The Role of Fundamentals in the Price of Housing: Theory and Evidence
The Role of Fundamentals in the Price of Housing: Theory and Evidence

Pedro Álvarez-Lois†, Galo Nuño-Barrau‡

BBVA Economic Research Department

20 July 2007

Abstract

Despite the extensive research in the field of housing economics, the role of some ‘fundamental’ economic variables, such as income, interest rates or stock of houses per capita, on the real price of housing is still not fully understood. In this paper we develop a dynamic general equilibrium model to micro-fundament the price of housing in a market without renting. In this framework we underline the dual role of housing both as a good that produces valuable services and as an investment asset that can be resold in a future date. To test the theoretical results obtained, we analyze the Spanish housing market from 1995 to 2006 as it seems to satisfy the theoretical assumptions in practice. We examine the extent to which real house prices at the regional level are driven by fundamentals by applying Panel Cointegration methods such as Common Correlated Effects, Dynamic Ordinary Least Squares and Vector Error Correction. Results are fully consistent with the theory and underline the importance of both long-run adjustment and persistence processes to explain the dynamic behaviour of prices.

JEL Classification: C23, R21, R31

Keywords: Panel Cointegration, housing prices, adjustment dynamics, macroeconomic model

† E-mail: pedro.a_lois@grupobbva.com
‡ E-mail: galonuno@grupobbva.com
* The authors wish to acknowledge the helpful comments and criticism of Sebastian Guevara, Tomás Riestra and José Luís Escrivá.
1. Introduction

Most economists agree that one of the main challenges when analyzing the housing price is the dual nature of houses. The fact that a house is both a good that renders valuable housing services and an asset, in the sense that it is a durable good that may be sold in the future, implies that the housing market is composed by the whole range of individuals that go from the pure speculator, who buys the house only as an investment, to the pure owner, who is only concerned about the utility that he obtains from the possession of the house. Most of the literature departs from this fact to construct asset pricing models à la Campbell and Shiller (1988) that exploit some intertemporal links between house prices, rents and interest rates. The final aim of these models is to shed some light on the relative importance of ‘expectations’ and ‘fundamental’ variables in understanding a number of stylized facts observed in episodes of surge in house prices, including the positive correlation of the price with the volume of housing purchases for investment motives or the negative correlation with real interest rates.

But how does causality goes between purchasing and rental prices? Most empirical papers based on these models, as for example Ayuso and Restoy (2006) or Nagahata et al. (2004), consider the rental price as exogenous, and try to verify if the purchasing price behaves consistently with the theory. The problem is that for countries where rental markets are neither liquid nor deep, this assumption might be implausible and causality could run reversed, i.e. rental prices could be determined by purchasing prices. In consequence, purchasing price should be either determined by some economic fundamentals or to be assumed to behave erratically. As housing is a key component of the wealth of families and an essential factor to understand expenditure patterns (BIS 2002), applied economists have been trying to find some consistent relationship between housing prices and fundamentals such as family incomes, demographic trends, unemployment rates, restrictions on building or interest rates.

The aim of this paper is to analyze the relationship between house prices and their fundamentals, especially in cases when the rental market is underdeveloped, so approaches à la Campbell and Shiller may not be appropriated to obtain purchasing prices. In order to do so, in Section 2 we develop a theoretical model that departs from the literature in Dynamic General Stochastic Equilibrium (DGSE) models such as Heathcote and Davis (2005) to derive a supply-demand equation for the housing market. We consider the economy to be populated by identical households that consume final goods and housing services and who may hire their labour to firms and invest both in capital and in housing assets. Firms belong to one of the two productive sectors: manufacturing and housing-producing. The solution of the model allows for multiple (‘sunspot’) equilibria and takes into account both the expectation about future housing prices and fundamentals such as interest rates, consumption or stock of houses per capita. In consequence, we obtain an equilibrium relationship between housing prices and fundamentals, which may be tested empirically. This is an important issue, as this is not a theoretical paper, but an empirical one based on a general equilibrium theoretical model.

In Section 3 we present the stylised facts of the Spanish housing market. There are diverse reasons to choose Spain to conduct empirical work. The first reason is the small relative size of its rental market (Ruiz and San Martín 2004), which validates our initial assumption. The second reason is the profusion of regional statistics, allowing the construction of a panel for the period 1995-2006 for the 50 Spanish provinces. Additionally Spain is one of the countries where the surge in housing prices has been more spectacular in the period under consideration, and it has also suffered different shocks, such as economic and demographics ‘boom’ since the late nineties. In Section 3 we also analyse. As a matter of comparison, the literature of panel cointegration analysis of the housing price also includes the US market (Holly, Pesaran and Yamagata 2006) and the Japanese one (Nagahata et al. 2004).

In Section 4 we proceed to perform the econometric analysis of the relationship between price and fundamentals. Different methods can be applied; the Error Correction Model (ECM) approach is probably the most extended one (Capozza et al. 2002, Jacobsen and Naug 2005, McCarthy and Peach 2002) as it is simpler to estimate and may offer additional information about evidences of backward-looking expectations (Capozza and Seguin, 1996). A second approach is to apply a Vector Autoregression (VAR) o Structural VAR (SVAR) framework to model the dynamic interaction between the housing price and some of its fundamentals (Tsatsaronis and Zhu 2004). The third approach is less extended, and it consists of applying some kind of dynamic factor model to capture the co-movement between a group of observable time series (the price and its fundamentals) and a small number of unobservable variables, which try to capture global and local environment in the housing market and in the economy as a whole (IMF 2004).
In our case, due to the panel structure of the data, we have preferred to employ a fourth approach. Panel cointegration methods apply cointegration techniques to find long-run relationships in panel data with a substantial number of cross-section units (Breitung and Pesaran 2007). They capture heterogeneity due to region-specific characteristics by considering that the short-run dynamics may differ across regions, whereas the long-run relationship is the same for all. The reason for assuming a homogeneous long-run (or equilibrium) relationship is that the underlying economic principles that are employed to establish the equilibrium and derive from the theoretical foundations should apply similarly in all regions, whereas the adjustment process may differ due to behavioural and institutional features.

In Section 5 we present our main conclusions. The existence of a cointegration relationship between real housing prices and fundamental variables seem to support our theoretical model, which indicates that the evolution of the price in Spain is not due to ‘Irrational Exuberance’ but to the surge in family income, the arrival of more than 4 million immigrants and the highest ever reduction in real interest rates. Results are robust to a reduction in the sample size by eliminating outliers where other effects may be playing an important role, such as the acquisition of residential houses by foreigners. Additionally, the theoretical model seems to support the hypothesis that real housing prices tend to grow in the long-term, as an especial case of the ‘Baumol’s Cost disease’ (Baumol. 1967). Further research should extend this framework to other countries and periods.
2. A Model of Housing and the Macro-Economy

In this section, we introduce a dynamic macro-economic model that takes explicit account of the housing market. The objective is two-fold: firstly, to show the conditions under which the real price of housing is non-stationary and, secondly to analyse in a coherent way the determinants of the real price of housing and its interaction with the rest of variables in the economy.

In the model, the economy is populated by identical, infinitely-lived households. The population grows at a constant gross rate $\eta$ so that in what follows all variables are in per-capita terms. The economy consists of two productive sectors. One sector produces final goods that can either be consumed or invested in productive capital. The other sector produces houses. Both sectors differ in the level technology as well as in the quantity of inputs used for production. For simplicity we do not consider the government in the model. Moreover, we assume perfect competition in all markets and absence of real as well as nominal frictions.

2.1 Firms

We consider two types of inputs, namely labour and physical capital. Let $P^h_t$ be the price of houses in terms of the final good. Let $w_t$ and $r^k_t$ be, respectively, the competitive wage and rental rate on capital measured in the same units of the final good. The representative housing-producing firm maximises profits according to

$$\max P^h_t Y^h_t - w_t N^h_t - r^k_t K^h_t$$

s.t.

$$Y^h_t = A^h_t \left( K^h_t \right)^{\alpha_h} \left( N^h_t \right)^{1-\alpha_h}$$

where $A^h_t$ is the level of total factor productivity in the construction sector. We assume that this variable follows a non-stationary unit-root process

$$A^h_t = A^h_{t-1} \exp \left( \Phi^h + \epsilon^h_t \right)$$

where $\epsilon^h_t$ is a white noise shock and $\Phi^h$ is the long-run rate of growth of the increase in the technology factor. The first order conditions associated with (2.1) imply that the optimal demand of inputs equalises their price to the marginal productivities:

$$\alpha_h P^h_t A^h_t \left( K^h_t \right)^{\alpha_h-1} \left( N^h_t \right)^{1-\alpha_h} = r^k_t$$

and

$$\left( 1 - \alpha_h \right) P^h_t A^h_t \left( K^h_t \right)^{\alpha_h} \left( N^h_t \right)^{1-\alpha_h} = W_t$$

Similarly, firms in the final-goods sector choose optimal demand of production factors in order to maximise their profits, that is,

$$\max Y^c_t - w_t N^c_t - r^k_t K^c_t$$

s.t.

$$Y^c_t = A^c_t \left( K^c_t \right)^{\alpha_c} \left( N^c_t \right)^{1-\alpha_c}$$

where $A^c_t$ is the level of total factor productivity in the final-goods sector. As before, we assume that this variable follows a non-stationary unit-root process

$$A^c_t = A^c_{t-1} \exp \left( \Phi^c + \epsilon^c_t \right)$$

where $\epsilon^c_t$ is a white noise shock and $\Phi^c$ is the long-run rate of growth of the increase in the technology factor. The first order conditions associated with (2.6) are

$$\alpha_c A^c_t \left( K^c_t \right)^{\alpha_c-1} \left( N^c_t \right)^{1-\alpha_c} = r^k_t$$

and

$$\left( 1 - \alpha_c \right) A^c_t \left( K^c_t \right)^{\alpha_c} \left( N^c_t \right)^{1-\alpha_c} = W_t$$

Now, notice that (2.5) combined with (2.2) can be expressed as

$$\left( 1 - \alpha_h \right) P^h_t \frac{Y^h_t}{N^h_t} = W_t$$

Similarly, combining (2.10) with (2.7) we get

$$\left( 1 - \alpha_c \right) \frac{Y^c_t}{N^c_t} = W_t$$

Now, from (2.11) and (2.12) we obtain a relationship for the real price of houses:

$$P^h_t = \left( 1 - \alpha_h \right) \frac{Y^h_t}{N^h_t} \frac{N^c_t}{Y^c_t}$$

It can be shown that the production of $Y^h_t$ has a trend growth rate of $\Phi^h/(1 - \alpha_h)$ and that trend growth of final goods production is $\Phi^c/(1 - \alpha_c)$. Moreover, employment has no trend growth in this model. Hence, taking into account (2.13), the real price of housing would be non-stationary whenever the trend growth rates of the two productive sectors differ. For instance, if the final good sector is more capital intensive than the construction sector, that is, $\alpha_c > \alpha_h$ and that technology grows more slowly in the latter sector, that is, $\Phi^h < \Phi^c$ then the trend growth rate of the real price of housing would be positive. The intuition behind this result is that the supply of housing would be insufficient to meet the demand.
This result might be considered as a special case of the so-called *Baumol's disease* (Baumol, 1967). Baumol’s disease predicts that, in the long-run, the prices of low-productivity goods and services (like education or housing) should rise relative to the prices of manufactured goods (like automobiles or telephone calls). In the same line, the multisector model of growth of Ngai and Pissarides (2007) states the rate of change of the relative price of good \( i \) to good \( j \) is equal to the difference between the Total Factor Productivity growth rates of sector \( j \) and sector \( i \). In consequence, provided that productivity in the construction sector grows slower than in the manufacturing one, which seems *a priori* quite a plausible assumption, real housing prices are expected to grow in the long run. Historical data for OECD countries, presented in Figure 1, show how the average growth rate has been above 2% for the period considered.

### 2.2 Households

Next, we introduce households in the model. Specifically, a representative household supplies homogeneous labour and rents capital to the two productive sectors in the economy. The representative household derives utility each period from per-capita consumption \( C_t \), from per-capita housing owned \( H_t \), and from leisure. The amount of per-household member labour supplied plus leisure cannot exceed the period endowment of time, which is normalised to 1. Period utility per household member at date \( t \) is assumed to be given by

\[
U(C_t, H_t, (1-N_t)) = \log C_t + \gamma \log H_t + \rho \log (1-N_t) \tag{2.14}
\]

where \( \gamma > 0 \) is the utility weight of housing and \( \rho > 0 \) is that of leisure. The functional form of the instantaneous utility is motivated by theoretical results in Ngai and Pissarides (2004). These authors show that necessary and sufficient conditions for the existence of an aggregate balanced growth path in a multi-sector economy are logarithmic preferences and a non-unit price elasticity of demand.

At date 0, the expected discounted sum of future period utilities for the representative household is given by

\[
E_0 \sum_{t=0}^{\infty} \beta^t \eta U(C_t, H_t, (1-N_t)) \tag{2.15}
\]

where \( \beta < 1 \) is the discount factor.\(^3\) Households receive income from the supply of labour and capital. They also receive income from selling houses. Income is divided between consumption, spending on new capital that will be rented out next period, and spending on new housing that will be occupied next period. The depreciation rate for capital is given by \( \delta_k \) and \( \delta_h \) represents that of housing. Thus, the household’s budget constraint is:

\[
C_t + K_t + P_{i,t} H_t = W_t N_t + \frac{(1-\delta_k)}{\eta} P_{i,t} H_{i,t} + \frac{(1-\delta_h)}{\eta} K_{i,t-1} \tag{2.16}
\]

The representative household chooses state-contingent values for consumption, hours, capital and housing for all \( t \geq 0 \) to maximize expected discounted utility (2.15) subject to a sequence of budget constraints (2.16) and a set of inequality constraints \( C_t, N_t, H_t, K_t \geq 0 \) and \( N_t \leq 1 \). The household takes as given prices, a probability distribution over future possible states, and the initial stocks of capital and housing. The first order conditions are the following:

\[
U_{i,j} = \beta E_t \left[U_{1,t+i} \left(1 + r_{i+1}^k - \delta_k \right) \right] \tag{2.17}
\]

where \( U_{i,j} \) is the partial derivative of the instantaneous utility function \( U(t) \) with respect to its first argument, that is, consumption. Equation (2.17) is the standard Euler condition for capital accumulation. The first order with respect to housing accumulation is given by

\[
P_{i}^h U_{i,j} = U_{2,j} + \beta E_t \left[U_{1,t+i+1} P_{i,t+1}^h \left(1-\delta_k \right) \right] \tag{2.18}
\]

The intuition behind this equation is the following: investing in one additional unit of housing has a cost equal to the price of housing, \( P_{i}^h \), times the number of consumption goods that cannot be consumed. On the other hand, investing in one unit of housing yields a direct utility to the household in the period the investment is made. The next period, the household can sell a portion \( (1-\delta_k) \) of housing at a price \( P_{i+1}^h \) and buying with this consumption goods, that would report each a utility of \( U_{i+1,t+i+1} \). The optimal amount of housing is the one that equalises the costs and benefits. Finally, the supply of labour is governed by the condition that the real wage rate is equal to the marginal rate of substitution between leisure and consumption, that is,

\[
W_t = \frac{U_{2,t}}{U_{1,t}} \tag{2.19}
\]

### 2.3 Equilibrium

Equality between supply and demand of production inputs leads to equilibrium in these two markets, that is,

\[
N_t^k + N_t^c = N_t \tag{2.20}
\]

and

\[
K_t^h + K_t^c = K_{t-1} \tag{2.21}
\]

\(^3\) Notice that the flow of utility that households receive from occupying housing they own will constitute an implicit rent that is untaxed.
Moreover, in equilibrium, the total stock of housing that may be enjoyed by households evolves according to

\[ H_t = Y_t^h + \left( \frac{1 - \delta_h}{\eta} \right) H_{t-1} \quad (2.22) \]

Next, equilibrium in the final goods market is

\[ Y_t^c = C_t + K_t + \left( \frac{1 - \delta_k}{\eta} \right) K_{t-1} \quad (2.23) \]

Finally, gross domestic product in this economy is given by

\[ Y_t = Y_t^h + Y_t^c \quad (2.24) \]

### 2.4 Qualitative Analysis

Once we have presented the model and stated the equilibrium of the economy, the natural next step would be to provide numerical values to the parameters and proceed to solve and simulate the model. However, given that the empirical analysis offered in Section 5 is based on a single equation approach, it might be interesting to first present some partial equilibrium results. To that end, we shall focus on equation (2.18), which relates the real price of housing to fundamental variables in the economy and which for convenience is restated here:

\[ P_t U_{1,t} = U_{2,t} + \beta E_t \left[ U_{1,t+1} P_{t+1}^h (1 - \delta_h) \right] \]

Next, given the assumption that the instantaneous utility adopts a logarithmic form, we have that

\[ P_t^h = \gamma \frac{C_t}{H_t} + \beta E_t \left[ \frac{C_t}{C_{t+1}} P_{t+1}^h (1 - \delta_h) \right] \quad (2.25) \]

The next step is to log-linearise (2.25) around a steady state. Given that there is growth in the model due to the non-stationary technology factors, we have to de-trend the variables. Hence, we define the de-trended real price of housing as

\[ p_t^h = P_t^h Z_t^h / Z_t^c, \quad \text{where} \quad Z_t^c = \left( A_t \right)^{\frac{1}{1 - \delta_k}} \]  

and

\[ Z_t^h = \left( A_t^h \right)^{\frac{1}{1 - \delta_h}} \]; similarly, de-trended consumption

\[ c_t = C_t / Z_t^c \]  

and the de-trended stock of houses

\[ h_t = H_t / Z_t^h \]. Equation (2.25) thus becomes:

\[ p_t^h = \gamma \frac{C_t}{H_t} + \beta E_t \left[ \frac{c_t}{c_{t+1}} p_{t+1}^h (1 - \delta_h) \right] \quad (2.26) \]

where, \( z_{t+1}^h = Z_t^h / Z_t^h \) which by the definition of \( Z_t^h \) given in (2.3) is equal to \( z_{t+1}^h = \left[ \exp \left( \Phi + \epsilon_{t+1}^h \right) \right]^{1/(1-\delta_h)} \). Next, we denote a variable in log deviation from its steady state value with a “^\cdot” and proceed to log-linearise (2.26), leading to

\[ p_t^h = \gamma \frac{c_t^\cdot}{c_t^\cdot} - \gamma \frac{h_t^\cdot}{h_t^\cdot} + (1 - \delta_h) \beta E_t \left[ p_{t+1}^h \right] + \frac{1 - \delta_h}{\beta} E_t \left[ c_t^\cdot - c_t^\cdot \right] \quad (2.27) \]

In order to gain some intuition on equation (2.27), notice that in this economy we can price any asset using the Stochastic Discount Factor, that is, the Euler equation associated with the Lagrange multiplier corresponding to the Household’s budget constraint. It is thus straightforward to show that

\[ r_t = E_t \left( c_{t+1}^\cdot - c_t^\cdot \right) \quad (2.28) \]

where \( r_t \) is the net return on a risk-less one-period discount bond. Hence, after combining (2.28) with (2.27) we arrive to

\[ p_t^h = \gamma c_t^\cdot - \gamma h_t^\cdot + \frac{(1 - \delta_h)}{\beta} E_t \left[ p_{t+1}^h \right] + (1 - \delta_h) \beta E_t \left[ c_t^\cdot - c_t^\cdot \right] \quad (2.29) \]

This expression shows that, ceteris paribus, the real de-trended price of housing depends positively on consumption and on its expected future value; whereas there is a negative relationship with respect to the stock of houses and with the real interest rate. It would be interesting to analyse the role of expectations in the dynamic of the real price of housing. To that end, we keep with the partial equilibrium analysis and focus the attention on Equation (2.29), which we reformulate as

\[ p_t^h = (1 - \delta_h) \beta E_t \left[ p_{t+1}^h \right] + \Gamma u_t \quad (2.30) \]

where \( u_t \) is a vector of fundamental variables. For simplicity, we assume that the fundamental variables follow a non-persistent exogenous process. Next, we re-write (2.30) as a system of first order matrix difference equations in the two endogenous variables \( p_t^h \) and \( E_t \left[ p_{t+1}^h \right] \):

\[ \begin{bmatrix} 1 & -\beta(1-\delta_h) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{p}_t^h \\ E_t \left[ \hat{p}_{t+1}^h \right] \end{bmatrix} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{p}_t^h \\ E_t \left[ \hat{p}_{t+1}^h \right] \end{bmatrix} + \begin{bmatrix} \Gamma u_t \\ \omega_t \end{bmatrix} \quad (2.31) \]

where \( \omega_t \) is a non-fundamental term related to expectations errors and is defined as the difference between the observed price in period \( t \) and the expected price in \( t-1 \). Pre-multiplying (2.31) by the inverse of the left-hand side matrix, we obtain:
where \( a = \beta (1 - \delta_h) \). The dynamics of the system would depend on the roots of the right hand side matrix. The characteristic polynomial is

\[
\begin{bmatrix}
\hat{\hat{p}}_t^h \\
E(\hat{\hat{p}}_{t+1}^h)
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
0 & 1/a
\end{bmatrix} \begin{bmatrix}
\hat{\hat{p}}_t^h \\
E(\hat{\hat{p}}_{t+1}^h)
\end{bmatrix} + \begin{bmatrix}
0 \\
-1/a
\end{bmatrix} u_t + \begin{bmatrix}
1 \\
1/a
\end{bmatrix} \omega_t \tag{2.32}
\]

In this case, there are two roots which we call \( \theta \) and \( \lambda \). One of this roots, let’s say \( \theta \), is equal to zero and the other one, namely \( \lambda \) is equal to \( 1/a \). Notice that if the equilibrium is unique, there must be one un-stable root that allows one to pin-down the non-predetermined variable \( E_t(\hat{\hat{p}}_{t+1}^h) \) as a function of the lagged state variable \( \hat{\hat{p}}_{t-1}^h \) and the fundamental variable \( u_t \). For this condition to be satisfied \( |\lambda| > 1 \), that is, \( |1/\beta (1 - \delta_h)| > 1 \) which is satisfied in this model since we have assumed that \( \beta > 1 \) and \( 0 \leq \delta_h \leq 1 \). Accordingly, the model satisfies the Blanchard and Kahn (1980) conditions. Finally, applying the methods developed by Beyer and Farmer (2006), it is possible to show that the solution to (2.31) is

\[
\hat{\hat{p}}_t = \Gamma u_t \tag{2.34}
\]

and

\[
E_t(\hat{\hat{p}}_{t+1}^h) = 0 \tag{2.35}
\]

Hence, the real de-trended price of housing is a linear function of the fundamentals. In this analysis we have implicitly assumed for simplicity that the fundamentals are not persistent. Hence, the economy is expected to return to its steady state immediately after a shock. That is why the expectation in (2.35) is equal to zero. Notice that the analysis so far has been of a partial character. In the case of full system dynamics, there might be combinations of the parameters such that the solution to the model would be indeterminate and thus non-fundamental or expectations-driven dynamics could not be precluded. We left the analysis of possible “sunspot” dynamics for future research.
3. Stylized Facts about the Spanish Real Estate Market

As commented in the introduction, there are diverse reasons to choose Spain to conduct empirical work about the role of economic fundamentals on the price of housing. The first reason is the small relative size of its rental market (Ruiz and San Martin 2004), as it can be seen in Figure 2. As less than 80% of the houses are rented, most of the agents are assumed to face decisions similar to those described in Section 2 (either invest in house or in real assets).

The second reason is the profusion of regional statistics, allowing the construction of a panel for the period 1995-2006 for the 50 Spanish provinces. We have constructed panel data for the Spanish real house prices and its fundamentals for the 50 Spanish provinces. The real house price has been obtained as the nominal house price deflated by the Consumer Price Index CPI of each province. The nominal house price is the average price per square meter of houses in the market4, as reported by the Spanish Ministry of Housing, the CPI has been obtained from the Spanish National Statistics Institute/Instituto Nacional de Estadística (INE). Data on the stock of houses has been constructed by combining the number of new houses in the market provided by the Ministry of Housing and the stock of houses in 2001 from the INE5. Regional GDP (also deflated by the CPI) and population figures include the effect of immigration (quite important in Spain since the late 1990s) and also come from the INE. Nominal interest rates are the average rates for new mortgage credits as reported by the Spanish Central Bank. Real interest rates are constructed with the nominal rates and the CPI. A list of the variables under consideration is presented in Table 1.

Thirdly, Spain is one of the countries where the surge in housing prices has been more spectacular in the period under consideration: the total increment in the nominal housing price in the period 1995:2006 was of 187% and in the real one was 106%. At the same time, the Spanish economy has experienced a complete overhaul in this decade, which makes it an interesting case to analyse. The massive arrival of immigrants has spurred the economy, increasing the population in more than 4 million people (around a 10% of the initial population). The Spanish economy has been able to absorb this entire workforce by promoting labor-based growth, as presented in Figure 3. In Figure 4 it can be seen how participation rates (especially among women) have gone up while unemployment was significantly reduced, which has raised average household incomes. Additionally, the entrance of Spain in the Euro Area in 1999 has helped to reduce both nominal and real interest rates (which have even reached negative levels as presented in Figure 5), making the access to credit cheaper and promoting the investment in real assets such as housing. During this period, investment in housing has boosted the amount of wealth in real-estate, as shown in Figure 6. The reduction in interest rates, the growth of adult population and the increase in household incomes are the fundamental economic variables that should drive the price of housing, according to the theoretical model presented above, so the aim of the next section is to check if this model correctly explains the evolution in the price of housing.

Maza and Villaverde (2007) have analyzed the shocks affecting the Spanish regions between 1975 and 2005. Their most relevant conclusion is that, during this period of ever increasing globalization, the Spanish regions have been mainly affected by symmetric shocks. As a result, one should expect a great deal of cross-section dependence in the data. At the same time, during the period 1995-2006, Spanish regions have experienced a process of convergence, as shown in Figure 7, which has induced more growth in the initially ‘poorer’ regions. These features have important implications on the estimation method, as commented next.

4 We distinguish between houses in the market and public houses, the latter provided by the State at prices well below their market value. When considering the stock of houses we consider both types, as they both provide similar housing services.

5 The correct magnitude should have been the number of square meters (as this is the item being priced). We are implicitly assuming that new houses have the same number of square meters than previous ones. This could be wrong due to demographical issues, such as the reduced number of family members in the new generations, but it may be considered as a first-order approximation. An additional problem is that we are not accounting for depreciation of the houses.
4. Empirical Analysis based on Panel Cointegration

4.1 Panel Unit Root Tests and Cross-Dependence

As commented in the introduction, we have decided to apply an array of panel cointegration models to validate our theoretical conclusions, in an effort to exploit the maximum amount of information contained in the regional data of Spanish provinces. Panel cointegration methods apply cointegration techniques to find long-run relationships in panel data with a substantial number of cross-section units\(^6\). They capture heterogeneity due to province-specific characteristics by considering that the short-run dynamics may differ across regions due to institutional characteristics, whereas the long-run relationship is the same for all. The general model is as follows: consider the \( m \) time series \( z_{ijt} = (z_{i1t}, z_{i2t}, \ldots, z_{imt}) \) where \( z_{ijt} = p_i \) is the housing price and \( x_{ijt} = (z_{i1t}, z_{i2t}, \ldots, z_{imt}) \) is a vector of fundamentals observed on the \( j \)th region \( i = 1, 2, \ldots, N \), over the period \( t = 1, 2, \ldots, T \). Suppose that for each \( i \)

\[
z_{jt} \sim I(1), \quad j = 1, 2, \ldots, m \tag{4.1}
\]

Then \( z_{jt} \) is said to form at least one cointegration relation if there are linear combinations of \( z_{jt} \) for \( j = 1, 2, \ldots, m \) that are \( I(0) \) i.e. if there exists an \( m \times r \) matrix \( \beta \) such that

\[
\begin{bmatrix}
\beta' & z_{jt} & \epsilon_{jt}
\end{bmatrix} \sim I(0) \tag{4.2}
\]

First of all, it is convenient to test whether the price and fundamentals display significant cross section dependence, as should be expected according to the previous section. Following Holly, Pesaran and Yamagata (2006), we compute a test of error cross section dependence (CD) developed by Pesaran (2006a) that is applicable to short \( T \) and large \( N \) panels. These CD test estimates of the \( p \)th-order Augmented Dickey-Fuller test statistics for \( p_i, y_i, \) and \( r_i \), reported in Table 2, clearly show that the cross correlations are statistically significant. Cross-section dependence can arise due to a variety of factors, such as omitted observed common factors, spatial spill over effects, unobserved common factors, or general residual interdependence that could remain even when all the observed and unobserved common effects are taken into consideration.

The next step is to check whether the variables are \( I(1) \). Due to the existence of error cross section dependence, it is necessary to apply second generation tests of integration, such as the CIPS one proposed by Pesaran (2006b)\(^8\). The CIPS test results, summarized in Table 3, show that for \( p_i, y_i, \) and \( h_i \), the unit root hypothesis cannot be rejected if the trended nature of these variables is taken into account. The problem arises as the test is not able to reject the unit root hypothesis either for the real interest rates or for the first difference of the variables. A potential explanation for this is the reduced length of the time period, which is inferior to a whole economic cycle of the Spanish economy. Notwithstanding, due to the economic considerations commented in Section 2, we assume \( p_i, y_i, \) and \( h_i \) to be \( I(1) \) and \( r_i \) to be \( I(0) \).

4.2 Residual-based Approaches

Residual-based approaches are estimation techniques appropriate when \( r_i = 1 \), and \( z_i \) can be partitioned such that \( z_i = (p_i, x_i') \) with no cointegration among the \( m - 1 \) fundamental variables of \( x_i' \). Residual-based approaches usually consider the following regression:

\[
p_i = \delta'd_i + x'_i \beta + u_i, \quad i = 1, 2, \ldots, N \tag{4.3}
\]

where \( \delta'd_i \) represent the deterministic trend. The innovations in \( \Delta x_i \) denoted by \( \epsilon_i = \Delta x_i - \hat{E} (\Delta x_i) \), are allowed to be cointegrated with \( u_i \). It is assumed that the vector of coefficients \( \beta \) is the same for all regions (homogeneous cointegration relationship). Applying a sequential limit theory it can be shown that the OLS estimator of \( \beta \) is \( T^{1/2}N \) consistent and, therefore, the time series dimension is more informative than the cross-section one on the long-run coefficients. However, the OLS estimator is inefficient in the model with endogenous regressors.

To obtain an asymptotically efficient estimator, Pedroni (1995) and Philips and Moon (1999) proposed the FMOLS approach that adjusts for the effects of endogeneity and short-run dynamics of the errors (Philips and Hansen 1990). An alternative is the Dynamic OLS (DOLS) of Kao and Chiang (2000) (which performs somewhat better than FMOLS). The problem with both estimators (FMOLS and DOLS) is that they may be severely biased in data with reduced number of time samples (short- \( T \) panels). Another shortcoming of these methods is that they do not fully capture contemporaneous correlation among cross section units as the one presented in the data considered (Breitung, 2005).

\(^6\) A review of the panel counter part of the classical literature on cointegration techniques developed by Engle and Granger (1987), Johansen (1995) and Philips (1991) has been recently presented in Breitung and Pesaran (2007).

\(^7\) Due to the reduce number of time periods, we should reduce the analysis to \( p = 2 \).

\(^8\) Due to the reduce number of time periods, we should reduce the analysis to \( p = 1 \).
DOLS Estimator

The DOLS estimator discomposes the error term as

\[ u_t = \sum_{k=-q}^{q} y_{ik} \Delta x_{i,t+k} + v_{it}, \tag{4.4} \]

where \( y_{ik} \) is orthogonal to all leads and lags of \( \Delta x_{it} \). In practice the infinite sums are truncated at some small number of \( k \) denoted as \( q \). From (4.3) and (4.4) we obtain a general expression (assuming no time trend and allowing \( x_{i1} = 1 \forall i \))

\[ y_{it} = \beta' x_{it} + \sum_{k=-q}^{q} y_{ik} \Delta x_{i,t+k} + v_{it}. \tag{4.5} \]

Kao and Chiang (2000) show that in the homogeneous case the FMOLS and the DOLS have the same limiting distribution. For the problem under consideration, our aim is to test if there is a cointegration relationship between the price \( p_{it} \) and its fundamentals \( y_{it}, h_{it}, \) and \( r_{it} \). So we estimate by DOLS a model derived from Equation (2.34):

\[ p_{it} = \alpha_i + \beta_i y_{it} + \gamma_i h_{it} + \delta_i r_{it} + u_{it}, \quad i=1,2,\ldots,N; \quad t=1,2,\ldots,T \tag{4.6} \]

where the error term \( u_{it} \) follows a truncated version of with \( q = 1 \) due to sample size limitations\(^{10}\).

Results presented in Table 4 show a coefficient on income of 1.44, which is consistent with previous literature where it approximately ranges from 0.3 to 3 according to the review of Girouard et al. (2006), under different definitions, data and econometric methods. Holly, Pesaran and Yamagata (2006) propose a value of 1. This would mean that housing is a luxury good, as its income-elasticity is greater than one. The coefficient on stock of housing per capita is -1.38 (Girouard et al. 2006): the elasticity relatively to housing stock ranges from -0.5 to -8). The coefficient on real interest rates is -8.43 (Girouard et al. 2006): from -0.1 to -9.4). This value is relatively high in comparison to previous studies. All variables are significant\(^{11}\) above the 1 per cent level.

The CD statistic demonstrates that there is a high degree of cross-dependence: as it was mentioned above, this method does not explicitly account for cross-section correlation. Nevertheless, the CIPS test rejects the null hypothesis of unit root below the 1 per cent; so Equation (4.6) seems to be a valid cointegration relation to describe the real price of housing.

\[^6\text{Income per capita (GDP per capita) is employed as a proxy for consumption due to the lack of data about regional consumption.}\]

\[^7\text{This is not as short as it may seem: for example Stock and Watson (1993) chose } q \text{ equal to 2 for the period 1900-1989, which is significantly longer than 1995-2005.}\]

\[^8\text{It is important to mention that the standard errors presented have been re-scaled by the method presented in Stock and Watson (1993) so the } t \text{ statistic tends asymptotically to a } N(0,1) \text{ distribution.}\]

Short-term Dynamics

Having established by DOLS a panel cointegration relation between the price and its fundamentals, we may turn our attention to the dynamics of the adjustment and estimate the panel error correction model:

\[ \Delta p_{it} = \alpha_i + \phi_i (u_{it-1}) + \theta_i W \Delta p_{i,-1} + \delta \Delta x_{it} + v_{it}, \tag{4.7} \]

\[ \Delta x_t = \begin{bmatrix} \Delta y_{it} \\ \Delta h_{it} \\ \Delta r_{it} \end{bmatrix} \tag{4.8} \]

The coefficient \( \phi_i \) provides a measure of the speed of adjustment of house prices to a shock. The half life of a shock to \( p_{it} \) is approximately \(-\ln(2)/\ln(1+\phi_i)\). \( u_{it} \) is the lagged residual from Equation (4.6) and \( v_{it} \) is the error term. \( W \) is a \( 1 \times N \) vector of weights \( w_i \) so that

\[ w_i = \frac{GDP_{1205}/Pop_{1205}}{\sum_{i=1}^{N} GDP_{1205}/Pop_{1205}} \tag{4.9} \]

and \( p_t = [P_{it}, P_{it}, \ldots, P_{it}]' \) is the vector of real prices. The term \( \theta_i W \Delta p_{i,-1} \) tries to capture the possibility of spill-over effects of the price, as if the price of the houses rises in the richest per capita regions, it may be logical to assume that some of their residents will purchase houses in cheaper regions, both as a capital investment and as a second residence. Equation (4.7) is estimated by OLS regressions separately for each provincia.

Results are presented in Table 5. We display the mean value of the coefficients for the 50 provinces. It is interesting to note how the mean value of coefficient for the residual \( \phi_i \) takes the value -0.2, so the half-life of shocks is around 3.1 years. Signs are not surprising for the short-term responses to the fundamentals. However, the mean response of the price of a province to a general increase in prices is positive (0.4), possibly reflecting a persistence component due to spill-over effects between provinces. Test statistics still show a high degree of cross-dependence. The null hypothesis of unit root can be rejected with significance below the 1 per cent.

CCE Estimator

In an attempt to overcome the problem of cross-section dependence, Pesaran (2006a) proposed the Common Correlated Effects (CCE) estimator, conceived to work in panel data models with a multifactor error structure where the unobserved common factors are correlated with exogenously given individual-specific regressors. The
basic model follows Equation (4.3) where errors $u_i$ have the multifactor structure

$$u_i = \gamma^T f_i + \eta_i$$  \hspace{1cm} (4.10)

In which $f_i$ is the $m \times 1$ vector of unobserved common effects and $\eta_i$ are the region-specific errors assumed to be independently distributed of $(d, x_i)$. A further extension of this model includes the potential correlation of $(d, x_i)$ and $f_i$ by considering the general model for the individual specific regressors

$$x_i = a_i + \Gamma_i f_i + v_i$$  \hspace{1cm} (4.11)

where $a_i$ is the $k \times 1$ vector of individual effects, $\Gamma_i$ is a $k \times m$ factor loading matrix with fixed components and $v_i$ are the specific components of $x_i$, distributed independently of the common effects and across $i$, but assumed to follow general covariance stationary processes. Under this multifactor model, Pesaran (2006a) demonstrate that it is possible to obtain CCE estimators of the cointegration vector for both the homogeneous and heterogeneous cases that are consistent regardless of whether the common factors $f_i$ are stationary or nonstationary.

We have estimated by CCE Equation (4.6) assuming that the errors follow (4.10). Results are shown in Table 6. They totally lack significance, as most elasticities are close to zero. In consequence the estimator fails to find a successful cointegration relationship. A possible reason of this is the short $T$ dimension of the sample: to be able to estimate 3 cointegration coefficients, the estimator generates a total of 8 coefficients per cross-section unit that should be estimated with 11 samples.

### 4.3 A Panel Vector Error Correction Model

As discussed above, although DOLS and FMOLS approaches are an elegant way to estimate non-stationary panel data models, they may be problematic especially in fairly small samples. In particular, the FMOLS estimator may be severely biased in empirically relevant sample sizes (Pedroni 2000). Another problem with these methods is that they are based on a single equation approach. Consequently, feed-back effects cannot be modelled in this set-up. For all these reasons, a parametric approach may be a promising alternative, in particular, for panels with a small number of time periods. Pesaran, Shin and Smith (1999) have suggested estimation procedures for cointegrated panel data based on a vector error correction (VECM) format. However, these methods are based on a ML estimator and may have problematic small–sample properties as well as convergence issues.

Following Breintug (2005), we apply a simple asymptotically efficient two-step estimation procedure. The individual specific parameters are estimated in the first step, whereas, in a second step, the common long-run parameters are estimated from a pooled regression. The resulting estimator is asymptotically efficient and normally distributed.12 Moreover, since the second step of the parametric approach is based on an ordinary least-squares regression, it is straightforward to account for possible contemporaneous correlation among the errors.

We consider a cointegrated VAR(1) model with individual short-run dynamics and deterministic terms. As usual in the panel cointegration framework, we assume that the mean (or trend) and the short-run dynamics may differ across provinces, whereas the long-run relationship is the same for all provinces.13 The model takes the following general form:

$$\Delta z_{it} = \Psi d_{it} + \alpha_i \beta z_{i,t-1} + \Gamma_i \Delta z_{i,t-1} + u_{it}$$  \hspace{1cm} (4.12)

Where $z_{it}$ is the $m \times 1$ vector of endogenous variables that include the real price of housing14, per-capita GDP, number of houses per capita and the real interest rate presented in Equation (4.1); $d_{it}$ is a vector of deterministic variables (a constant in this case) and $\Psi_i$ is a $m \times j$ matrix of unknown coefficients. $u_{it}$ is an $m$-dimensional white noise error vector with $E(u_{it}) = 0$ and positive definite covariance matrix $\Sigma_i = E(u_{it}u_{jt})$. The term $\beta \Delta z_{i,t-1}$ captures long-run relationships amongst the variables in the model. In this specification, the $m \times r$ $(0 < r < m)$ cointegration matrix $\beta$ is the same for all cross section units, whereas the $m \times r$ loading matrix $\alpha_i$ and the error covariance matrices $\Sigma_i$ are allowed to vary across $i$.

Following Breintug (2005) we implement a two-step estimation procedure. Since the information matrix of the Gaussian likelihood is asymptotically block diagonal with respect to the “short-run parameters” $(\alpha_i, \Sigma_i)$ and the matrix of cointegration vectors $\beta$, the latter can be estimated efficiently based on some consistent initial estimator of $\alpha_i$ and $\Sigma_i$ $(i = 1, \ldots, N)$. Hence, in the first step, we compute a consistent estimator (as $T \rightarrow \infty$) of $\alpha_i$ from estimating separate models for all $N$ cross section units. In our empirical analysis, we restrict ourselves to just one cointegration relationship and, thus, use the two-step estimator suggested by Engle and Granger (1987).15

---

11 The results of some Monte Carlo experiments suggest that the two-step estimator performs better than the FMOLS and DOLS estimator in typical sample sizes.

12 The reason for assuming a homogenous long-run relationship is that the underlying economic principles that are employed to establish the long-run equilibrium should apply similarly in all provinces, whereas the adjustment process towards the long-run equilibrium may differ due to behavioural and institutional characteristics.

13 To ease visualization, in this sub-section we do not employ bold letters to denote vectors.

14 The ML estimator of Johansen (1991, 1995) could also be used.
At the first estimation stage, the restriction that the cointegration vectors are the same for all cross section units is ignored, but this does not affect the asymptotic properties of the final estimator of $\beta$.

At the second stage, the system is transformed such that the cointegration matrix $\beta$ can be estimated by the ordinary least-squares of the pooled regression

$$\hat{\Delta} z_{it} = \hat{\beta} \cdot z_{i,t-1} + \hat{\gamma}_i + \hat{\delta}_t \quad i = 1, \ldots, N \text{ and } t = 1, \ldots, T,$$

where $\hat{\Delta} z_{it} = (\hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i) \Delta \hat{z}_{it}$ where $\hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i$, and $\hat{\delta}_t$ denote the residual vectors from the least squares regression of $\Delta z_{it}$ on $\Delta z_{i,t-1}, \ldots, \Delta z_{i,T-1}, \Delta z_{i,T-1}$ and $d_i$. Finally, $\hat{\gamma}_i$ is defined as $\hat{\gamma}_i = (\hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i) \Delta \hat{z}_{i,t-1}$. Based on a sequential limit theory, shows that the two-step estimator has a normal limiting distribution. From this result, it follows that the long-run parameters are asymptotically normally distributed and, therefore, inference on the cointegration parameters involves standard limiting distributions. To estimate the covariance matrix of $\hat{\beta}$, the finite sample moments of $\hat{\Delta} z_{it}$ can be used and, therefore, the ordinary $t$-statistics of the elements of $\hat{\beta}$ have a standard normal limiting distribution.

In the present analysis we decided to keep the real interest rate as an exogenous explanatory variable in the Panel-VECM. Given that this latter variable is formed by a common nominal interest rate minus the province specific inflation rate, we consider it as weakly exogenous. Given that the number of endogenous variables is 3, there might be at most 2 cointegration relationships. Even thought it is possible to test the exact dimension of the cointegration space (see Theorem 2 in Breitung, 2005), we assume it is of dimension 1. Next, and taking into consideration this restriction, we normalise the cointegration vector such that the coefficient associated with the real price of houses is fixed to one.

We apply the two-step estimator suggested by Engle and Granger (1987) to each province. Table 7 shows the estimated loading factors, that is, the vector of parameters $\alpha_i$. For concreteness, we just show the elements corresponding to the house price equation. The parameters are thus an indicator of the speed of adjustment of house price to the corresponding “equilibrium” values. To facilitate the comparison, we show the ratio of each loading factor with respect to the mean of the provinces. The results show that there is not a clear pattern. Some provinces, such as Baleares, Barcelona, Malaga and Navarra, appear to be very dynamic, whereas Albacete, León, Gerona, Valladolid are relatively static.

Next, we proceed with the second step of the estimation procedure and compute the long-run coefficients. Table 8 shows the estimates of the common coefficients with the associated $t$-statistics. The numbers can be regarded as long-run elasticity. The estimated coefficient of real GDP is positive and significant. In particular, it is 1.74 which is a value in line with the estimates of Holly et al (2006) for the U.S. cities and relatively similar to the one obtained by DOLS. Regarding the elasticity of the number of houses per capita -1.72, the sign is negative but the level of significance is not very high, although the value is similar to the -1.38 obtained by DOLS. Finally, the price of houses appears to respond negatively to the real interest rate and very strongly. The estimated coefficient is -3.50 and the level of significance is high. This value is more consistent with the literature than the -7.75 obtained by DOLS.

The analysis so far has assumed that there is one co-integration relationship between the variables in the model. A full assessment of this assumption would requires the application of properly constructed tests, a task that is currently out of the scope of the present paper. Nevertheless, we might gain some confidence in the results by analysing the stationary properties of the residuals of the long-run equation. To that end, we first study the correlation properties of the residuals. The cross-section dependence test yields a value of -2.05. Given that the CD-test is asymptotically normally distributed, we might reject the hypothesis of no cross-section dependence. Hence we should apply the CIPS panel unit-root test. The estimated CIPS is -2.5, whereas the truncated CIPS is -2.2. These numbers indicate that we can confidently reject the null of I(1) and thus a stationary relationship seems to exist among house prices, real GDP, the number of houses and the real interest rate. This result tends to reinforce the conclusion that the data seem to support the existence of a log-term relationship between the housing price and its fundamental economic variables.

An interesting exercise is to compute the contributions, in the short-run dynamic equation, of the different variables in the model. In this regard, Figures 8 and 9 show the contributions to the growth in the real price of housing for Madrid and Barcelona respectively. In these cases, the disequilibrium or error-correction term has contributed positively to the increase in the price of housing between the years 1997 and 2003. In the last part of the sample, the contribution has been less pronounced, even negative. The results for all the provinces show the important role of GDP growth as a source to explain the observed pattern of house prices in the period analysed. It is also noticeable the contribution
of real interest rates, particularly in the last year of the sample. The small contribution of the stock of houses might reflect that the production of houses has been relatively elastic to its demand during this period. In consequence, part of the increase in the demand of housing would have been satisfied by producing more houses, thus reducing the pressure on the prices.

Finally, Figure 10 shows an indicator of the relative disequilibrium of the price of housing for each Spanish province corresponding to the year 2006. The indicator shows how large (or small) is the gap of each province with respect to the gap of the Spanish economy. One can observe that those provinces in the Mediterranean Coast, as well as those around large metropolitan areas, such as Madrid, turn out to feature large positive disequilibrium, that is, the observe price is well above the fundamental. These results should, nevertheless be taken with caution. For instance, the role of second residences, which is clearly an important issue in some provinces, has to be taken into account. The results are robust to this feature of the Spanish housing market16.

---

16 To check that we have re-estimated the model for a subsample of provinces where the effect of second residences is small (mainly inner provinces). Resulting coefficients change by less than 15%. 
5. Concluding Remarks

This paper tries to shed some light in the determinants of the price of housing. Its first contribution is to develop a general equilibrium model of the economy where agents should choose among investing in capital goods, consuming or investing in real estate. This framework indicates that, provided that the productivity growth in the construction sector is smaller than in the manufacturing one, the price of housing relative to the price of composite good will rise in the long term. This feature seems to be supported by the empirical evidence of the last 30 years in OECD countries.

A second feature of the model is that it allows defining the deviations of the price from its steady-state as a linear function of the deviations of its economic fundamentals: consumption, stock of houses per capita and real interest rates. An additional term is the expectation about the future price, which may produce “sunspot-equilibria”. However under some constraints we have shown that the expectations might be consistent with the economic fundamentals, so the price is completely determined by economic variables.

To test this proposition we have decided to apply panel cointegration methods to a panel of data for the 50 Spanish regions in the period 1995:2006. The reasons to choose Spain has been the small size of its rental market (as assumed in the theoretical framework), the availability of regional data and the considerable changes that the main variables has experiences (price, income, population, interest rates), which make optimal the use of panel methods for non-stationary series. We have decided to apply a whole array of methods in order to give more consistency to our conclusions.

Single-equation methods such as FMOLS or DOLS have the problem that they do not cope well with cross-section dependence, a feature of the problem under consideration due to the existence of symmetrical regional shocks. Nevertheless, our results applying DOLS strongly support the existence of a cointegration relationship between the real price and the economic fundamentals. This rules out the so-called “bubble-theory” in Spanish real market by displaying that the increase in price has been mainly a consequence of the surge in family income and the reduction of interest rates as a consequence of the entry of Spain in the Euro-area. CCE methods fail to find any relationship, probably due to the short temporal length of the series.

The application of a vector model VECM ratifies the results obtained with single-equation approaches. A cointegration relation is also found with similar coefficients, although it reduces the impact of interest rates compared to the DOLS case. Both methods indicate that housing is a luxury good, with elasticity between 1.4 and 1.8. A short term analysis shows that, despite the fact that the price is fully justified by fundamentals, the current price of housing in Spain is above its long-term trend, which may justify a correction in the close future. Both estimation methods underline the heterogeneity of the regional responses and the existence of some persistence in the determination of the price, i.e. in many provinces the price tend to rise in the long-term as a consequence of a general rise in the price level.

Notwithstanding rigour, the analysis has several flaws that should be commented. The most important is the short temporal length of the series. Data does not cover a full price cycle, which might bias results. Another problem is the existence of second residences; a feature does not taken into account by the theory. Second residences distort the relationship between regional incomes, population and housing demand. In the case of Spain, Mediterranean coast is an important holiday zone for many Europeans and Spaniards. In order to make a simple check about the influence of this, we have re-run the estimations for a subset of provinces that exclude the coast without seeing big changes in the results.

This paper should begin a more structured line of analysis of the real estate market that goes beyond the basic theoretical model à la Campbell and Shiller. Thanks to the advances in general equilibrium macro-models and the new econometric tools for panel developed in the last decade, we hope that this line of research will be extended in the near future to capture more features (such as rental markets) and more countries.
References


IMF (2004), World Economic Outlook, pp. 71-89, September.


15
Appendix

Table 1: List of Variables and their Description

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{rt}$</td>
<td>Regional Consumer Price Index ($2005=1$)</td>
</tr>
<tr>
<td>$P_{it}$</td>
<td>Regional housing price per square meter</td>
</tr>
<tr>
<td>$GDP_{it}$</td>
<td>Regional GDP</td>
</tr>
<tr>
<td>$POP_{it}$</td>
<td>Regional population</td>
</tr>
<tr>
<td>$RB_{it}$</td>
<td>Spanish average mortgage interest rates</td>
</tr>
<tr>
<td>$HOU_{it}$</td>
<td>Stock of houses</td>
</tr>
<tr>
<td>$p_{it}$</td>
<td>Natural logarithm of the real house price</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>Natural logarithm of the real per capita GDP</td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>Natural logarithm of houses per capita</td>
</tr>
<tr>
<td>$r_{it}$</td>
<td>Real interest rates</td>
</tr>
</tbody>
</table>

Notes: The estimation method is Panel DOLS with $q=1$. Standard errors are given in parenthesis and have been rescaled by the method presented in Stock and Watson (1993) so the $t$ statistic tends asymptotically to a $t(0,1)$. $R^2$ refers to the adjusted $R$ squared statistic. CD is the cross-dependence test statistic with ADF(1), which tends to $t(0,1)$ under the null hypothesis of no error cross section dependence. CIPS and CIPS* refer respectively to the standard and truncated cross section averages of Cross-sectionally Augmented Dickey-Fuller test statistics CADF(1) with no intercept and no trend. Relevant lower 1% critical values for the CIPS and CIPS* statistic are -1.78 and -1.77 respectively. *** means that the test is significant at the 1 and 10 per cent level.

Table 2: Residual Cross Dependence of ADP(p) Regressions

<table>
<thead>
<tr>
<th>CD Test Statistic</th>
<th>ADF(1)</th>
<th>ADF(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{p}$</td>
<td>7.09</td>
<td>6.75</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>20.95</td>
<td>12.00</td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>15.79</td>
<td>8.60</td>
</tr>
<tr>
<td>$r_{it}$</td>
<td>88.29</td>
<td>80.63</td>
</tr>
</tbody>
</table>

Notes: $p$-order Augmented Dickey-Fuller test statistics, ADF(p) are computed for each cross-section unit separately. For all the variables except $r_{it}$ an intercept and a linear trend are included. $r_{it}$ includes only an intercept. CD - $\sqrt{\sum_{i=1}^{N}(N-1)\sum_{j=1}^{N-1}c_{ij}^2}$ tends to $t(0,1)$ under the null hypothesis of no error cross section dependence. $\rho_{p}$ is the pair-wise correlation of the residuals from the ADF.

Table 3: Pesaran’s CIPS Panel Unit Root Test Results

<table>
<thead>
<tr>
<th></th>
<th>CIPS</th>
<th>CIPS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept and Trend</td>
<td>-2.10</td>
<td>-2.05</td>
</tr>
<tr>
<td>$y_{it}$</td>
<td>-2.12</td>
<td>-1.99</td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>-1.64</td>
<td>-1.64</td>
</tr>
<tr>
<td>Only Intercept</td>
<td>-0.64</td>
<td>-0.64</td>
</tr>
<tr>
<td>$\Delta\rho_{p}$</td>
<td>-1.97</td>
<td>-1.97</td>
</tr>
<tr>
<td>$\Delta y_{it}$</td>
<td>-1.98</td>
<td>-1.83</td>
</tr>
<tr>
<td>$\Delta h_{it}$</td>
<td>-1.39</td>
<td>-1.39</td>
</tr>
<tr>
<td>No Intercept No Trend</td>
<td>-0.73</td>
<td>-0.73</td>
</tr>
<tr>
<td>$\Delta r_{it}$</td>
<td>-2.40**</td>
<td>-2.40**</td>
</tr>
<tr>
<td>$\Delta y_{it}$</td>
<td>-2.89**</td>
<td>-2.63**</td>
</tr>
<tr>
<td>$\Delta h_{it}$</td>
<td>-1.77**</td>
<td>-1.77**</td>
</tr>
</tbody>
</table>

Notes: The reported values are CIPS(1) statistics, which are cross section averages of Cross-sectionally Augmented Dickey-Fuller test statistics CADF(1). CIPS* refers to the truncated version of the test statistic, more suitable for small 7 panels. Relevant lower 5%/10% critical values for the CIPS statistic are -2.86/-2.71 for the intercept and trend case, -2.19/-2.07 for the only intercept case and -1.58/-1.46. Relevant lower 5%/10% critical values for the CIPS* statistic are -2.75/-2.73 for the intercept and trend case, -2.16/-2.05 for the only intercept case and -1.57/-1.46 for the no intercept no trend case. *** means that the test is significant at the 5 and 10 per cent level respectively.

Table 4: Estimation Results: Cointegration Relation Estimated by DOLS

<table>
<thead>
<tr>
<th></th>
<th>Estimated Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{i}$</td>
<td>1.44</td>
</tr>
<tr>
<td>(0.18)</td>
<td></td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>-1.38</td>
</tr>
<tr>
<td>(0.25)</td>
<td></td>
</tr>
<tr>
<td>$r_{it}$</td>
<td>-7.75</td>
</tr>
<tr>
<td>(0.77)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.94</td>
</tr>
<tr>
<td>CD</td>
<td>20.53</td>
</tr>
<tr>
<td>CIPS</td>
<td>-1.98***</td>
</tr>
<tr>
<td>CIPS*</td>
<td>-1.98***</td>
</tr>
</tbody>
</table>

Notes: Estimated model is Equation (4.6). “Without $h_{it}$” refers to the same model without including houses per capita. The estimation method is Panel DOLS with $q=1$. Standard errors are given in parenthesis and have been rescaled by the method presented in Stock and Watson (1993) so the $t$ statistic tends asymptotically to a $t(0,1)$. $R^2$ refers to the adjusted $R$ squared statistic. CD is the cross-dependence test statistic with ADF(1), which tends to $t(0,1)$ under the null hypothesis of no error cross section dependence. CIPS and CIPS* refer respectively to the standard and truncated cross section averages of Cross-sectionally Augmented Dickey-Fuller test statistics CADF(1) with no intercept and no trend. Relevant lower 1% critical values for the CIPS and CIPS* statistic are -1.78 and -1.77 respectively. *** means that the test is significant at the 1 and 10 per cent level.

Table 5: Estimation Results: Error Correction Model (OLS)

<table>
<thead>
<tr>
<th></th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{it}$</td>
<td>-0.20</td>
</tr>
<tr>
<td>$W_{it}p_{it}$</td>
<td>0.40</td>
</tr>
<tr>
<td>$\Delta y_{it}$</td>
<td>0.33</td>
</tr>
<tr>
<td>$\Delta h_{it}$</td>
<td>-0.56</td>
</tr>
<tr>
<td>$\Delta r_{it}$</td>
<td>-1.22</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.73</td>
</tr>
<tr>
<td>CD</td>
<td>16.48</td>
</tr>
<tr>
<td>CIPS</td>
<td>-2.81***</td>
</tr>
</tbody>
</table>

Notes: Estimated model is Equation (4.7). $\mu_{it}$ is the residual Coefficients errors are mean values of the regional ones whereas test statistics are for the whole sample. The estimation method is OLS. $R^2$ refers to the adjusted $R$ squared statistic. CD is the cross-dependence test statistic with ADF(1), which tends to $t(0,1)$ under the null hypothesis of no error cross section dependence. CIPS and CIPS* refer respectively to the standard and truncated cross section averages of Cross-sectionally Augmented Dickey-Fuller test statistics CADF(1) with no intercept and no trend. Relevant lower 1% critical values for the CIPS and CIPS* statistic are -1.78 and -1.77 respectively. *** means that the test is significant at the 1 and 10 per cent level.

Table 6: Estimation Results: Cointegration Relation Estimated by CCE

<table>
<thead>
<tr>
<th></th>
<th>Pooled</th>
<th>CCE Results</th>
<th>Mean Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{i}$</td>
<td>0.45</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>(0.23)</td>
<td>(0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_{it}$</td>
<td>-0.28</td>
<td>-0.71</td>
<td></td>
</tr>
<tr>
<td>(0.61)</td>
<td>(0.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{it}$</td>
<td>-0.41</td>
<td>-0.74</td>
<td></td>
</tr>
<tr>
<td>(0.54)</td>
<td>(0.46)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The estimation methods are Panel Common Correlated Effects, both Pooled and Mean Group. Standard errors are given in parenthesis.
### Table 7: Speed of Adjustment Price of Housing to Disequilibrium

<table>
<thead>
<tr>
<th>Province</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alava</td>
<td>6.00</td>
</tr>
<tr>
<td>Albacete</td>
<td>0.01</td>
</tr>
<tr>
<td>Alicante</td>
<td>0.75</td>
</tr>
<tr>
<td>Almeria</td>
<td>0.23</td>
</tr>
<tr>
<td>Asturias</td>
<td>0.22</td>
</tr>
<tr>
<td>Avila</td>
<td>1.16</td>
</tr>
<tr>
<td>Badajoz</td>
<td>1.53</td>
</tr>
<tr>
<td>Baleares</td>
<td>2.22</td>
</tr>
<tr>
<td>Barcelona</td>
<td>2.31</td>
</tr>
<tr>
<td>Burgos</td>
<td>2.56</td>
</tr>
<tr>
<td>Caceres</td>
<td>0.78</td>
</tr>
<tr>
<td>Cádiz</td>
<td>0.60</td>
</tr>
<tr>
<td>Cantabria</td>
<td>0.87</td>
</tr>
<tr>
<td>Castellon</td>
<td>0.24</td>
</tr>
<tr>
<td>Ciudad Real</td>
<td>0.40</td>
</tr>
<tr>
<td>Cordoba</td>
<td>0.11</td>
</tr>
<tr>
<td>A Coruña</td>
<td>0.18</td>
</tr>
<tr>
<td>Cuenca</td>
<td>0.63</td>
</tr>
<tr>
<td>Gerona</td>
<td>0.02</td>
</tr>
<tr>
<td>Granada</td>
<td>2.14</td>
</tr>
<tr>
<td>Guadalajana</td>
<td>0.56</td>
</tr>
<tr>
<td>Guipúzcoa</td>
<td>3.29</td>
</tr>
<tr>
<td>Huelva</td>
<td>0.08</td>
</tr>
<tr>
<td>Huesca</td>
<td>0.11</td>
</tr>
<tr>
<td>Jaén</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Notes: The numbers indicate the ratio of the loading factor \( \hat{\alpha}_i \) with respect to the mean of the provinces.

### Table 8: Estimation Results: Cointegration Relation

<table>
<thead>
<tr>
<th>Estimated Model</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_t )</td>
<td>1.74</td>
</tr>
<tr>
<td>( \hat{h}_t )</td>
<td>-1.72</td>
</tr>
<tr>
<td>( \hat{r}_t )</td>
<td>-3.50</td>
</tr>
<tr>
<td>( CD )</td>
<td>-2.05</td>
</tr>
<tr>
<td>( CIPS )</td>
<td>-2.5***</td>
</tr>
<tr>
<td>( CIPS^* )</td>
<td>-2.2***</td>
</tr>
</tbody>
</table>

Notes: Estimated coefficients in the second step of the Breitung (2005) Panel-VECM procedure. Standard errors are given in CIPS and CIPS* refer respectively to the standard and truncated cross section averages of Cross-sectionally Augmented Dickey-Fuller test statistics CADF(1) with no intercept and no trend. Relevant lower 1% critical values for the CIPS and CIPS* statistic are -1.78 and -1.77 respectively. ***means that the test is significant at the 1 and 10 per cent level.

### Figure 1: Average Real Growth of Housing Prices 1971-2006

Source: Authors’ calculation based on OECD data

### Figure 2: Share of home-ownership in 2002

Source: BIS and BBVA

### Figure 3: Real Per Capita GDP Growth

Source: Authors’ calculation based on AMECO data
Fernando C. Ballabriga, Sonsoles Castillo: BBVA-ARIES: un modelo de predicción y simulación para la economía de la UEM.

Rafael Doménech, María Teresa Ledo, David Taguas: Some new results on interest rate rules in EMU and in the US

Carmen Hernansanz, Miguel Sebastián: The Spanish Banks’ strategy in Latin America.

Jose Félix Izquierdo, Angel Melguizo, David Taguas: Imposición y Precios de Consumo.

Rafael Doménech, María Teresa Ledo, David Taguas: A Small Forward-Looking Macroeconomic Model for EMU

Jorge Blázquez, Miguel Sebastián: ¿Quién asume el coste en la crisis de deuda externa? El papel de la Inversión Extranjera Directa (IED)

Jorge Blázquez, Javier Santiso: Mexico, ¿un ex-emergente?

Angel Melguizo, David Taguas: La ampliación europea al Este, mucho más que economía.

Manuel Balmaseda: L’Espagne, ni miracle ni mirage.

Alicia García Herrero: Emerging Countries’ Sovereign Risk: Balance Sheets, Contagion and Risk Aversion

Alicia García-Herrero and María Soledad Martínez Pería: The mix of International bank’s foreign claims: Determinants and implications

Alicia García Herrero, Lucía Cuadro-Sáez: Finance for Growth: Does a Balanced Financial Structure Matter?

Rodrigo Falbo, Ernesto Gaba: Un estudio econométrico sobre el tipo de cambio en Argentina

Manuel Balmaseda, Ángel Melguizo, David Taguas: Las reformas necesarias en el sistema de pensiones contributivas en España.

Ociel Hernández Zamudio: Transmisión de choques macroeconómicos: modelo de pequeña escala con expectativas racionales para la economía mexicana

Alicia Garcia-Herrero and Daniel Navia Simón: Why Banks go to Emerging Countries and What is the Impact for the Home Economy?


0703 Ociel Hernández y Cecilia Posadas: Determinantes y características de los ciclos económicos en México y estimación del PIB potencial

0704 Cristina Fernández, Juan Ramón García: Perspectivas del empleo ante el cambio de ciclo: un análisis de flujos.

0801 Alicia García-Herrero, Juan M. Ruiz: Do trade and financial linkages foster business cycle synchronization in a small economy?

0802 Alicia García-Herrero, Eli M. Remolona: Managing expectations by words and deeds: Monetary policy in Asia and the Pacific.

0803 José Luis Escrivá, Alicia García-Herrero, Galo Nuño and Joaquín Vial: After Bretton Woods II.

0804 Alicia García-Herrero, Daniel Santabárbara: Is the Chinese banking system benefiting from foreign investors?

0805 Joaquín Vial, Angel Melguizo: Moving from Pay as You Go to Privately Manager Individual Pension Accounts: What have we learned after 25 years of the Chilean Pension Reform?


0807 Ociel Hernández, Javier Amador: La tasa natural en México: un parámetro importante para la estrategia de política monetaria.

0808 Patricia Álvarez-Plata, Alicia García-Herrero: To Dollarize or De-dollarize: Consequences for Monetary Policy


0902 Alicia García-Herrero, Jacob Gyntelberg and Andrea Tesei: The Asian crisis: what did local stock markets expect?

0903 Alicia García-Herrero and Santiago Fernández de Lis: The Spanish Approach: Dynamic Provisioning and other Tools

0904 Tatiana Alonso: Potencial futuro de la oferta mundial de petróleo: un análisis de las principales fuentes de incertidumbre.

0905 Tatiana Alonso: Main sources of uncertainty in formulating potential growth scenarios for oil supply.

0906 Ángel de la Fuente y Rafael Doménech: Convergencia real y envejecimiento: retos y propuestas.

Alicia García-Herrero, Philip Woolbridge and Doo Yong Yang: Why don’t Asians invest in Asia? The determinants of cross-border portfolio holdings.

Alicia García-Herrero, Sergio Gavilá and Daniel Santabárbara: What explains the low profitability of Chinese Banks?.


J.E. Boscá, R. Doménech and J. Ferri: Search, Nash Bargaining and Rule of Thumb Consumers

Angel Melguizo, Angel Muñoz, David Tuesta and Joaquín Vial: Reforma de las pensiones y política fiscal: algunas lecciones de Chile


Angel Melguizo, Angel Muñoz, David Tuesta and Joaquín Vial: Pension reform and fiscal policy: some lessons from Chile.

Alicia García-Herrero and Tuuli Koivu: China’s Exchange Rate Policy and Asian Trade


Alicia García Herrero y Daniel Santabárbara García: Una valoración de la reforma del sistema bancario de China

C. Fung, Alicia Garcia-Herrero and Alan Siu: A Comparative Empirical Examination of Outward Direct Investment from Four Asian Economies: China, Japan, Republic of Korea and Taiwan

Javier Alonso, Jasmina Bjeletic, Carlos Herrera, Soledad Hormazábal, Ivonne Ordóñez, Carolina Romero and David Tuesta: Un balance de la inversión de los fondos de pensiones en infraestructura: la experiencia en Latinoamérica

Javier Alonso, Jasmina Bjeletic, Carlos Herrera, Soledad Hormazábal, Ivonne Ordóñez, Carolina Romero and David Tuesta: Proyecciones del impacto de los fondos de pensiones en la inversión en infraestructura y el crecimiento en Latinoamérica
The analyses, opinions and findings of these papers represent the views of their authors; they are not necessarily those of the BBVA Group.

The BBVA Economic Research Department disseminates its publications at the following website: http://serviciodeestudios.bbva.com