usually available in econometric packages. For example, Breitung "t" is usually implemented but may lead to biased results when T is small.

[44] Levin Lin and Chu (2002) and Hadri (2000) assume a common unit root process whereas Im, Pesaran and Shin (1995) assume an individual unit root process. In Levin, Lin & Chu "t*"and Im, Pesaran and Shin "W" the null hypothesis is "there is a unit root" whereas in Hadri (2000) the null hypothesis is "there is not a unit root".

and r_{it} are I (1) and I (0), respectively. An explanation of these unexpected results may be the fact that we are not considering the possible existence of cross section dependence[45], since is generally accepted that, under normal economic conditions, prices, income and population are I (1) variables. Consequently, we proceed taking p_{it}, y_{it}, d_{it} and h_{it} as I (1) and r_{it} as I (0) variables.

Next, to properly find a panel cointegration relationship between prices and their fundamentals, following Holly et al. (2006), a two-stage procedure should be taken using the Common Correlated Effects (CCE) estimator[46]. Since this is not possible to do with the usual econometrics packages we have to assume that the theory is correct and consequently a long-run equilibrium will exist along with a cointegration relationship among the variables.

4.2. Static and dynamic panel estimation

Having considered panel cointegration among the variables we now proceed to estimate the panel model. First we estimate a static panel using OLS (see table 3) with and without the real interest rate (r_{it}). When the real interest rate is considered (equation 1 in table 3), the rest of the explanatory variables present the expected sign and are strongly significant, however the real interest rate is not significant (and does not present the expected sign). If we do not consider the real interest rate in the estimated equation, the results hardly change (equation 2 in table 3). Nevertheless, two relevant econometric caveats should be done regarding these estimations. First, we were not able to really proof the existence of a long-run relationship between the explanatory variables and the dependant variable and, in addition, the estimations are based in data in levels. Both questions considerably increase the probability that these estimations are just a result of spurious correlation.

In order to tackle this possible weakness, we estimate a dynamic panel using GMM according to Arellano (2003) (see table 4)[47]. This estimator is based on the approach of the general instrumental variables estimator (IV), considers data in first differences and improves consistency with N for a small fixed T. Moreover, is robust with regard to the model's initial conditions and presents consistency and normality even in very weak conditions, Wooldridge

[45] "the Im, Pesaran and Shin test procedure is not valid when the errors, u_{it}, are dependent across i, and its use in the case of the house price data can lead to spurious inference." Holly et al (2006, p. 8).
[46] See Holly et al. (2006) for a detailed explanation.
[47] We have considered different lags, however, only the first one is significant. Results for additional lags are not reported.

(2002)[48]. In addition, in order to control for cross section dependence we have add time dummies to the specified equation (see table 5)[49]. The new results, considerably more reliable than the previous ones, present coefficients with the expected sign according to the theory and which are, all of them, highly significant when time effects are not considered (table 4). However, when time dummies are included, the only two coefficients that are robust to the different specifications are those referred to the lagged change in real house price and to the change in real income per capita. It is worth highlighting equation (3) in table 5, which presents the highest explanatory power among those equations which have been estimated using GMM. In equation (3), all but one (year 2006) of the time dummies coefficients are significant and neatly depict the change of trend in real house prices in Spain (see figure 2) starting in 2007 with a change of sign in such coefficients. Moreover, these results are clear evidence of strong serial correlation in real house prices in Spain between 1996 and 2008, with a coefficient for the lagged change in real prices of around 0.6, and also show that income elasticity of real house prices in Spain between 1996 and 2008 was around 0.3. That is to say that, for example, an annual economic growth of 5 percent caused a growth in prices of 1.5 percent[50].

5. Spatial structure in the data

Although results presented in table 5 are really reliable they are not completely satisfactory since in the process of estimation the possible existence of cross section dependence has not been explicitly taken into account. In this section we intend to show that such dependence really exists in the dynamics of real house prices in Spain and that, consequently, in subsequent work on the analysis of such dynamics an effort has to be done in order to explicitly consider the spatial structure in the data. First, we present and intuitive consideration of the spatial pattern tabulating correlations coefficients between the spatial units, in the same manner as Holly et al. (2006). Second, we validate our intuition testing our data with the best known spatial independence test in the literature of spatial econometrics, the test of Moran. We are going to carry out this analysis with real house price and real

[48] For a detailed analysis of the GMM estimator see Arellano (2003).
[49] Although we acknowledge, as Holly et al. (2006) point out, that "local housing market shocks are likely to be correlated in ways that are not captured by simple time effects" (p. 10).
[50] This is lower than other estimations of income elasticities for countries like US, see Capozza et al. (2004).

income data, since only this fundamental turn out to be clearly robust in the estimation process developed in the previous section.

To begin with, we consider simple correlations between each province[51] and between correlations for a group of five geographical regions named: North, Ebro, East, South and Centre[52]. Regarding real income per capita, one of the main results is that the highest coefficient of correlation (0.93) is the one between the provinces of Madrid and Barcelona. These two provinces are at the same time highly correlated with other developed provinces such as the Balearic Islands and Navarre. This result reveals that, probably, are advanced services developed in different parts of the country the main factors that stimulated Spanish economic growth during the period 1996-2008 and consequently that there is not a clear spatial pattern in the behaviour of this variable. However, what it seems clear, when we look at the between correlations at the level of geographical regions (results shown in table 7), is that, in general, real income growth is correlated across the country. Moreover, these results are not clearly rejecting a possible, although weak, spatial component in Spanish economic growth since North is higher correlated with closer regions, as Ebro and Centre, than with distant ones, as East. The same happens in the case of East, which is higher correlated, in terms of economic growth, with Ebro and South than with Centre.

In the case of real house prices the results point at a more perceptible spatial pattern. When we look at simple coefficient correlations, the one between Madrid and Barcelona (0.69) is lower than those between Madrid and its neighbours: Toledo (0.93), Guadalajara (0.91) and Segovia (0.88), the same happens with Barcelona and its neighbours: Gerona (0.84), Lerida (0.83) and Tarragona (0.78). However, real house prices in Madrid are also highly correlated to that in some coastal provinces, where a high number of second residences of nationals and foreigners are located, such as Castellón (0.93), Cádiz (0.91), Almería, Huelva, Málaga and Murcia (0.90) or Alicante (0.88). Clearly, the latter correlations are not of spatial nature. With regard to the correlations at geographical region level (results shown in table 8), the spatial pattern is confirmed and more clear than in the case of real income. As can be seen, correlations on average decline with distance.

[51] Due to space problems province level correlation coefficients are not completely reported. In the appendix there is a selection of these coefficients, see tables A1 and A2.

[52] See table 6 for the grouping of provinces and real regions in these five geographical regions. A map of Spanish provinces can be found in the Appendix (figure A1)

To sum up, this previous analysis shows that both, real income per capita and real house prices, are correlated along the country (see figures 5 and 6)[53] and that, at least, in the case of the latter, it seems, that this correlation has a specific spatial pattern.

In addition, to gather more clear evidence in favour of the existence of a spatial structure in our data we report provincial maps for real house prices and real income per capita for different years (1995, 2000, 2005 and 2008) along the period of analysis (see figures 7 and 8). In these maps the different intensity of colour shows different levels of price and income. As can be seen, a clear spatial pattern is depicted for the two variables along the period 1995-2008[54], with the exception of year 2008 for real house prices probably due to the change of trend in this variable initiated in the preceding year.

Finally, to verify econometrically these intuitive results we compute the Moran test[55] for our data. This is a test of spatial independence or to be precise of no correlation between the series and its spatial lags (where the null hypothesis is no correlation). The test results are reported in table 9, where 6 spatial lags[56] are considered. The evidence is strongly in favour of the existence of cross section dependence in both variables[57]. However, in the case of real house prices it seems that something happened at the end of the period probably reflecting, as explained before, the change of trend.

---

[53] The movements of the two variables are shown in terms of logarithm changes in order to allow the possible heterogeneity to be seen more clearly. However, it can be seen that Spanish geographical regions experienced broadly comparable movements during the period 1996-2008, particularly in the case of real house prices.

[54] Under this approach, it seems that the spatial pattern is more clear for real income than for real house prices, quite the opposite that under the correlation coefficients approach.

[55] See Moran (1950) for a detailed explanation of the test.

[56] These can be understood as contiguity levels or degree of neighbourhood relation.

[57] Notice that the fact that the sign of the statistic change as the number of lags increased is additional evidence that space matters since that change of sign is reflecting that the structure in the behaviour of the variable change as we increase distance.

6. Conclusions and further research

This paper studies the dynamics of house prices in Spain at the level of provinces between 1995 and 2008. In order to do this, first, we develop the usual theoretical framework, from which we derive a specific relationship between house prices and fundamentals. Next, we give a detailed description of the Spanish housing market during the period studied, paying special attention to the behaviour of the fundamentals identified in the theoretical analysis. Then, we test empirically the theoretical relationship obtained previously. Before we estimate this relationship, we try to test for panel cointegration, but it is not possible since the required methods are not implemented in the usual econometric packages. However, we provide a detailed panel unit root analysis of the variables. A static and dynamic panel are estimated with results according to the theory. There is clear evidence of serial correlation in Spanish house prices and of income elasticity of house prices of around 0.3.

These results are econometrically sound and reliable. However, as pointed out along the estimation process, the results can be improved if cross section dependence is explicitly considered in the analysis. In the last section we provide clear evidence that this type of dependence exists in Spanish real house prices and main fundamental real income per capita. Consequently, a subsequent paper is required where spatial structure in the data is explicitly considered in the analysis, basically following the different tests proposed in Breitung and Pesaran (2008) and Presaran (2006a) which have to be computed *manually*.

References

Anselin, L. (1988) *Spatial Econometrics*. *Methods and models*. Kluwer, Dordrecht

Arellano M. (2003) *Panel data econometrics*. Oxford University Press, Oxford

Breitung, J., and H. Pesaran (2008) Unit roots and cointegration in panels, in *The econometrics of panel data: fundamentals and recent developments in theory and practice (3rd edition)*, Mátyás, L. and P. Sevestre (eds.), Springer.

Cameron, G., J. Muellbauer, and A. Murphy (2006) Was there a British house price bubble? Evidence from a regional panel. *Economics Series Working Papers* 276, University of Oxford

Capozza, D., P. Hendershott and C. Mack (2004) An anatomy of price dynamics in illiquid markets: analysis and evidence from local housing markets. *Real Estate Economics* 32: 1-32

Case, K. and R. Shiller (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity* 2: 299-342

Colegio de registradores de la propiedad, bienes muebles y mercantiles de España (2009) *Estadística registral inmobiliaria. Anuario 2008,* Madrid

Cho, M. (1996) House price dynamics: a survey of theoretical and empirical issues. *Journal of Housing Research* 7: 145-172

Debelle G. (2004) Macroeconomic implications of rising household debt. *Bank for International Settlements Working Papers* 153

Dougherty, A. and R. Van Order (1982) Inflation, housing costs, and the consumer price index. *American Economic Review* 72: 154-165

Engle, R and C. Granger (1987) Co-integration and error correction: representation, estimation, and testing. *Econometrica* 55: 251-276

European Central Bank (2003) *Structural factors in the EU housing market*

Girouard N., M. Kennedy, P van den Noord and C. André (2006) Recent house price developments: The role of fundamentals. *OECD Economics Department Working Papers* 475

Hadri, K (2000) Testing for stationarity in heterogeneous panel data. *Econometrics Journal* 3: 148-161

Hendry, D. (1984) Econometric modelling of house prices in the UK, in *Econometrics and quantitative economics*, Hendry, D. and K. Wallis (eds.), Blackwell, Oxford

Hilber, C. (2007) *The determinants of homeownership across Europe: panel data evidence*. London School of Economics

Hilber, C. and C. Mayer (2009) Why do households without children support local public schools? Linking house price capitalization to school spending. *Journal of Urban Economics* 65: 74-90

Holly, S., M. Pesaran and T. Yamagata (2006) A spatio-temporal model of house prices in the US. *Cambridge Working Papers in Economics* 0654, University of Cambridge

Im, K., M. Pesaran and M. Shin (1995) Testing for unit roots in heterogeneous panels. *Department of Applied Economics Working Papers* 9526, University of Cambridge

Jacobsen, D. and B. Naug (2005) What drives house prices? *Norges Bank Economic Bulletin* 1: 29-41

Johansen, S. (1991) Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica* 59: 1551–1580

Larraz, B. and J. Alfaro (2008) Asymmetric behaviour of Spanish regional house prices. *International Advances in Economic Research* 14: 407-421

Levin, A., C. Lin and C. Chu (2002) Unit root tests in panel data: Asymptotic and finite sample properties. *Journal of Econometrics* 108: 1-24

Malpezzi, S. (1999) A simple error correction model of house prices. *Journal of Housing Economics* 8: 27–62

Martínez, J. and L. Maza (2003) Analysis of house prices in Spain. *Bank of Spain Working Paper Series* 0307

McAvinchey, I. and D. Maclennan (1982) A regional comparison of house price inflation rates in Britain, 1967-76. *Urban Studies* 19: 43-57

McCarthy, J. and R. Peach (2004) Are home prices the next bubble? *Federal Reserve Bank of New York Economic Policy Review* (December): 65-78

Meen, G. (1990) The removal of mortgage market constraints and the implications for econometric modelling of UK house prices. *Oxford Bulletin of Economics and Statistics*: 52: 1-24

Meen, G. (1996) Ten propositions in UK housing macroeconomics: An overview of the 1980s and early 1990s. *Urban Studies* 33: 425–44

Meen, G. (1999) Regional house prices and the ripple effect: a new interpretation. *Housing Studies* 14: 733-753

Moran, P. (1950) Notes on continuous stochastic phenomena. *Biometrika* 37: 17-23

Muth, R. (1960) The demand for nonfarm housing, in Harberger, A. (ed.), *The demand for durable goods*, University of Chicago Press, Chicago: 29-96

Nagahata, T., Y. Saita, T. Sekine and T. Tachibana (2004) Equilibrium land prices of Japanese prefectures: A panel cointegration analysis. *Bank of Japan Working Paper Series*: 04-E-9

Pain, N. and P. Westaway (1997) Modelling structural change in the UK housing market: a comparison of alternative house price models. *Economic Modelling* 14:587-610

Pesaran, M. (2006a) Estimation and inference in large heterogeneous Panels with a Multifactor Error Structure. *Econometrica* 74: 967-1012

Pesaran, M. (2006b) *A simple panel unit root test in the presence of cross section dependence*. University of Cambridge

Phillips, P. (1991) Optimal inference in co-integrated systems. *Econometrica* 59: 282-306

Poterba, J. (1984) A tax subsidies to owner-occupied housing: An asset market approach. *Quarterly Journal of Economics* 99: 729-52

Reid, M. (1962) *Housing and income*, University of Chicago Press, Chicago

Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge (Mass)

Table 1. Summary statistics on the data series

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
Real average price per square metre (euros)	700	977.93	415.05	427.53	2492.51
Real GDP per capita (thousands of euros)	700	15.501	3.719	7.558	27.320
Stock of houses per total population	700	0.5547	0.1027	0.3866	0.9582
Working adults per total population	700	0.4843	0.0338	0.4144	0.5608
Real interest rate	700	0.0262	0.0205	-0.0343	0.0772

Table 2. Panel unit root tests results

	Levin, Lin & Chu t*	Im, Pesaran and Shin W-stat	Hadri Z-stat
With an intercept and a linear trend			
p_{it}	4.28	2.98	9.51***
y_{it}	4.58	4.50	13.26***
d_{it}	8.73	8.12	13.69***
h_{it}	-4.56***	0.78	9.90***
Only with an intercept			
Δp_{it}	1.92	-0.45	1.41*
Δy_{it}	3.23	-2.44***	6.26***
r_{it}	-11.23***	-4.42***	13.17***
Δd_{it}	3.73	2.30	8.00***
Δh_{it}	-7.18***	-4.51***	3.28***

Note: The superscripts "***", "**" and "*" mean that the test is significant at the 1, 5 and 10 per cent level respectively.

Table 3. Static panel estimation results (Least Squares). Cross-section and period fixed effects (dummy variables)

Explanatory variables	Dependent variable: p_{it}	
	(1)	(2)
y_{it}	0.53 (0.10)	0.54 (0.10)
r_{it}	0.24 (0.47)	------
d_{it}	1.72 (0.24)	1.72 (0.24)
h_{it}	-0.92 (0.11)	-0.93 (0.11)
α_i	6.05 (0.36)	6.03 (0.36)
Observations	700	700
Adjusted R^2	0.96	0.96

Note: Standard errors (corrected for White heteroscedasticity) are given in parenthesis.

Table 4. Dynamic panel estimation results (Generalized Method of Moments). Cross-section fixed effects (first differences)

Explanatory variables	Dependent variable: Δp_{it}
	(1)
Δp_{it} (-1)	0.66 (0.01)
Δy_{it}	0.77 (0.07)
Δr_{it}	-0.65 (0.07)
Δd_{it}	0.75 (0.10)
Δh_{it}	-0.56 (0.17)
Observations	600
Adjusted R^2	0.43

Note: Standard errors(corrected for White heteroscedasticity) are given in parenthesis.

Table 5. Dynamic panel estimation results (Generalized Method of Moments). Cross-section fixed effects (first differences) and period fixed effects (dummy variables, in table)

Explanatory variables	Dependent variable: Δp_{it}		
	(1)	(2)	(3)
Δp_{it} (-1)	0.61 (0.04)	0.59 (0.03)	0.64 (0.03)
Δy_{it}	0.35 (0.12)	0.36 (0.10)	0.30 (0.11)
Δr_{it}	0.01 (0.19)	------	-0.01 (0.16)
Δd_{it}	0.87 (0.53)	0.98 (0.51)	------
Δh_{it}	-0.10 (0.21)	-0.19 (0.17)	------
1997	-0.005 (0.007)	-0.004 (0.007)	0.002 (0.005)
1998	0.009 (0.008)	0.007 (0.007)	0.022 (0.005)
1999	0.023 (0.009)	0.022 (0.008)	0.032 (0.006)
2000	0.009 (0.012)	0.009 (0.012)	0.025 (0.005)
2001	0.018 (0.009)	0.018 (0.008)	0.028 (0.006)
2002	0.013 (0.009)	0.014 (0.009)	0.022 (0.004)
2003	0.017 (0.007)	0.017 (0.007)	0.028 (0.004)
2004	0.038 (0.007)	0.040 (0.006)	0.040 (0.005)
2005	0.011 (0.008)	0.011 (0.007)	0.020 (0.006)
2006	0.003 (0.007)	0.004 (0.005)	0.003 (0.004)
2007	-0.025 (0.008)	-0.023 (0.005)	-0.024 (0.005)
2008	-0.049 (0.004)	-0.049 (0.004)	-0.045 (0.003)
Observations	600	600	600
Adjusted R^2	0.59	0.58	0.62

Note: Standard errors (corrected for White heteroscedasticity) are given in parenthesis.

Table 6. Geographical regions

NORTE	EBRO	ESTE	SUR	CENTRO
P. ASTURIAS	ARAGÓN	CATALUÑA	ANDALUCÍA	CASTILLA Y LEON
Asturias	Huesca	Barcelona	Almería	Ávila
CANTABRIA	Teruel	Girona	Cádiz	Burgos
Cantabria	Zaragoza	Lleida	Córdoba	León
GALICIA	LA RIOJA	Tarragona	Granada	Palencia
A Coruña	La Rioja	C. VALENCIANA	Huelva	Salamanca
Lugo	C. NAVARRA	Alicante	Jaén	Segovia
Ourense	Navarra	Castellón	Málaga	Soria
Pontevedra		Valencia	Sevilla	Valladolid
PAIS VASCO		I. BALEARES	EXTREMADURA	Zamora
Álava		Islas Baleares	Badajoz	CASTILLA - MANCHA
Guipúzcoa			Cáceres	Albacete
Vizcaya			I. CANARIAS	Ciudad Real
			Las Palmas	Cuenca
			S. C. Tenerife	Guadalajara
			R. MURCIA	Toledo
			Murcia	C. MADRID
				Madrid

Table 7. Correlation coefficients between geographical regions: first difference of logarithm of real GDP per capita

	NORTH	EBRO	EAST	SOUTH	CENTRE
NORTH	*				
EBRO	0.82	*			
EAST	0.58	0.85	*		
SOUTH	0.82	0.90	0.79	*	
CENTRE	0.81	0.78	0.61	0.87	*

Table 8. Correlation coefficients between geographical regions: first difference of logarithm of real house prices

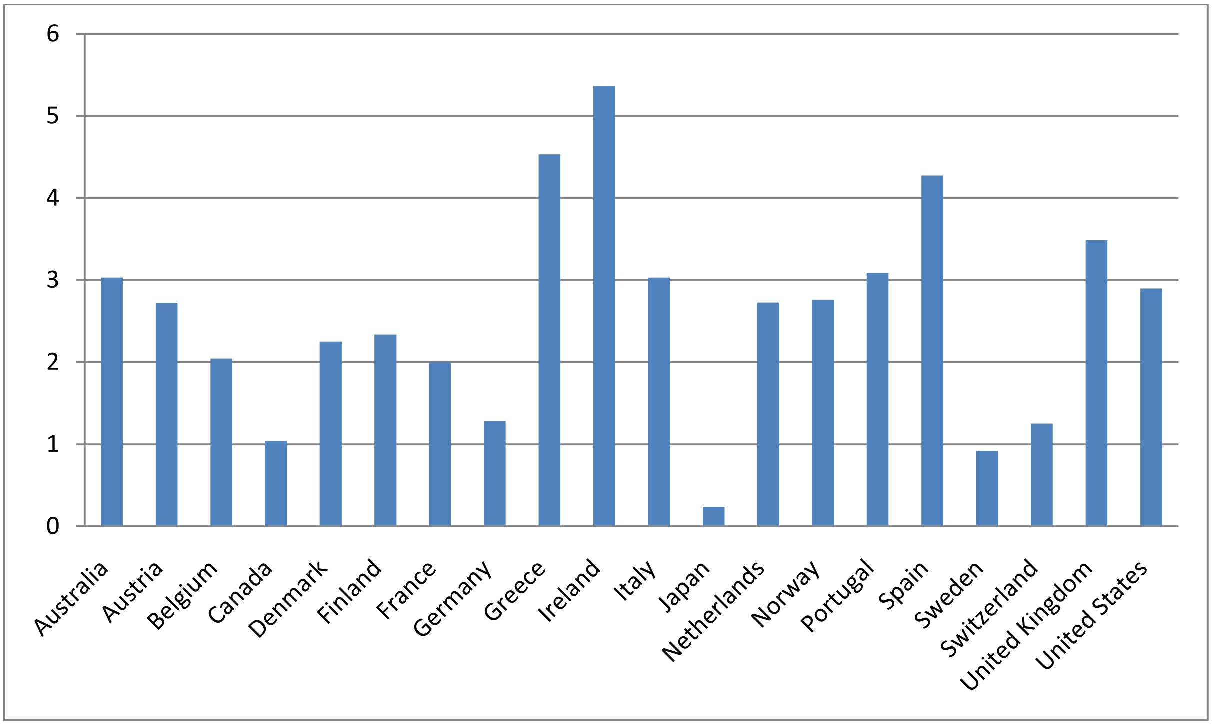
	NORTH	EBRO	EAST	SOUTH	CENTRE
NORTH	*				
EBRO	0.94	*			
EAST	0.93	0.96	*		
SOUTH	0.90	0.91	0.95	*	
CENTRE	0.91	0.85	0.91	0.96	*

Table 9. Moran spatial independence test results

Lags	1995	2000	2005	2008
Variable: p_{it}				
1	0.4763***	0.4638***	0.3168***	0.3122***
2	0.2410***	0.2111***	0.0230	0.0099
3	0.0783*	-0.0018	-0.0758	-0.0806
4	-0.1223**	-0.1183**	-0.0615	-0.0042
5	-0.2398**	-0.2345**	-0.0635	-0.0848
6	-0.3219***	-0.2439**	-0.1324*	-0.1723**
Variable: y_{it}				
1	0.7476***	0.6836***	0.6948***	0.6904***
2	0.5184***	0.428***	0.4062***	0.4162***
3	0.1514**	0.1154**	0.1126**	0.1404***
4	-0.2397***	-0.2164***	-0.2185***	-0.1950***
5	-0.4917***	-0.4356***	-0.4321***	-0.4400***
6	-0.4719***	-0.3968***	-0.4013***	-0.4361***

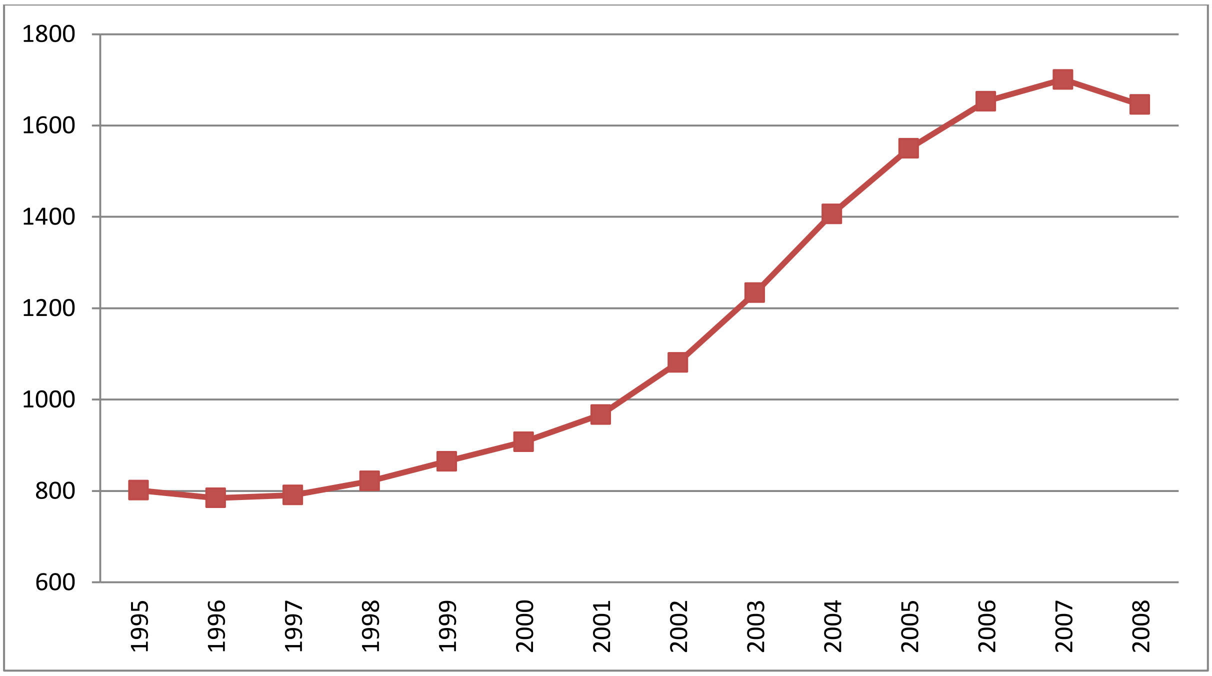
Note: The superscripts “***”, “**” and “*” mean that the test is significant at the 1, 5 and 10 per cent level respectively.

Figure 1: Average annual real growth of housing prices 1995-2008



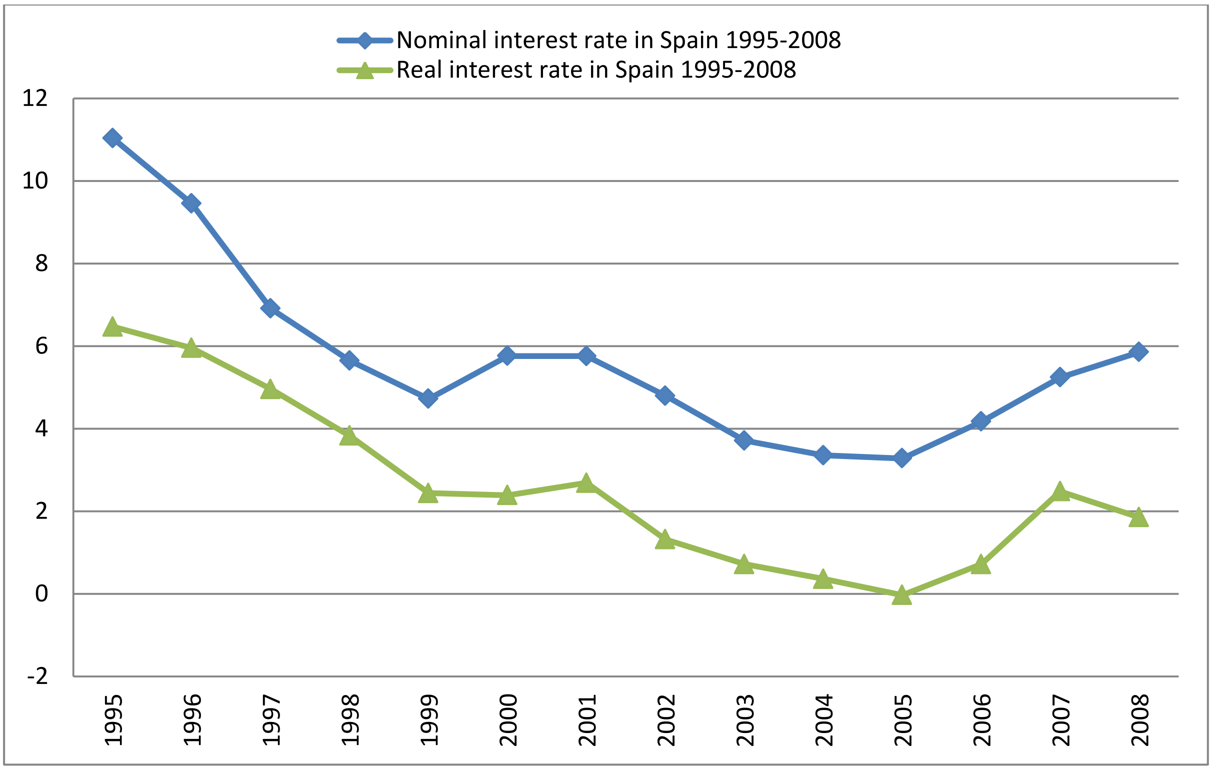
Source: Own calculations based on OECD Main Economic Indicators data.

Figure 2: Real average price per square metre (Euros). Spain 1995-2008



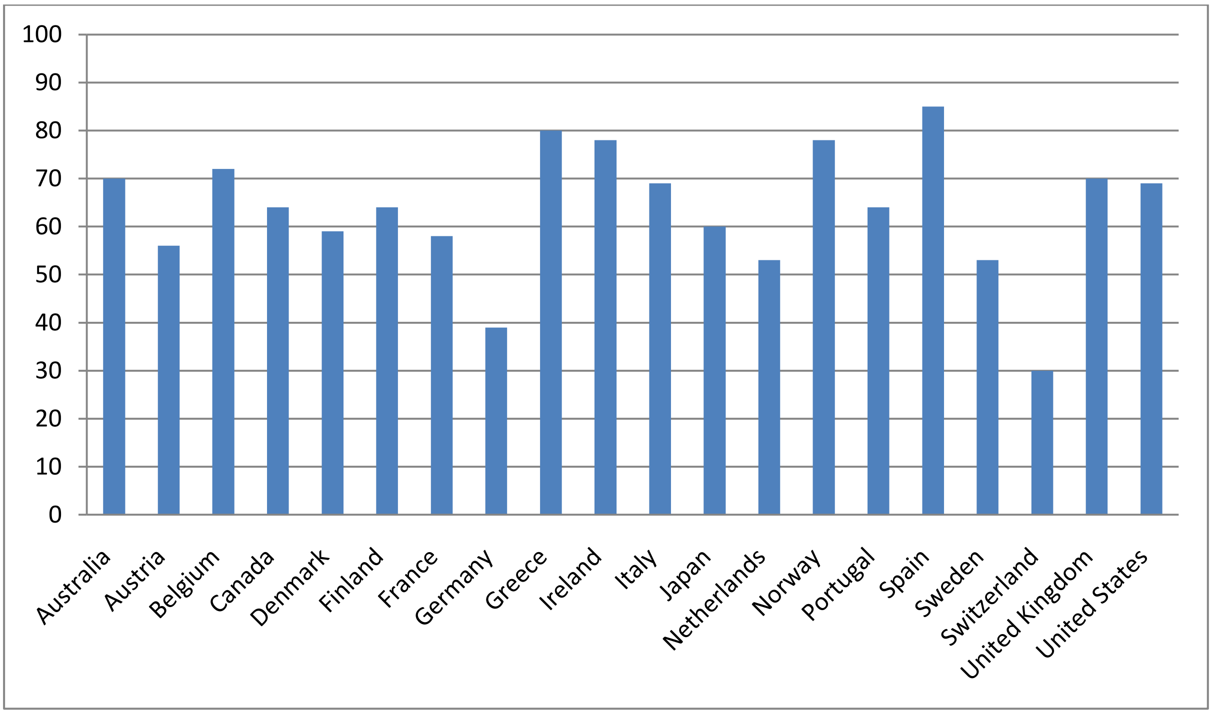
Source: Spanish Department of Housing and INE (Instituto Nacional de Estadística).

Figure 3: Nominal and real Spanish interest rates



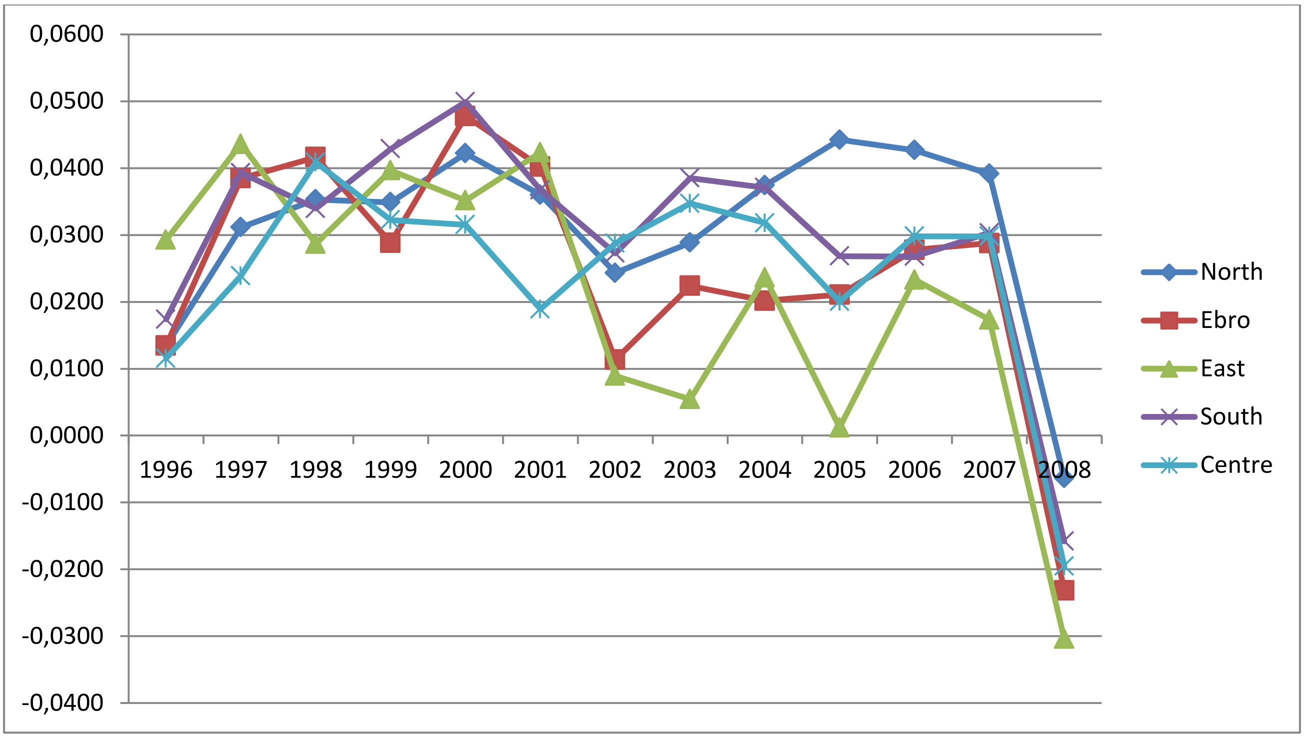
Source: Banco de España and INE.

Figure 4: Share of home-ownership in 2002



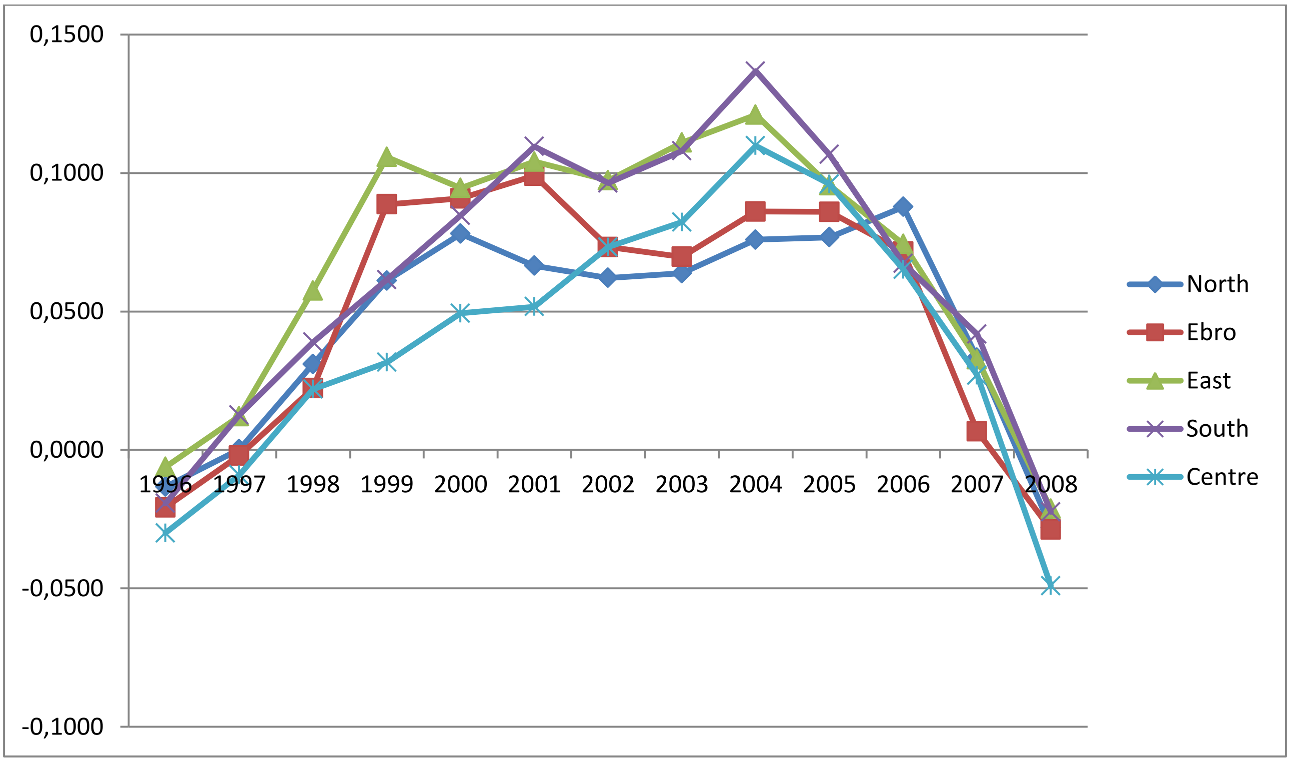
Source: Debelle (2004), Table 1.

Figure 5: Logarithm changes in real income per capita. Geographical regions 1996-2008



Source: Own calculations based on INE data.

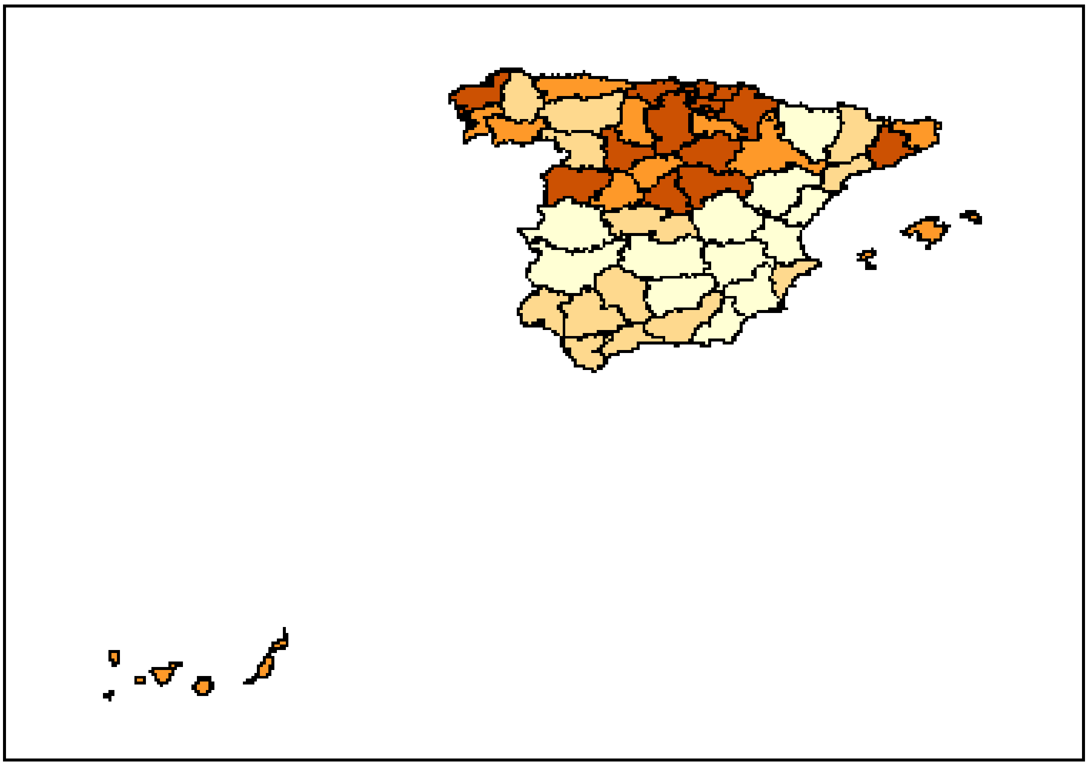
Figure 6: Logarithm changes in real house prices. Geographical regions 1996-2008



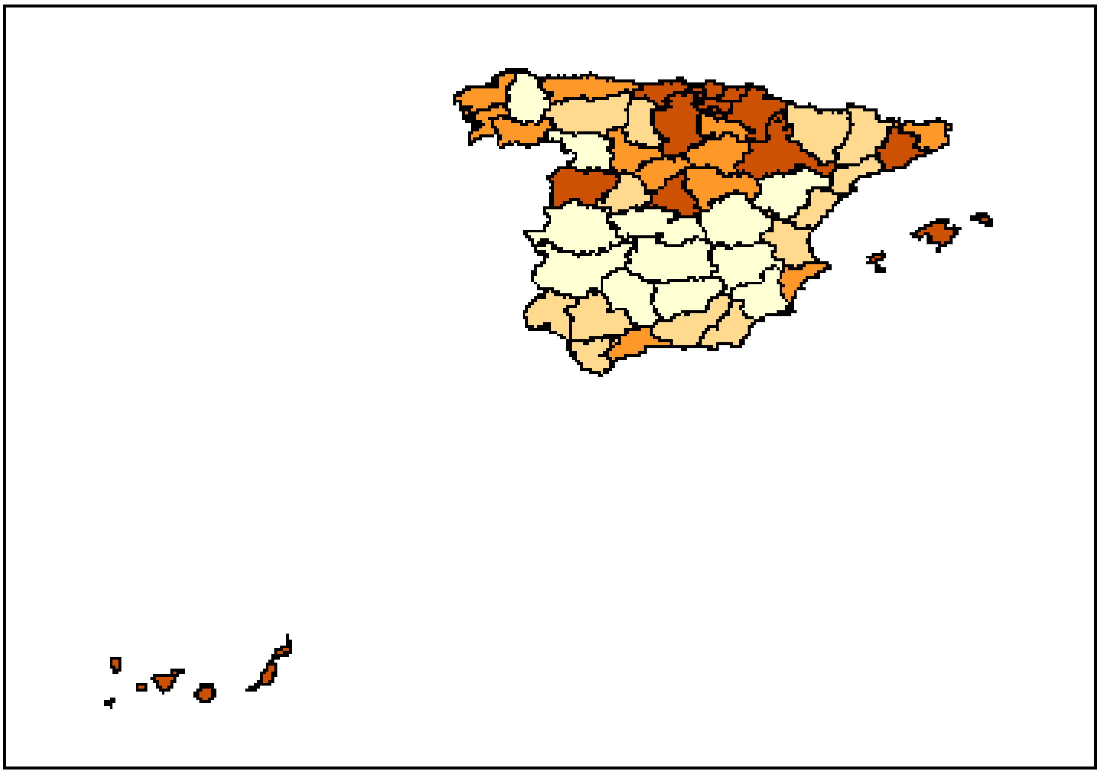
Source: Own calculations based on Spanish Department of Housing and INE data.

Figure 7. Logarithm of real house prices. Provincial maps for different years.

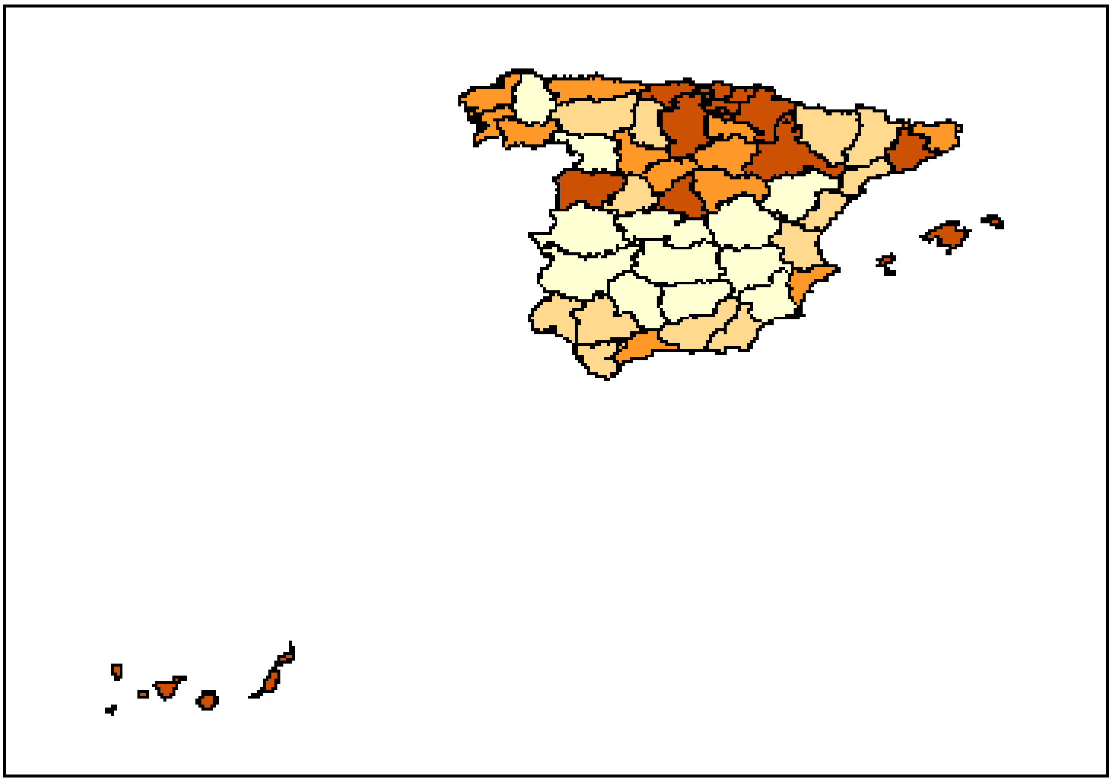
1995



2000



2005



2008

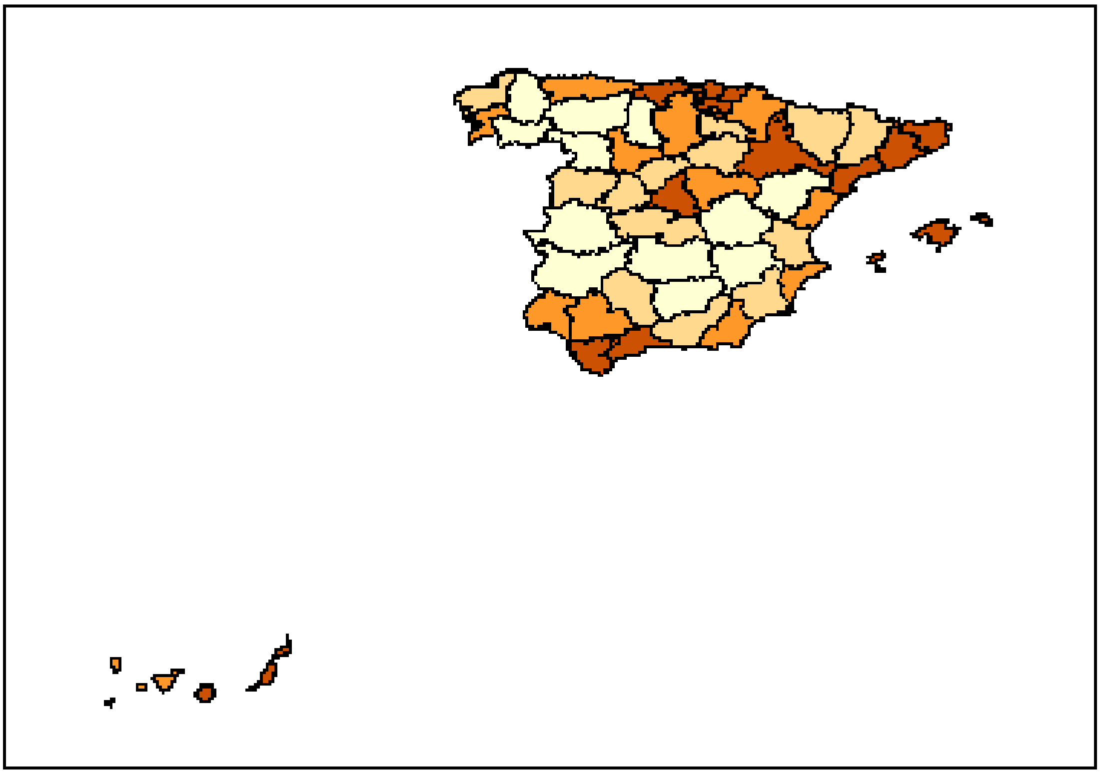
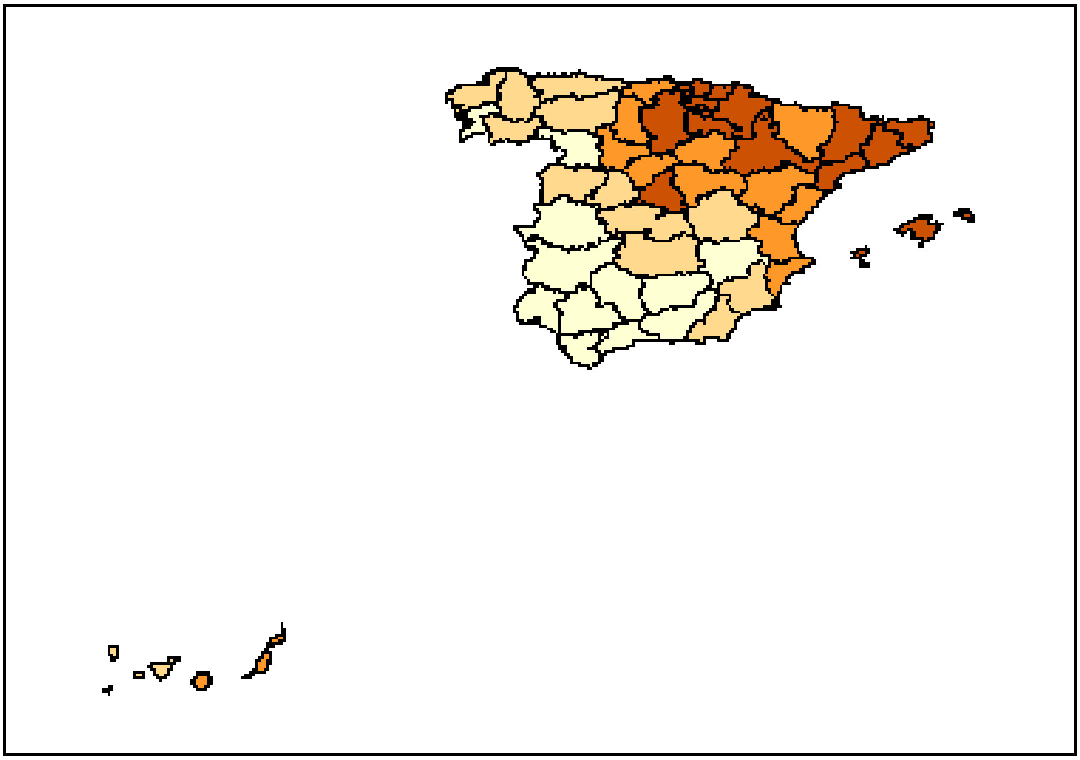
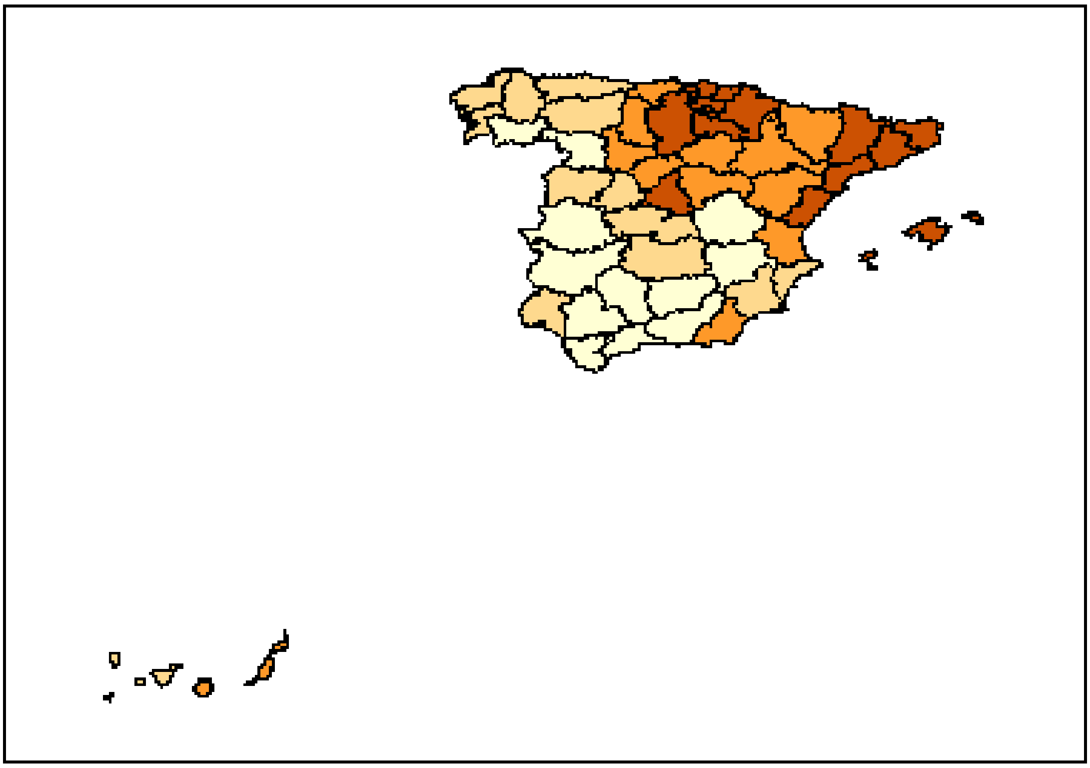


Figure 8. Logarithm of real GDP per capita. Provincial maps for different years.

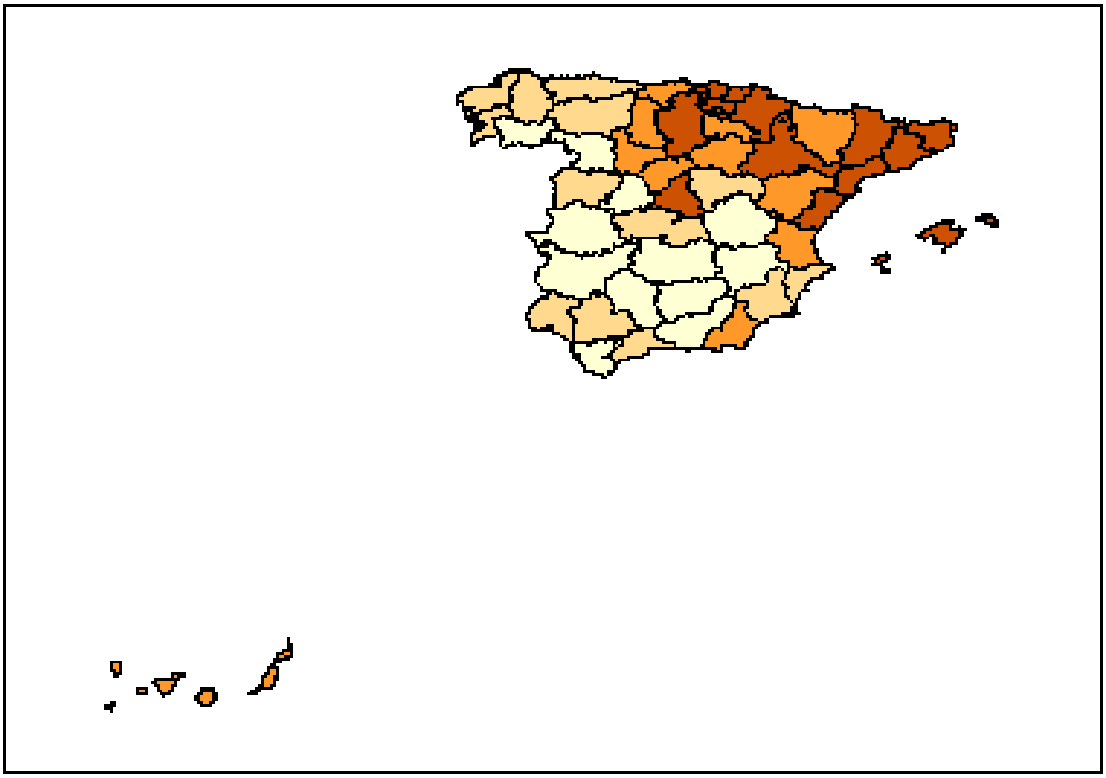
1995



2000



2005



2008