Empirical Research: Assessing Macroeconomic Drivers of House Prices

A comparison Across Europe

By: **ERIC GUERRA CIVIT**

Universitat de Barcelona

Facultat d'Economia i Empresa

June 2018

Supervisor: Dr. Vahagn Jerbashian

(Department of Economic Theory)

ABSTRACT & KEYWORDS

In this empiric research I analyze the relationship between house prices and the following macroeconomic indicators: real GDP per capita, real long-term interest rates and unemployment rate. The analysis is applied to 5 different countries that represent the 4 economic models within Europe: Spain and Italy as a Mediterranean economies, Norway as a representative of the Nordic model, Germany represents the continental model and the United Kingdom that represents the Anglo-Saxon economy. The research focuses to find long run relationship between variables using a VECM methodology. The VECM can only be applied to Spain and Italy since they are the only cases where cointegration relationships can be found. For the rest of the countries only short run can be studied. The resulting explanatory power differs according to the country selected. While the model captures well the Mediterranean countries and the UK ($R^2 > 50\%$) it doesn't fit as well for the Nordic and Continental economies ($R^2 < 50\%$). Although the signs of the coefficients prove to be the equal across all countries, they vary when dynamics are added into the models.

Keywords: House Price Index, Cointegration, Vector Error Correction Method, Price Dynamics, Residential Real Estate, Macroeconomic Drivers, Structural Break Dummies.

Investigació empírica: Analitzant els determinants macroeconòmics dels preus residencials: Una comparativa europea

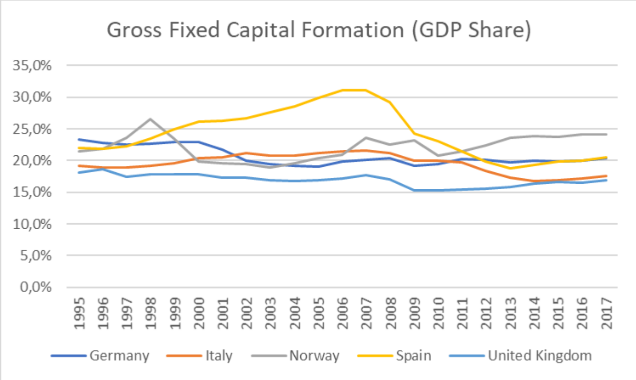
En aquesta investigació empírica em centro en analitzar les relacions entre el preus immobiliaris residencials amb els següents indicadors macroeconòmics: PIB real per càpita, tipus d'interès real a llarg termini i la taxa de d'atur. L'anàlisi s'aplica a 5 països diferents que representen cadascun dels 4 models econòmics Europeus: España i Itàlia com a economies mediterrànies, Noruega com a representant dels països nòrdics, Alemanya que representa el model continental i per últim el Regne Unit que representa al model Anglosaxó. La investigació es centra en trobar relacions a llarg termini entre les variables emprant la metodologia *Vector Error Correction Mechanism* (VECM). Dita metodologia ha resultat viable per a Espanya i Itàlia ja que son els únics casos en els que es troben relacions de cointegració. Per a la resta de països únicament s'analitza el curt termini. El poder explicatiu del model varia en funció del país seleccionat, mentre que el model descriu gran part del comportament dels preus residencials per als països mediterranis i per al Regne Unit ($R^2 > 50\%$), no es capaç de reproduir aquest poder predictiu per als models nòrdic i continental ($R^2 < 50\%$). Tot i que els signes dels coeficients mostren ser iguals per a tots els països, aquests canvien en quant s'introdueixen dinàmiques en el model

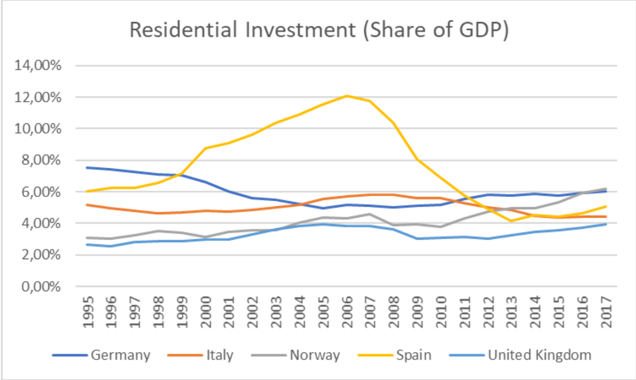
Paraules clau: Index de preus residencials, Vector Error Correction Mechanism, dinàmiques de preus, immobiliària residencial, determinants macroeconòmics, variable fictícia de canvi estructural.

INDEX

I. INTRODUCTION

Real estate plays a significant role in the economy and it is deeply linked with the financial sector. The economic development process has always required major investments in infrastructure and capital creation. Real Estate is characterized to be immobile, hence considered as a local market. Scarce. It is non-homogeneous since it heavily depends on physical and geographical characteristics. It is illiquid (even if demand is high there are high transaction costs, including timeframe between willing to sell and sale realization). It is durable, which imply that its value can be enhanced via CAPEX (known as investment permanence); It requires high initial costs, due to its nature the initial disbursement required is usually higher than other investments which usually imply the need of financing. Typical costs include land acquisition, rehabbing, tax and registry, fees, financing… Risk, like any other investment real estate has an associated risk component which can be very volatile affecting also range of profitability. Housing sector has also the characteristic of being a very important component of either consumption and investment. The Following figures show the share of GDP of gross fixed capital formation and residential sector specifically. During the analyzed period, Spain is the country whose GDP has relied more on fixed capital formation and housing investment specifically, reaching levels of 30% and 12% respectively of GDP share just before the arrival of the financial crisis. After the 2008 crisis, the housing bubble for Spain crashes and making converge the residential investment's GDP share to similar levels as other European countries. Germany, on the contrary, exhibit almost an opposite behavior as Spain, having its maximum relevance of housing investment during the 90's with an 8% of GDP and consistently decreasing during the pre-crisis period to later increase after the crash. Italy follow a similar behavior as Spain but with much less volatility and less relative levels, reaching a peak of 23% in gross fixed capital formation's share and a 6% residential investment's share reached before the financial crisis. Norway have a very distinctive behavior. Its residential sector shows a constant growing pattern but few volatility and relative levels (from 3% of GDP share to 6% considering whole period). On the other hand, Norway has higher share of gross fixed capital formation compared to other analyzed countries. Is especially relevant that during the 90's and post-crisis period, Norway is the country with higher levels of gross fixed capital formation as a share of GDP. The reasons for having such a gap between the two indicators in the Norwegian case are the high saving propensity (High relevance of pensions + cultural propensity to save) and its economic structure based of natural resources very intense in physical capital like: Petroleum extraction, fishing (Norway has the fourth large fleet in the world), Metal industry… The UK is the country with lower relative levels of either residential investment and fixed capital investment. The Anglo-Saxon country and specially England have historically had problems with its residential sector. The speculation, scarcity of land, high fixed costs or the fact that there is a high level of asset reform instead of new construction are some of the factors explaining lack of new construction in the UK.





Real Estate has traditionally been considered as capital stock and it can be divided into housing or residential real estate and non-residential, which include commercial, logistic, industrial… Residential sector differs from non-residential since it is considered to be more volatile, with more chances to lead to a business cycle and make use of different technology (Cooley y Prescott 1995). The residential sector is not tradable and international markets and overall supply do not allow to reduce volatility by simple arbitrage. Housing is a basic need and as a consequence, it has a relevant weight on the total consumption share (Average of 35-65% depending on the country). On the other hand, housing is like other capital stock, a significant component of investment, especially due to its profitability (with its corresponding risk element) and potential growing demand.

Residential sector is especially susceptible to overrated expectations, demand and high investment. Some factors like increase of disposable income, increase of population, low interest rates, lax mortgage policies, lack of financial expertise of buyers, or the believes that price tendencies are expected to behave consistently upwards are key factors for bubble creation with more or less impact depending on the country. For instance, Mediterranean countries such as Spain had, during the pre-crisis period, more growth in its housing stock than Germany and France together, causing a great impact on the economy when the crash occurred in 2008. The expectations play an important role on either demand and supply and are a relevant factor of the bubble creation. According to Nakajima (2011) house price factors can be divided into: Supply: which is affected by land price, construction cost, financing costs, land/urban regulations… Demand is determined among other factors by population growth (Mankiew and Weil 1989) and income growth. Expectations: As previously described, the expectations have a large role on the formation of speculative bubbles, on part due to self-fulfilling expectations: an increase of current price affect the believe that future prices will increase too, economic agents react in consequence and forcing the future price to go up ($\uparrow P_t \rightarrow \uparrow E[P_{t+1}] \rightarrow \uparrow P_{t+1}$) (Piazzesi y Martin 2009). Also, if an economy output or productivity is expected to grow, it will affect long term house prices ($E[\Delta y, \Delta Productivity] \rightarrow \uparrow E[House\ Price]$) (Kahn 2008). The understanding of price dynamics is a key component to be able to distinguish between the regular cycle and the speculative component of housing sector driven by demand and potentially causing bubbles.

According to Wheaton (1999) and empirically evidenced by Leung and Chen (2006) on the long run, the reservation equilibrium price is equal to the actual value of the property income (potential rent + residual value). For that reason, the expectations have such an influence on price determination.

I consider relevant to study how the sector has historically interacted with other main macroeconomic components since residential sector has such an impact on consumption, investment and by consequence the general economy,. The topic could be approached from several perspectives such as focusing on the microeconomics (price determination, market structure, individual behavior, demand and supply formation…), historic evolution (analyzing impact of relevant factors such as regulations, past events and evidence) and many other disciplines. I decided to focus on the macroeconomic linkage with house prices because we would expect to find some relation between such a relevant economic sector and the main macroeconomic proxies such as output, inflation or demographics. I consider the research may be of some relevance, especially after digging on the extensive literature related to the topic, which indicates there exists a significant interest on the issue. In order to find those relationship I will use a VAR methodology relating house prices, GDP per capita, real interest rates and unemployment. I will focus on analyzing short run vs long run equilibrium as well as studying the effects of price dynamics into the model.

II. LITERATURE

There are a lot of research regarding the macroeconomic impact of house prices and the possible global effect in case of business cycle synchronization. For instance, Green (1997) uses several time-series specifications to find that housing investment causes growth of GDP but is not caused by it, it is also considered to lead the business cycle. While non-housing investment do not cause output growth but is caused by it and is considered, as with other investments, to lag the business cycles. Green (1997) suggests the idea of channeling investment from housing to other capital stock and infrastructure to avoid severe short run imbalances. Coulson and Kim (2000) show that a shortcoming of the exercise performed by Green (1997) is that he didn't consider the influences of other GDP components other than residential sector might have in the determination of GDP. For that reason, they use a multivariate VAR models to test and compare the effect of housing and non-housing prices on output and its components. They find that residential real estate investments shocks are more relevant in the determination of GDP than non-housing investment, which is a similar conclusion to Green (1997). Reinhart and Rogoff (2009) show that financial crisis are usually associated with output recession and house price downturns stretched over long periods of time.

Many papers also analyze whether the house prices can have a global trend due to dynamic synchronization. In an early study Renaud (1995) provides a comprehensive descriptive analysis of the international cycle in advanced economies between '85 and '94 finding that the cycle synchronization of house prices was caused by the general liberalization of financial markets in the late '80s. Hirata, et. al. (2013) analyze how, in the past two decades, while the relative importance of global factor was declining, there has been some convergence of business cycle fluctuations wihin AEs (Advanced Economies) and EMEs (Emergin Economies) separately. Consistently with this view, some EMEs have become resilient to shocks originated in AEs. This phenomenon is known as *Regionalization Hypothesis*. This results indicate that house prices can be globally correlated to a certain extend, or at lest there are common trend that are detected when jointly analyzed. I would expect to find at lest some degree of correlation between selected european countries.

To asses the relation of housing investment and GDP, many researchers have opted to use Vector Auto Regression methodologies and its variants. For instance, Otrok and Terrones (2005) use a VAR specification and find "a large degree of synchronization or comovement between the growth rate of real house prices and macroeconomic aggregate such as real output, consumption and residential investment". Both the relationship and the methodology haven't been only applied to the U.S. Xiao (2015) uses same VAR methodology to study whether the relation holds for China, concluding that House Price expectations, money supply, household income and real estate investment are dominant factors to explain house price evolution. Kurita (2010) studies the long run relationship for Japan's house prices and macroeconomics aggregates searching for cointegration to run a Vector Error Correction Model. Meidani Ali (2011) apply a VAR methodology in order to study Granger Causality between Iranian house prices, economic output and inflation. Confirming that either GDP and CPI Granger cause house prices.

The research is not restricted to one country. Many researchers wanted to carry a panel data analysis, focusing on the topic with an international scope. Cesa-Bianchi (2012) make use of a Global Vector AutoRegresion (GVAR), originally proposed Pesaran, Schuermann and Weiner (2004) to investigate the international transmission of housing shocks. Specifically the study analyzes: housing demand shocks originated in the US (to analyze how the shock could be propagated to the rest of the world, triggering the financial crisis), demand housing shock originated in all advanced economies, and finally equity price shocks originated in all advanced economies (these two shocks are analyzed to understand the effect of "common regular shocks"). Vanstreenkiste (2007) uses GVAR methodology to find that California house price shocks appear to be an important factor driving prices in other states. While on a latter study Hiebert and Vansteenkiste (2009) conclude that house price shocks play a minor role in explaining house spillovers in the euro area. The drawback is that GVAR methodology doesn't allow to structurally identify shocks, which, according to Cesa-Bianchi (2012) "imply that there is no economic interpretation of the housing shocks in those studies" and in addition "it is difficult to understand how country weights affect the influence of individual country variables in the transmission of shocks across borders" since the methodology characterizes cross-border linkages by averaging variables into a global aggregate.

III. TECHNICAL BACKGROUND

I will analyze the link between a key variable of any market, its price, with some main aggregated macroeconomic variables such as GDP, Interest rate and Unemployment to understand the impact and relevance that the real estate sector has on the overall economy. To establish comparisons and be able to contextualize the results, I will run the analysis for 5 different countries: Spain, Italy, Norway, Germany and United Kingdom. Those specific countries were selected in order to have a representative of each idiosyncratic and economic framework: Nordic, Anglo-Saxon, Continental, and Southern-European. I decided to also include Italy, having two countries from the same area to compare across the same wide economic model.

The distinction of those economy-wide frameworks is often related to the study of welfare state, market labor and social components. The real estate market and specially the residential sub-sector is such a transversal and relevant sector that I consider it can be interesting to check for differences between models and within them (such as Spain-Italy case) and look for possible correlation between price evolution and regulation dimension of the economic block represented by individual countries.

At the first stage of the research, the intention was not only to study the influence and correlation of prices and macroeconomic variables, but rather to analyze real state price specific factors of Spanish economy and its influence on price dynamics, and speculative behavior in depth. Using variables such as price-to-income ratio, housing stock, dynamics in construction's employment, real estate transactions, number of mortgages… The inclusion of those variables and the enforcement of their analysis would have allowed to depict a more consistent outlook of the topic. Even that those variables are available for the Spanish economy to a certain degree, when more countries are added to the analysis the scarcity of data becomes a problem. For that reason, I decided to drop that research line and center the focus on the comparison of several countries using the same model with more general variables instead of analyzing the factors more in depth.

To study the relation between real estate residential prices I center my analysis on the possible long-run relationship between the stated variables. In order to do so, the variables must be stationary. As a reminder we say that a stochastic process is stationary when its unconditional joint probability distribution (Probability of the events happening at the same time) is independent of time. For this analysis we only need the process to be weak stationary by having its mean, variance and autocorrelation structure independent of time.

If the variables are I(0) the process is already stationary. In case that they are I(1) a difference with its previous value is needed in order to achieve stationarity on that particular variable (e.g. $\Delta X_t = X_t - X_{t-1}$). By applying differences, we only can study the short run effects between the variables unless there exists a cointegrating relationship. There are several approaches for checking the order of integration such as the Augmented Dickey-Fuller test or the KPSS test.

This procedure of detecting the order of integration is considerably relevant in order to detect possible spurious correlations. If X_t and Y_t are time series that are entirely independent of each other, we can expect that a simple linear regression between these two would usually produce an insignificant estimate of the coefficients. However, this may not be the case if the variables behave like random walks, which are I(1) processes. In that case, the estimates of the parameters in the regression do not have Student's *t* standard distribution, even asymptotically. As $n \to \infty$, it is possible to reject the null hypothesis that the coefficient is insignificant ($\beta_i = 0$) with probability 1. Moreover, the R^2 converge to functionals of Brownian motions, resulting in a high R-squared when in should be close to 0. A quick approach to know whether a regression is spurious without checking for unitary roots is to take a look to the *Durbin-Watson* statistic of the regression (by approximation $DW = 2(1-\rho)$, where ρ is the first autocorrelation coefficient). When $DW \sim 0$ it is indicative that the model is capturing a spurious effect, while if $DW \sim 2$ the model is not result of spurious regression, but more tests are recommended to confirm both cases. To know more about the issue check Granger & Newbold (1974)

A cointegrating relationship is a linear combination of two or more non-stationary variables that give as a result a stationary process. If cointegration relationship is found is still possible to analyze the long run effect of the process. The main requisite for cointegration to exist is that all variables must be the same order of integration. Since the goal is to get the long-run relationship we need either all variables I(0) or all variables I(1) and find cointegration relationships. To check for possible cointegration relationships the main methodologies are:

- Engle-Granger which is based on checking unitary roots for the variables (must be the same order of integration) and on the error of the regression in levels (should be one degree less than the variables for cointegration to exist). It is based on ADF test. For further details more issues are articulated by Engle & Granger (1987)
- Johansen Test: computes the maximum likelihood estimator of the reduced rank model. Contrary to Engle-Granger, it can be applied to vectoral space, being able to find several cointegrating relationships. Is divided into Trace test and Lmax test. It determines the matrix rank of cointegration vectors that exist for a specific vector of variables. For further details check Johansen (1995)

The requisites for cointegration can be summarized with the Granger Representation Theorem which states that systems with cointegrated I(1) variables have three equivalent representations: Common trend, Moving Average and ECM specification.

The long-run relationship can be studied by applicating Error Correction Mechanism (ECM) if all variables are I(1) and there exists cointegration relationship among them,. The ECM term is the lagged error of the regression in levels and its coefficient represent the speed of the adjustment towards the steady state after a short run shock.

$$\Delta y_t = \sum_{i=0}^{k} \beta_n \, \Delta x_{n,t-i} + \sum_{j=1}^{k} \beta_m \, \Delta y_{m,t-i} - \gamma ECM + u_t \qquad (1)$$

$$ECM = y_{t-1} - \sum_{i=0}^{k} \beta_n \, x_{n,t-i}$$

$$where:\ -1 < \gamma < 0$$

As it can be seen in equation (1), we study short run relationship between I(1) variables by analyzing their differences. If a cointegration relationship is found among them, the ECM can be computed (notice that the error correction part is in levels). The ECM captures the deviation from the long run after a short-run shock. Its negative coefficient indicates that is actually acting like a pivot, neutralizing the shock and converging to the steady state.

This concept can also be applied for a vector of variables, becoming known as Vector Error Correction Model (VECM). The main advantage of the VECM is that all equations can be tested at the same time, creating r cointegrating vectors, in that case, equation 1 looks like:

$$\Delta y_{i,t} = \alpha + \sum_{j=1}^{k-1} \Gamma_j \, \Delta y_{i,t-j} - \Pi y_{i,t-1} + u_t \qquad (2)$$

$where:\ \Gamma_j = (A_{J+1} + \cdots A_k)$; $\Pi = (I - A_1 - \cdots A_k)$ and y_i is the vector of independent variables

Π is the cointegrating matrix and its rank represent the total number of cointegration relations. The rank of the Π matrix can be computed by applying Johansen tests (either Trace or Lmax). The rank of the matrix can be:

- Rank(Π) = 0: which indicates that there exist no cointegration relation and the model can be only studied in first differences
- Full rank *k:* Being k the number of endogenous variables. This result imply that variables cannot be I(1) and the model should be estimated in levels without VECM
- Rank(Π) = m, where 0 < m < k as is the case of cointegration, being m the number of cointegrating relations between the variables.

IV. THE DATA

Before deciding the final model to be used, I decided to do a broad research regarding data availability. In order to study the behavior of house prices (henceforth HPI), many variables were considered. First, I considered the capital market including number of mortgages, mortgage rate, rate of arrears and house price-to-income ratio. Then I contemplated both the demand side, with variables like total household debt, real net disposable income per capita, tax values on property or deductions… and the supply side with variables like housing stock, Gross value added on construction by total GVA, employment in construction sector, construction costs, number of housing transactions… And finally, broader macroeconomic variables were considered such as total population, unemployment, GDP, inflation and interest rates. Is relevant to add that, instead of using raw data, I would have adapted the variables in order to apply the model by adjusting them to inflation, population… More detail will be found in the data treatment section.

When searching for the data I found that I could find most of the variables but neither for all countries nor for all the period. The institutions in charge of recollecting data don't provide the same figures on every country. For instance, the Spanish ministry of public works and transport provide a full array of data related to real estate sector like number transactions, housing stock, issued building licenses … they even provide annual data on transferred surface, and transactions divided by hedonic characteristics of asset. While Spain provide broad and useful information to carry on a proper research, other countries, by the contrary, stand out for offering scarce details on their real estate sector. For instance, there is not a single institution providing housing stock data for the UK. Instead, each member country provides their own data, having different available periods.

Another problem found during the research was the timeframe available and the periodicity. Many considered variables were available in annual frequency and since this research is based on quarterly data, my only choices were either to drop the specific variables or to apply a transformation assuming linear quarterly increase. Since the transformation wouldn't reflect the real shifts in inter-quarterly variations, I decided to choose for the former options and drop the variables in case of having only annual availability.

As for the studied period, it was originally set to study the evolution of prices since the 2008 financial crisis. But after selecting the data and start running some models I realized that the during the period there is considerable structural changes that make the data have some integration problems. After doing some technical research I end up with two solutions on how to fix the behavior of the data since 2008: Use a wider timeframe in order to appreciate a more general tendency and diminishing structural break influence, or by adding dummy variables that capture the periods on which the structural break is affecting the data. I decided to support the first choice and keep the second solution in case the break is strong enough to affect the wider period such as will be the case of Spain. By extending the time period the number of observations increase, allowing for better asymptotic behavioral response and increasing the degrees of freedom. The final considered period is 1995Q1-2017Q4, having a total of 92 observations. The period allows to study the behavior and dynamics of prices after the 90's crisis due to oil price shocks, restrictive monetary policy, and exhausting of construction boom caused by an overbuilding during the 80's. although, each country reacted differently to the 90's crisis, I think that is relevant to study the behavior during a generalized state of recovery followed by a global crisis (2008) and its consequent readjustment.

The final variables to be used are based on what the literature provides. Most researches analyzing the topic only select main macroeconomic variables such as Interest Rates, GDP, inflation… For instance (Hirata, et al. 2013) select house prices, interest rates, reserves, credit spread, GDP, or default rate. (Lourenço & Rodrigues, 2014) use interest rates, disposable income and labour force to estimate a ECM model and compare house prices and its factors between countries. (2011) consider Real GDP, CPI, house price index and exchange rate to stablish a VAR model and check for Granger causality. Kurita (2010) make use of HPI, Real GDP, GDP deflator to monitor inflation, growth rate of housing prices and Government Bond Yield as a proxy for interest rates and mortgages rates.

4.1 Data Treatment

The final variables included in the model are: Gross Domestic Product, House Price Index, Long term interest rates and unemployment. The variables were transformed in order to exclude the influence of population changes and prices. For that reason, GDP was divided by CPI and Population, obtaining Real GDP per capita (henceforth RGDPPC). House Price Index was divided by the CPI, getting Real House Price Index (RHPI). I subtracted growth rate of consumer price index from long-term interest rate to obtain real long-term interest rate (RLTIR). Finally, I divided the unemployment by active population in order to get the unemployment rate (UNEMP). It would have been relevant to include a variable that could reflect the supply side of the market such as the housing stock, but as previously mentioned, it was not possible to find the data for all the countries on a quarterly basis from 1995Q1-2017Q4.

The data for all the following variables have quarterly variation are for the period 1995Q1-2017Q4

4.1.1 Real GDP per capita:

The data for GDP are from the OECD website. GDP series are measured in US$ and in current prices. They are seasonally and purchase power parity adjusted. I divide GDP by CPI and population in order to obtain real GDP per capita. The GDP per capita is one of the main macroeconomic variable to monitor the overall dynamics of the economy and its widely used as a proxy for economic development. Spain and Italy have the lowest levels of GDP, being their average levels during the period of 27.746 and 31.293. Norway the country with the highest average with 48.153 but is also the one with highest variability, with a standard deviation of 14.335 compared to the other countries' range of 5150-8614. The next figure shows nominal GDP of the analyzed countries. The GDP/CPI can be found on the appendix section. I would expect real GDP per capita to positively affect house prices. If the relation between doesn't hold, it may be due to the presence of a housing bubble where prices are more driven by the expectations than by macroeconomic drivers.

GDP per capita, current prices, current PPPs, Seasonally adjusted					
Country	**Nº Observations**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	38.008	15.981	27.746	6604,71
UK	92	43.449	20.235	32.489	6919,60
GERMANY	92	50.885	23.368	35.611	8614,00
ITALY	92	39.878	22.054	31.293	5150,72
NORWAY	92	67.594	23.607	48.153	14335,53

source: OECD

4.1.2 Unemployment rate:

The unemployment variable that can monitor labor force dynamics and the saving capacity of households. Generally, in order to get a mortgage, the lenders require either a constant stream of income or some collateral asset in order to concede the loan, so it can also depict credit availability for households. The unemployment data was extracted from OECD data website. Originally in aggregate levels, I divided it by active population in order to get the unemployment rate. The intention was not only to capture changes in unemployment but also include the dynamics of active population, which is relevant considering the tendency of young people to extend their education period or, as is especially relevant in Spanish case, long-term unemployed that abandon their job seeking due to factors such as mismatching between labor demand and supply profiles, shifting of economic activity, automation (Vivarelli 1997) or lack of migration capacity (Pissarides & Wadsworth 1989). During the period the country with a highest average of unemployment rate is Spain with a 16,35% average compared to Norway with the lowest average of 3,72%. I would expect that unemployment negatively affects to house prices. According to Okun's law one-point increase in the cyclical unemployment is associated with two percentage points of negative growth. I expect that to be especially relevant in the Mediterranean economies due to its large structural unemployment and high elasticity of cyclical unemployment with respect to GDP.

Unemployment rate, percentage of active population, seasonally adjusted					
Country	Nº Observations	Max	Min	Average	Standard deviation
SPAIN	92	26,20	8,00	16,35	5,57
UK	92	8,70	4,30	6,13	1,31
GERMANY	92	11,20	3,60	7,60	2,09
ITALY	92	12,80	6,00	9,66	1,95
NORWAY	92	5,90	2,40	3,72	0,75
source: Eurostat					

4.1.3 Real Long-term interest Rate:

The Real long term interest rate is assumed to contain information about future economic conditions and is strongly related with mortgage rates and for depicting credit availability, discount factors when analyzing investments, and affecting future expectations on investment. So, it is considered to be relevant when studying house price dynamics (Sutton, Mihaljek and Sub 2017). On average, the countries with higher levels of Real long term interest rate during the period are Spain (4,08) and Italy (4,33) while UK and Norway have both similar averages of 3,7 and Germany having the lowest with 3,1. I would expect the interest rate to negatively affect housing prices. An increase of interest rates would lower demand (more expensive financing) and supply (lower investment, due to lower gap between financing cost and profit)

Real Long Term Interest Rates					
Country	Nº Observations	Max	Min	Average	Standard deviation
SPAIN	92	10,62	-0,32	4,08	2,16
UK	92	7,96	-0,05	3,78	1,93
GERMANY	92	6,60	-0,59	3,11	1,86
ITALY	92	11,25	0,79	4,33	2,07
NORWAY	92	7,52	0,05	3,77	1,86
Source: OECD and own computations					

4.1.4 House Price Index:

House price index are computed by analyzing transaction price or by appraisal value. For the transaction price, it can be used the repeated sales methodology, single sale methodology, the hedonic price approach or a mixture between all. We would expect that past lags of price affect positively to current price. The measurement of house prices can affect the interpretation of their behavior because there are conceptual differences between transaction prices and appraisal values. For that reason, the following section will detail several methods of house price computation. For my analysis I selected appraisal value due to data availability.

- Repeated Sales: It considers the variation across multiple sales of the asset, Initially proposed by Bailey, Muth and Nourse (1963), it computes the Price of the *n^th* property as a function of a general regional price level, a Gaussian random walk that represents the trend in individual value over time and a homoscedastic error term. The methodology set a first milestone on price index computation, and it was further developed by (Case & Shiller 1987) by considering that the error term was, in fact, heteroskedastic. To address the issue, they used weighted least Squares (WLS) and added a time trend. Case & Shiller methodology was adopted by Standard and Poor's to compute their house price index: The S&P/C-S which instead of using the logarithm of prices it uses the transaction price for easier interpretation. C-S was also adopted by the Office of Federal Housing Enterprise Oversight (OFHEO) but applying different weights to the observation according to the variance error terms. Although the S&P/C-S and OFHEO are the most used worldwide repeated sales indices, new developments have been made to get a more efficient estimate, such as the case of the autoregressive index. Originally computed by Nagaraja, Brown & Zhao (2011) the autoregresive index takes into consideration the gap of time between sales to adjust for depreciation, it also adds a hedonic random variable proxied by ZIP code, a log of general price index and introduces autoregresive component by introducing 1 lagged error term. This method have proved to be more accurate and with more predictive capacity than the other repeated sales approaches. The main advantages of the repeated-sales method is its ease of computing and the data availability (only need the each transaction price and its transfer date). As a drawback, the method doesn't take into account the properties that are not sold more than once deriving in a possible selection bias problem. Also the method can't provide separate prices for plots and buildings. As an example, equation (4) show the Nagaraja, Brown y Zhao (2011) methodology

$$\begin{aligned} y_{i,1,z} &= \mu + \beta_{t(i,1,z)} + \tau_z + \varepsilon_{i,1,z} && j = 1 \\ y_{i,j,z} &= \mu + \beta_{t(i,j,z)} + \tau_z + \phi^{\gamma(i,j,z)}\varepsilon_{i,j-1,z} + \varepsilon_{i,j,z} && j > 1 \end{aligned} \qquad (4)$$

Where: $\beta_{t(i,1,z)} \rightarrow Log\ price\ index\ at\ t(i,j,z)$

$\phi \rightarrow Autoregressive\ coefficient\ and\ |\phi| < 1$

$\tau_z \rightarrow Random\ effect\ of\ ZIP\ code\ z,\ and\ \tau_z \sim N(0,\sigma_\tau^2)$

$j = sale\ number,\ i = House,\ z = ZIP\ code,\ y = Price$

- Hedonic Characteristics: The methodology consists on regressing transaction prices as a function of physical characteristics of the asset like surface, location, quality... in order to study the marginal contributions or shadow prices of those characteristics on the transaction price. Originally formulated by Court (1939), the hedonic pricing was properly theorized and popularized by Rosen (1974) who argued that "an item total price can be considered as the sum of each homogeneous attributes prices, where each attribute has a unique implicit price in an equilibrium market" being able to regress the item price as on the attributes to determine how each characteristic marginally affects the overall price. Other scholars criticized Rosen's work arguing that the estimated coefficients were not strictly equal to the willingness to pay. The Hedonic model has an implicit difficulty when selecting functional form. Many scholars have presented their own versions including time dummies, semilogarithmic models, quality adjustment parameters (to consider depreciation)… The advantages of hedonic regression are that is an efficient method, similar to other price indices in its computation, allows for more stratification, etc. While it has some drawback like: Very data intensive which can be difficult to obtain, the model can suffer from specification issues like the difficulty of setting a functional form or because of technical econometric problems like Heteroskedasticity, autocorrelation, multicollinearity… An example of hedonic method is the one used to compute the Spanish house price index by the National Institute of Statistics. The regression is formed by discrete and binary variables that take values according to the specific characteristics of the asset such: New/Used, Apartment/individual house, presence of parking, basement, cooperative… Surface (range of 10 possible values), province, population, tourism relevance, postal code. When the price is computed it is indexed by chained Laspeyres Index.

- Appraisal value: The method is based on observing the valuation price of the property. Although the appraisal value of the asset is very correlated with its transaction price, the literature agrees on the existence of a gap between transaction price and valuation known as *Appraisal Smoothing,* which is a systematic bias characterized by having lower volatility. Gertner (1989) defined appraisal smoothing as the situation when the ratio of transaction price index to the appraisal standard deviation is higher than 1. Fisher, Miles and Webb (1999) observed that the transaction price is usually higher than appraisal value when the market is on a growing trend. By the contrary, if the market is on a downward trend the appraisal values are usually higher than the observed transaction prices. Originally Quan y Quigley (1991) computed a model known as Partial adjustment method to fix some of those issues. The model explains appraisers behavior by applying weighted average of reservation price and offer price, assumes that volatility is exogenous, can't be observed and follow a random walk. Also, following Ibbotson y Siegel (1984) and Gertner (1989) stated that previous values of either transaction price and appraisal value of the asset have a significant incidence on the actual valuation, causing autocorrelation and being one of the causes of appraisal smoothing (e.g. Sales comparison method by comparable analysis). Equation (5) shows how Quan and Quigley presented their model.

$$P_t^* = E[P_t | P_t^T, \Omega_{t-1}] \tag{5}$$

Where: $\Omega_{t-1} \rightarrow Available\ information\ set\ at\ t-1$

$P_t^T = P_t + \upsilon_t \rightarrow Long\ term\ equilibrium\ (\upsilon_t \sim N(0;\ \sigma^2))$

If equation (5) is expanded, it ends up being:

$$P_t^* = K \cdot P_t^T + (1 - K) \cdot P_{t-1}^* \quad (6)$$

Where: $K \rightarrow Appraiser's\ confidence\ parameter$
$to\ the\ information$ (Clayton, Geltner y Hamilton 2001)

For further information on the computation of house price indexes I strongly recommend the Eurostat methodological guide: *Handbook of Residential Properties Prices Indices, Eurostat 2013*

The idea of discarding the construction of a house price index came after realizing the difficulty of data availability. In Spain, the only institution that have transaction level data is the illustrious college of notaries. They don't provide the data to the overall public, they just share quarterly or yearly aggregates. Therefore, I decided to get an already constructed index.

The European Union set in 2008 a standard method to compute the HPI based on hedonic regression. Although this is the data I would like to use, the fact that it started in 2008 make impossible to study previous price evolutions. In order to avoid mixing data, I decided to get the values form the ministry of housing of every country which is based on an appraisal methodology. Even though appraisal method can suffer from bias compared to observed transaction price due to *Appraisal Smoothing* factor (Gertner 1989), its evolution and dynamics are very correlated. For that reason, I consider valuation value as a proxy of observed transaction price.

The following table provides the main statistics of nominal house price index, I computed the Real House Price Index by dividing by CPI, the table of RHPI can be found in the appendix section. Germany is the country where the house price has less variability and was less affected by the bubble crash of 2008, followed by Italy. Norway and UK are the countries with more price volatility with a deviation of 35,48 and 30,04 respectively.

Nominal House Price Indices, seasonally adjusted					
Country	**Nº Observations**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	111,82	32,63	70,64	25,50
UK	92	132,57	32,85	82,91	30,04
GERMANY	92	134,35	95,34	104,81	9,26
ITALY	92	105,94	50,96	80,91	18,59
NORWAY	92	147,99	29,31	82,27	35,48

source: OECD

V. EMPIRICAL RESEARCH

I apply logs to House Price Index and Real GDP Per Capita to reduce volatility and variance. The variables unemployment and Real Long Term Interest Rate are not logged since it would not make economic sense.

5.1. Checking for unitary roots

5.3.1 Spanish Case

In order to find possible long run relationship between the house prices and macroeconomic indicators, the first step is to check for order of integration. Knowing if the variables are stationary is a key part to know whether the long run relationship can be studied. In case all variables are stationary, the model can be run in levels, allowing to study long run relationship as usual. In case there are presence of unit roots, it will be important to know whether the variables have the same order of integration to be able to search for possible cointegration relationships to analyze the long run steady state with a vector error correction method. Finally, in case there are different orders of integration it will be only possible to analyze short run effects, working with the model in first differences. The following figure illustrate the initial results for Spanish Case, and the rest of countries results can be found in the appendices. I apply the Augmented Dickey Fuller Tests and the KPSS tests to check the presence of unitary root. To find technical details and methodology specification, check the technical background section.

House Price and GDP appear as I(2), which can be a problem to find possible cointegrating relationship since in this research I won't apply ARDL test to deal with cointegration with different integration orders. Also, in the literature GDP is considered to be I(2) when is in nominal and aggregate terms, but per capita real GDP is usually I(1). As a consequence I will assume that there is a specification error in the model.

In the Spanish case, the structural break of 2007/2008 is clearly causing specification errors, causing the tests to detect higher orders of integration that it really should be. The reason is that the stationarity tests detect the structural change as a shock that is not able to return to the steady state, causing the series to appear as non-stationary.

Spain		ADF		KPSS		Conclusion
		Level I(1)	Δ I(2)	Level I(1)	Δ I(2)	
Variables	l_RHPI	T -1.94973 0,6281 (2)	NC -2,89949* 0,056* (0)	T 0,52233*** 0,00001*** (3)	- 0,551005** 0,037** (3)	I(2)
	l_RLTI	T -3,0167 0,1275 (6)	NC -0,72720*** 4,4e-14*** (0)	T 0,211706** 0,012** (3)	- 0,551005** 0,037** (3)	I(1)
	l_UNEMP	C -1,78509 0,3884 (5)	NC -2,82528*** 0,004595*** (4)	NC 0,426749*** 0,000001*** (3)	- 0,325908 p > 0,1 (3)	I(1)
	l_RGDPPC	T -2,534 0,3115 (1)	C --1,74548 0,4052 (0)	T 0,545551*** 0,000001*** (3)	- 0,7821*** 0,000001*** (3)	I(2)

() The number in brackets indicate The optimal number of lags for ADF and KPSS

NC	→	No Constant
C	→	Constant
T	→	Time Trend

*	→	Refuse H_0 at 10%
**	→	Refuse H_0 at 5%
***	→	Refuse H_0 at 1%

In order to omit structural break effect and detect the right order of integration of the variables, two solutions where considered. Both regarding the inclusion of time dummies that account for trend changes in the data

Potential Solutions:

- Including yearly dummies to cancel yearly trends:

$$l\_RHPI_t = \alpha + \beta_1 RLTIR_t + \beta_2 UNEMP_t + \beta_3 l\_RGDPPC_t + \beta_4 l\_RHPI_{t-1} + \sum_{year=1996}^{2017} \beta_{(Q1+Q2+Q3+Q4),year} + \varepsilon_t \tag{7}$$

The idea is to introduce a time dummy for each year of the series, allowing to cancel the yearly trends and focusing on inter-quarter variation. Even though this method is effective in terms that avoid the issues with structural break, allowing all variables to be I(1), the final results prove that the variables are too powerful in terms of capturing the trend. With this specification, most relationships are captured by the dummies, causing the error correction mechanism to have a higher impact than it should. Also, the R-squared is not representative since most its explanatory power is caused by the dummies

- Including period dummies to cancel period trends.

$$l\_RHPI_t = \alpha + \beta_1 RLTIR_t + \beta_2 UNEMP_t + \beta_3 l\_RGDPPC_t + \beta_4 l\_RHPI_{t-1} + \beta_5 Break2007 + \beta_6 Break2013 + \beta_7 Time + \varepsilon_t \quad (8)$$

$$l_{RHPI_t} = \alpha + \beta_1 RLTIR_t + \beta_2 UNEMP_t + \beta_3 l_{RGDPPC_t} + \beta_4 l_{RHPI_{t-1}} + (1 + \beta_5 Break2007 + \beta_6 Break2013) * Time + \varepsilon_t \quad (9)$$

In the first equation the dummies have been introduced following an additive scheme, where the constant stand for the period from 2013Q3-2017Q4, Break2007→ 1995Q1-2007Q2, Break2013 → 2007Q2-2013Q2. With the additive scheme the model can take into account changes on the constant while the following scheme can capture changes in trends. The final procedure selected is the multiplicative scheme since it allows to stand for slope changes, and if time series plot is considered a change of slope can be observed. Check the appendix to see the time series plots. Even that the approach is similar to the one that introduces yearly trends, the fact that it cancel the effects of the 2 structural breaks instead of cancelling yearly trends allow for a better analysis on the period.

The idea of introducing time variables to the initial regression to later test residuals using ADF test was introduced by Perron (1989). In his paper, exposes different ways to introduce time dummies to avoid specification errors when checking for unitary root. The only difference with the standard ADF test is the inclusion of those dummies on the initial regression to latter normally check the error term unitary root (apply engle-granger test) searching for cointegration with special critical values (the values can be found on Perron's paper)

After applying Perron's methodology the unitary roots results prove to be more in line with what the literature states in general. Even that are several discussions on whether interest rates should be I(0) or I(1), and there is consensus on the fact that GDP should, in general, be I(1), is important to keep in mind that the studied period has several structural changes that can affect the results, contradicting many assumptions. The considered period starts after a crisis to later grow at a higher pace until half the period, where the 2008 crisis hits and makes the economy plumber like it wasn't observed since the '30s crisis. Afterwards the recovery takes place until 2013 where another recession takes place (in this case the impact of it have very different implication depending on the analyzed country). For that reason, the reader may find some results to be different from what it should be expected according to traditional literature. Due to that facts I removed structural breaks.

Spain		ADF		KPSS		Conclusion
		Level I(1)	Δ I(2)	Level I(1)	Δ I(2)	
Variable	l_RGDPPC	T -2,70522 0,2344 (6)	C -3,0636** 0,02939** (2)	T 0,43481*** p < 0,01*** (3)	- 0,309062 p > 0,1 (3)	I(1)
	l_RHPI	T -2,0113 0,5946 (2)	NC -2,34246** 0,01851** (6)	T 0,5215*** p < 0,01*** (3)	- 0,542773** 0,038** (3)	I(1)
	RLTIR	T -3,0294 0,124 (3)	NC -3,98945*** 6,76e-05*** (3)	T 0,26194*** p < 0,01*** (3)	- 0,159305 p > 0,1 (3)	I(1)
	UNEMP	T + Period Dummies -1,940* 0,0559* (2)	NC -3,19091*** 0,00170*** (0)	T 0,43481*** p < 0,01*** (3)	- 0,365782* 0,093* (3)	I(1)

() The number in brackets indicate The optimal number of lags for ADF

Since all the variables prove to be I(1), it will be possible to look for cointegration relationship to study the long run effect.

5.1.2 Other Countries

For the other countries the structural break wasn't significant enough to require applying Perron's methodology to check for unitary roots. As it can be seen in the appendix, Germany proves to have all its variables as I(1) except for Real GDP per Capita which is I(0). Italy has all its variables as I(1). Norway prove to have all its variables I(1) except for the interest rate which is I(0). Since GDP and Interest rate are often considered to be I(0), I will consider that is their correct order of integration preventing to check for cointegration by Engle-Granger procedure (but not with Johansen methodology as we will see later). Finally the UK results show contradictory results on the order of integration of GDP stating that it could possibly be I(2), since it make no economic sense, I will consider it to be I(1) (actually the null hypothesis of GDP being I(2) is only accepted with a p-value of 0.06). Real interest rates prove to be I(0) and the unemployment and House Price Index are I(1)

Even though we must properly apply cointegration test on the countries, we already can have an idea that, except of the Italian case, we might have some problems finding cointegration relationship for the rest of the countries. Cointegration relationship can only be found in the cases where all variables are integrated of the same order, but we still have to apply further testing to confirm that hypothesis. For my research I will only consider that to find cointegration relationships, all variables should be integrated of the same order, so in the case of Norway, Germany and UK I will only focus on the short run effect. To study long run relationship on models with different integration orders, an ARDL bound method (Pesaran, 2001) should be applied, which won't be considered in this research.

Keep in mind that in case of contradiction between KPSS test and ADF test I gave priority to the ADF test

5.2. Lag Structure

Before applying Johansen test to check for possible cointegration relationships, it is needed to determine the lag structure of the Vector Auto Regression for each model. This will allow to determine the number of lags that should be included according to BIC, AIC and HQC criterions. In my analysis I prioritized the results of the Bayesian Information Criteria (BIC) since is more robust asymptotically compared to other indicators.

The following table shows Spanish results when lag selection is applied. The results of other countries can be found in the appendices.

VAR system, maximum lag order 8

The asterisks below indicate the best (that is, minimized) values of the respective information criteria, AIC = Akaike criterion, BIC = Schwarz Bayesian criterion and HQC = Hannan-Quinn criterion.

lags	loglik	p(LR)	AIC	BIC	HQC
1	419.47708		-9.225645	-8.299619	-8.853390
2	509.01696	0.00000	-10.976594	-9.587556*	-10.418213*
3	527.80875	0.00173	-11.043065	-9.191015	-10.298557
4	550.57116	0.00012	-11.204075	-8.889012	-10.273439
5	561.55775	0.14406	-11.084708	-8.306632	-9.967945
6	599.22123	0.00000	-11.600506	-8.359416	-10.297615
7	616.11187	0.00581	-11.621711	-7.917609	-10.132693
8	637.19166	0.00037	-11.742659*	-7.575544	-10.067514

There are contradictory results between AIC against BIC an HQC criterions. BIC criterion will be the reference, so 2 lags will be considered on the model.

For the rest of the countries we also have contradictory results between information criterions, the final result will be the one stated by BIC. Italy, Norway, UK and Germany BIC's select 2 lags to be considered by Johansen cointegration test and VECM.

5.3. Cointegration Analysis:

5.3.1 Engle-Granger Approach

To check for cointegration relationships I considered two different tests. The Engle-Granger procedure and the Johansen test. Since the Engle-Granger method is based on identifying the order of integration of the variables to then regress them and check the unitary root of the residual, the first step is to identify the order of integration of the model's variables. The results can be observed on section 3.1. Since only Italy and Spain have all their variables having the same order of integration, it makes no sense to apply Engle-Granger approach to Germany, Uk and Norway since their variables aren't the same order of integration.

For Spain, the initial regression was computed using the following model:

$$l\_RHPI_t = \alpha + \beta_1 RLTIR_t + \beta_2 UNEMP_t + \beta_3 l\_RGDPPC_t + (\beta_4 Break2007 + \beta_5 Break2013) * Time + \varepsilon_t$$

Which is composed by I(1) variables and the time dummies have been introduced to reduce structural break issues. In order to detect possible correlation relationships, the error term of the previous regression should be I(0) using ADF test with MacKinnon (1991) special critical values

The following figure show the results for the Spanish model. As it can be seen, even that the null hypothesis of the error term being I(1) is rejected with a p-value of 0.00166, the test is using standard ADF critical values. The critical value for Engle-Granger test with 4 variables, no constant and 87 observations is -4.2 at 5% significance levels which is higher than the obtained statistic of -3.13, for that reason we accept null hypothesis of the errors being I(1).

```
Augmented Dickey-Fuller test for uhatprob
testing down from 11 lags, criterion BIC
sample size 87
unit-root null hypothesis: a = 1

  test without constant
  including 4 lags of (1-L)uhatprob
  model: (1-L)y = (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.182656
  test statistic: tau_nc(1) = -3.13848
  asymptotic p-value 0.00166
  1st-order autocorrelation coeff. for e: 0.032
  lagged differences: F(4, 82) = 8.663 [0.0000]
```

For the Italian case, we apply a standard regression in levels of the house prices as a function of the explanatory variables in levels, all the variables being I(1). I further test the presence of a unitary root on the regression residual like in the Spanish case. The next figure provide the results for the Italian case. As it can be observed, the obtained statistic is -1.95 which is lower than MacKinnon (1991) critical values of -4.2. That means that the null hypothesis is not rejected, considering the residuals as I(1).

```
Augmented Dickey-Fuller test for uhat
testing down from 4 lags, criterion AIC
sample size 88
unit-root null hypothesis: a = 1

  model: (1-L)y = (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.0406632
  test statistic: tau_nc(4) = -1.95846
  asymptotic p-value 0.7444
  1st-order autocorrelation coeff. for e: -0.045
  lagged differences: F(3, 84) = 6.799 [0.0004]
```

According to Engle-Granger tests there are no cointegration relationships in both Spain and Italy. But according to Bilgili (1998) Engle-Granger test is more suitable to study bi-variate models to find one cointegration relationship while Johansen may be more adequate to study cointegration relationships on a vectorial context. For that reason, it can be possible to obtain different results from Engle-Granger and Johansen Tests.

5.3.2 Johansen tests

The following figure shows the results for either Trace tests and Maximum eigenvalue tests. Again, since Spain and Italy are the only countries whose variables are all I(1) I only applied Johansen test to them because it would make no sense to search for cointegration relationships on models that have different orders of integration.

Lmax Test	**Spain**	**Italy**
None	40,078 (0,0016)	28,628* (0,0885)
At most 1	17,431* (0,3204)	20,282 (0,1592)
At most 2	13,163 (0,1747)	10,257 (0,3809)
At most 3	3,0372 (0,0814)	4,7302 (0,0296)

* indicates the cointegrating rank of π

Trace Test	**Spain**	**Italy**
None	73,710 (0,0004)	63,897 (0,0062)
At most 1	33,631* (0,0729)	35,269 (0,0491)
At most 2	16,201 (0,990)	14,987* (0,1427)
At most 3	3,0372 (0,0860)	4,7302 (0,0325)

* indicates the cointegrating rank of π

As the results show, the Trace test and Lmax test provide different conclusions on the number of cointegrating relationships. According to Saikkonen, Lütkepohl and Trenkler (2000) in general both tests are similar although asymptotically, Trace test exhibit higher power performance specially when detecting two or more cointegration relationships, that is why many researchers choose Trace test over maximum eigenvalue.

For Spain, both tests detect the presence of one cointegration relationship. The Trace test for the presence of 1 cointegration relationship against the alternative of the presence of 2 or more is only rejected at 10%. Since I set the significance level at 5% there is no enough evidence to reject the null hypothesis of the presence of at most 1 cointegrating relationship.

For Italy, the Trace test indicates the presence of two cointegrating relationships: The null hypothesis on cointegrating rank is 2 is accepted against the alterative hypothesis of π = 3). The Lmax test, by the contrary, exhibit no presence of any cointegrating relationship since initial null hypothesis of rank(π) = 0 is accepted (only rejected at 10% level).

Following the analysis of Saikkonen, Lütkepohl and Trenkler (2000) I will give priority to Trace test over Maxim Eigenvalue.

5.4 Long Run Estimation

Since I could only find cointegration relationships for Spain and Italy, the long run analysis will only be applied to those countries. The first step in the process is to compute the vector of cointegration relationships that will form the VECM. In order to compute it, we follow the procedure described in the technical background sector. The following figures show the alfa and beta vector for Spain (1 vector) and Italy (2 vectors). In the appendix section there is the plot for the VECMs. Since we center the analysis on explaining the evolution of house prices the cointegrating vectors have been normalized to get house price coefficients equal to 1

```
beta (cointegrating vectors, standard errors in parentheses)

l_RHPI          1.0000
              (0.00000)
l_RGDPPC       -13.719
               (2.9132)
RLTIR          0.18689
             (0.054417)
Unemp         -0.11986
             (0.038151)

alpha (adjustment vectors)

l_RHPI       -0.012304
l_RGDPPC     0.0033924
RLTIR         -0.55459
Unemp        -0.016917
```

Cointegrating Vector for Spain

```
beta (cointegrating vectors, standard errors in parentheses)

l_RHPI          1.0000        0.00000
              (0.00000)      (0.00000)
l_RGDPPC       0.00000         1.0000
              (0.00000)      (0.00000)
UNEMP        -0.068910       0.054639
              (0.16635)     (0.050863)
RLTIR         -0.94090        0.28496
              (0.17043)     (0.052109)

alpha (adjustment vectors)

l_RHPI       -0.019640      -0.074228
l_RGDPPC     -0.032373       -0.11092
UNEMP         -0.99773        -3.0658
RLTIR           2.0980         6.0561
```

Cointegrating Vectors for Italy

During the computation of the VECM I included a constant and their respective dummy trends as exogenous variables. In the case of Spain we can identify three different trends. The variable timebreak 2007 captures the trend form 1995Q1-2007Q2 the variable timebreak 2013 compress the period 2007Q3-2013Q2 and the standard variable trend captures the remaining period 2013Q3-2017Q4. Since Italy have the structural breaks on different periods, the dummies were adapted to cover the exact shifts. The Italian trend can be broken down to 3 periods (1995Q1-2003Q2 → timebreak2003, 2003Q3-2012Q2 → timebreak2012, 2012Q3-2017Q4 → Time). Is important to add that the trend variables are only included on the VECM, since adding them to the long run regression would imply to over constrain the model.

The following tables show the results for the long run model for the house prices as endogenous variable. Since I applied a vectorial framework, the regressions for the other variables being endogenous can be found on the appendix section.

Model 2: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

	coefficient	std. error	t-ratio	p-value	
const	0.0227824	0.0790465	0.2882	0.7739	
d_l_RGDPPC_1	0.139768	0.227887	0.6133	0.5413	
d_RLTIR_1	-0.00697790	0.00142619	-4.893	4.74e-06	***
d_Unemp_1	-0.00373980	0.00281185	-1.330	0.1871	
ErrorCorrection	0.000284443	0.000988304	0.2878	0.7742	
d_l_RHPI_1	0.720561	0.0775199	9.295	1.51e-014	***

Mean dependent var	0.004975	S.D. dependent var	0.024337
Sum squared resid	0.016962	S.E. of regression	0.014210
R-squared	0.678233	Adjusted R-squared	0.659081
F(5, 84)	35.41174	P-value(F)	2.52e-19
Log-likelihood	258.2416	Akaike criterion	-504.4832
Schwarz criterion	-489.4844	Hannan-Quinn	-498.4348
rho	-0.107367	Durbin's h	-1.503175

Excluding the constant, p-value was highest for variable 97 (ErrorCorrection)

Long Run Estimation for Spain

Model 1: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

	coefficient	std. error	t-ratio	p-value	
const	0.420773	0.108812	3.867	0.0002	***
d_l_RGDPPC_1	-0.0859063	0.0744405	-1.154	0.2518	
d_UNEMP_1	-0.00453703	0.00224116	-2.024	0.0461	**
d_RLTIR_1	0.00105485	0.00119182	0.8851	0.3787	
ErrorCorrection1	-0.0182105	0.00487464	-3.736	0.0003	***
ErrorCorrection2	-0.0672156	0.0174923	-3.843	0.0002	***
d_l_RHPI_1	0.651610	0.0698591	9.327	1.44e-014	***

Mean dependent var	0.000057	S.D. dependent var	0.013053
Sum squared resid	0.002116	S.E. of regression	0.005049
R-squared	0.860438	Adjusted R-squared	0.850349
F(6, 83)	85.28626	P-value(F)	2.22e-33
Log-likelihood	351.9031	Akaike criterion	-689.8062
Schwarz criterion	-672.3075	Hannan-Quinn	-682.7497
rho	-0.071498	Durbin's h	-0.905781

Excluding the constant, p-value was highest for variable 17 (d_RLTIR_1)

Long Run Estimation for Italy

Observing the results we can observe that Italy has both error correction terms significant and lower than one. After apliyng differences and having structural break corrected, the model most probably does not estimate causal relationships. The only significant variables for Italy are the lag of the differences of house prices with a coefficient of 0.65 (both in logs, so an increase of last quarter's house prices imply an increase of 0.65% on the present) and the unemployment with a coefficient of -0.0045 (since the house prices are in logs, the correct interpretation is that increasing preavoius quarter unemployment rate by 1% the present house prices growth is reduced by 0.45%) The lagged differences of the interest rate and GDP per capita turn out to be non-significat.

For the spanish case, it can be ovserved that the error correction is neither negative nor significant. That may be due to the behaviour of the differences of house prices. Even that we have added the trend dummies on the error correction term, the price growth still have some trend components in its behaviour. For that reason, we run the model again adding the trend dummies into the long run regression

```
Model 3: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

                  coefficient    std. error    t-ratio   p-value
  ---------------------------------------------------------------
  const           -0.944079      0.371520      -2.541    0.0130   **
  d_l_RGDPPC_1    -0.455388      0.239462      -1.902    0.0608   *
  d_RLTIR_1       -0.00144846    0.00167895    -0.8627   0.3908
  d_Unemp_1       -0.000967116   0.00294172    -0.3288   0.7432
  ErrorCorrection -0.0123039     0.00483028    -2.547    0.0128   **
  break2007time    0.000367754   0.000115272    3.190    0.0020   ***
  break2013time   -0.000417899   0.000112023   -3.730    0.0004   ***
  time            -0.000660878   0.000295058   -2.240    0.0278   **
  d_l_RHPI_1       0.276776      0.108616       2.548    0.0127   **

Mean dependent var   0.004975   S.D. dependent var   0.024337
Sum squared resid    0.012652   S.E. of regression   0.012498
R-squared            0.759998   Adjusted R-squared   0.736294
F(8, 81)             32.06211   P-value(F)           4.55e-22
Log-likelihood       271.4347   Akaike criterion    -524.8693
Schwarz criterion   -502.3710   Hannan-Quinn        -515.7967
rho                 -0.050599   Durbin-Watson        2.100940

Excluding the constant, p-value was highest for variable 75 (d_Unemp_1)
```

Long Run Estimation for Spain with trend dummies

After applying the trend dummies the model is much more responsive. The error correction term is significant, negative and lower that one. In the Spanish case the variables d_RLTIR_1 and d_Unemp_1 are not significant. D_l_RHPI_1 is significant with a coefficient of 0.27 which is much lower than the Italian case. That would indicate that in the Italian case the effect of past prices tends to have a greater impact on the growth of current prices. The d_l_RGGPPC_1 is also significant with a negative coefficient of -0.4553. the fact that both Spain and Italy have negative coefficients for the lagged GDP per capita growth (even that wasn't significant for Italy) may indicate a price rigidity where prices adapt on the next period (check sign interpretation section). If the model is considered without dynamics, we can find a positive correlation between gdp per capita and house prices as it will be shown in the short run model, but when dynamics are added into the model the relation gets inverted. This phenomenon will be properly explained on the interpretation section. The exogenous variables are lagged to properly apply the long run estimation following VECM procedure described on the Technical background section.

Before analyzing the results, we need to make sure that there are no specification errors on the model. For that reason, we run several tests on the model:

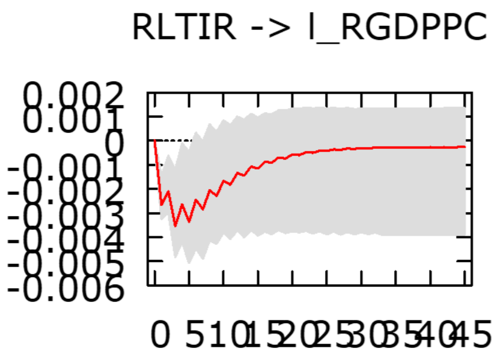
- Residual Normality: To study residual behavior on a vectorial context, we use Doornik-Hansen test for residual normality. The null hypothesis of residuals ~ $(0, \sigma^2)$ is not rejected for both Spain (p-value = 0.2716) and Italy (p- value = 0.0562) with a significance level of 5%. Is relevant to add that, in the case of Spain, residuals only behave normally when dummy trends are added to the long run model. Since their omission would be captured by the residuals breaking normality assumption. Italy on the contrary have its residuals behaving properly only if the time dummies are added on the VECM but is not necessary to include them on the long run model.

- Model linearity: We run a Reset test for non-linear specification considering squares, cubes and squares+cubes on both models. The results indicate that the linear specification is correct on both models (check results on the appendix section). Again, the models show linearity issues if trends are not considered.

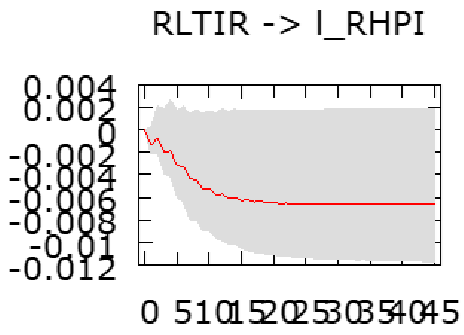
- Heteroskedasticity: To analyze possible heteroskedasticity issues we apply white test and Breusch and Pagan test on both models. For Spain, both tests do not reject the null hypothesis of No Heteroskedasticity. The Italian model accepts null hypothesis on white test but the Breusch and Pagan test indicate the presence of Heteroskedasticity on the model which would imply a loss of efficiency of the estimators.

- Autocorrelation: we find no autocorrelation issues on either model when applying the tests. The tests considered a total of 4 lags to account a full year instead of a single quarter.

Since the models seems to be correctly specified we will analyze the effect of a random shock on the variables to analyze the behavior of the vector of variables. On the appendix section I provide the plots for the shock effects on the 4 variables. In both cases we get an interesting behavior. On one hand we can check how the Error Correction term reacts to the shock, minimizing their effects and stabilizing them. On the other hand, we can see how the shock doesn't converge to 0 in both cases. That phenomenon may be caused to endogenous effects of the shock, having an impact on the long run by shifting the steady state. The results are relevant since we can observe both effects: the stabilizing effects of the Error Correction terms and the shifting of the long run steady state when a shock is applied on any of the variables. In a more complex analysis we could disaggregate the behavior by analyzing the effects separately. To illustrate the effect the following figure shows, for Spain, how a positive shock on the interest rate has a negative effect of GDP per capita. Initially the shock has a great impact but the EC term stabilize the shock. Is important to notice that the after the initial shock, the GDP per capita doesn't return to its initial level but to a lower one. The reason of it is that the initial shock on the interest rate have an effect on the long run steady state by lowering the long run levels of GDP per capita. Similarly the second figure illustrate the same effect on a shock on Spanish long term interest rate on House Prices



EC stabilizes the shock. Initial l_RGDPPC = 0, final l_RGDPPC = -0.0003



EC stabilizes the shock. Initial l_RHPI = 0, final l_RHPI = -0.0063

5.5 Short Run Estimation

The short run analysis has been computed according to the following model:

$$\Delta \mathrm{l\_RHPI}_t = \alpha + \beta_1 \Delta \mathrm{l\_RGDPPC}_t + \beta_2 \Delta \mathrm{UNEMP}_t + \beta_3 \Delta \mathrm{RLTIR}_t + \beta_1 \Delta \mathrm{l\_RHPI}_{t-1} + u_t$$

Apart from not having the EC term, notice that the explanatory variables introduced are from the same period as the prices, so we do not introduce dynamics except for the lagged prices. In the appendix section there are the estimation results for all the countries. Notice that except for Germany, all the coefficients have the same sign which indicates that we are capturing similar behaviors.

Before analyzing the results, we need to make sure that the model doesn't suffer from specification issues. For this reason we test the following (Results on the appendix section):

- Linearity: We apply Reset test on the models considering squares, cubes, and squares + cubes. Any of the cases suffer from linearity issues according to the results.

- Heteroskedasticity: I apply white test and Breusch and pagan test to detect possible heteroskedasticity issues. UK, Spain do not refuse null hypothesis of homoscedasticity on both tests. Italy accepts the null on BP test and rejects it with White test which may indicate an efficiency loss of the estimators. Germany and Norway reject the null hypothesis on both tests due to the presence of an influential observation (outlayer) on 2007Q1 (Germany) and 2009Q1 (Norway). If those observations are removed, the issues are fixed and homoscedasticity is achieved. But for comparison results I will stay with the full range of observations even considering the resulting efficiency loss.

- Residual Normality: Chi squared test is applied on the residuals to check if they follow a normal distribution. Spain and Norway accept the null hypothesis of errors behaving normally while Italy, Germany and United Kingdom rejects it. On the Italian and German case, the issue is caused by an influential observation just before the structural break (2003Q2 and 2007Q1 respectively) while on the UK case the issue is caused by an autocorrelation with 4th lag of house prices growths ($\Delta \text{l\_RHPI}_{t-4}$). If influential observations are removed (Italy and Germany) and $\Delta \text{l\_RHPI}_{t-4}$) is introduced on UK model, the residuals behave following a normal distribution. For comparison purposes the analysis is done without correcting for those issues causing a efficiency loss of the estimators.

- Autocorrelation: To test autocorrelation we apply LM, Breusch and Godfrey tests and L-jung Box Q to detect AR (1) and AR (4). Quarterly data is very sensitive to lag 4 due to yearly differences apart from regular quarterly differences (AR(1)). All countries except UK accept null hypothesis of no autocorrelation on the two tests for either AR(1) and AR(4). The UK accepts the null hypothesis of AR(1) on both tests but it rejects it for AR(4) on both tests. The autocorrelation on UK model is causing the residuals to not behave normally. The problem can be fixed by adding $\Delta \text{l\_RHPI}_{t-4}$ on the UK model but again for comparison purposes the model will remain as it is.

The first thing to notice is that the behavior of German prices' growth may exhibit a time trend. If the trend is not added, any of the variables is significant and the R-squared remains at 0.019. For the sake of explaining significance of the parameters I will assume that the trend is added into the German model. As for the significance of the parameters we find that Real GDP per capita is only significant for Italy and UK. The real long term interest rate is significant for all countries except for Germany. The unemployment is only significant for Germany and the lag of house price growth is significant for all countries. As for the explanatory power of the model, we can see how Italy and Spain have very high R-squares (0.85 and 0.71 respectively). UK and Norway have lower but still relevant R-squares (0.60 and 0.42 respectively) and German model is the less predictive one with only an R-squared of 0.28 (with time trend) and 0.019 (without time trend). That fact could explain the differences in coefficients' sign on the German case, but the discussion will be properly developed on the sign interpretation section.

Estimation Results	SPAIN		ITALY		GERMANY	NORWAY	UK
	Short Run	Long Rung	Short Run	Long Run			
Constant	0,000469	-0,944**	0,000277	0,468**	-0,0139***	0,0059***	0,00273*
Δl_RGDPPC$_t$	0,3483	-	0,1632**	-	0,1203	0,05117	0,4251**
Δl_RGDPPC$_{t-1}$	-	-0,4553*	-	-0,1174	-	-	-
ΔRLTIR$_t$	0,00503***	-	0,00358***	-	0,00102	0,00736***	0,00464**
ΔRLTIR$_{t-1}$	-	-0,00144	-	-0,00046	-	-	-
ΔUNEMP$_t$	-0,00345	-	-0,16327	-	0,0111*	-0,00841	-0,00686
ΔUNEMP$_{t-1}$	-	-0,00096	-	-0,00548**	-	-	-
Δl_RHPI$_{t-1}$	0,7213***	0,2767**	0,85312***	0,5853***	-0,2393***	0,5327***	0,6031***
EC 1	-	-0,01230**	-	-0,0182***	-	-	-
EC 2	-	-	-	-0,0672***	-	-	-
R^2	0,710	0,759	0,853	0,860	0,284	0,425	0,6076
Durbin Watson	2,297	2,101	1,917	2,740	1,974	1,898	1,904

(***) Significant at 10%, (**) Significant at 5%, (*) Significant at 1%

5.6 Sign Interpretation

As it can be seen on the results, the long run and short run coefficients seem to affect very differently on growth of house price index but these discrepancies have not to do with the temporal horizon set but more with the dynamics within the model. I've computed the short run analysis of all countries but adding first order lags to properly study their behavior. The next figure summarizes the coefficients' sign and their significance for both the long run and short run analysis of the difference of real house prices.

Short Run					
	SPAIN	ITALY	UK	GERMANY	NORWAY
Δl_RGDPPC_t	+	+*	+*	+	+
Δl_RGDPPC_{t-1}	-	-	-	-	-
$\Delta RLTIR_t$	+*	+*	+*	+	+*
$\Delta RLTIR_{t-1}$	-*	-	-*	+	-
$\Delta UNEMP_t$	-	-	-	+*	-
$\Delta UNEMP_{t-1}$	-	-*	+	+	+
Δl_RHPI_{t-1}	+*	+*	+*	-*	+*
* indicates significance at 10%					

Long Run		
	SPAIN	ITALY
Δl_RGDPPC_t		
Δl_RGDPPC_{t-1}	-*	-
$\Delta RLTIR_t$		
$\Delta RLTIR_{t-1}$	-	-
$\Delta UNEMP_t$		
$\Delta UNEMP_{t-1}$	-	-*
Δl_RHPI_{t-1}	+*	+*
* indicates significance at 10%		

Having a look onto the coefficients it can be seen how the discrepancies of the coefficients' sign has not to do with the short run against long run analysis, but more with the model's dynamics. The coefficients of Spain and Italy on the long run have the have the same sign than the ones on the short run model if we consider lags of the explanatory variables. The only country that show opposite signs on the short run is Germany. Considering its low R-squared and its overall lack of explanatory power we won't rely on the German results, treating them as non-significant. The only discrepancy can be found on the first lag of unemployment and will be properly explained on the following section:

5.6.1 Understanding the dynamics

Δl_RGDPPC_t : All coefficients have the same positive sign on the contemporary House Price Levels. The reason is that when output is increased, so it does the aggregated demand formed by consumption, investment and public expenditure. When demand increases the prices automatically increase since we consider the supply constant on the short run. This behavior can be significantly observed on UK with an elasticity of 0.49 and Italy whose coefficient is equal to 0.16. For the other countries the variable is not significant but still positive.

$$\uparrow Y_t \rightarrow \uparrow (C_t, I_t, G_t) \rightarrow \uparrow AD_t \rightarrow \uparrow \pi_t \rightarrow \uparrow House\ Prices_t$$

Δl_RGDPPC_{t-1} : When considering past increases of output, we face the continuation of the previous effect which stated that an increase of output has a positive impact on the contemporaneous house prices. Considering that stock of houses doesn't significantly change across a single quarter, the increase of past output that causes an increase of past house prices affects negatively to present aggregated demand, causing the current price to decrease. The variables don't appear as significant in any of the considered models but all provide the same sign

$$\uparrow Y_{t-1} \rightarrow \uparrow(C_{t-1}, I_{t-1}, G_{t-1}) \rightarrow \uparrow AD_{t-1} \rightarrow \uparrow \pi_{t-1} \rightarrow \uparrow \text{House Prices}_{t-1} \rightarrow \downarrow AD_t, \uparrow AS_t \rightarrow \downarrow \text{House Prices}_t$$

$\Delta RLTIR_t$: The variable appears positive on all models, being Germany the only model on which is not significant. Since the variable is not in logs, to get the real elasticities we have to multiply the coefficient by 100, ordered from highest to lowest: 0.73 (Norway), 0.5 (Spain), 0.46 (UK), 0.35 (Italy) and 0.1 (Germany) This case is an example of price rigidities. From an investor's point of view, the expected return of the investment is set with the available information. When there is a sudden increase on the real long run interest rate it also increases the opportunity cost for the investment (alternative profitability) and produce an increase of the financing cost (which reduces investment profitability). Due to this fact, the current house price has to be increased to maintain original profitability levels. The increase of real interest rates also imply a decrease of investment (negatively correlated) causing a decrease in the growth rate of house stock (supply) and therefore increasing House Prices.

$$\uparrow r_t \rightarrow \downarrow I_t \rightarrow \downarrow \Delta \text{Housing Stock}_t \text{ (Supply)} \rightarrow \uparrow \text{House Prices}_t$$

$\uparrow r_t$ ↗ $\uparrow$Financial $\text{Costs}_t \rightarrow \downarrow$ Investment Profitability_t ↘

$\uparrow r_t$ ↘ $\uparrow$Opportuniy $\text{Costs}_t \rightarrow \uparrow$ Alternative Profitability_t ↗

If we want actual profitability constant: $\uparrow$House Prices_t

To ilustrate this effect I will provide a numerical example: Suppose that on a single quarter, an investor is considering to buy a house from Sareb's stock at a price of 800.000€. The investor knows that if it reforms the property and make use of its extensive commercial network and know how he can achieve a profitability of 50%. He ask for financement to purchase the house, develop it and re-sellit within a month. If interest rates are currently at 25% he will have financial costs of 200.000€. Total cost for the investor is 1.000.000€ and if he want to get a 50% profit he will have to re-sell the property for 1.500.000. Suposse now that there is a sudden increase of the interest rate to 50%. Now the investor have to face a cost of 400.000€, having a total cost of 1.200.000€. If he sell the property at the same price as before, he will only get a 25% profitability. See how the increase of the interest rate is not only increasing its financial costs (lowering investment profitability) but also increases the opportunity cost for him (why would he do an investment that offer a 25% profitability when the new risk-free interest rate is now offering a 50% return?). If he want to obtain a profitability of 50% on the investment he will have to sell the house at 1.800.000 € (now he would be technically indiferent between doing the investment and not doing it). Since this operation has occurred on the same quarter, we can see how an increase of current interest rates can positively affect current house prices.

$\Delta RLTIR_{t-1}$: All countries except for Germany have a negative sign for the lagged real long term interest rate. It is significant just for Spain and the UK. The relation is negative due to lack of price rigidities when more than one period is considered. When interest rates are increased, its effects may not immediately affect current investment since the increase may be partially transferred to the current price. But it will have a notably effect on future investments (now the increase is not sudden and the expectations are re-computed accordingly).

$$\uparrow r_{t-1} \rightarrow \downarrow I_t \rightarrow \downarrow AD_t \rightarrow \downarrow \pi_t \rightarrow \downarrow House\ Prices_t$$

$\Delta UNEMP_t$: The unemployment appears to be negative and non-significant for all countries except for Germany which by the contrary is positive and significant. Without considering German case due to lack of explanatory power, the unemployment will have a negative impact on current house price's growth. When unemployment is increased, the current aggregated demand will decrease (Okun's Law) and so will do the inflation (Philips curve) so it makes economic sense to find that relationship

$$\uparrow UNEMP_t \rightarrow \downarrow Y_t \rightarrow \downarrow \pi_t \rightarrow \downarrow House\ Prices_t$$

$\Delta UNEMP_{t-1}$: The lag of the unemployment rate provide mixed results. Is only significant for Italy and is negative only for Italy and Spain. On the contrary, is positive for UK, Germany and Norway. This discrepancy may be related on the labor structure of each country. Italy and Spain are characterized for having a high rate of structural unemployment due to labor legislation, mismatch between supply and demand and productive structure. If a worker gets unemployed in Italy or Spain he has a much higher possibility of remaining unemployed for more than 3 months, that could explain why the previous unemployment has the similar effect on house prices than the current one for those countries. Thanks to the labor regulation's flexibility on the UK and Germany and the active policies on labor creation in the case of Norway, the chance of being unemployed for more than one quarter is significantly lower than the Mediterranean counterparts. Because of that, the coefficient isn't negative for those countries.

$$\uparrow UNEMP_{t-1} \rightarrow \downarrow Y_t \rightarrow \downarrow \pi_t \rightarrow \downarrow House\ Prices_t \text{ (Spanish and Italian Cases)}$$

Δl_RHPI_{t-1}: The lagged price growth appears significant on all countries and is positive for all except for Germany (again due to lack of explanatory power of the model on German prices). The effect is positive since it captures the cyclical component of house prices. As stated on the data section, the appraisal value of a house tend to be very correlated with past realizations of the same variable, causing what is known as Appraisal Smoothing. Also, the house price tend to have longer cycles than a quarter, for that reason it makes economic sense that the present house price is very influenced by the past realization of it. This is notably relevant on periods where the demand is high and supply remains relatively constant due to time constrain. Another reason is the speculative component of residential sector depicted by the generalized housing bubble that took place during most of the period.

$$\uparrow \textit{House Prices}_{t-1} + \textit{High } AD_t + {\sim}AS_t \rightarrow \uparrow \textit{House Prices}_t$$

VI. CONCLUSIONS

After the empirical research, there are three main conclusions to extract. The first one is related with how the variables Real House Price Index, Real GDP per capita, Unemployment and Real Long Term Interest Rate interact with each other on the long run. After finding cointegration relationships for Spain and Italy that allowed us to study the long run equilibrium using VECM procedure, we discovered that a shock on any of the variables can have a permanent effect on the model by shifting the steady state. Even with the introduction of the Error Correction term, the shock gets stable but it doesn't fade out completely. This could mean that the initial shock was not exogenous as originally assumed but rather endogenous. The fact that Germany, Norway and UK variables are not the same order of integration caused that no possible cointegration relationships could be found, forcing to analyze only the short run equilibrium by applying the model in first differences.

The second conclusion has to do with how sensitive Real House Prices are to the macro-economic conditions. By regressing the log of house price growth as a function of real long term interest rates, log of real GDP per capita, unemployment rate and the first lag of the log of house prices growth we obtain very different results on the explanatory power of the model. For instance, the model explains a 85% of the evolution of Italian house price's growth, 71% for the Spanish ones, 60% for the UK, 42% for the Norwegian and only 1%-28% (depending on if a trend is added or not) for the German case. The differences in predictability can be related to institutions and idiosyncrasies of each model. In the German case is the less responsive to the model due to housing regulation that promotes residential lease with low yields, which cause the German market to be non-as profitable and attractive to investors. Also, the fact that German house prices follow an opposite behavior than most countries (lower prices before the crash) affects the model explanatory power. In the Nordic case, the model is affected by the fact that Norway has control over its monetary policy, controlling its interest rates. As a result, their price levels become independent from other European countries who share a common monetary policy. The low unemployment of Norway and its low volatility also has an effect on the explanatory power of the model. Finally, the Norwegian house bubble also breaks the relation between GDP per capita and house prices because prices are driven by the growth expectations rather than linked with increases of real GDP per capita. The model has a better explanatory power in the case of the UK due to the liberalization of the sector. The fact that Anglo-Saxon economies have a lax regulation allow prices to adjust freely and therefore become more sensitive to macroeconomic conditions. Finally, in the Italian and Spanish cases is where the model has a greater explanatory power. The southern-European economies where very affected by the expansionary monetary policies from the beginning of the period until the crash. The ease of borrowing has been a crucial factor in the housing bubble created, and this behavior is captured by the model (both countries have significant impact of interest rates during the period). From the beginning of period interest rates and real GDP per capita grow affecting also the evolution of house prices (which have a faster growth due to speculative component). After the 2008 crash, the sudden drop on income, increase of unemployment and drecrease of investment cause the housing bubble to break, allowing prices to fall following the trend of the macroeconomic conditions.

The final conclusion is that the results do not differ as much between long run estimation and short run estimation, but rather on the impact that the exogenous variables have on house price index growth when dynamics are considered. Real GDP per capita, real long term interest rates and unemployment (in case of Spain and Italy) have a an inverse effect when past realizations of the variables are introduced into the model. The case of real long term interest rates is especially relevant since we can find that in the very short run the price rigidities of the residential sector may break the most basic economic intuitions. A sudden increase of the long term rates doesn’t imply a price decrease but rather an increase on the current house prices. When the timeframe is widened the effect reverses, becoming more in line with what the traditional economic literature suggests.

VII. BILBIOGRAPHY

Bailey, Martin, Richard: Muth, and Hugh Nourse. 1963. "A regresion Method for Real Estate Price Index Construction." *Journal of the American Statistical Association* 933-942.

Bilgili, Faik. 1998. "Stationarity and cointegration tests: Comparison of Engle - Granger and Johansen methodologies." *Munich Personal RePEc Archive.*

Case, K.E, and R.J Shiller. 1987. " Prices of single-family homes since 1970: new indexes." *New England Economic Review* 45-56.

Cesa-Bianchi, Ambrogio. 2012. "Housing Cycles and Macroeconomic Fluctuations: A Global Perspective." *Inter-American Development Bank.*

Clayton, J, D Geltner, and S,W Hamilton. 2001. "Smoothing in Commercial Property Valuation: Evidence from Individual Appraisals." *Real Estate Economics* 337-360.

Cooley, Thomas, and Edward Prescott. 1995. "Economic Growth and Business Cycles."

Coulson, N.E,, and M.S. Kim. 2000. "Residential Investment, Non-residential Investment and GDP." *Real Estate Economics* 233-247.

Court, M. 1939. "Hedonic Prices Indexes with Automotive Examples." *The Dynamics of Automobile Demand* 98-119.

Engle, Robert, and C Granger. 1987. "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55 (2): 251-276.

Eurostat. 2013. *Handbook on Residential Property Prices Indices.* Luxembourg: Publications Office of the European Union.

Fisher, J, M Miles, and R Webb. 1999. "How Reliable are Commercial Property Appraisals? Another Look." *Real Estate Finance* 9-15.

Gertner, D. 1989. "Bias in Appraisal-Based Returns." *Journal of American Real Estate and Urban Economics Association* 338-352.

Granger, and Newbold. 1974. "Spureous Regression in econometrics." *Journal of Econometrics* (Journal of econoemtrics).

Green, R.K. 1997. "Follow the Leader: How Changes in Residential and Non-residential Investment Predict Changes in GDP." *Real Estate Economics* 253-270.

Hirata, Hideaki, Ayhan Kose, Otrok Christopher, and Marco Terrones. 2013. "Global House Price Fluctuations: Synchronization and Determinants." *IMF Working Paper.*

Hirata, Hideaki, Ayhan Kose, Otrok Christopher, and Marco Terrones. 2011. "Global House Price Fluctuations: Synchronization and Determinants." *IMF Working paper.*

Ibbotson, R, and L.B Siegel. 1984. "Real Estate Return: a Comparison with Other Investment." *AREUEA* 219-242.

Johansen, Søren. 1995. *Likelihood-based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press.

Kurita, Takamitsu. 2010. "Econometric Modeling of the Interactions between Japan's Housing Investment and GDP." *International Journal of Applied Economics* 56-71.

Leung, and Chen. 2006. "Intrinsic Cycles of Land Price: A Simple Model." *Journal of Real Estate Research* 293-320.

Lourenço, Rita, and Paulo Rodrigues. 2014. "Dynamics and Contrast of House Prices in Portugal and Spain." *Banco de Portugar, Economic Bulletin.*

MacKinnon, J. G. 1991. *Critical values for cointegration tests.* Oxford, Oxford University Press.

Mankiew, Gregory, and David Weil. 1989. "The Baby Boom, The Baby Bust and the Housing Market." *Regional Science and Urban Economics* 235-258.

Meidani Ali, Zabihi Maryam, Ashena Malihe. 2011. "House Prices, Economic Output and Inflation Interactions in Iran." *Research in Applied Economics.*

Nagaraja, Chaitra H., Lawrence D. Brown, and Linda H. Zhao. 2011. "An Autoregressive Approach to House Price Index." *The Annals Of Applied Statistics* 124-149.

Nakajima, Makoto. 2011. "Understanding House-Price Dynamics." *Business Review* 20-28.

Otrok, Christopher, and Marco Terrones. 2005. "House Prices, Interest Rates and Macroeconomic Fluctuations: International Evidence." *IMF working papers.*

Perron, Pierre. 1989. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica* 1361-1401.

Pesaran, M.H, t Schuermann, and M Weiner. 2004. "Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model." *Journal of Business & Economic Statistics* 129-162.

Piazzesi, Monika, and Schneider Martin. 2009. "Momentum Traders in the Housing Market: Survey Evidence and a Search Model." *American Economic Review* 406-411.

Pissarides, Christopher, and Jonathan Wadsworth. 1989. "Unemployment and the Inter-regional Mobility of Labour." *The Economic Journal* 739-755.

Quan, D,C, and J,m Quigley. 1991. "Price Formation and the Appraisal Function in Real Estate Markets." *Journal of Real Estate Finance and Economics* 127-146.

Reinhart, Carmen S., and Kenneth S. Rogoff. 2009. *This Time is Different: Eight centuries of financial folly.* Princeton University Press.

Renaud, B. 1995. "The 1985-94 global real estate cycle : its casues and consequences." *Policy Research Working Paper.*

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 34-55.

Saikkonen, Lütkepohl, and Trenkler. 2000. *Maximum Eigenvalue Versus Trace Tests for the Cointegrating Rank of a VAR process.* ECONSTOR.

Sutton, D. Gregory, Dubravko Mihaljek, and Sub. 2017. "Interest rates and house prices in United States and around the world." *Bank for international settlements working papers.*

Vanstreenkiste, Isabel. 2007. *Regional housing market spillovers in the US - lessons from regional divergences in a common monetary policy setting.* Working Paper, European Central Bank.

Vivarelli, Marco. 1997. "The Economics of Technology and Employment: Theory and Empirical Evidence." *The Economic Journal* 239-241.

Wheaton, C William. 1999. "Real Estate “Cycles”: Some Fundamentals." *Real Estate Economics.*

Xiao, P. 2015. "An empirical analysis of factors influencing China's real estate prices-based on the VAR model." *Lingnan Journal of Banking, Finance and Economics.*

VIII. Appendices

8.1. Results Summary Table

Estimation Results	SPAIN		ITALY		GERMANY	NORWAY	UK
	Short Run	Long Rung	Short Run	Long Run			
Constant	0,000469	-0,944**	0,000277	0,468**	-0,0139***	0,0059***	0,00273*
Δl_RGDPPC_t	0,3483	-	0,1632**	-	0,1203	0,05117	0,4251**
Δl_RGDPPC_{t-1}	-	-0,4553*	-	-0,1174	-	-	-
$\Delta RLTIR_t$	0,00503***	-	0,00358***	-	0,00102	0,00736***	0,00464**
$\Delta RLTIR_{t-1}$	-	-0,00144	-	-0,00046	-	-	-
$\Delta UNEMP_t$	-0,00345	-	-0,16327	-	0,0111*	-0,00841	-0,00686
$\Delta UNEMP_{t-1}$	-	-0,00096	-	-0,00548**	-	-	-
Δl_RHPI_{t-1}	0,7213***	0,2767**	0,85312***	0,5853***	-0,2393***	0,5327***	0,6031***
EC 1	-	-0,01230**	-	-0,0182***	-	-	-
EC 2	-	-	-	-0,0672***	-	-	-
R^2	0,710	0,759	0,853	0,860	0,284	0,425	0,6076
Durbin Watson	2,297	2,101	1,917	2,740	1,974	1,898	1,904

(***) Significant at 10%, (**) Significant at 5%, (*) Significant at 1%

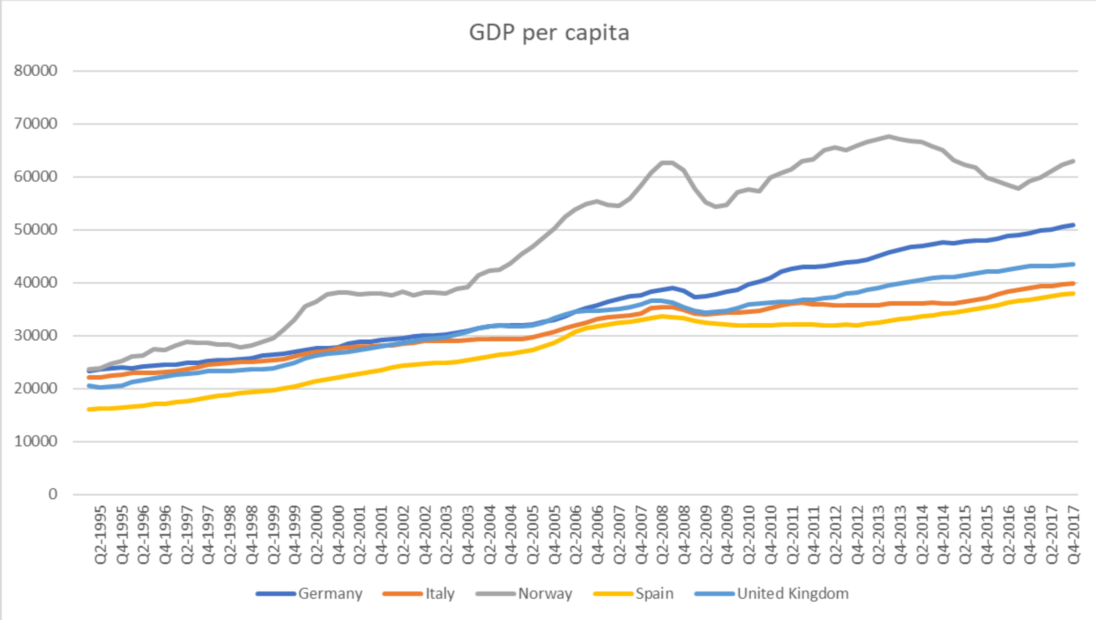
8.2. Data

8.2.1 Nominal GDP per capita

GDP per capita, current prices, current PPPs, Seasonally adjusted

Country	Nº Observations	Max	Min	Average	Standard deviation
SPAIN	92	38.008	15.981	27.746	6604,71
UK	92	43.449	20.235	32.489	6919,60
GERMANY	92	50.885	23.368	35.611	8614,00
ITALY	92	39.878	22.054	31.293	5150,72
NORWAY	92	67.594	23.607	48.153	14335,53

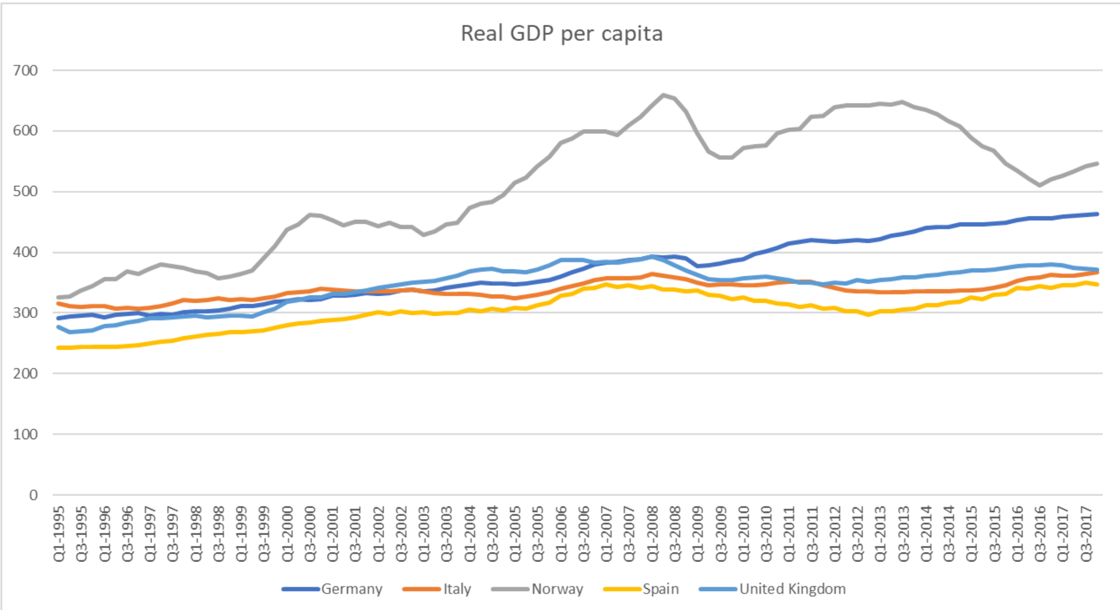
source: OECD



8.2.2 GDP/CPI

GDP per capita/CPI					
Country	**Nº Observations**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	349	242	304	31,02
UK	92	394	269	346	34,54
GERMANY	92	462	291	371	54,31
ITALY	92	367	307	337	15,15
NORWAY	92	660	325	511	101,53

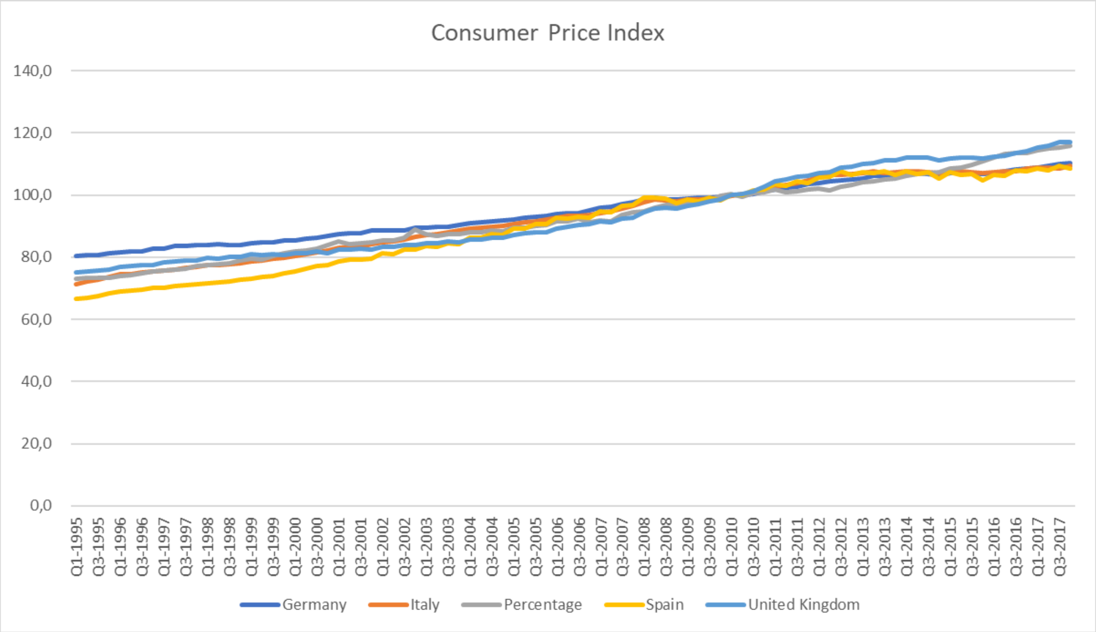
source: OECD and own computations



8.2.3 *CPI*

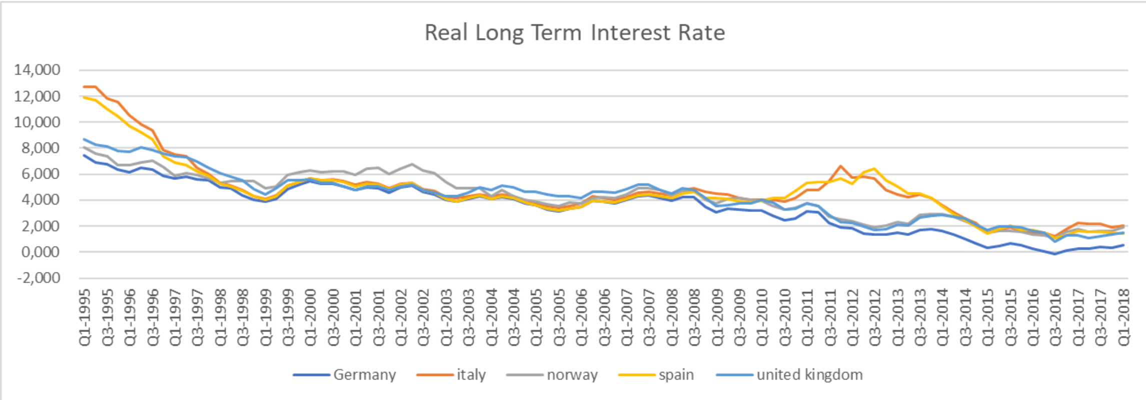
Consumer Price Index, 2010=100					
Country	**Nº Observations**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	109,48	65,88	90,15	14,22
UK	92	116,92	74,13	93,08	13,19
GERMANY	92	110,03	80,17	94,70	9,23
ITALY	92	108,93	70,00	92,38	12,13
NORWAY	92	115,18	72,61	92,30	12,22

source: OECD



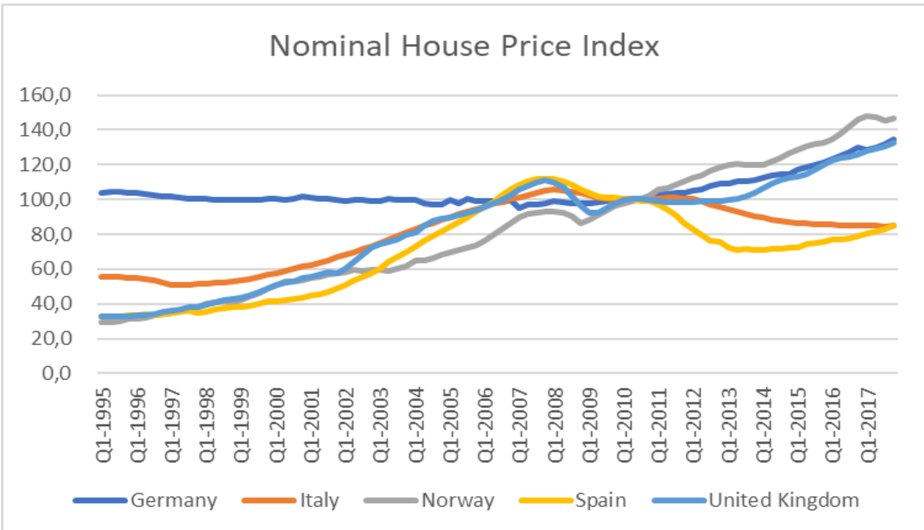
8.2.4 *Real Long Term Interest Rate*

Real Long Term Interest Rates					
Country	**Nº Observations**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	10,62	-0,32	4,08	2,16
UK	92	7,96	-0,05	3,78	1,93
GERMANY	92	6,60	-0,59	3,11	1,86
ITALY	92	11,25	0,79	4,33	2,07
NORWAY	92	7,52	0,05	3,77	1,86



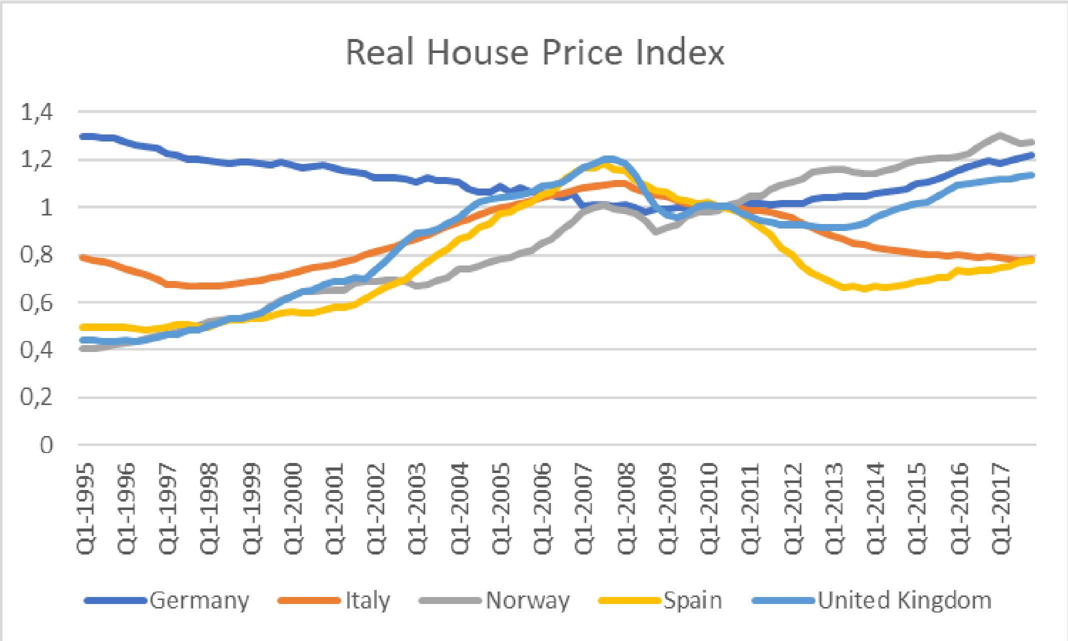
8.2.5 Nominal House Price Index

Nominal House Price Index					
Country	**Observatior**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	111,82	32,63	70,64	25,50
UK	92	132,57	32,85	82,91	30,04
GERMANY	92	134,35	95,34	104,81	9,26
ITALY	92	105,94	50,96	80,91	18,59
NORWAY	92	147,99	29,31	82,27	35,48



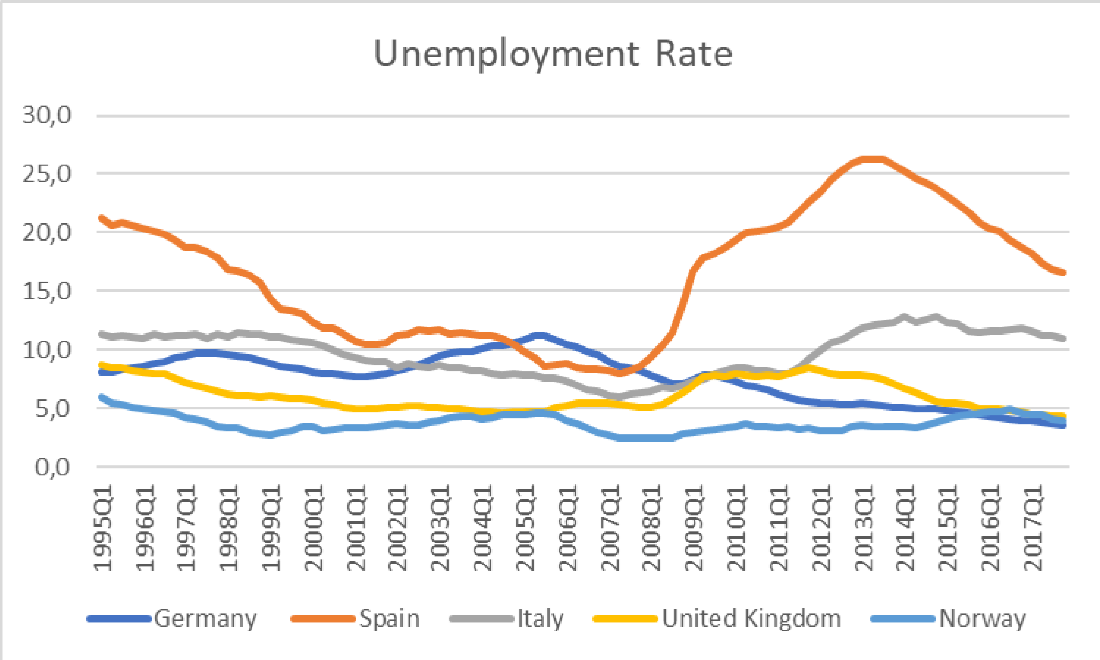
8.2.6 Nominal House Price Index/CPI

HPI/CPI					
Country	**Nº Observations**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	1,18	0,48	0,77	0,22
UK	92	1,20	0,43	0,87	0,24
GERMANY	92	1,30	0,98	1,11	0,09
ITALY	92	1,10	0,67	0,87	0,13
NORWAY	92	1,30	0,40	0,86	0,27
source: OECD					



8.2.7 Unemployment rate

Unemployment rate, percentage of active population, seasonally adjusted					
Country	**Nº Observations**	**Max**	**Min**	**Average**	**Standard deviation**
SPAIN	92	26,20	8,00	16,35	5,57
UK	92	8,70	4,30	6,13	1,31
GERMANY	92	11,20	3,60	7,60	2,09
ITALY	92	12,80	6,00	9,66	1,95
NORWAY	92	5,90	2,40	3,72	0,75
source: Eurostat					



8.3. Unit Root Analysis

8.3.1 Spain

Spain		ADF		KPSS		Conclusion
		Level I(1)	Δ I(2)	Level I(1)	Δ I(2)	
Variable	l_RGDPPC	T -2,70522 0,2344 (6)	C -3,0636** 0,02939** (2)	T 0,43481*** p < 0,01*** (3)	- 0,309062 p > 0,1 (3)	I(1)
	l_RHPI	T -2,0113 0,5946 (2)	NC -2,34246** 0,01851** (6)	T 0,5215*** p < 0,01*** (3)	- 0,542773** 0,038** (3)	I(1)
	RLTIR	T -3,0294 0,124 (3)	NC -3,98945*** 6,76e-05*** (3)	T 0,26194*** p < 0,01*** (3)	- 0,159305 p > 0,1 (3)	I(1)
	UNEMP	T + Period Dummies -1,940* 0,0559* (2)	NC -3,19091*** 0,00170*** (0)	T 0,43481*** p < 0,01*** (3)	- 0,365782* 0,093* (3)	I(1)

() The number in brackets indicate The optimal number of lags for ADF

8.3.2 *Germany*

Germany		ADF		KPSS		Conclusion
		Level I(1)	Δ I(2)	Level I(1)	Δ I(2)	
Variable	l_RGDPPC	T -4,4595*** 0,0017*** (3)	C -5,5939*** 1,06e-06*** (7)	T 0.07223 p > 0,1 (3)	- 0.04084 p > 0,1 (3)	I(0)
	l_RHPI	T 1,6460 1 (1)	T -12,7822*** 6,57e-15*** (0)	T 0.52213*** p < 0,01*** (3)	- 0.19479** 0,022** (3)	I(1)
	RLTIR	T -2,37707 0,3916 (3)	NC -7,8998*** 8,66e-14*** (2)	T 0.23181*** p < 0,01*** (3)	- 0.04364 p > 0,1 (3)	I(1)
	UNEMP	T -1,97946 0,612 (2)	NC -2,5522** 0,01037** (9)	T 0.43233*** p < 0,01*** (3)	- 0.40478* 0,075* (3)	I(1)

() The number in brackets indicate The optimal number of lags for ADF and KPSS

8.3.3 *UK*

Uk		ADF		KPSS		Conclusion	
		Level I(1)	Δ I(2)	Level I(1)	Δ I(2)		
Variable	l_RGDPPC	T -1.92823 0.6395 (7)	NC -1.8340* 0.06351* (6)	T 0.5115*** p < 0.01*** (3)	- 0.43590* 0.062* (3)	I(1)	Result show posible I(2) but in economic literature RGDP is usually I(1)
	l_RHPI	T -1.55843 0.8093 (1)	NC -3.43283*** 0.00077*** (0)	T 0.5334*** p < 0.01*** (3)	- 0.49432** 0.045** (3)	I(1)	
	RLTIR	T -3.5648** 0.03286** (4)	NC -4.1523*** 3.434e-5*** (3)	T 0.0527625 p > 0,1 (3)	- 0.065706 p > 0.1 (3)	I(0)	Economically can be I(0) or I(1), in this case is clearly I(0)
	UNEMP	T[2] -1.4205 0.9558 (1)	NC -4.19563*** 5.094e-5*** (0)	T 0.28537*** p < 0.01*** (3)	- 0.285777 p > 0.1 (3)	I(1)	

he number in brackets indicate The optimal number of lags for ADF selected by

8.3.4 Norway

Norway		ADF		KPSS		Conclusion
		Level I(1)	Δ I(2)	Level I(1)	Δ I(2)	
Variable	l_RGDPPC	T -1.4860 0.8347 (1)	NC -4.80248*** 3.842e-6*** (0)	T 0.43185*** p < 0.01*** (3)	- 0.37306* 0.090* (3)	I(1)
	l_RHPI	T -2.12828 0.5294 (1)	NC -5.58277*** 7.363e-6*** (0)	T 0.43988*** p < 0.01*** (3)	- 0.36142* 0.095* (3)	I(1)
	RLTIR	T -6.82714*** 4.227e-7*** (0)	NC -9.7398*** 9.83e-19*** (2)	T 0.11262 p > 0,1 (3)	- 0.0407136 p > 0.1 (3)	I(0)
	UNEMP	T^2 -2.96011 0.3155 (2)	NC -5.63921*** 3.971e-8*** (0)	T 0.2185*** p < 0.01*** (3)	- 0.30916 p > 0.1 (3)	I(1)

Economicall y can be I(0) or I(1), in this case is clearly I(0)

() The number in brackets indicate The optimal number of lags for ADF selected by BIC

8.3.5 Italy

Italy		ADF		KPSS		Conclusion
		Level I(1)	Δ I(2)	Level I(1)	Δ I(2)	
Variable	l_RGDPPC	T -2.4918 0.3323 (1)	NC -4.8804*** 2.66e-06*** (0)	T 0.2338*** p < 0.01*** (3)	- 0.08152 p > 0.1 (3)	I(1)
	l_RHPI	T^2 -2.9098 0.3408 (1)	NC -2.20472** 0.02721** (0)	T 0.5211*** p < 0.01*** (3)	- 0.57133** 0.034** (3)	I(1)
	RLTIR	T -3.2480* 0.0818* (0)	NC -6.8817*** 3.48e-11*** (1)	T 0.25924*** p < 0.01*** (3)	- 0.38192* 0.086* (3)	I(1)
	UNEMP	T^2 -2.0528 0.796 (2)	NC -3.4035*** 0.00065*** (1)	T 0.52631*** p < 0.01*** (3)	- 0.38674* 0.083* (3)	I(1)

() The number in brackets indicate The optimal number of lags for ADF selected by BIC

8.4 Lag Structure

LAG SELECTION	SPAIN	ITALY	GERMANY	NORWAY	UK
BIC	2	2	2	2	2
AIC	8	2	2	4	3
HQC	2	2	2	2	2

The number indicates the optimal number of lags according to specific criteria

8.5 Cointegration analysis (Johansen)

Johansen Test	Trace Test				Lmax Test				Number of cointegrating relationships (Trace test)
	0	1	2	3	0	1	2	3	
Spain	73,71 (0,0004)*	33,63 (0,0706)	16,20 (0,098)	3,037 (0,0814)	40,078 (0,0016)*	17,43 (0,3204)	13,163 (0,174)	3,0372 (0,0814)	1
Italy	30,89 (0,0062)*	35,26 (0,0470)*	14,98 (0,141)	4,73 (0,0296)*	28,62 (0,088)*	20,28 (0,159)	10,25 (0,38)	4,73 (0,029)*	2

*The number in brackets indicates the p-value of the test. *indicates refuse at 5%. Only Spain and Italy since are the only cases where all variablas are the same order of integration.*

8.6 Model Estimation: VECM

8.6.1 Spain

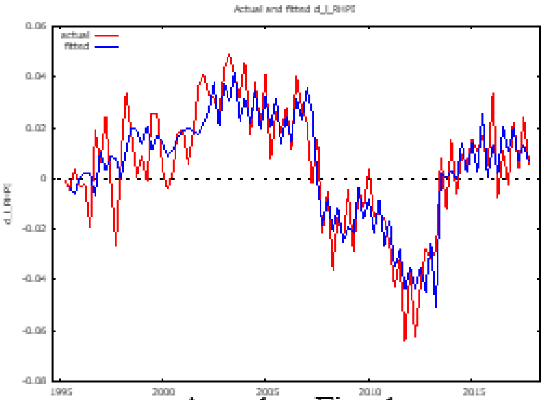
```
Equation 1: d_l_RHPI

                  coefficient    std. error    t-ratio   p-value
  -------------------------------------------------------------
  const           -0.944079      0.371520      -2.541    0.0130   **
  d_l_RHPI_1       0.276776      0.108616       2.548    0.0127   **
  d_l_RGDPPC_1    -0.455388      0.239462      -1.902    0.0608   *
  d_RLTIR_1       -0.00144846    0.00167895    -0.8627   0.3908
  d_Unemp_1       -0.000967116   0.00294172    -0.3288   0.7432
  break2007time    0.000367754   0.000115272    3.190    0.0020   ***
  break2013time   -0.000417899   0.000112023   -3.730    0.0004   ***
  time            -0.000660878   0.000295058   -2.240    0.0278   **
  EC1             -0.0123039     0.00483028    -2.547    0.0128   **

Mean dependent var   0.004975   S.D. dependent var   0.024337
Sum squared resid    0.012652   S.E. of regression   0.012498
R-squared            0.759998   Adjusted R-squared   0.736294
rho                 -0.050599   Durbin-Watson        2.100940
```

```
Equation 2: d_l_RGDPPC

                  coefficient    std. error    t-ratio   p-value
  -------------------------------------------------------------
  const            0.261213      0.215970       1.209    0.2300
  d_l_RHPI_1      -0.0205510     0.0631397     -0.3255   0.7457
  d_l_RGDPPC_1     0.368714      0.139202       2.649    0.0097   ***
  d_RLTIR_1       -0.00629334    0.000975995   -6.448    7.66e-09 ***
  d_Unemp_1       -0.00316195    0.00171006    -1.849    0.0681   *
  break2007time    0.000109337   6.70094e-05    1.632    0.1066
  break2013time   -3.78347e-05   6.51202e-05   -0.5810   0.5629
  time             0.000224432   0.000171521    1.308    0.1944
  EC1              0.00339237    0.00280791     1.208    0.2305

Mean dependent var   0.003995   S.D. dependent var   0.011669
Sum squared resid    0.004275   S.E. of regression   0.007265
R-squared            0.647221   Adjusted R-squared   0.612379
rho                  0.160438   Durbin-Watson        1.668593
```

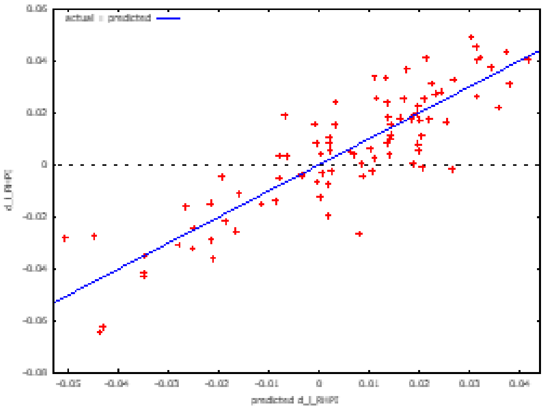
```
Equation 3: d_RLTIR

                  coefficient    std. error    t-ratio   p-value
  -------------------------------------------------------------
  const           -43.1831       22.4474       -1.924    0.0579   *
  d_l_RHPI_1      -12.6286        6.56260      -1.924    0.0578   *
  d_l_RGDPPC_1    -58.3178       14.4684       -4.031    0.0001   ***
  d_RLTIR_1        -0.494314      0.101443     -4.873    5.38e-06 ***
  d_Unemp_1        -0.328066      0.177740     -1.846    0.0686   *
  break2007time     0.0229807     0.00696480    3.300    0.0014   ***
  break2013time    -0.00260004    0.00676845   -0.3841   0.7019
  time             -0.0265755     0.0178275    -1.491    0.1399
  EC1              -0.554588      0.291848     -1.900    0.0610   *

Mean dependent var  -0.116132   S.D. dependent var   1.643495
Sum squared resid    46.18724   S.E. of regression   0.755124
R-squared            0.807870   Adjusted R-squared   0.788894
rho                  0.221675   Durbin-Watson        1.538947
```

```
Equation 4: d_Unemp

                  coefficient   std. error    t-ratio    p-value
  -------------------------------------------------------------
  const           -1.36861      11.5923       -0.1181    0.9063
  d_l_RHPI_1       0.157124      3.38906       0.04636   0.9631
  d_l_RGDPPC_1    -6.37345       7.47176      -0.8530    0.3962
  d_RLTIR_1        0.0491770     0.0523871     0.9387    0.3507
  d_Unemp_1        0.569311      0.0917884     6.202     2.22e-08 ***
  break2007time    0.00194735    0.00359676    0.5414    0.5897
  break2013time    0.00677395    0.00349536    1.938     0.0561   *
  time            -0.00231511    0.00920647   -0.2515    0.8021
  EC1             -0.0169175     0.150716     -0.1122    0.9109

Mean dependent var  -0.044444   S.D. dependent var   0.667060
Sum squared resid    12.31767   S.E. of regression   0.389962
R-squared            0.688965   Adjusted R-squared   0.658246
rho                  0.105774   Durbin-Watson        1.755585
```

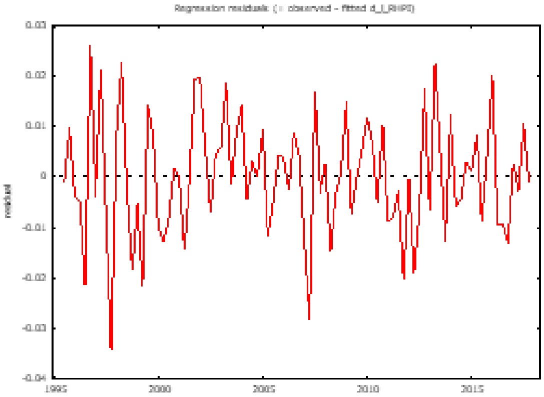
```
beta (cointegrating vectors, standard errors in parentheses)

l_RHPI          1.0000
              (0.00000)
l_RGDPPC       -13.719
               (2.9132)
RLTIR          0.18689
             (0.054417)
Unemp         -0.11986
             (0.038151)

alpha (adjustment vectors)

l_RHPI       -0.012304
l_RGDPPC     0.0033924
RLTIR         -0.55459
Unemp        -0.016917
```



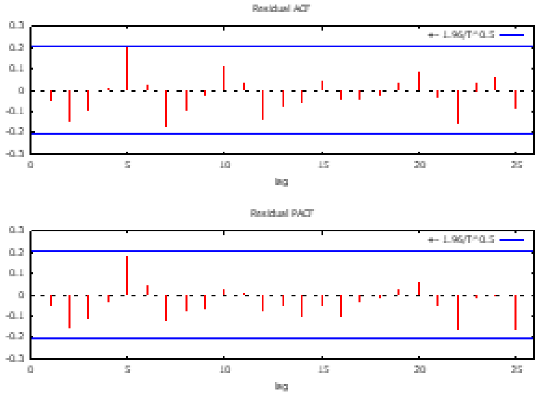
Actual vs Fitted (Time)



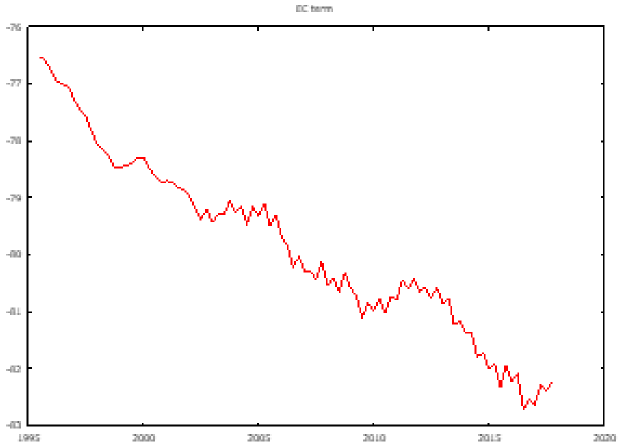
Actual vs Fitted (Values)



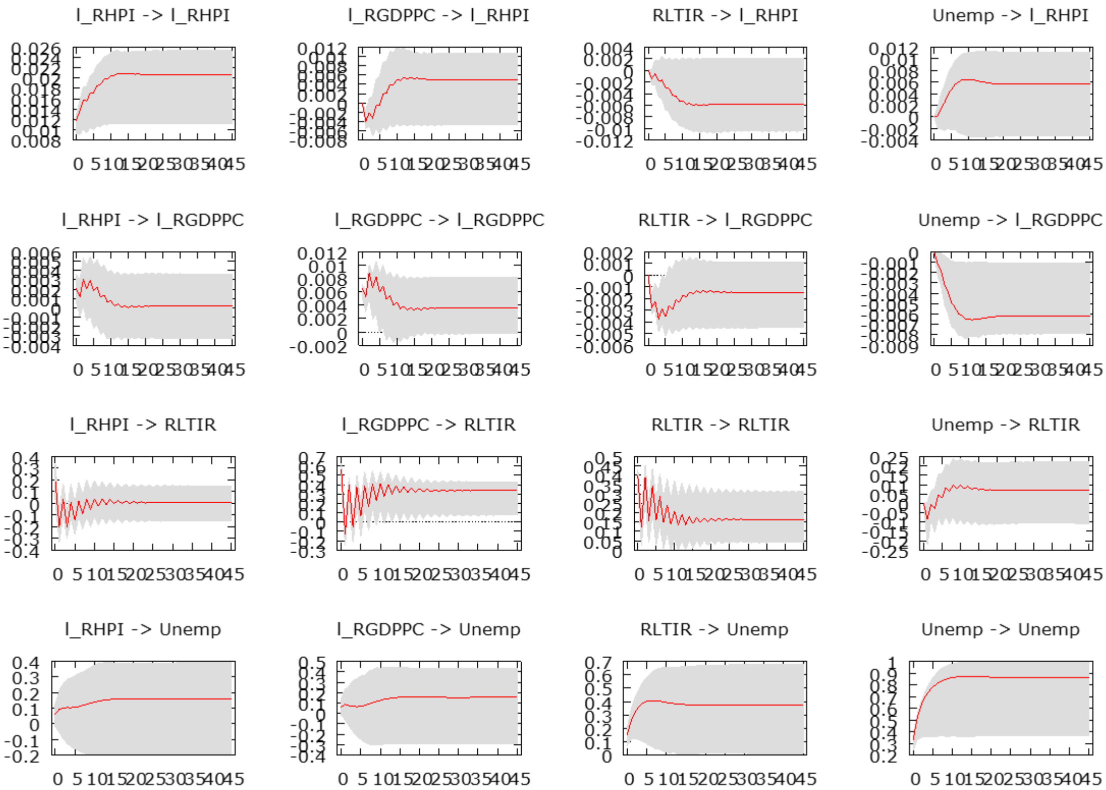
Residual Time Series Plot



Residual Correlogram



EC TERM



8.6.2 Italy

Equation 1: d_l_RHPI

	coefficient	std. error	t-ratio	p-value	
const	0.468074	0.220957	2.118	0.0373	**
d_l_RHPI_1	0.585316	0.0748340	7.822	1.94e-011	***
d_l_RGDPPC_1	-0.117471	0.0745635	-1.575	0.1191	
d_UNEMP_1	-0.00319624	0.00224283	-1.425	0.1581	
d_RLTIR_1	0.00227383	0.00120856	1.881	0.0636	*
timebreak2012	-5.75229e-05	2.72728e-05	-2.109	0.0381	**
timebreak2003	8.51378e-05	8.29174e-05	1.027	0.3077	
time	-7.63609e-06	0.000107359	-0.07113	0.9435	
EC1	-0.0196395	0.0114219	-1.719	0.0894	*
EC2	-0.0742277	0.0370833	-2.002	0.0488	**

Mean dependent var	0.000057	S.D. dependent var	0.013053
Sum squared resid	0.001817	S.E. of regression	0.004796
R-squared	0.880140	Adjusted R-squared	0.864968
rho	-0.147807	Durbin-Watson	2.074044

Equation 2: d_l_RGDPPC

	coefficient	std. error	t-ratio	p-value	
const	0.679219	0.272438	2.493	0.0148	**
d_l_RHPI_1	-0.117646	0.0922696	-1.275	0.2060	
d_l_RGDPPC_1	0.475893	0.0919361	5.176	1.68e-06	***
d_UNEMP_1	-0.00483751	0.00276539	-1.749	0.0841	*
d_RLTIR_1	0.000740224	0.00149015	0.4967	0.6207	
timebreak2012	-8.92158e-05	3.36271e-05	-2.653	0.0096	***
timebreak2003	-0.000177524	0.000102236	-1.736	0.0864	*
time	0.000155212	0.000132373	1.173	0.2445	
EC1	-0.0323729	0.0140831	-2.299	0.0242	**
EC2	-0.110921	0.0457234	-2.426	0.0175	**

Mean dependent var	0.001850	S.D. dependent var	0.007799
Sum squared resid	0.002763	S.E. of regression	0.005914
R-squared	0.489573	Adjusted R-squared	0.424962
rho	0.007666	Durbin-Watson	1.964421

Equation 3: d_UNEMP

	coefficient	std. error	t-ratio	p-value	
const	17.6475	10.8676	1.624	0.1084	
d_l_RHPI_1	-5.70308	3.68067	-1.549	0.1253	
d_l_RGDPPC_1	-4.37262	3.66736	-1.192	0.2367	
d_UNEMP_1	0.0934835	0.110312	0.8474	0.3993	
d_RLTIR_1	0.0199721	0.0594425	0.3360	0.7378	
timebreak2012	0.00358265	0.00134140	2.671	0.0092	***
timebreak2003	0.00762619	0.00407824	1.870	0.0652	*
time	0.0125739	0.00528040	2.381	0.0197	**
EC1	-0.997730	0.561781	-1.776	0.0796	*
EC2	-3.06577	1.82392	-1.681	0.0967	*

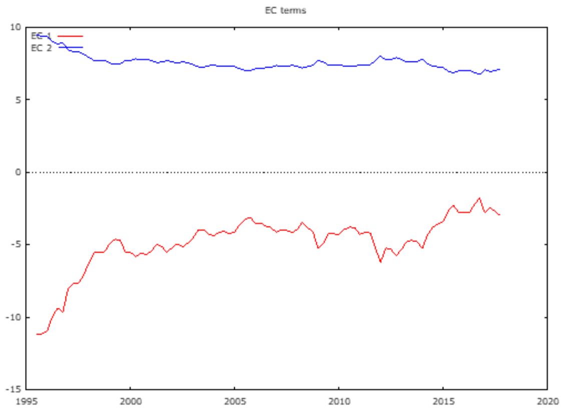
Mean dependent var	-0.001111	S.D. dependent var	0.283039
Sum squared resid	4.396608	S.E. of regression	0.235909
R-squared	0.383355	Adjusted R-squared	0.305299
rho	-0.027516	Durbin-Watson	2.041609

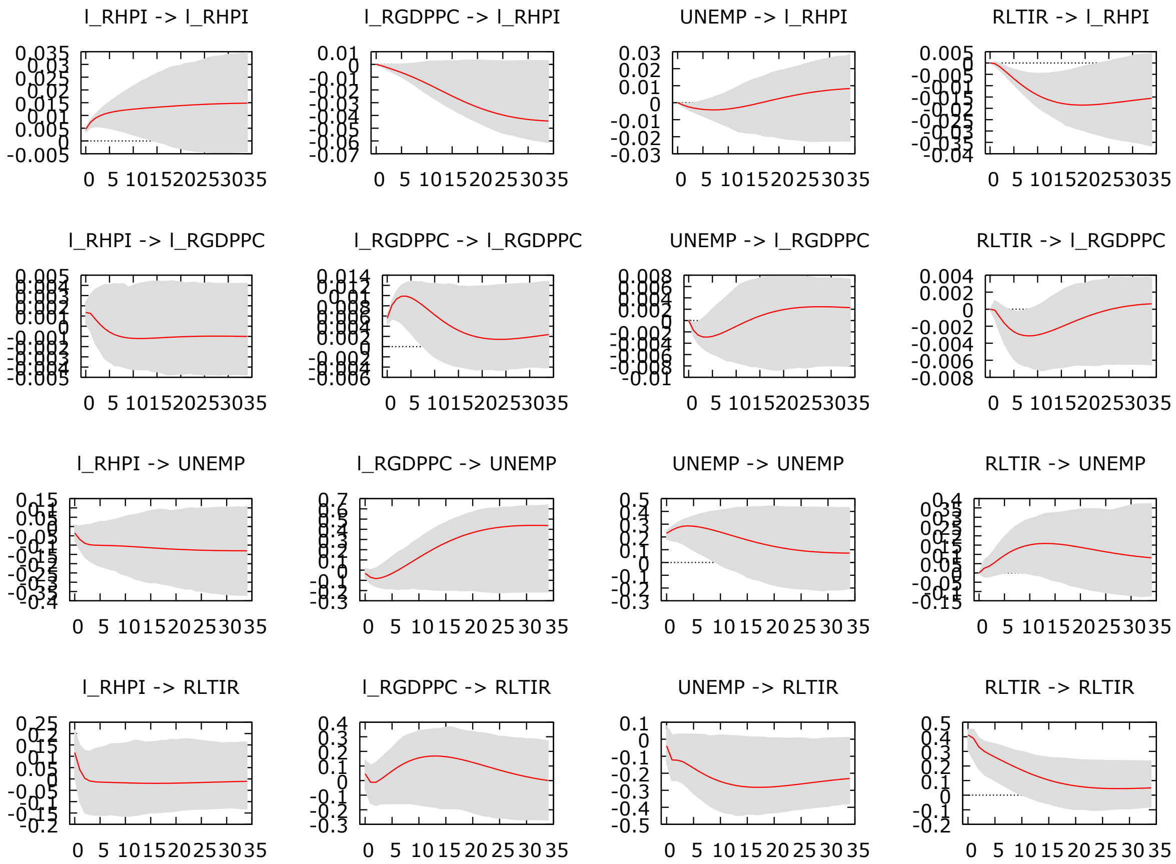
Equation 4: d_RLTIR

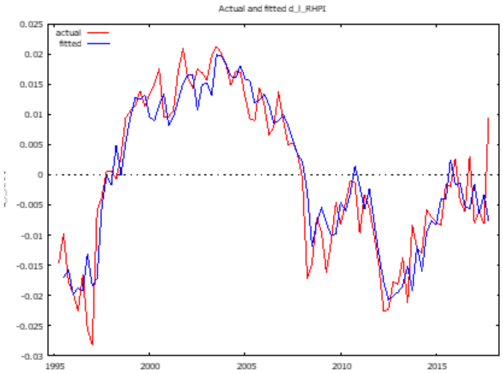
	coefficient	std. error	t-ratio	p-value	
const	-34.5665	20.4743	-1.688	0.0953	*
d_l_RHPI_1	-16.0736	6.93427	-2.318	0.0230	**
d_l_RGDPPC_1	-15.4377	6.90920	-2.234	0.0283	**
d_UNEMP_1	-0.362413	0.207825	-1.744	0.0851	*
d_RLTIR_1	0.151761	0.111988	1.355	0.1792	
timebreak2012	0.00304639	0.00252715	1.205	0.2316	
timebreak2003	0.00189386	0.00768329	0.2465	0.8059	
time	-0.0258919	0.00994811	-2.603	0.0110	**
EC1	2.09797	1.05838	1.982	0.0509	*
EC2	6.05614	3.43621	1.762	0.0819	*

Mean dependent var	-0.096138	S.D. dependent var	0.497923
Sum squared resid	15.60506	S.E. of regression	0.444446
R-squared	0.292785	Adjusted R-squared	0.203265
rho	-0.028813	Durbin-Watson	2.040260

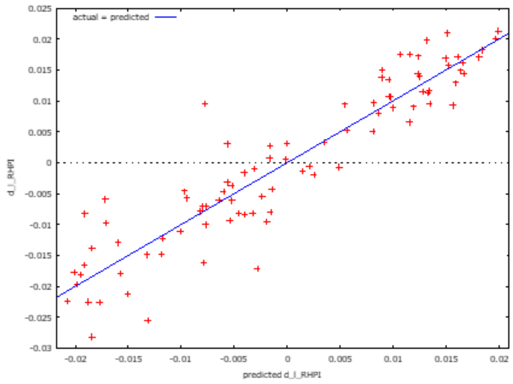
```
beta (cointegrating vectors, standard errors in parentheses)

l_RHPI          1.0000      0.00000
              (0.00000)    (0.00000)
l_RGDPPC       0.00000       1.0000
              (0.00000)    (0.00000)
UNEMP        -0.068910     0.054639
              (0.16635)   (0.050863)
RLTIR         -0.94090      0.28496
              (0.17043)   (0.052109)

alpha (adjustment vectors)

l_RHPI       -0.019640    -0.074228
l_RGDPPC     -0.032373     -0.11092
UNEMP         -0.99773      -3.0658
RLTIR           2.0980       6.0561

Log-likelihood = 668.22175
Determinant of covariance matrix = 4.1792868e-012
AIC = -13.7827
BIC = -12.4495
HQC = -13.2451
```



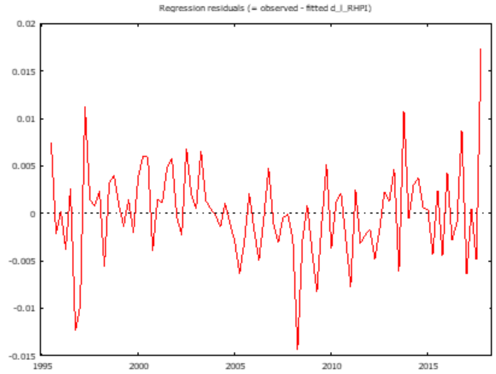




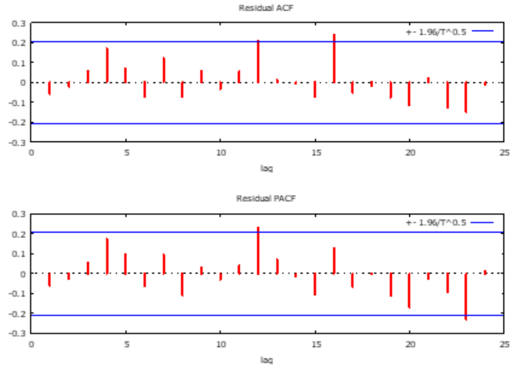
Actual vs. Fitted (Time)



Actual vs. Fitted (Values)



Error Term Time Series Plot

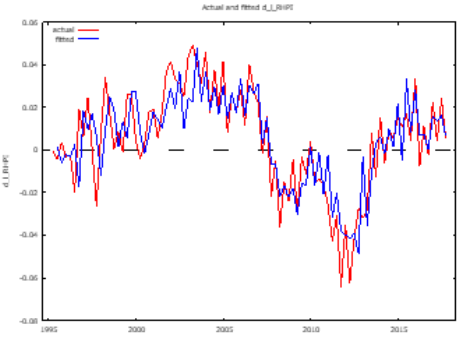


Residual Correlogram:

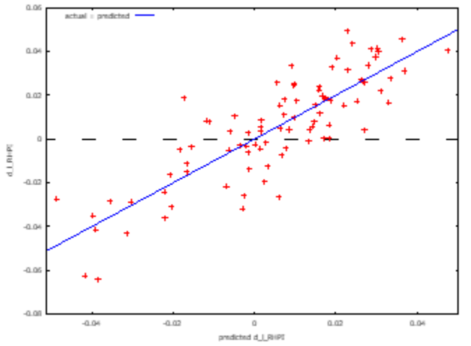
8.7 Short Run Estimation

8.7.1 Spain

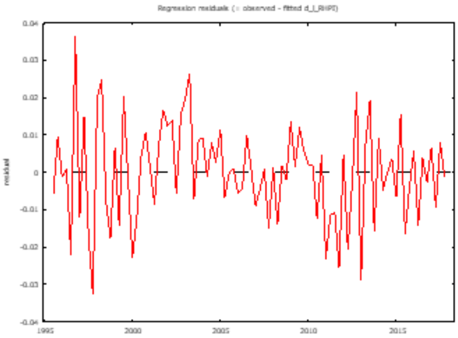
```
Model 2: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

                coefficient    std. error    t-ratio    p-value
  --------------------------------------------------------------
  const          0.000469983   0.00166693    0.2819    0.7787
  d_l_RGDPPC     0.348338      0.214971      1.620     0.1089
  d_RLTIR        0.00503813    0.00148314    3.397     0.0010     ***
  d_Unemp       -0.00345813    0.00267202   -1.294     0.1991
  d_l_RHPI_1     0.721351      0.0701460    10.28      1.38e-016 ***

Mean dependent var    0.004975   S.D. dependent var    0.024337
Sum squared resid     0.015286   S.E. of regression    0.013410
R-squared             0.710031   Adjusted R-squared    0.696385
F(4, 85)              52.03368   P-value(F)            4.40e-22
Log-likelihood        262.9240   Akaike criterion     -515.8480
Schwarz criterion    -503.3489   Hannan-Quinn         -510.8076
rho                  -0.150093   Durbin's h           -1.907625

Excluding the constant, p-value was highest for variable 36 (d_Unemp)
```



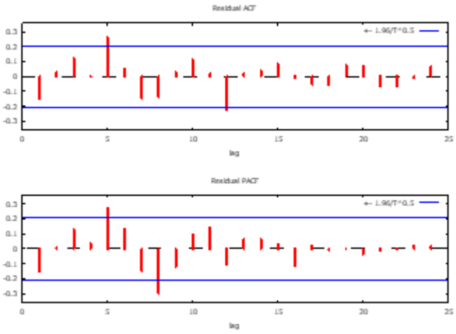
Actual vs Fitted (Time)



Actual vs Fitted (Values)



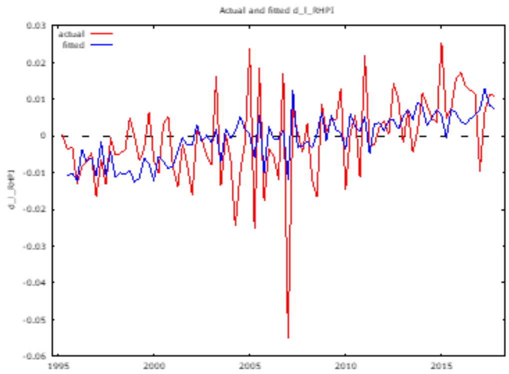
Residual Time Series Plot



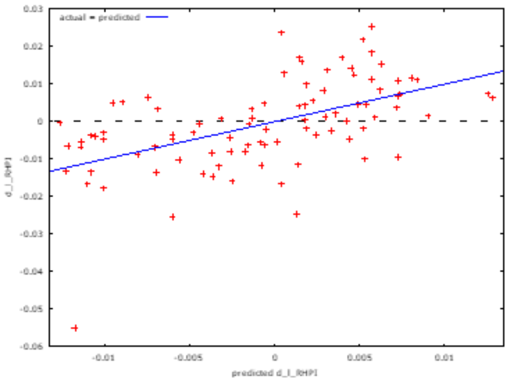
Residual Correlogram

8.7.2 Germany

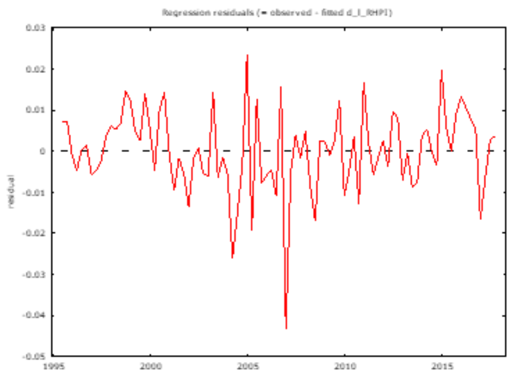
```
Model 1: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

                 coefficient    std. error    t-ratio    p-value
  ------------------------------------------------------------
  const          -0.0139894     0.00279968    -4.997     3.13e-06 ***
  d_l_RGDPPC      0.120375      0.166027       0.7250    0.4704
  d_UNEMP         0.0111242     0.00611520     1.819     0.0725   *
  time            0.000275950   4.95600e-05    5.568     3.02e-07 ***
  d_RLTIR         0.00102986    0.00240852     0.4276    0.6700
  d_l_RHPI_1     -0.293737      0.105720      -2.778     0.0067   ***

Mean dependent var   -0.000677   S.D. dependent var   0.011992
Sum squared resid     0.009164   S.E. of regression   0.010445
R-squared             0.284039   Adjusted R-squared   0.241422
F(5, 84)              6.664956   P-value(F)           0.000028
Log-likelihood        285.9489   Akaike criterion    -559.8978
Schwarz criterion    -544.8989   Hannan-Quinn        -553.8494
rho                   0.009344   Durbin-Watson        1.974322

Excluding the constant, p-value was highest for variable 8 (d_RLTIR)
```



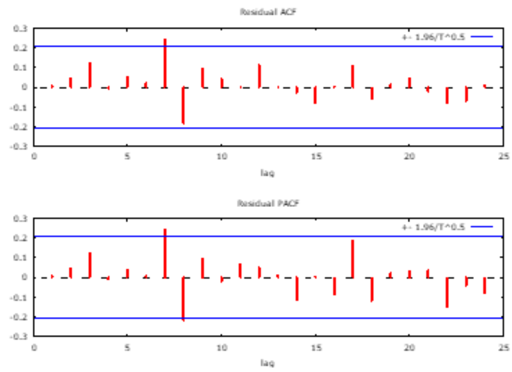
Actual vs. Fitted (Time)



Actual vs. Fitted (Values)



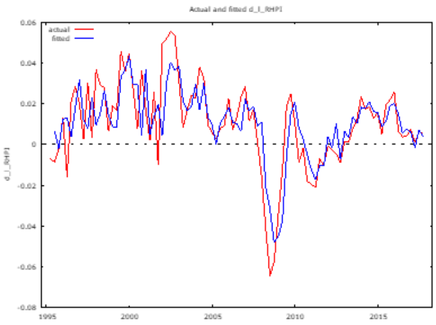
Error Term Time Series Plot



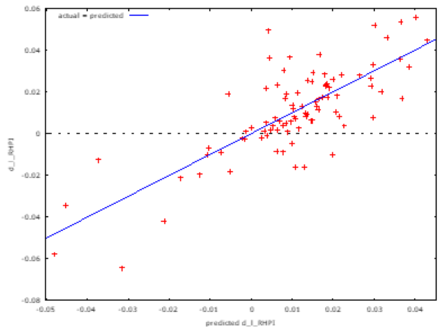
Residual Correlogram:
Residual correlogram may indicate a moving average on lag 7

8.7.3 UK

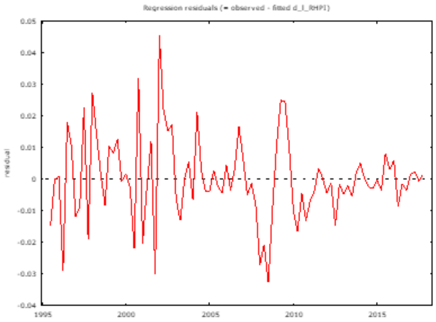
```
Model 2: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

              coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const        0.00273531   0.00163540    1.673    0.0981   *
  d_RLTIR      0.00464623   0.00229331    2.026    0.0459   **
  d_UNEMP     -0.00686569   0.00815895   -0.8415   0.4024
  d_l_RGDPPC   0.425120     0.196272      2.166    0.0331   **
  d_l_RHPI_1   0.603140     0.0904187     6.671    2.44e-09 ***

Mean dependent var   0.010517   S.D. dependent var   0.021525
Sum squared resid    0.016181   S.E. of regression   0.013797
R-squared            0.607618   Adjusted R-squared   0.589153
F(4, 85)             32.90635   P-value(F)           1.45e-16
Log-likelihood       260.3634   Akaike criterion    -510.7267
Schwarz criterion   -498.2277   Hannan-Quinn        -505.6864
rho                  0.040782   Durbin's h           0.752694

Excluding the constant, p-value was highest for variable 8 (d_UNEMP)
```



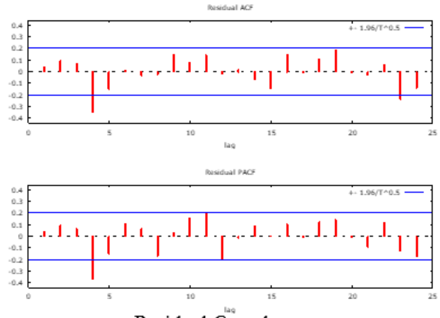
Actual vs. Fitted (Time)



Actual vs. Fitted (Values)



Error Term Time Series Plot



Residual Correlogram:

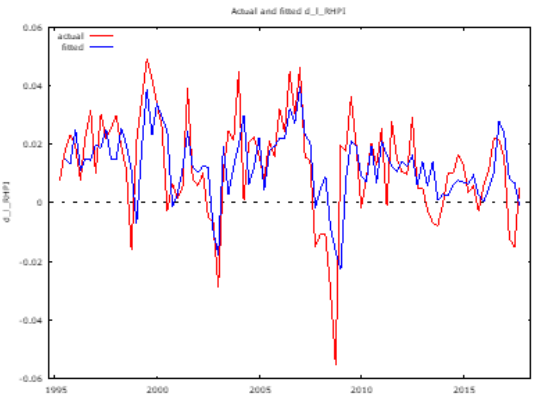
8.7.4 Norway

Model 1: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

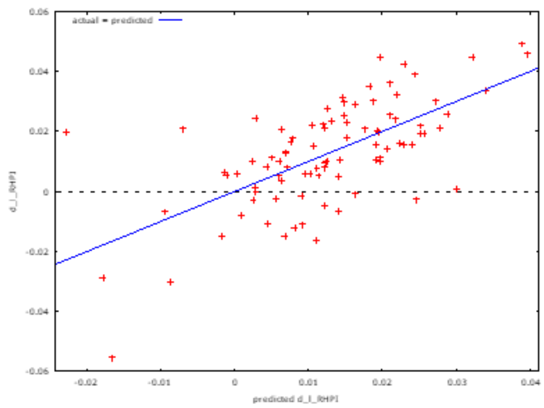
	coefficient	std. error	t-ratio	p-value	
const	0.00593508	0.00179186	3.312	0.0014	***
d_l_RGDPPC	0.0511727	0.0695790	0.7355	0.4641	
d_UNEMP	−0.00841435	0.00809904	−1.039	0.3018	
d_RLTIR	0.00736418	0.00154604	4.763	7.77e-06	***
d_l_RHPI_1	0.532785	0.0933724	5.706	1.65e-07	***

Mean dependent var	0.012684	S.D. dependent var	0.017597
Sum squared resid	0.015821	S.E. of regression	0.013643
R-squared	0.425886	Adjusted R-squared	0.398869
F(4, 85)	15.76354	P-value(F)	1.09e-09
Log-likelihood	261.3747	Akaike criterion	−512.7494
Schwarz criterion	−500.2503	Hannan-Quinn	−507.7090
rho	0.049525	Durbin's h	1.012467

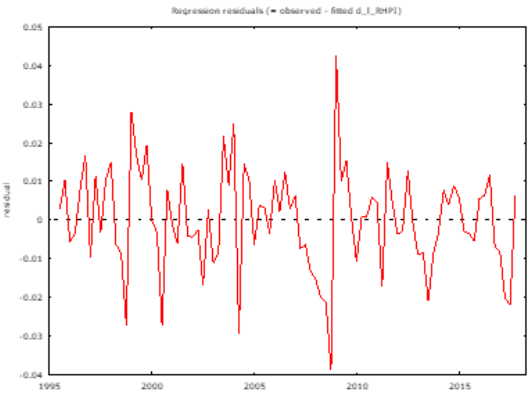
Excluding the constant, p-value was highest for variable 7 (d_l_RGDPPC)



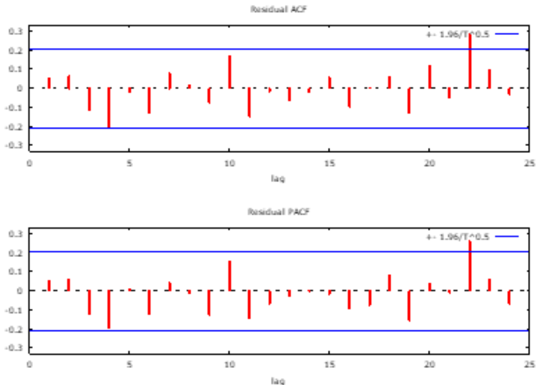
Actual vs. Fitted (Time)



Actual vs. Fitted (Values)



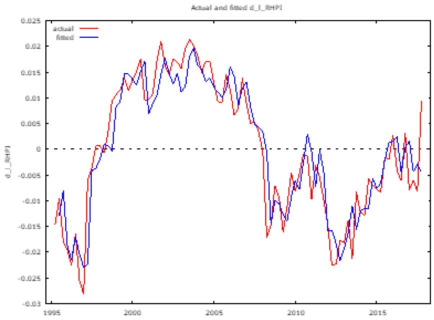
Error Term Time Series Plot



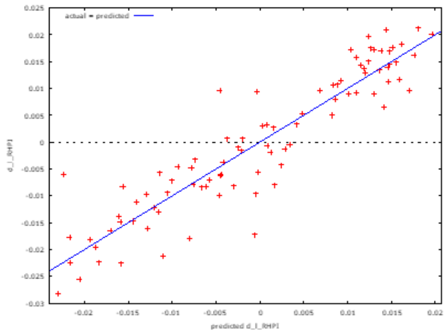
Residual Correlogram:

8.7.5 Italy

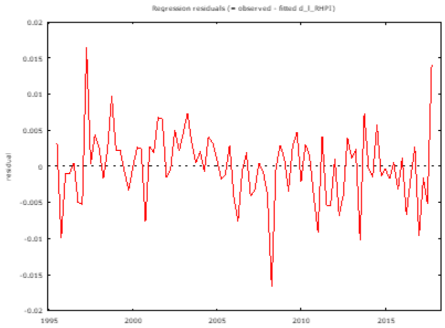
```
Model 1: OLS, using observations 1995:3-2017:4 (T = 90)
Dependent variable: d_l_RHPI

                 coefficient    std. error    t-ratio    p-value
  ------------------------------------------------------------
  const          0.000277677   0.000569402    0.4877    0.6270
  d_l_RGDPPC     0.163279      0.0754681      2.164     0.0333     **
  d_UNEMP       -0.00203878    0.00225702    -0.9033    0.3689
  d_RLTIR        0.00358022    0.00110159     3.250     0.0017     ***
  d_l_RHPI_1     0.853120      0.0457824     18.63      9.22e-032 ***

Mean dependent var   0.000057   S.D. dependent var   0.013053
Sum squared resid    0.002229   S.E. of regression   0.005120
R-squared            0.853030   Adjusted R-squared   0.846114
F(4, 85)             123.3372   P-value(F)           1.51e-34
Log-likelihood       349.5757   Akaike criterion    -689.1513
Schwarz criterion   -676.6523   Hannan-Quinn        -684.1110
rho                 -0.006290   Durbin's h          -0.066243

Excluding the constant, p-value was highest for variable 9 (d_UNEMP)
```



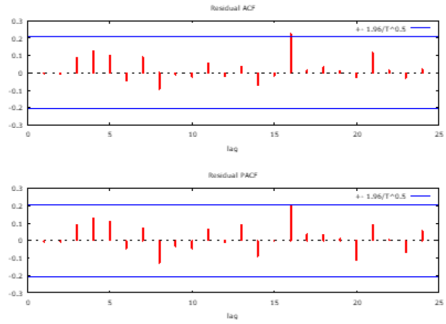
Actual vs. Fitted (Time)



Actual vs. Fitted (Values)



Error Term Time Series Plot



Residual Correlogram:

8.8 Further Testing

8.8.1 Linearity

Linearity	Squares and cubes	Squares	Cubes
SPAIN	0,063 (0,939)	0,00323 (0,955)	0,114 (0,736)
ITALY	1.239 (0,2995)	1,609 (0,208)	1,681 (0,198)
GERMANY	0,849 (0,431)	0,0016 (0,968)	1,672 (0,199)
NORWAY	0,361 (0,698)	0,398 (0,529)	0,048 (0,82)
UK	0,056 (0,94)	0,1125 (0,738)	0,0041 (0,949)

All models are linear

8.8.2 Heteroskedasticity

Heteroskedasticity	White Test	Breusch and Pagan	Coments
SPAIN	47,75 (0,217)	13,091 (0,1087)	Homoskedastic
ITALY	24,34* (0,0415)*	7,50 (0,111)	White test indicate heteroskedasticity, we might expect a decrease in efficiency of the estimators
GERMANY	49,537* (0,0002)*	17,4872* (0,0036)*	Heteroskedasticity caused by influential observation 1Q2007. When removed, the issue is fixed
NORWAY	52,57* (2*10^-6)*	22,232* (0,00018)*	Heteroskedasticity caused by influential observation 1Q2009. When removed, the issue is fixed
UK	8,6364 (0,853)	5,869 (0,2091)	Homoskedastic

* Indicates refuse Ho at 5%

8.8.3 Residual Normality

Residual Normality	Chi Squared test (Short Run)	Doornik-Hansen test (Long Run)	Coments
SPAIN	0,293 (0,863)	9,907 (0,2716)	Residuals are normal
ITALY	14,184* (0,00083)*	11,352 (0,1825)	Residuals are normal in the long run. In the short run the models suffer from efficiency loss of the estimators.
GERMANY	15,377* (0,00046)*	-	residual not normal due to influential observation 1Q2007. When removed, the issue is fixed
NORWAY	4,563 (0,10216)	-	Residuals are normal
UK	6,606* (0,03677)*	-	residual not normal due to influential Autocorrelation with lag 4. When taken into acount, the issue is fixed

* Indicates refuse Ho at 5%, number in parenthesis indicates p-value

8.8.4 Autocorrelation

Autocorrelation	AR (1) Quarterly correlation			AR (4) Yearly correlation			Coments
	Lagrange Multiplier F	Breusch and Godfrey	Ljung-Box Q	Lagrange Multiplier F	Breusch and Godfrey	Ljung-Box Q	
SPAIN	0,6445 (0,424)	0,7193 (0,396)	0,2381 (0,626)	1,40425 (0,241)	6,1189 (0,19)	0,1792 (0,528)	No Autocorrelation
ITALY	2,2285 (0,134)	2,44035 (0,118)	1,2762 (0,259)	1,4493 (0,226)	6,1532 (0,188)	2,363 (0,669)	No Autocorrelation
GERMANY	0,00861 (0,926)	0,00934 (0,923)	0,00076 (0,978)	0,28388 (0,888)	1,2596 (0,868)	1,2714 (0,866)	No Autocorrelation
NORWAY	0,574 (0,451)	0,6108 (0,434)	0,2269 (0,634)	1,5022 (0,209)	6,215594 (0,184)	5,78855 (0,216)	No Autocorrelation
UK	0,01713 (0,896)	0,0186 (0,891)	0,009138 (0,924)	3,476133* (0,0113)*	13,185966* (0,0104)*	13,0355* (0,011)*	AR (4)

*Refuse H_0 at 5%. Numbers in brackets are the p-vale