UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES
DEPARTAMENTO DE FUNDAMENTOS DE ANÁLISIS ECONÓMICO II

TESIS DOCTORAL

Quantitative analysis of commercial and residential real estate markets (an approach from cointegration and spatial econometrics)

Análisis de mercados inmobiliarios (un enfoque desde cointegración y econometría espacial)

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Ramiro J. Rodríguez Ramírez

DIRECTOR

Simón Sosvilla Rivero

Madrid, 2017

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES
Departamento de Fundamentos de Análisis Económico II
(Economía Cuantitativa)

QUANTITATIVE ANALYSIS OF COMMERCIAL AND RESIDENTIAL REAL ESTATE MARKETS
(AN APPROACH FROM COINTEGRATION AND SPATIAL ECONOMETRICS)

Análisis de mercados inmobiliarios
(Un enfoque desde cointegración y eonometría espacial)

PHD THESIS

Presented by

Ramiro J. Rodríguez Ramírez

Thesis Advisor

Prof. Dr. Simón Sosvilla

Madrid, 2016

Table of Contents

List of tables

List of illustrations

This work is dedicated to family and friends. Their support and understanding has been fundamental to undertake this research

Acknowledgement

I feel deeply grateful to my wife, Susy, who has shared a great burden of my work. My mentors at BNP Paribas Real Estate, Dr. Christophe Pineau and Dr. Samuel Duah who, with all their patience and wisdom, have encouraged me to keep on in these late five years. I'm really thankful to Simón Sosvilla who has guided and offered me a boundless supply of knowledge and state-of-the-art research on which this thesis is based.

Part of the action has also been played by my beloved sister Clara and my extraordinarily sapient mother. I'm eternally in indebted to them.

Since the beginning of this project in 2011 I have been honored to receive emotive and empathetic expressions of support from my friends to whom now I feel closer, although this endeavor has in one way or another put some physical distance at certain moments. A humble expression of thankfulness to Celia Rodríguez, Arturo García, Ricardo Serrano, Eva Núñez, Lorena Arciniegas, Emilie Gradassi, Pau Blasi, Jose Ramón Monsalve, Guillaume Delattre, Javier Mérida, Antonio Bello, David Rodríguez, Kallum Pickering and my best and smart friend Edgard Rodríguez.

Throughout their thoughts I have accomplished this project but also personally experienced the warmth sensation of someone caring for you.

QUANTITATIVE ANALYSIS OF COMMERCIAL AND RESIDENTIAL REAL ESTATE MARKETS

Ramiro J. Rodríguez[1]

INTRODUCTION

Executive summary

The first chapter of this thesis examines the formation process of residential prices in Spain (1995Q1 – 2012Q4). We propose two models to compare their performance in the context of comparative dynamics and predictive capacity. A structural model is derived from an eclectic theoretical framework in which we review published literature on the housing market and select a set of variables representative of this literature. We used GDP pre-capita, interest rates, the supply of new residential buildings and the gross residential-capital formation as explanatory variables for the average house price per square meter in Spain. The other model is generated by an algorithm known as GASIC[2]. Using our review of the literature we select a set of 46 variables, we form the respective database and let the algorithm to select the best model out the 2^{46G}(70 trillion) nested models. The condition imposed on the algorithm is to be parsimonious, i.e. having only 4 regressors. Annual theoretical effort of families to pay for their residence, the apparent concrete consumption, the mortgage interest rate and the real GDP are selected by GASIC to explain the average residential price in Spain; a similar model to the structural one.

Our analytical framework is cointegration. Therefore, we assessed the integration order of both models' variables. We identified all variables have order of integration of first degree (some with a structural break in the recent economic crisis). This leads us to test the hypothesis of cointegration. Proving such an existence, two error correction models (ECM) were estimated (one for the structural approach and one for the algorithmic) to calculate price and income elasticities, and produce dynamic forecasts.

The long-term equations in both models behave similarly and give a good idea of the long-term equilibrium relationship between housing prices and their fundamentals. It is in the short term specification where the structural model and the algorithmic model differ. The model generated with GASIC has got a non-significant error correction mechanism, implying that the gap between the change in housing prices and long-term path is not traced. The consequence of such failure generates less accurate house price forecasts. However, the analysis of elasticities remains valid in both long and short term price equations.

For its part, price dynamics of the structural model is adequate, with the expected signs for the regressors as well as a negative error mechanism, correctly bounded between minus one and zero. The dynamic forecasting also has high performance, given the low forecast residuals.

---

[1] Universidad Complutense de Madrid, PhD in economics candidate. European Economist - International Research, BNP Paribas Real Estate.

[2] Acosta-Gonzalez E, Fernandez-Rodriguez F. Model Selection via genetic algorithms illustrated with Cross-country growth data. Empirical Economics 2007;33; 313-337.

Other findings of this research are that prices adjust quickly when out of long-term path and that during the property boom and bust in Spain housing prices were not so far from economic fundamentals.

In the second chapter we revisit cointegration techniques but for Madrid's office market. We study three endogenous variables, i.e. average real office rent, vacancy rate and office stock. Our database is provided by BNP Paribas Real Estate having a quarterly structure (2001Q1 to 2015Q2). Following Englund et al., 2008[3], we estimate a system of equations that also depends on an exogenous economic driver. In the literature such exogenous variable ranges from national to regional activity indicators. Consequently, we compare how well the modelling fits to both, Spanish GDP and Service Sector Employment. We also assess the performance of single equation error correction models (SEECM), as our literature review yields no commercial property research made so far with this approach and the preferred approach is the two stage error correction modelling (2SECM).

The equations used to model the dynamics of Madrid's office market comprise average rent, vacancy rate and stock variation. Each short run equation depends on its own lags, as well as lags of the other endogenous variables. We included the error correction term of the long run rent equation and the vacancy rate. However, as the literature suggests, the long run value of the vacancy rate is a natural-constant level. Therefore, in the short run equations of the endogenous variables the impact of the gap between short run vacancy (its actual level) and its long run value (a constant figure, actually embedded in the constant term of the short run equation) is the actual coefficient of the vacancy rate level. The equation system was estimated by the method of seemingly unrelated regression (SUR) in order to control for possible feedback across the three variables' residuals, and therefore increase the efficiency of our estimations compared to independent OLS regressions.

For estimations using Employment or GDP as well as for estimations using the SEECM and the 2SECM, the correction mechanisms are negative, as expected, and their magnitude signal a mild pace of adjustment. In particular, we found that rent variations correct each quarter between 11% to 20% long term rent deviations, depending on if we model them with GDP, Employment or with SEECM or 2SECM. Rent growth responds to vacancy rate deviations from its natural level with a 3% correction each quarter regardless of the approach selected (SEECM or 2SECM). Vacancy rate adjusts between 5% to 11% to its own gap, depending on the modelling approach. For the case of office stock, it slowly adjusts to rent and vacancy gaps; around 1% of correction each quarter, as expected, given the inelastic properties of real estate supply. Other important finding is the estimated natural level of vacancy rate. In most of the estimated models we arrive to the conclusion that the natural vacancy rate level is around 6% to 8%.

Restricting the sample to 2001Q1 to 2010Q4 we dynamically projected our endogenous variables for the period 2011Q1 to 2015Q2. We conclude that for predicting rents and vacancy the least forecast error is obtained using 2SECM and GDP as economic proxy. However, when predicting the full system (or stock only) we may keep the estimation method, but move to employment as demand proxy.

---

[3] Englund, Peter, Åke Gunnelin, Patric H. Hendershott, and Bo Söderberg. 2008. Adjustment in Property Space Markets: Taking Long-Term Leases and Transaction Costs Seriously. Real Estate Economics 36 (1): 81–109.

In our third chapter we set the goals of 1) studying hedonical decomposition of office rents; 2) the utilization of spatial econometrics; 3) conform a rental index relying on the geo-hedonic rent level estimation of an archetypical office for the Madrid office letting market. We used a detailed database of office lease contracts with a semi-annual structure provided by BNP Paribas Real Estate. From this database we obtain date of the transaction, headline lease rent, business sector of the tenant and business district to which the leased office belongs. We match this database with other extended database from the Spanish Cadastre. This latter dataset, gives us added hedonic characteristics such as date of construction, geographic coordinates and technical quality of the building. A third database of geographic coordinates of underground station entrances was used to calculate another variable comprising distance of the leased office to closest metro entrance. We used a OLS benchmark model to compare the results of the spatial econometrics. The spatial model employed is the Spatial Lag, which fits the idea that in real estate markets the price reached in my neighbour's transactions may impact the price of my transaction. Moreover, the price reached by my closest neighbours will have more impact on the price of my transaction than the prices achieved by distant neighbours. We compared the Spatial Lag's explanatory capacity, the properties of the residuals and the estimated endogenous variable against the OLS approach. We found better results with the spatial approach in virtually all comparisons. In terms of elasticities, we find that most of the price decomposition is incorporated in the business district the let office is located in, the age of the property and the technical quality of the construction. We also found strong evidence of spatial feedback across the Madrid office market and that estimation should take it into consideration as it is an unseen characteristic of the transaction and ignoring it may lead to biased rent estimations.

ANÁLISIS CUANTITATIVO DE LOS MERCADOS INMOBILIARIOS RESIDENCIAL Y COMERCIAL

Resumen ejecutivo

El primer capítulo de esta tesis analiza el proceso de formación del precio medio residencial por metro cuadrado en España (T1 1995 – T4 2012). Proponemos dos modelos para comparar su rendimiento en los contextos de estática comparativa y capacidad predictiva. Un modelo es estructural, derivado de un marco teórico ecléctico en el cual revisamos la literatura publicada en el sector inmobiliario residencial y seleccionamos un conjunto de variables representativo de esta literatura. Utilizamos el PIB per cápita, las tasa de interés, las entregas de los nuevos edificios residenciales y la formación bruta de capital inmobiliario como variables explicativas del precio residencial medio por metro cuadrado en España. El otro modelo es generado por un algoritmo conocido como GASIC [4]. De nuestra revisión de la literatura seleccionamos un conjunto de 46 variables, formamos la respectiva base de datos y dejamos que algoritmo conforme el mejor modelo posible de los 2^{46G}(70 billones) modelos anidados. La condición impuesta al algoritmo es que sea parsimonioso, o sea, que tenga solo 4 regresores. El esfuerzo teórico anual de las familias para pagar su residencia, la producción aparente de concreto, el tipo de interés hipotecario y el PIB real son seleccionados por GASIC para explicar el precio medio residencial en España; un modelo similar al estructural.

Nuestro marco analítico es de cointegración. Por lo tanto, evaluamos el orden de integración de las variables de ambos modelos. Se ha identificado que todas tienen orden de integración de primer grado (algunas de ellas con un shock estructural en la reciente crisis económica). Esto nos da pie para probar la hipótesis de cointegración. Demostrando tal existencia, se han estimado dos modelos de corrección del error (ECM) para calcular elasticidades precio e ingreso y producir previsiones dinámicas.

Las ecuaciones de largo plazo en ambos modelos se comportan de forma similar dan buena idea de la relación de equilibrio de largo plazo entre el precios de la vivienda y sus variables fundamentales. Es en la especificación de corto plazo cuando el modelo estructural y el modelo algorítmico difieren. En el modelo generado por GASIC, el mecanismo de corrección del error es no significativo, lo que implica que la brecha entre la variación de precios de viviendas y su senda de largo plazo no es capturada por el modelo. La consecuencia de tal falta genera previsiones menos precisas de los precios del inmobiliario residencial. Sin embargo, el análisis de las elasticidades sigue siendo válido para ambas especificaciones de largo y corto plazo.

Por su parte, la especificación de la dinámica de precios del modelo estructural es adecuada, con los signos esperados para los regresores y un mecanismo de corrección del error negativo y acotado entre menos uno y cero. La previsión dinámica presenta un alto rendimiento, dados los bajos errores de previsión.

[4] Acosta-Gonzalez E, Fernandez-Rodriguez F. Model Selection via genetic algorithms illustrated with Cross-country growth data. Empirical Economics 2007;33; 313-337.

Otros hallazgos de esta investigación son que los precios se ajustan con rapidez cuando están fuera de la ruta a largo plazo y que durante el auge de la propiedad y caída del sector inmobiliario en los precios residenciales en España, no estaban tan lejos de sus fundamentos económicos.

En el segundo capítulo vamos a retomar las técnicas de cointegración, pero para el mercado de oficinas de Madrid. Estudiamos tres variables endógenas, es decir, precio medio real del alquiler de oficinas, las tasas de disponibilidad y el parque de oficinas. Nuestra base de datos, proporcionada por BNP Paribas Real Estate, cuenta con una estructura trimestral para el periodo T1 2001 a T2 2015. Tomando como referencia el trabajo de Englund, et. Al, 2008[5], se estima un sistema de ecuaciones que depende de un impulsor económico exógeno. En la literatura, tal variable exógena oscila entre indicadores de orden nacional hasta de orden regional. En este trabajo se compara el ajuste del modelo tomando como referencia el PIB español (variable nacional) y el empleo del sector servicios (variable regional). También evaluamos el rendimiento de los modelos de corrección del error de ecuaciones individuales (single equation error correction mechanism, SEECM), dado que nuestra revisión de la literatura ha indicado la inexistencia del uso de este enfoque en la investigación del inmobiliario comercial hasta la fecha y el enfoque preferido es el modelado de dos etapas de corrección de errores (two stage error correction mechanism, 2SECM).

Las ecuaciones usadas para modelar la dinámica del mercado de oficinas de Madrid comprenden la variación de la renta media, la variación de la tasa de disponibilidad y la variación del parque construido. Cada ecuación de corto plazo depende de sus propios desfases, así como también los desfases de las otras variables endógenas. Se incluyeron el término de corrección del error de la ecuación de largo plazo de la renta y de la tasa de vacío. Sin embargo, como sugiere la literatura, el valor a largo plazo de la tasa de vacío es un nivel natural constante. Por lo tanto, el impacto de la brecha entre la disponibilidad (su nivel real) y su valor de largo plazo (una cifra constante, en realidad embebida en la constante de la ecuación de corto plazo) es el coeficiente de la disponibilidad, en niveles. El sistema de ecuaciones se estimó por el método de regresión aparentemente no relacionada (SUR) con el fin de controlar la posible retroalimentación a través de los residuos de las tres variables endógenas, y por lo tanto aumentar la eficiencia de nuestras estimaciones en comparación con regresiones MCO independientes.

Para las estimaciones utilizando Empleo o PIB, así como para las estimaciones utilizando el SEECM y la 2SECM, los mecanismos de corrección son negativos, como se espera, y su magnitud de la señales de un ritmo suave de ajuste de las rentas. En particular, hemos encontrado que las rentas se corrigen cada trimestre entre el 11% y el 20% ante desviaciones del precio del alquiler de largo plazo, dependiendo de si los modelos se estiman con el PIB o el empleo o con SEECM o 2SECM. El crecimiento de los alquileres responde a las desviaciones de la tasa de vacío de su nivel natural con una corrección del 3% cada trimestre, independientemente del método elegido (SEECM o 2SECM). La tasa de vacío se ajusta entre 5% a 11% ante desviaciones de valor de largo plazo, dependiendo del enfoque de modelado. Para el caso del parque de oficinas, este se ajusta muy suavemente cada trimestre, ante desviaciones de renta y disponibilidad. El cambio es de alrededor de 1% cada

[5] Englund, Peter, Åke Gunnelin, Patric H. Hendershott, and Bo Söderberg. 2008. Adjustment in Property Space Markets: Taking Long-Term Leases and Transaction Costs Seriously. Real Estate Economics 36 (1): 81–109.

trimestre, como era de esperar, dada las propiedades inelásticas de la oferta de bienes raíces. Otro hallazgo importante es el nivel natural estimado de tasa de vacío. En la mayoría de los modelos estimados se llega a la conclusión de que el nivel de tasa natural de disponibilidad natural es de entre 6% y 8%.

Restringiendo la muestra a T1 2001 – T4 2010 realizamos la previsión dinámica de nuestras variables endógenas para el período T1 2011 a T2 2015. Llegamos a la conclusión de que para la predicción de los alquileres y de disponibilidad se obtiene el menor error de pronóstico usando 2SECM y el PIB como impulsor económico. Sin embargo, a la hora de predecir el sistema completo (o la disponibilidad solamente) podemos emplear de nuevo el método 2SECM, pero usando el empleo como proxy de la demanda.

En nuestro tercer capítulo establecemos los objetivos de 1) el estudio de la descomposición hedónica de los alquileres de oficinas; 2) la utilización de la econometría espacial; 3) conformar un índice rentas que provenga de un modelo de estimación hedónica-espacial de la renta de oficinas de Madrid. Se utilizó una base de datos detallada de los contratos de arrendamiento de oficinas con una estructura semestral proveída por BNP Paribas Real Estate. A partir de esta base de datos se obtiene la fecha de la transacción, titular del contrato (ocupante), sector de actividad del ocupante y distrito de negocios al que pertenece la oficina alquilada. Hemos extendido esta base de datos cruzándola con otras bases de datos del Catastro español. Este último conjunto de datos nos da características hedónicas como la fecha de construcción del inmueble, coordenadas geográficas y la calidad técnica del edificio. Se utilizó una tercera base de datos de las coordenadas geográficas de entradas de la estación de metro para calcular la distancia entre la oficina alquilada y la entrada del metro más cercano. Se utilizó un modelo de referencia OLS para comparar los resultados de la econometría espacial. El modelo espacial empleado es el retardo espacial, que se ajusta a la idea de que en los mercados de bienes raíces el precio alcanzado en las transacciones de los vecinos afecta el precio de mi transacción. Por otra parte, el precio alcanzado por mis vecinos más cercanos tendrá más impacto en el precio de mi transacción que los precios fijados por los vecinos más distantes. Se compara la capacidad predictiva, las propiedades de los residuos y la variable endógena estimada entre el modelo espacial y el MCO. Encontramos mejores resultados con el enfoque espacial en prácticamente todas las comparaciones. En términos de elasticidades, nos encontramos con que la mayor parte de la descomposición de precios se incorpora en el distrito de negocios de la oficina, la edad del edificio y la calidad técnica de la construcción. También se encontró una fuerte evidencia de retroalimentación espacial a través del mercado de oficinas de Madrid y que la estimación debe tomarlo en consideración, ya que es una característica no visible de la transacción y hacer caso omiso de ello puede conducir a estimaciones sesgadas de las rentas de alquiler.

Motivation for the research

The candidate's main personal drivers for making this research were twofold:

1) Exploring and learning how scientific knowledge is produced. This is by far his main finding: knowledge seekers endeavour to test their ideas and intuitions, even feelings and hunches, with creative and innovative tools. They humbly open their findings to sceptic scrutiny by most-of-the-time unknown peers who traditionally depart from the premise that what the researcher statement is false and after careful study, at its best, they cannot claim as untrue. This is one fundamental and beautiful

feature of our society's knowledge. It is a system that expands itself by the constant interplay of researchers and referees on the back of hypothesis rejection and non-rejection, but never on the complete acceptance of truth. After this finding, the candidate went on to try to produce his own small contribution. He has tested his own hypotheses and happily puts below his late academic venture.

2) Gain strong competences in forecasting with econometric tools. He wanted to develop skills to deliver the insights of likely future developments of the markets. His learnings seem to work and their reflexion is this document. Far from a perfect forecasting exercise, what is more important here is the proposal of modelling methodologies and new applications of current technologies on real estate markets, in general, and on Spanish markets, in particular. The candidate hopes those who get to read this thesis find it at least as a stand point to impel the discussion on applied economics on property markets.

At an academic level we may start saying that real estate markets have a deep relationship with economic growth and welfare. Moreover, economic developments in Europe since 2007 deeply changed the backdrop of space markets and, accordingly, prices have altered dramatically in peripheral economies as the Spanish one and implications in today's markets are conspicuous rendering a scenario, in the view of the candidate, worth to investigate. In this context, with this research we scrutinize different techniques of econometric analysis on the price formation process under both the long and short term perspectives aiming to validate existing techniques of model selection, estimation and forecasting as well as their cogency in the recent market developments in Spain.

The assessment of this thesis author is that economic literature is quite scant in terms of Spanish commercial property research. The most important reason may be the virtually complete absence of official sources of statistical information on commercial property. Actually, for this research we use private database for office rents, vacant space and stock. Therefore, we give for the first time light from the stand point of academic research to the issue of commercial property in Spain. Compared to the pioneering markets in terms of commercial property research (London and US cities) the lag, before this work, has got to some 20 years.

There may be some spill over effects of office markets research and its lessons may be extended to other commercial property markets such as retail and logistics property markets, as their fundamentals and market dynamics are related to each other as it is the firm who is making the decisions. Learning lessons in the commercial property market from the housing market is less articulated as the latter is based on person's decisions rather than businesses'. That is the reason why in this thesis we wanted to cover these two markets: Residential property which may be business-to-consumer market and office market which normally is business-to-business.

Objectives of the thesis

In a series of three papers we intend to analyse the price formation process in property markets with three types of datasets and give new references of research on both the residential and commercial property markets. Namely the objectives of this work are (1) to analyse the residential property market in Spain using as the main endogenous variable for the models utilised the average price. Besides studying different impacts of the selected exogenous variables, we wanted to explore the accuracy of automatic model selection techniques that, may come in hady when having extensive datasets with a great amount of 'candidate' exogenous variables. These techniques may be useful when theory does not outline a particular model specification, such as in the residential markets literature where, as pointed by our eclectic approach, models greatly vary in terms of type and nature of exogenous variables. Ideally, an automatic modelling technique finds a parsimonious specification from several combinations of several candidate regressors. Point in case, given the extension of our database, some 73 trillion nested models were necessary to 'visit' in order to get a final parsimonious model of 5 variables to explain average residential price in Spain. We also have the objective of implementing, to the best of our knowledge, for the first time an automatic modelling technique to the real estate market. (2) To employ cointegrating forecasting techniques for a system of equations implying supply, demand and rents for the Madrid's office space market. This approach compares two cointegrating techniques and two economic exogenous drivers therefore, making several modelling scenarios to study long and short run leasing prices determinants and also forecasts of rental values in the commercial market. This approach is innovative in two senses: a) It is the first time error correction mechanisms are applied in a Spanish market under the framework of a system of equations and b) we have tested for the first time a single equation error correction mechanism to property markets. (3) To contribute to real estate price index estimation by means of hedonical models that take into consideration the geo-localization of the comparable transactions participating in the price model dataset by means of spatial econometrics. This is the first time that spatial econometrics has been brought to commercial real estate, as far as we have been able to find.

The three aforementioned objectives will be reached in their corresponding chapter. Therefore, in each chapter there are further objectives, more specific than the so far commented. Each thesis chapter is essentially an individual research paper with its own structure fitted to that of a 'publishable' paper in a scientific-peer-reviewing journal and has its own hypotheses, methodology and dataset, but all the time orbiting around property markets analysis. Below we make a more detailed description of each chapter's objectives:

Our fist paper has a double purpose: a) investigating the main drivers of housing prices in Spain under the light of long and short term dynamics and b) compare the performance of structural modelling and automatic model selection methods. We used an eclectic approach to make our modelling, meaning that we have collected a great extent of literature and extracted a set of variables and their respective proxies and collected our own version of such proxies for the Spanish residential market. The outcome of this effort has been a panel of 53 variables, including several definitions of housing price. The search for the maximum number of variables discussed in the real estate literature was in order to gather the sufficient amount of series that allowed a blind data-driven modelling technique to be applied. One benefit of such a sizable dataset is that we have material for further research. As a by-product, in a further iteration of this thesis, we are to extend our research to analysing price formation processes in bubble conditions as well as with system of equations instead of a single price equation. Since the beginning of our research we wanted to try for the first time automatic algorithms for modelling selection to the real estate market. We studied the GASIC (Genetic Algorithm with Schwarz Information Criterion) technique [6], consisting in a computer algorithm that selects a parsimonious model of, in this case, 4 variables, out of 48 candidate regressors. The GASIC algorithm actually 'visits', but does not estimate, nearly 73 trillion nested models, as it relays of combinations of 'parent' models to produce better fitting 'offspring' than their parents. The selected model happens to be the best fitted model for the endogenous variable and uses a genetic procedure, meaning that combination of explanatory variables has to beat their parents in terms of information criterion that in this case is the Schwartz Information Criterion. Further, the automatically selected model is comparted with a structural model. This model is proposed as the most common used model in literature to predict real estate prices giving our work its eclectic approach, which we find quite fitted to residential property given the great extent of different explanatory variables in the literature. We use an error correction model (ECM) to estimate variables' elasticities and make a dynamic forecast performance comparison. Our structural model outperforms the automatic selected model in its dynamic forecasting properties. However, complementary learnings from the estimated elasticities of each model can be extracted.

Our second paper also uses the ECM approach but studies a commercial property market. In this case we selected Madrid office market and this work is innovative in several ways. We build a system of three equations to model short term dynamics. We take into consideration one equation for rent change, one for vacancy rate change (ratio of available stock to total stock) and one for office stock change. We use an error correction framework to incorporate deviation gaps from the long term equilibrium rent and vacancy rate levels. Consequently, we also estimate long term expressions of average real rents and vacancy rate. As commented before, we wanted to compare results using two different versions of the error correction mechanism techniques. On the one hand we employ the two-stage error correction mechanism[7] as well as the single equation error correction mechanism [8]. We also search for exogenous variables' impact measurement. In this sense we have followed theory and practice choosing Spanish gross domestic product as the main demand proxy for office space demand in Madrid. At the same time, we have compared such modelling with that resulting from selecting a different, but related, driver for space demand: Madrid's service sector employment. As a result, we got to two perspectives of how macro variables impact the office market. Our main finding in this sense is that as Madrid is one of the main economic hubs in Spain, national GDP actually works for modelling rental levels and in terms if impact measurement and forecasting works almost as good as regional service sector employment. This conclusion has been extracted by using several methods of forecast error measurement, in all modelled equations of rents, vacancy rate and office stock. We have also concluded that when

[6] Acosta-Gonzalez E, Fernandez-Rodriguez F. Model Selection via genetic algorithms illustrated with Cross-country growth data. Empirical Economics 2007;33; 313-337.
[7] Engle, Robert F., and C. W. J. Granger. 1987. Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica 55 (2).
[8] Banerjee, Anindya, Juan J. Dolado, John W. Galbraith, and David Hendry. 1993. Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data. OUP Catalogue. Oxford University Press.

comparing cointegration techniques, the two-stage error correction mechanism fits better than the single equation error correction mechanism.

Our third paper also explores commercial property market but from a different perspective: market rents estimation. The initial purpose of this paper was to apply, for the first time, hedonical rent estimation techniques to property markets in Spain. It was a simple but valid enterprise, as so far no such practice has been made for any Spanish office market from the academic research. Therefore, a classical OLS hedonical estimation seemed enough to produce a PhD chapter. Departing from this objective the authors proceeded to build the hedonical database with which Madrid's office markets rent levels would be explained. Two, say, serendipities ensued: 1) The need of our current and classical hedonic OLS estimation to increase its explanatory capacity combined with the increasing references in recent literature on Spatial Econometrics and its adjustment to hedonical modelling. 2) The possibility of usage of geographic coordinates to apply spatial econometrics and the fact that our source of hedonical information had actually the geographical coordinates of each property of the city, later crossed with our transaction and prices database. So we set a new objective for this chapter and was testing hedonical estimation including spatial econometrics. Our first goal was to produce estimates of the letting rent of an ideal or typical office using the average hedonical characteristics. Plenty of literature in real estate uses price decomposition with hedonical modelling; less are references integrating spatial feedback. This set our second goal, being that of search for evidence of interplay of rental prices through unseen characteristics such as physical approximation of the transactions. A third goal came 'by default' with the second: if actually the spatial feedback was proven, measuring the size of such impact. A fourth goal was to compare the results of estimation with and without spatial econometrics, both in terms of estimated hedonic variables' elasticities and estimated rents. Our results point to effective improvement in terms of explanatory capacity and forecast error when using the spatial approach.

MODELLING RESIDENTIAL PRICES WITH COINTEGRATION TECHNIQUES AND AUTOMATIC SELECTION ALGORITHMS[9]

[9] This work has been presented in the European Real Estate Society Congress 2014 in Bucharest, Romania. It has been awarded with the Doctoral Prize as "Best Peper presented at the European Real Estate Society 2014 Conference". The candidate much appreciates ERES organization for their support.

1.1. Introduction

On the one hand, housing is both an investment as well as consumption good. On the other, it is a key sector for any economy as it has inter-linkages with other industries: construction, renovation, maintenance and those related to trading, financing, mortgage banking, real estate agents, appraisers, movers, notaries, etc. Moreover, housing sector is impacted by both monetary and fiscal policy, macro prudential norms and labour policy prevalent in the economy (Hilbers *et al.,* 2008). House prices vary in response to changes in both housing demand and housing supply. A number of empirical studies establish that key determinants of housing prices are income levels, interest rates, supply conditions, demographic changes, number and size of households, maintenance costs, property taxes, and speculative pressures [see Olsen (1987) and Whitehead (1998) for broad reviews of the early empirical literature].

In Spain, house prices have been growing at very high rates within the period 2002-2008, thereby providing a significant support to economic activity, through wealth effects, and raising concern that real estate markets could be subject to speculative waves that could eventually trigger sharp corrections and generate macroeconomic and financial instability.

The last boom of the Spanish housing market, which ended with the bust of the bubble in 2008, has offered great opportunities in real estate research to gain insights of price formation processes in an economy with a renewed institutional framework. Since the integration of the Spanish economy to the European single market, owned housing has seen a noteworthy boost, as income, credit access and cost of debt played advantageously to this end. However, as expectations on swift housing price growth were formed, off-setting forces as stock increase and grater shares of income dedicated to house acquisition were disengaged and an ever increasing trend in prices followed suit. This was the signal of the existence of a bubble in that particular market as well as of the estrangement from the long term trend of fundamental variables (i.e., house price). By analysing the residential property market from the scope of long term relationships and short term adjustment processes, we use a cointegrating framework to analyse the main forces driving aggregate house prices in Spain. The major findings of this paper are: 1) with the data used in this work it is possible to represent a long run equilibrium path for the house price, throughout the fully-modified procedure suggested by Phillips and Hansen (1990). The result is also used to estimate an Error Correction Model (ECM) in a short run expression for housing price dynamics which conforms a structural modelling of prices for Spain. 2) Long term house price responds positively to purchasing capacity and negatively to interest rate and new residential stock added each quarter. Nevertheless, capital formation seems to positively affect prices, suggesting improvement in properties increase property values. 3) Short term price levels oscillate along the long term path: nearly 22% of the price deviation is corrected each quarter in the Spanish market.

The ECM has been used as a benchmark against which we have compared the forecasting accuracy of an algorithmic model selection technique. In particular the model selection technique used here has been that of a Genetic Algorithm (GA) known as GASIC, developed by Acosta-Gonzalez, and Fernandez-Rodriguez (2007). The main finding of this exercise is twofold: 1) the automatically selected model has good properties for forecasting as fitted as the structural model. 2. Although the selected variables (from a pool of 46 candidate regressors) not always have the expected signs, the ECM estimated regressors have high significance.

The remainder of this paper is structured in this manner: Section 2 presents the data used for modelling, their definitions and adaptations to this work. Section 3 describes the econometric methodology adopted in this study and the empirical results obtained. Section 4 offers some concluding remarks.

1.2. Data description and database creation

As the aim of this paper is twofold (i.e. generate a structural model with an ECM framework and test it against automatic modelling processes) we have created a comprehensive collection of real estate variables. In a first step we have made a profound review of the economic literature in real estate aiming to pin down the greatest extent of variables participating in recent economic real estate related literature. In this stage a set of 167 variables was created (see Appendix). To better capture the dynamics of the market and maximize the number of observations, we decided to build a quarterly database. From this point we began to construct the database with a thorough selection of sources. One point of reference has been the 'Síntesis de Indicadores de la Vivienda' (SIV), a gathering of 86 indicators of the residential real estate sector in Spain from different official sources collected by the Spanish Central Bank. The structure of this database is monthly, but its indicators have varied frequencies ranging from monthly to decennial. Though the first observation starts in 1960, little of the dataset is that long and we decided to set the beginning of the streamlined database in 1995, a year when 50 out of 86 variables started being measured and we conveniently capture two complete economic and property cycles [see, e. g., Berge and Jordà (2013, or Economic Cycle Research Institute (2014)].

Monthly observations had to be arranged for quarterly data: Flow variables had to be aggregated for the three months of each quarter, stock variables were taken in the final month of the quarter and other variables such as interest rates and stock market index were averaged.

To the prevailing variables from SIV dataset we added data coming from other sources as Spanish National Statistics Institute, Ministry of Public Works and Bank of Spain, among others.

All monetary variables have been deflated by the implicit Gross Domestic Product (GDP) deflator that, in its turn, has been calculated as the ratio of Quarterly Nominal GDP to Quarterly Real GDP.

1.3. Econometric methodology

We use two types of models that are to be compared in terms of both estimation accuracy and forecasting capacity. Below we describe the ECM approach and the automatic model selection techniques employed in this study.

A structural ECM

We follow the previous literature and investigate the long-term and short-term determinants of house price movements using a two-step approach [see, e. g., Abraham and Hendershott (1996), Malpezzi (1999), Capozza *et al.*(2002), and Meen (2002)]. In a first step, the fundamental value of housing is calculated. In the second step, the short-term dynamics of house prices are determined by a mean reversion process to their fundamental values and by a serial correlation movement.

Long-run equilibrium

It is assumed that in each period there is a fundamental value of housing that is largely determined by economic conditions in the form of

$$P_t^* = f(X_t) \qquad \text{(Eq. 1.1)}$$

Where P_t^* is the log of the real fundamental value of house prices at time *t*, *f (·)* is a function and X_t is a vector of macroeconomic variables conforming house price fundamentals.

Assuming a log-linear relationship between the dependent variable and its determinants, we can obtain the following long-run equilibrium equation suitable for estimation:

$$\log(P_t^*) = \alpha_0 + \sum_{i=1}^{q} \alpha_i \log(X_{i,t}) + \xi_t \qquad \text{(Eq. 1.2)}$$

Where the unobservable variable P_t^* has been substituted by the log of (observed) real house prices (P_t).

Short -run dynamics

Arguably, equilibrium is rarely observed in the short-run due to the inability of economic agents to adjust instantaneously to new information. According to Granger Representation Theorem (see, Engle and Granger, 1987), a cointegrated system of variables can be represented as an ECM, and vice versa. Therefore, in a second-step, the following ECM including the lagged residuals from the cointegrating regression (Eq. 1.2) as an error-correction term can be postulated in order to model the short-run dynamics:

$$\Delta\log(P_t) = \sum_{i=1}^{q} \sum_{j=1}^{n} \gamma_{ij} \Delta \log(X_{i,t-j}) + \beta \xi_{t-1} + \varepsilon_t \qquad \text{(Eq. 1.3)}$$

where Δ denotes first difference.

The ECM captures the short-run dynamics towards long-run equilibrium in the form of gradual adjustment and incorporating the information provided by past disequilibria. In equation (1.3), given that housing is a slow-clearing durable asset, it is reasonable to expect that current price changes are partly governed by the deviation from the fundamental value (0<β<1) and partly by contemporaneous adjustment to changes in fundamentals ($0<\gamma_i<1$). Therefore, estimates of γ_i provide us with short term effects of X_{it} on P_t, while estimates of β offer the speed at which P_t returns to equilibrium after a deviation has occurred.

1.4. Database and data structure

a. Data

We have gathered five variables for the residential market such as real house price per square meter (*HPM2*) as endogenous variables: gross domestic product per capita (*GDPPC*), mortgage interest rate (*MORTRATE*), free market residential buildings starts (*BSFREE*) and real gross capital formation in dwellings (*GCFDWELL*) as regressors (see Appendix for a full description of the data). This specification has been adopted in line with Gattini and Hiebert (2010), Iacoviello and Minetti (2008) and Iacoviello (2002), who investigate effects of monetary policy and business cycles on residential house prices by the means of parsimonious specifications.

i. *GDPPC* is used as a proxy of households' purchasing capacity. It is a demand-side factor: We posit that higher income tends to encourage greater demand for housing, therefore pushing up house prices. It is the measure of the quarterly value of national output at constant euros of 2008.

ii. *MORTRATE* is used as a proxy of (opportunity) cost of resources invested in housing for households. The higher the cost of issuing debt for house acquisition, the lower the hose price. It is the weighted average of more than three years mortgage credit rate.

iii. *BSFREE* is used as a proxy of change of level of supply. It is a supply-side factor: In the long run, an increase in housing stock tends to bring down house prices. It is the number of new residential units delivered to the market at a national level.

iv. *GCFDWELL* is used as proxy of value added in the economy invested in dwelling instead of being consumed, measured in constant euros of 2008.

b. Integration tests

A previous step in cointegration analysis consists of testing the order of integration of the variables. To that end, we tested for the order of integration by means of the Augmented Dickey-Fuller (ADF) tests (Dickey and Fuller, 1981). Following Carrion-i-Silvestre *et al.*'s (2001) suggestion, we confirm this result using the Kwiatkowski et al. (1992) (KPSS) tests, where the null is a stationary process against the alternative of a unit root. The three versions of the ADF test and the two versions of KPSS test were calculated for each variable. The decision rule was observing if three out of the five tests ran yielded non-stationarity or stationarity. The results for *MORTRATE* and *BSFREE*, not shown here to save space but available from the authors upon request, decisively reject the null hypothesis of non-stationarity in the first regressions. They do not reject the null hypothesis of stationarity in first differences, but strongly reject it in levels, in the second ones. So, they suggest that *MORTRATE* and *BSFREE* can be treated as first-difference stationary (i. e., I(1) variables). As for *HPM2*, *GDPPC* and *GCFDWELL*, the results indicate that they are second difference stationary, perhaps due to a long lasting bubble bust process in Spain. Constant variations in the same direction (e.g. permanent discounts of property prices) do not allow these three series to lose their trend when first differentiated. However, a strong change in any series trend and/or level signals structural breaks, therefore we resort to check stationarity under structural breaks, based on Perron (1997) and Zivot-Andrews (1992) unit root tests. These tests check for unit root with a break in the intercept, trend or both at an unknown time on any given series. Both tests have a null hypothesis of existence of unit root with a structural break and endogenously select the date of the break. The joint use of the tests will give additional support to our assumption of stationarity of the first differences of the three above mentioned variables. Details can be seen in Tables 1.1 and 1.2.

Table 1.1. Integration order with structural breaks using Perron test

Variable	Is	With a structural break in	In quarter
HPM2	I(1)*	Trend	2 – 2004
GDPPC	I(1)*	Intercept	1 – 2008
GCFDWELL	I(1)**	Intercept	1 – 2008

Notes:
* denotes significance at the 5% of level of confidence
** denotes significance at the 10% of level of confidence

Table1.2. Integration order with structural breaks using Zivot-Andrews test

Variable	Is	With a structural break in	In quarter
HPM2	I(1)*	Trend	2 – 2004
GDPPC	I(1)*	Intercept	2 – 2008
GCFDWELL	I(1)*	Intercept	1 – 2007

Note:
* denotes significance at the 5% of level of confidence

Both tests yield results that can be interpreted as sound evidence of the first order of integration of our endogenous variable, as well as two of its regressors. If not for this procedure, cointegrating regressions could have not been utilised. Results of Tables 1.1 and 1.2 will be incorporated to our modelling, namely in the long term equation throughout a dummy variable.

1.5. Empirical results from ECM approach

Long-run Equilibrium

We initially followed the two-step estimation procedure for dynamic modelling suggested by Engle and Granger (1989). So, in a first step, we estimated the cointegration regression (2). Notice that, even though the estimation by ordinary least squares (OLS) of the cointegration regression yields superconsistent estimates, the joint dependence of most aggregate time series and their nonstationarity invalidate the routine application of many statistical procedures. To overcome this problem, in the first step of the Engle-Granger procedure, we alternatively applied the estimation method proposed by Phillips and Hansen (1990). This single-equation semiparametric method allows the direct estimation of the long run relationship in a two-step procedure, filtering the data in the first step using a nonparametric correction for serial correlation and second order endogeneity bias (see Banerjee *et al*,.1986).

The results of applying the Phillips-Hansen procedure to equation (2) are as follows:

$$\begin{array}{lll}
\log G PHM2G = G \underset{-15.14194G}{-10.1785G} + G \underset{20.39477G}{1.5644G} * \log G \quad PPG & -G & \\
-G \underset{-7.234887G}{0.1398G} * \log G MG & -G \; -G \underset{-10.78238G}{0.0735G} * \log G & -G \\
+G \underset{16.99713G}{0.5029G} * \log G & +G \underset{5.227061G}{0.0724G} * G \quad MMG & \text{(Eq. 1.4)}
\end{array}$$

$$= 0.9926G \qquad = 2.0098G$$

The figures in brackets below each coefficient are the standard *t*-statistics. Note also that, since the model is estimated in logs, the estimated coefficients denote elasticities. As can be seen in equation (Eq. 1.4), a first variable appearing marginally significant was households’ purchasing capacity, proxied by per-capita real gross domestic product. This result would be in line with Abraham and Hendershott (1996), Capozza *et al.* (2002) and Coleman *et al.* (2008), among others. This variable has a strong influence on the price variable and can be said that it is its main driver.

A negative and significant coefficient was obtained for the lagged interest rate for housing purchase, suggesting that in a declining interest rate environment, which keeps servicing costs of ever larger mortgages within the household budget limits imposed by current income, would have boosted the demand for residential real estate. The negative coefficient would also indicate a substitution effect between houses and other financial assets in investors’ portfolios, being consistent with the findings in Hofmann (2004) and Tsatsaronis and Zhu (2004).

The estimated coefficient on lagged housing stock variation showed a negative sign, as expected, giving evidence that new supply has been able in the period of analysis to counteract, at least in part, the positive pressure on house prices derived from a quite active demand with an increasing purchasing capacity as least since 2004 and until 2008. The idea of that housing stock would have been constrained in the short run as a result of the length of the planning and construction phases and the inertia of existing land planning schemes is not supported.

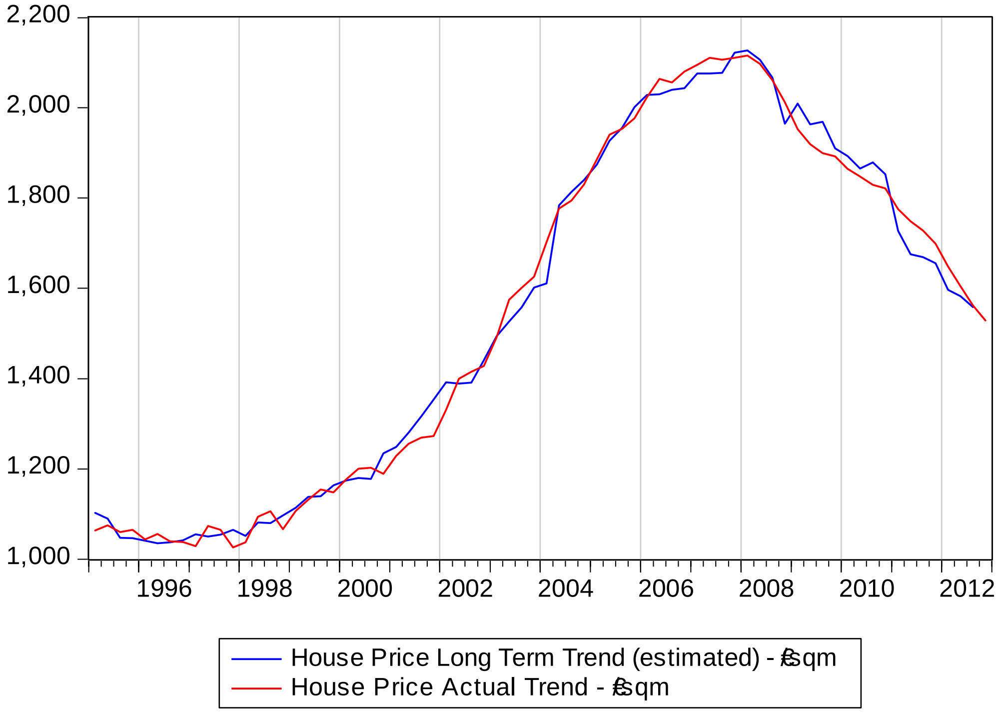
The estimated coefficient of gross capital formation resulted positive suggesting that investment in real estate assets, including refurbishments, increases its intrinsic value, therefore its market price.

The last estimator presented here corresponds to a dummy variable’s coefficient. This variable has a value of 0 if the observation belongs to a period before Q2 2004 and 1 if after. The presence of the dummy variable resolves two problems. It captures the structural shock to the market conveyed by the housing price boom in Spain as well as supports the assumption of first difference stationarity on GDP Per Capita as the Structural Break Test (Perron, 1989) confirms. Therefore, our model takes into account possible changes in market conditions, namely demand conditions, impelled by the price bubble.

As can be seen, the overall regression fit is very high, as measured by the value R^2. Additionally, the cointegrating regression Durbin Watson test statistic (*CRDW*) indicate that we can reject the null hypothesis of no-cointegration at least at the 5% level of significance, so equation (1.4) can be tentatively thought as representing a long-run relationship.

One useful application of the cointegrating regressions is that we can get acumens on what is the relative position of the actual price with respect to its theoretical long term trend as presented in Figure 1.1.

Figure 1.1. Actual and estimated long term trends of housing prices in Spain



Average, maximal and minimal percent deviation of the actual price from its long term level for the period analysed: 0.0%, 5.7% and 5.9%, respectively

In general terms, our model suggests that the housing prices in Spain do not get way too far from their fundamental value. Actually, while economy conditions are regular and no special price processes are undergoing, house prices fit quite close to their equilibrium level, such as in the period 1995-2000 (being 0% the average deviation from equilibrium price). Once the economy started to heat, came a period of upward drive in fundamentals (2001, with an average deviation from equilibrium price of -3%) followed suit by an over-reaction of actual prices in 2003 (average deviation from equilibrium price of 3%) when levels increased hastily. Then, we once again observe that prices caught-up their fundamental value in 2004-2005 (with an average deviation from equilibrium price of 0%) to remain above it, as expected, until the bubble bust in late 2007/early 2008 (being the average deviation from equilibrium price of 2%). Observed prices for a second time over-reacted to stand below their equilibrium in 2009-2010 (average deviation of. -2%). Finally, with the so called double-dip of the Spanish economy, another turn in the relation actual-equilibrium level developed in 2011-2012 (average deviation of 3%), when the economy set equilibrium prices below observed prices. Ending 2012, fundamental and actual prices met again, suggesting that the correction of the real economy permeated the property markets (deviation of. 0%).

We have run the Engel-Granger (1987) residual based cointegration test to assess the existence of cointegration in our single equation model. Table 1.3 reports the results for testing stationarity in the residuals of equation (1.4).

Table 1.3. Equation (1.4) series cointegration test

	Statistic name	Value	P-Value
Engel & Granger	Tau Statistic	-4.7316	0.0388
	Z-statistic	-34.0894	0.0344

The test employed calculates its statistics under the null hypothesis of non-existence of cointegration between the series. As can be seen in Table 1.3, both reject the null hypothesis of no cointegration (unit root in the residuals) at the 5% level, giving further support to the cointegrating equation (1.4). In this scenario, we can be sure we are not estimating spurious relationships among our variables and that equation (1.4) can be tentatively thought of as representing a long-run relationship.

Short-run dynamics

Regarding the short-run dynamics, we first followed the General-to-Specific (GETS) modelling approach (Hendry, 1995), initially over-parameterised ECM with four lags on the dependent as well as the explanatory variables [equation (3) with *m*=*q*=4] was continuously simplified and re-parameterised until a parsimonious representation of the data generation process was obtained. The OLS results are as following:

$$\Delta \log HPM2G = G \underset{-4.8559G}{-0.2878G} * \{ \log HPM2 \; -1G$$

$$-1.5644 * \log G \quad PPG \; -1G + 0.1398 * \log MG \qquad -1G$$

$$+ 0.0735 * \log G \qquad -1G - 0.5029 \log G \qquad -1G$$

$$-0.0724 * G \quad MM \quad + 10.1785 \}G + \underset{9.7248G}{0.6270G} * \Delta \log HPM2 \; -1G$$

$$- \underset{-2.3352G}{0.0404G} * \Delta \log MG \qquad -2 \quad + G \underset{2.1213G}{0.0185G} * \Delta \log G \qquad -4G$$

$$+ G \underset{2.1802G}{0.0182G} * \Delta \log G \qquad -2G \qquad \text{(Eq. 1.5)}$$

$$= 0.8294, G \qquad = 1.9105, G \quad = 1.7741, GMG = 1.3264, G \qquad H = 0.8974G$$

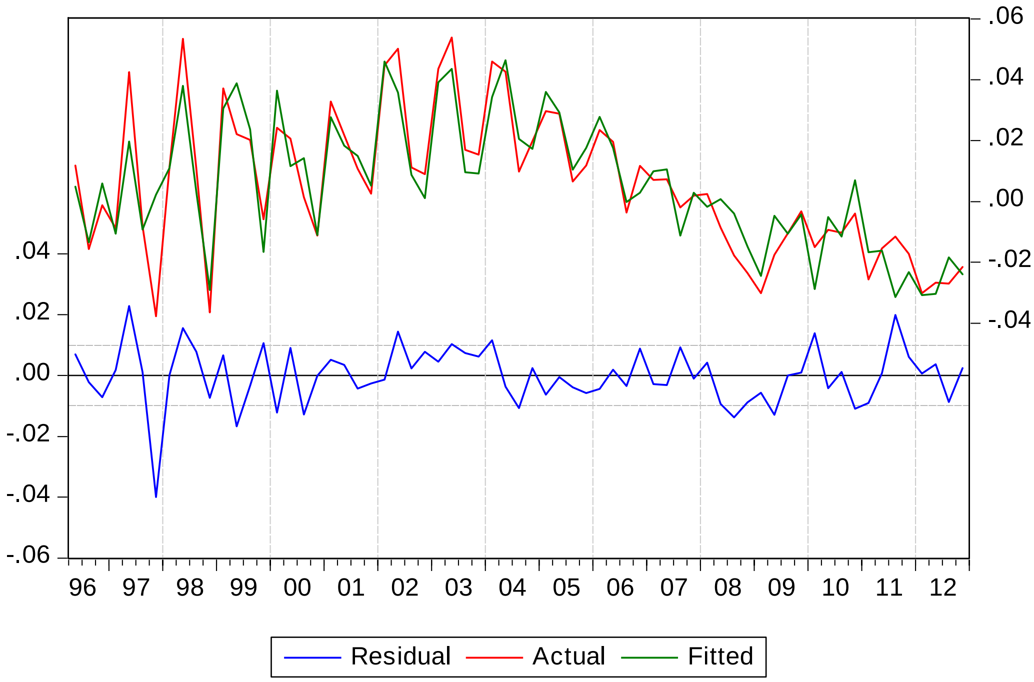
where Δ denotes first difference. Note that figures in brackets below each coefficient are *t*-statistics and that the first coefficient in (Eq. 1.5) is the estimator of the lagged residuals from equation (q.4).

As can be seen, the null hypothesis of no error correction term is rejected, giving further support to the cointegrating equation (1.4) as a long-run relationship (Kremers *et al.*, 1992). The estimated coefficients are statistically significant. In particular, we found that current price changes are positively affected by adjustments in house prices in the last quarter, negatively affected by changes in interest rates and positively affected by new increases in new deliveries. It should be noticed that we have not found a significant role for changes in GDP per capita nor Gross Capital formation in dwellings in the short-run equation. Finally, the estimated error correction term suggested that 29% of the disequilibrium is corrected each quarter. Therefore, ceteris paribus, once moved from equilibrium, in less than a year, Spanish house prices revert to their steady-state conduit.

When presented the results of equation (1.5), we also report some diagnostic test for normality, fourth-order residual autocorrelation and first-order autoregressive conditional heteroscedasticity (*N*, *LM* and *ARCH*, respectively), which do not show any sign of misspecification.

Figure 1.2 displays the actual and fitted values for the dependent variable in equation (1.5) $\Delta \log GHPM2G$, along with the residuals. As can be seen, the fitted values closely track the evolution of the observed variations in the residential property price per square meter, and the residuals remain inside the limits of one standard deviation.

Figure 1.3. Actual, fitted, and residuals from estimated equation (1.5)



1.6. Empirical results from automatic model selection techniques

We turn now to the analysis of house price dynamics with the GASIC algorithm. In a nutshell, automated model selection techniques pick out a particular model by avoiding assessing all sub-models. Previous literature (e.g. Lovell, 1983) used criteria such as t-ratio statistics to add a regressor to a particular model that contained the best regressor, which is a model using a unique regressor. After that new regressors are added one by one, as long as they are significant, in terms of their t-ratio at an already chosen level of significance. The process ends when all regressors not chosen are not significant. The inverse process is also possible, estimating an over-parameterized model with all candidate regressors available. The following step is eliminating, one by one, regressors with no significant coefficients given a chosen significant level against which the t-ratio is compared. The process keeps on until all participant regressors are significant. These methods - also known as data mining - are costly in terms of degrees of freedom and in terms of information, as the researcher has to increase her information on how the economy works as the model reduces its number of regressors (Lovell, 1983).

Different approaches using a search path appeared in the 1990's as Hendry (1995), Hoover and Perez (1999) and Hendry and Krolzig (1999) introduced the General to Specific model selection technique, also known as the LSE approach. This technique aims to find a parsimonious and encompassing model derived from a chosen General Unrestricted Model (GUM henceforth) in a sequence of steps. i) A GUM is chosen using researcher's intuition, theory, past evidence, etc. and controlling the parameterization is as orthogonal as possible (Hendry and Doornik, 2004) and may have several regressors. ii) After setting significance levels and miss-specification tests, the GUM is estimated using Instrumental Variables and some reduction tests ensue to eliminate irrelevant variables, therefore decreasing search complexity. iii) With this reduced GUM a path search procedure begins. A new model from the GUM is created by deleting surviving-from-step-two variables having the lowest and non-statistically significant *t*-ratios. The two models (with and without the variable or block of variables) are compared by means of some diagnostics test; if the reduced model outperforms the other, the next variable with the lowest *t*-ratio is selected and the variable is removed and a new round con comparisons initiate to check if the model without the chosen variable outperforms the other. If not, the variable is restored and the next variable with the lowest t-ratio is tested with the same procedure. The simplification process ends when all variables are significant and diagnostics tests fail to drop more variables, so a terminal specification is obtained. By choosing different critical values to set significance levels, new search-paths are created and new iterations of simplification processes commence, possibly yielding new terminal specifications. iv) Combinations of the competing models are formed and compared among the original Terminal Specification throughout *F*-tests. Chosen combined encompassing models become a new GUM and a new search-path process is started. When the combination of two models yield a non-encompassing model, the algorithm selects the best and final model using an information criterion (IC) that may be Akaike IC, Schwartz IC or Hannan-Quinn IC.

Lately, model selection techniques have tried to bypass the complicated LSE process, using only an information criterion as loss function, as proposed by Hansen (1999). However, this methodology, that uses the Schwartz IC (SIC), a Bayesian information criterion, does not always produce a reduced model and does not work with more than 10 candidate regressors (Acosta-González and Fernández-Rodríguez, 2007). Trying to solve these restrictions but leveraging on Hansen's claims that BIC rules can perform much better than the complicated Hoover and Perez (1999) algorithm, Acosta-González and Fernández-Rodríguez propose their Genetic Algorithm (GA). A GA is an optimization technique based on rules analogous to adaptive evolution of life, initially developed by Holland (1975). Using SIC as loss function Acosta-Gonzalez & Fernandez-Rodriguez designate their model selection technique with genetic algorithms as GASIC. The most appealing characteristics of this modelling technique are:

a) It performs as well as, or even better, than complicated selection techniques of the type of the LSE approach.
b) It can be used to undertake structural analysis in response to its parsimonious and robust model selection capabilities.

Selecting regressors with GA solves problems such as non-continuity or non-differentiability of the loss function. By starting with a randomly produced solution better approximations to an optimal solution are produced by applying the principle of fitness. Better performing solutions are combined in a cross-over binary breeding, in resemblance to Mendel's genetics. The objective of this crossover is to generate better fitted solutions to the optimization problem with respect to the solutions they were created from (evolutionary improvement). Some randomly generated 'mutations' are introduced to avoid local optima.

In particular, GASIC employs seven steps in order to automatically generate a model:

1. Selection of an initial population: Some models are randomly generated and each of them represents an approximation to the GUM. They are referred to as chromosomes. For example: 200 different models can be selected each of them with 5-tuple regressors.
2. Loss function ranking: SIC is calculated for every model and ordered from lowest to highest. Viability of using SIC in terms of the implicit significance of the estimated parameters is analysed by Campos, Hendry and Krolzig (2003).
3. Selection: In an analogy to Darwinian Natural Selection, most fitted chromosomes take hold while less fitted get extinct. In this case, the half of the models with the highest SIC gets erased.
4. Pairing: Surviving models from step 3 are randomly coupled
5. Origin of species: New models from the new couples are estimated. The exchange of genetic material is done placing a set of candidate regressors from one -mother- model onto the set of regressors of the other -father- model and vice versa, therefore producing two offspring chromosomes.
6. Random variation: This step is analogous to mutation. At this point some candidate regressors are randomly added or deleted from a randomly chosen subset of models. The aim of this step is to avoid local minimums in the loss function.
7. Satisfying a convergence criterion: Repeating the algorithm from Step 2 will engender successive generations of solutions. The end point of this process is defined by whether reaching a pre-set number of iterations or if the population come together to the same solution. In the case of this paper the convergence criterion has been the second.

We have used the GASIC algorithm to generate a parsimonious long term model from a group of possible exogenous variables. For this study, we have gathered 48 variables related to the Spanish residential market, with an eclectic approach, consisting in doing an extensive review of the variables utilized to model residential prices[10]. After collecting the variables used in the literature, we proceeded to build the data set for this variables. Using a public data base for the residential market in Spain called SIV from the Spanish Central Bank, we gathered most of the identified variables. We added other ones from trustable sources such as the Spanish National Statistics Institute and the Ministry of Public Works. A variable set (which actually is our GUM) was composed of 48 candidate regressors for modelling real house price in euros per square meter. The number of possible sub models using these 46 exogenous variables is 2^{46G}which equals to a little more than 70 trillion models. Off course computational demands are overwhelming and almost impossible to attend. Therefore, model selection algorithms such as GASIC are helpful to automatically opt

[10] An Appendix containing the papers used in this step, not presented in the interests of space, is available from the authors upon request.

for a restricted parsimonious and encompassing model, being efficient and consistent in its estimations. We capitalize the fact that SIC in the framework of linear regression and cointegration techniques are the same. Therefore, we state that GASIC remains valid to I(1) series as it may not be the of LSE approach, which relies on OLS estimations. Regarding the order of integration of the 48 candidate regressor, we ran ADF and KPSS tests for all of them. As mentioned before, the orders of integration ranged from 0 to 2 and in case I(2) variable were selected by the automatic algorithm, we have proceeded to test if they were I(1) with a structural break.

Long run equilibrium

We used different definitions of price in order to generate a long run expression for residential prices, yielding estimations that, besides being parsimonious and well estimated, corresponded well to economic intuition[11]. However, to maintain comparability with our structural model we kept the model selected through GASIC with the same endogenous variable (i.e. real average housing price for Spain, *HPM2*)).

The resulting auto-selected model used the following variables:

i. Theoretical annual effort of families (*EFFDED*): It is the share of the annual household income that is dedicated to pay the gross credit payments of a house financed in 80% of its value.
ii. Apparent concrete consumption (*ACC*): It is the difference in production (measured in metric tons) from one month to other at a national level, including concrete producers stocks and imports, excluding stock in hands of intermediaries. Statistics are reported by the Ministry of Public Works and collected by the Spanish Association of Concrete Producers. It acts as a proxy of housing supply.
iii. Mortgage rate (*MORTRATE*): used as a proxy of (opportunity) cost of resources invested in housing for households. The higher the cost of issuing debt for house acquisition the lower the hose price. It is the weighted average of more than three years mortgage credit rate.
iv. Gross Domestic Product Volume (*GDP2008*): It is the real value of GDP at 2008 prices. It works as a proxy of housing demand.

Apart from the endogenous variable, concrete consumption and real GDP were I(2), for the sample under study. As classical economic activity indicators these variables are and as intuition suggests, they have to be first order integrated. Hence, we again tested the stationarity of their first difference with a structural break, utilising Perron and Zivot-Andrews tests.

[11] Those estimations results are not shown here to save space, but they are available from the authors upon request.

Table 1.4. Integration order with structural breaks using Perron test

Variable	Is	With a structural break in	In quarter
GDP2008	I(1)*	Intercept	1 – 2008
ACC	I(1)*	Intercept	2 – 2008

* At a 5% of level of confidence

Table 1.5. Integration order with structural breaks using Zivot-Andrews

Variable	Is	With a structural break in	In quarter
GDP2008	I(1)*	Intercept	1 – 2008
ACC	I(1)*	Intercept	1 – 2008

* At a 5% of level of confidence

Results in Tables 1.4 and 1.5 suggest that the first difference of the two variables is stationary, with a structural break in the intercept at the beginning of 2008, time when the economy turned down to enter into the last recession. Therefore, these two variables can be used in the cointegration framework, with the caveat of such structural break[12].

The long run equation estimation in presented henceforth:

$$HPM2G = G{-}4.298316G + G0.639301G{*}G \quad -$$
$$-6.43833G \quad 21.63627G$$
$$+\ 0.160411G{*}G \quad - \quad -G\ 0.332009G{*}G \quad MG \quad -$$
$$17.22280G \quad -14.91011G$$
$$+G0.693200G{*}G \quad P2008G \qquad \text{(Eq. 1.6)}$$
$$11.46309G$$

$$= 0.9943G \qquad = 2.2802G$$

The figures in brackets correspond to the parameter estimators' t-ratios. These estimators reflect elasticities as long as the model was calculated in logarithms. In general, all the estimators are highly significant. The negative value of the constant means that in equilibrium, the house price is less than the combined weighted average of its regressors. The positive impact of the effort measure can be interpreted as an indication that as families dedicate greater shares of income to pay mortgages more pressure is put on housing buying therefore pushing prices up. Concrete consumption yielded a positive estimator, while it was expected negative. As expected, interest rate was negative, giving evidence of its role as credit access barrier as well as opportunity cost gauge. Finally purchasing capacity was captured by real GDP giving the positive expected sign. Regarding the obtained long term elasticities, all values estimated are below less than one percent, indicating that variation in the exogenous variables have no hyper-intensifying effects over house prices. Nevertheless, variations of one percent in family efforts, concrete consumption and GDP, imply positive prices variations of 0.6%, 0.2% and 0.7%, respectively. With a one percent increase in interest rate, there is a negative correction of house prices of 0.3%.

Having in mind that the Price, GDP and Concrete Consumption variables were first difference stationary with a structural break, we tried to take that into account in our estimation. We estimated a dummy variable having a value of cero before the shock

12 Full tests results are available from the authors upon request.

and value of one after it, mirroring the technique used in the long term estimation of our structural approach. Nonetheless, the dummy variable was non-significant and therefore we do not present the estimation results for this alternative specification, although it can be obtained from the authors upon request.

As in the case of the structural estimation the R^2 is close to 1 and the *CRDW* test suggests that the null hypothesis of no-cointegration is again rejected.

We have further checked for existence of cointegration among the auto-selected variables, using the Engel and Granger cointegration test, the same way it was used in the structural modelling (Table 1.3).

Table 1.6. Equation (1.4) series cointegration test

	Statistic name	Value	P-Value
Engel & Granger	Tau Statistic	-4.7603	0.0388
	Z-statistic	-34.7254	0.0344

Having the test the null hypothesis on non-existence of cointegration, it can be rejected at 5% of level of confidence. Having this in mind, we preceded to the specification of the short term dynamics.

Short Run Equilibrium

The short run dynamics were represented by the OLS estimation of the first difference of the logarithm of house price. The model was specified using again a GETS approach, estimating an over-parameterized model with the cointegrating vector lagged one period and four lags of the first difference of each of the exogenous variables selected by the GASIC algorithm. After eliminating non-significant variables the empirical results were the following:

$$\Delta\log HPM2G = G - \underset{-1.3179G}{0.0962G} * \{4.2983 + \log HPM2\ -1G$$

$$-0.6393 * \log G \quad -1G - 0.1604 * \log G \quad -1G$$

$$+ 0.3320 * \log \mathrm{MORTRATE}\ -1G - 0.6932 \log G \quad P2008\ -1G\}$$

$$+ \underset{8.1700G}{0.6835G} * \Delta\log HPM2\ -4G + G\underset{2.2621G}{0.1663G} * \Delta\log HPM2\ -3G$$

$$+G\ \underset{-2.9966G}{0.0490G} * \Delta G \quad -4G - G\ \underset{-2.8465G}{0.0510G} * \Delta G \quad MG \quad -2G$$

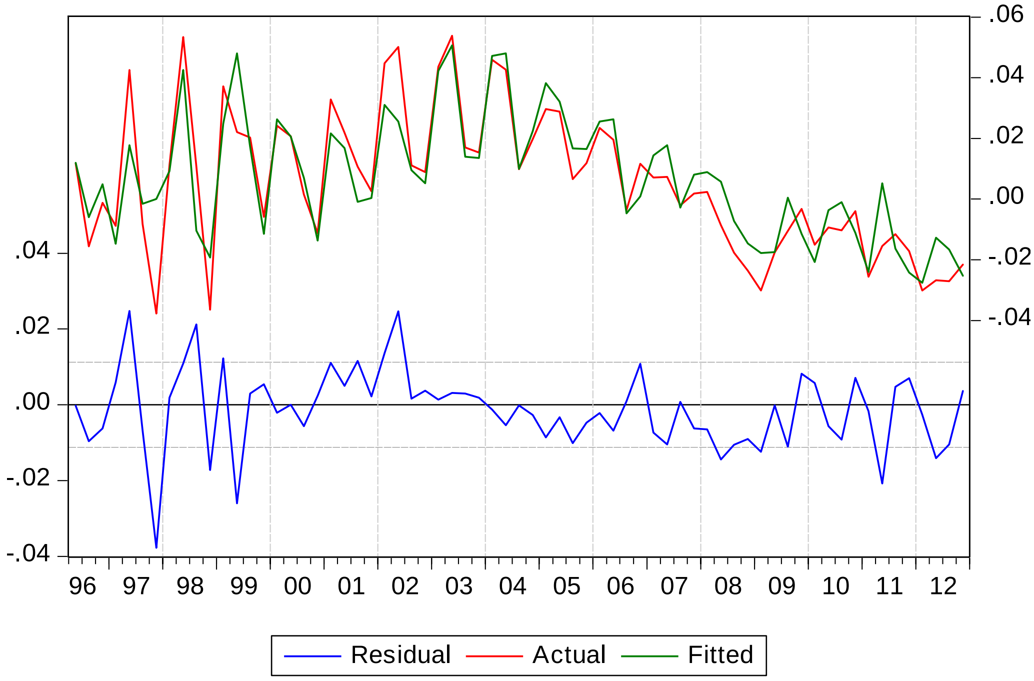
$$-G\ \underset{-2.2761G}{0.8004G} * \Delta G \quad P2008G - 3G + G\ \underset{-2.6769G}{0.9159G} * \Delta G \quad P2008G - 1G$$

(Eq. 1.7)

$$R^2 = 0.7872, \mathrm{DW} = 1.9861, G = 1.7741, GMG = 1.1506, G \quad HG = 0.5938G$$

where *t*-ratios are presented in brackets below of the estimators. The regressors hold a high explanation power as the R^2 figure is close to 0.8 and the DW statistic is closed to 2. Regarding the individual coefficients estimated, the first and most striking result is that for short run model the cointegrating term is non-significant[13], though holding a negative sing and being between 0 and 1, as expected from error correction mechanism. The remaining variables participating in the short run estimation were the third and fourth lags of the difference of the house price, the fourth lag of the difference of concrete consumption, the second lag of the difference of the mortgage rate and the lags one and three of the difference of GDP. Figure 1.3 depicts a graphical representation of the results of the short term estimation.

Figure 1.3. Actual, fitted, and residuals from short run equation (1.7) - GETS



Similar to results of the structural approach, residuals are inside confidence levels. At the same time the estimated values ('Fitted' in the chart) for the difference of house price seem to follow quite tightly the actual values. Goodness of fitness will be tested in the next chapter in a comparative framework against the structural modelling technique.

1.7. Forecasting performance of the structural and auto-selected models

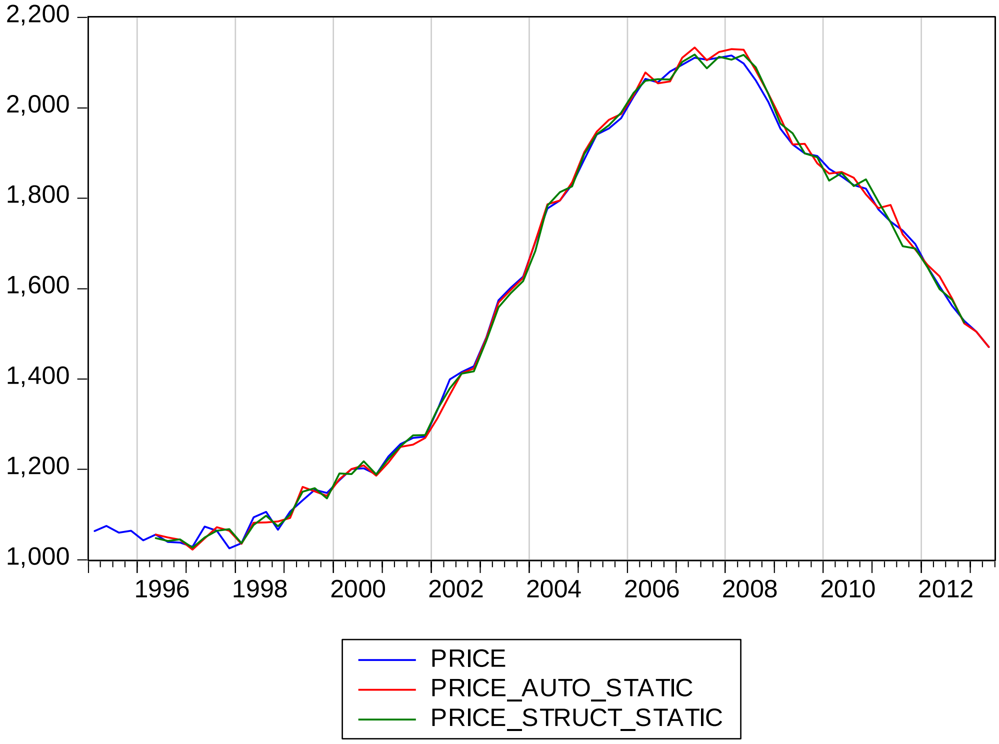
Figures 1.4 and 1.5 show the results of the inner-sample forecasts, using both a static and a dynamic approach[14] for the structural and automatic-select models.

[13] We have put the non-significant error correction term for illustrative proposes. The model without this term can be obtained from the authors upon request. It may be seen that results are quite similar and that the remaining regressors keep being significant, in the same way they appear in the expression with the error correction term.

[14] The static approach consists in making a one-step-ahead forecast only with historical data whilst the dynamic approach integrates the last forecast obtained as the last observation with which the next one-step-ahead forecast will be done. The latter approach is of special worth to test the forecast capability of a particular model, as it actually uses estimated results as arguments for the next forecasts.

Figure 1.4. Actual house prices per square meter vs. static forecasts

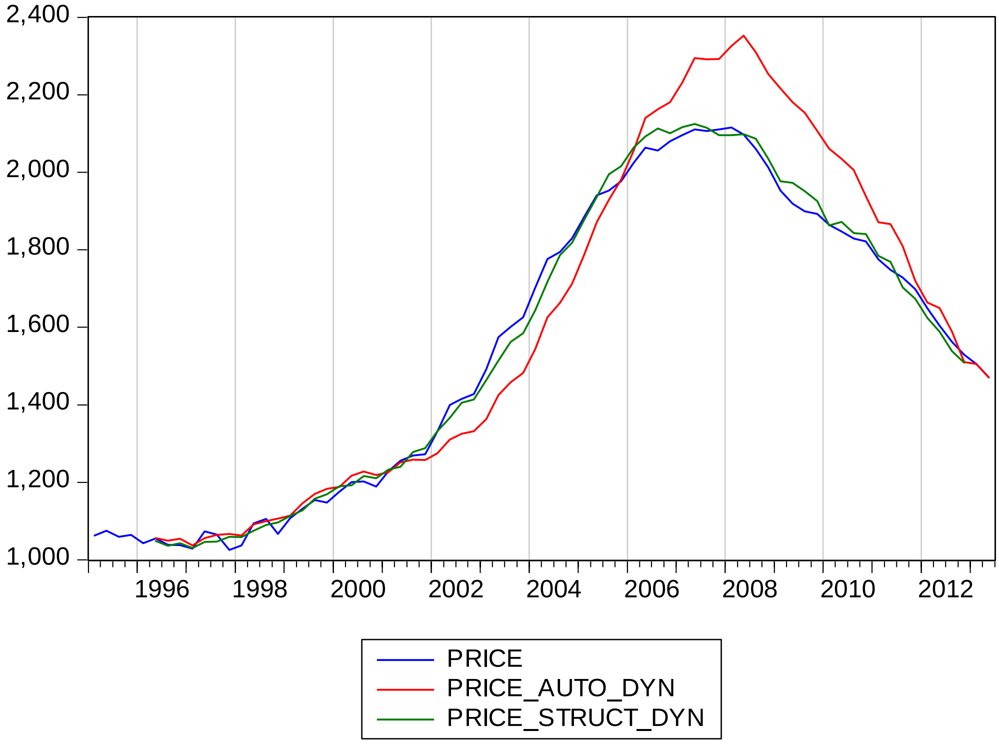
Simulation derived from the structural and automatic modelling, (€/m²)



When it comes to make one-step-ahead forecasts using only the historical data, the two models perform analogously (Figure 1.4). Checking in detail, the results from the auto selected model (red line) are more biased with respect to the actual value of the residential prices (blue line) in the peak of the boom in 2007, as well as more volatile than the forecast of the structural model (green line). The same happens in the period 2010-2011.

Figure 1.5. Actual house prices per square meter vs. dynamic forecasts

Simulation derived from the structural and automatic modelling, (€/m²)



As for Figure 1.5, the disparity of the dynamic forecasting capabilities of the two models becomes evident. Lacking an error correction mechanism, the automatic selected model (red line) strays from actual value of the house price. On the contrary, the forecast of the structural modelling follows quite fine the actual trend and is able to properly capture the turning point in bust of the bubble.

The dynamic forecasting properties of the two models have also analytically been assessed via some statics (Table 1.7). As can be seen, the performance of our structural model outpaces the forecasting capabilities of the auto-selected model. In particular, both the Root Mean Squared Error and the Mean Absolute Error are smaller for the structural model. The Mean Absolute Percent Error is more than three times greater for the case of the auto selected model. The Theil Inequality Coefficient - which has cero value when there's perfect fit - is closer to cero for the case of the structural model. Regarding the last three statistics, i) the Bias Proportion confirms that the forecast average of the structural model is closer to the actual average than in the case of the auto-modeling; ii) the Variance Proportion tells us that the forecast variation of the structural modelling is closer to the actual variation than in the case of the auto- modeling and iii) The Covariance Proportion shows us that a great proportion (94%) of the deviations of the forecasts comes from unsystematic forecasting error, in the case of the structural approach, while just a 61% of the forecast error comes from unsystematic factors.

Table 1.7. Dynamic forecast evaluation of the two modelling techniques

	Automatic	Structural
Root Mean Squared Error	121.50	26.40
Mean Absolute Error	90.42	21.50
Mean Abs. Percent Error	5.09	1.36
Theil Inequality Coefficient	0.04	0.01
Bias Proportion	0.10	0.00
Variance Proportion	0.28	0.06
Covariance Proportion	0.61	0.94

Forecast sample: 1995Q1 2012Q4
Adjusted sample: 1996Q2 2012Q4
Included observations: 67

1.8. Concluding remarks

The recent so-called property prices boom in Spain has conveyed special interest to real estate research as, apparently, observed price levels significantly got away from their equilibrium level. Besides, the boost of the house price bubble has driven Spanish economy to a long lasting economic crisis with deep implications to capital and labour markets as well as stern reductions in family income and welfare. All this has claimed for new insights on the dynamics of housing markets. This research tries to shed light on this issue with two state-of-the-art approaches that, at the same time, are compared in their forecasting performance.

In this paper, we have developed a structural modelling: A well fitted model for residential prices forecasting in Spain based on a cointegrating and error correction mechanism framework and an eclectic theoretical approach to select the fundamental variables that govern house price dynamics. This modelling reveals that the average residential price closely follows its long trend path. As a matter of fact, the short term price does not drift more than 6% away from the estimated equilibrium level. This suggests that economic fundamental variables actually supported such levels and that (irrational) speculative drivers were not as predominant as supposed to be. Our structural model captures interesting inflexion points in its estimation of the long-term equilibrium price path where the short term price level actually diverges, generating time spans of housing overvaluation and undervaluation. In particular, overvaluation periods (e.g. 2003-2004 and 2006-2008) follow spans of coincidence between short term level and long term path but, at the same time, increases in income. Conversely, ending 2008, landlords overreacted and average prices were an average of 2% below their equilibrium level. With the double dip of the Spanish economy fundamentals plummeted, and caused a new overvaluation period that finally was corrected ending 2012, where our last observation of housing price coincides with its long-term peer.

Our structural price model includes measures of opportunity costs, demand, supply and housing value-added drivers, all factually cointegrated. Added to the fact that all the proxies used resulted quite significant to determine residential prices, some had more prominent impacts on price than others. In particular, the income variable (proxied by GDP per Capita) happened to have the greatest impact on prices, with an elasticity of 1.5%, meaning that a 1% change in income per-capita increases residential equilibrium prices by a, more than proportional, 1.5%. As expected, increases in new building stock, and opportunity costs reduce the long term level of prices, but in a less than proportional manner. Valued added of the housing stock (proxied by Gross capital formation in real estate) increases the long term price of houses.

The so called house price boom-bust period actually claimed recognition in our research. The proxies for price, income and value added resulted to be first difference stationary with a structural break. We proceeded to estimate our long term price path including a proxy with value 0 until 2004 Q1 and 1 thereafter. The impact on price resulted positive and improved the explanation power of the model (adjusted R^2).

The estimation of house price dynamics suitably included the error correction mechanism derived from our long term estimation, with negative sing, and therefore:

- Confirms that prices have a ‘natural’ market driver that corrects their level to the equilibrium level and
- Reinforces the assumption of cointegration among the variables used in our model (Kremers et al. 1992)

The modelled short-run price variation actually follows quite well the observed price dynamics (proxied by the first difference of the observed price level), and the tests of goodness of fitness actually yield acceptable results.

Another important target of this study has been to test the usefulness of recently developed automatic modelling techniques for the real estate research. To address this target we needed two major constituents: An automatic model selection algorithm and a large database with several candidate regressors. The first one was the GASIC algorithm (Acosta-Gonzalez and Fernandez-Rodriguez, 2007), an automatic technique that auto selects models based on the Schwartz Information Criterion of nested competing models derived from a General Unrestricted Model (GUM). The second one was a variable set of real estate related variables that we created from a deep revision of the real estate research literature and several recognised statistics sources. This set actually acted as our GUM and comprised 46 candidate regressors. The resulting selected price model by GASIC was a parsimonious one including demand, opportunity costs and supply variables. The estimated cointegrating regression yielded reasonable estimators, in terms of expected impact size and sign. Unfortunately it was non-significant in the short term dynamics estimates, hindering the forecasting capabilities of the model.

The comparison of the forecasting performance of the structural and auto-selected models clearly indicated that the former overtook the latter in a (realistic) dynamic framework where the forecasts for subsequent periods are computed using information available at the start of the forecast sample.

As further steps for this research we propose the usage of other modern auto modelling selection such as the LSE approach or General to Specific with OxMetrics® software. We also suggest that the auto selection technique may be easily improved by assisting the algorithm with ‘guided’ process, in which the researcher has previously grouped variables by similarities in nature, for example, by theoretical proximity.

References

Abraham J. M., Hendershott P. H. Bubbles in metropolitan housing markets. Journal of Housing Research 1996;7; 191-207.

Acosta-Gonzalez E, Fernandez-Rodriguez F. Model Selection via genetic algorithms illustrated with Cross-country growth data. Empirical Economics 2007;33; 313-337.

Banerjee A, Dolado J, Hendry D F, Smith G. Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence. Oxford Bulletin of Economics and Statistics 1986;48; 253–277

Berg, T J, Jordà O. A chronology of turning points in economic activity, Spain, 1850–2011. SERIEs 2013;4; 1–34.

Campos J, Hendry D F, Krolzig H. Consistent model selection by an automatic GETS approach. Oxford Bulletin of Economics and Statistics 203;S1; 803-819.

Capozza D R, Hendershott P H, Mack C, Mayer C J. Determinants of real house price dynamics. Working Paper 9262, National Bureau of Economic Research, Cambridge, MA; 2002.

Carrión-i-Silvestre J L, Sansó-i-Roselló A, Ortuño M A. Unit root and stationarity tests wedding. Economics Letters 2012;70; 1–8.

Coleman M, LaCour-Little M, Vandell K D. Subprime lending and the housing bubble: Tail wags dog? Journal of Housing Economics 208:17; 272-290.

Dickey D A, Fuller W A. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 1981;74; 427-431.

Economic Cycle Research Institute. Business cycle peak and trough dates, 22 countries, 1948-2013. Economic Cycle Research Institute, London; 2014.

Engle R, Granger C W. Cointegration and errror correction: representation, estimation and testing. Econometrica 1987;55; 251-276.

Gattini L, Hiebert P. Forecasting and assesing Euro Area house prices through the lens of key fundamentals. Working Paper 1249, European Central Bank, Frankfurt am Main; 2010.

Hansen B E. Discussion of ‘Data mining reconsidered’. Econometrics Journal 1999;2; 192-201.

Hendry D F. Dynamic econometrics, Oxford University Press, Oxford; 1995.

Hendry D F, Krolzig H-M. Improving on ‘Data mining reconsidered’ by K.D. Hoover and S.J. Perez. Econometrics Journal 1999; 2; 202–219.

Hendry D F. Doornik J A. Empirical model discovery and theory evaluation. The MIT Press, Cambridge, MA: 2014.

Hilbers P, Hoffmaiste AW, Banerji A, Haiyan S. House price developments in Europe: A comparison. Working Paper 08/211, International Monetary Fund, Washington, DC; 2008.

Hofmann B. Bank lending and property prices: Some international evidence. Working Paper 22/2, Hong Kong Institute for Monetary Research, Hong Kong; 2004.

Holland J. Adaptation in natural and artificial systems. The MIT Press, Cambridge MA; 1975.

Hoover K D, Perez S J. Data mining reconsidered: Encompassing and the general-to-specific approach to specification search. Econometrics Journal 1999;2; 167-191.

Iacoviello M. House prices and business cycles in Europe: A VAR analysis. Working Paper in Economics 540, Boston College, Chestnut Hill, MA; 2007.

Iacoviello M, Minetti R. The credit channel of monetary policy: Evidence from the housing market. Journal of Macroeconomics 2008;30; 69-96.

Kremers J, Ericsson N, Dolado J. The power of cointegration tests. Oxford Bulletin of Economics and Statistics 1992;54; 325–348.

Kwiatkowski D, Phillips P C B, Schmidt P, Shin Y. Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics 1992;54; 159–178.

Lovell M. Data mining. Review of Economics and Statistics 1983;65; 1-12.

Malpezzi S. A simple error correction model of house prices. Journal of Housing Economics 1999;8; 27–62.

Meen G. The time-series behavior of house prices: A transatlantic divide? Journal of Housing Economics 2002; 11; 1–23.

Olsen E O. The demand and supply of housing service: A critical survey of the empirical literature. In: Mills E S (Ed), Handbook of Regional and Urban Economics, vol. 2, North-Holland, Amsterdam; 1987. p. 989-1022.

Perron P. The great crash, the oil price shock, and the unit root hypothesis. Econometrica 1989;57; 1361-1401.

Perron P. Further evidence on breaking trend functions in macroeconomics. Journal of Econometrics 1997;80; 355-385.

Phillips P C, Hansen B E. Statistical inference in instrumental variables regressions with I(1) processes. Review of Economic Studies 1990;57; 99-125.

Tsatsaronis K, Zhu H. What drives housing price dynamics: Cross-country evidence. BIS Quarterly Review 2004;March; 65-78.

Whitehead C M E Urban housing markets: Theory and policy. In: Cheshire P, Mills E S (Eds.), Handbook of Regional and Urban Economics, vol. 3, North-Holland, Amsterdam; 1998. p. 1559-1594.

Zivot E, Andrews D W. Further evidence on the great crash, the oil price shock and unit root hypothesis. Journal of Business and Economics Statistics 1992;10; 251-270.

OFFICE MARKET DYNAMICS IN MADRID: MODELLING WITH A SINGLE EQUATION ERROR CORRECTION MECHANISM[15][16]

[15] This work is to be presented in the main sessions of the Asian Real Estate Society Congress 2016 in Bangalore, India in July 2016. As of April 2016 submission has been approved.

[16] The results of this research have been submitted to the International Real Estate Review (IRER) in February 2016. As of April 2016 it is under review of for publication. The IRER is a double-blind refereed academic journal with an RePEC impact factor of 0.94.

2.1. Introduction

The study of commercial property markets (retail shops, warehouses and offices) has gained momentum in the economic literature since the 1980's, leveraging on previous work analysing the economics of residential real estate which began in the 60s and 70s the U.S. Research on the effects of economic cycles in the residential construction of Alberts (1962) and the price estimation of housing by Blanck & Winnick (1953), Pritchett (1977) and Ferri (1977) were the seeds of economic analysis of non-residential markets. In the last 20 years, certain conditions have been met prompting the investigation into the non-residential property markets (Ball, Lizieri, & MacGregor, 1998):

- The global economic boom of the late 80's and early 90's and its impact on development of offices, high street shops and shopping centres, and industrial warehouses and logistics
- The development and diffusion of new statistical analysis tools, including cointegration and error correction models
- Greater availability of longer time series of supply, prices and demand of property markets

In this context, the seminal works on cycles in office markets were born in the United Kingdom and the United States by Rosen (1984) and Wheaton (1987) analysing mechanisms of adjustment of real estate variables (rent, availability, absorption of space and construction) and their long and short run relationships with macroeconomic variables. Under the light of these works a substantial amount of literature has been developed, extending the analysis to other European markets since the late 90's.

Published research for the Spanish commercial property market is not abundant. It can be mentioned the work of Fuerst and McAllister (2008) and Brounen and Jennen (2009), that seek to explain the rents dynamics in different European cities (10 and 19 cities, including Madrid, respectively). Brounen and Jennen use an error correction model on maximum rents and Fuerst uses linear regression models to analyse the elasticity of supply.

The objective of this paper is, by using time series analysis (cointegration and error correction models), to describe the dynamics of office vacant space, deliveries of new office stock (office stock variation) and average rents in terms of elasticities as well as responses to long-term equilibrium deviations of rents and vacancy. We contribute proposing models capable of predicting future market developments, identify phases in which rents have been appreciated or depreciated against the long-run equilibrium and quantifying the possible overvaluation or undervaluation of the cyclic type of property. We also measure forecasting performance of the two stage error correction mechanism and the single equation error correction mechanism models and propose the best modelling system to analyse rents, vacancy rate and stock change.

In our study we have adapted the model developed by Hendershott et al. (2013, HJM hereafter) to the Madrid office market. Two models are estimated and compared: The error correction mechanism framework (Engle and Granger 1987); The single equation error correction mechanism (Banerjee et al., 1993). Our analysis yields that the best fitted model to do dynamic forecasting us the Engle and Granger approach.

After this introduction, the second section depicts a commercial property market model, the third section details the econometric models employed. The fourth section describes the data used and the fifth and sixth presents the econometric approach and the results of the estimated models. The seventh section compares the results of the two estimation methodologies and, finally, we present some concluding remarks.

2.2. Economics of the office property market

Non-residential real estate markets are composed of the interaction of four sub-markets (Ball et al., 1998):

- Final users, through which employers choose the locations where they develop their productive activity. They let space to owners of available office stock. In turn, these owners have acquired these properties recurring to the:
- Investment market, whereby institutional or private investors (or even occupants) acquire real estate assets based on their expected performance relative to other assets and their risk profile (opportunity cost). They may have bought their properties recurring to the second hand market or to the:
- Development market, through which new buildings are added to the existing stock. New office stock is activated when businesses require additional space, in a market with an inelastic short term supply. As a matter of fact, construction time may take years, explaining the inelasticity of the supply. The land on which new buildings are constructed is acquired in the:
- Development sites market, corresponding to the (limited) locations on which the new stock will be developed. The type of building to develop depends on the opportunity costs of alternative uses that may be chosen. Consequently, every possible activity (residential, commercial, industrial, offices, etc.) is competing with the others, thereby determining cost of the land.

This work aims to analyse the final user market, where new letting contracts reflect market's relative scarcity of office stock to the current demand and, therefore, give birth to letting rents. We procced now to describe the operation of this office rental market, giving support to our econometric specification and analysis.

Demand for offices is mainly derived from the need to use space as production input mainly of non-industrial economic activities, needing a specific location for that labour. Among the main activities demanding office space we can mention:

• Business services sector

• Financial, insurance and real estate

• Support for industrial production (management, human resources, etc.)

• Public Administration

The labour absorbed by these activities corresponds mainly to the service sector activities and may be housed in office buildings (Wheaton, 1987). Consequently, the occupied stock (and letting rentals) depends deeply on the cycle of the service sector employment.

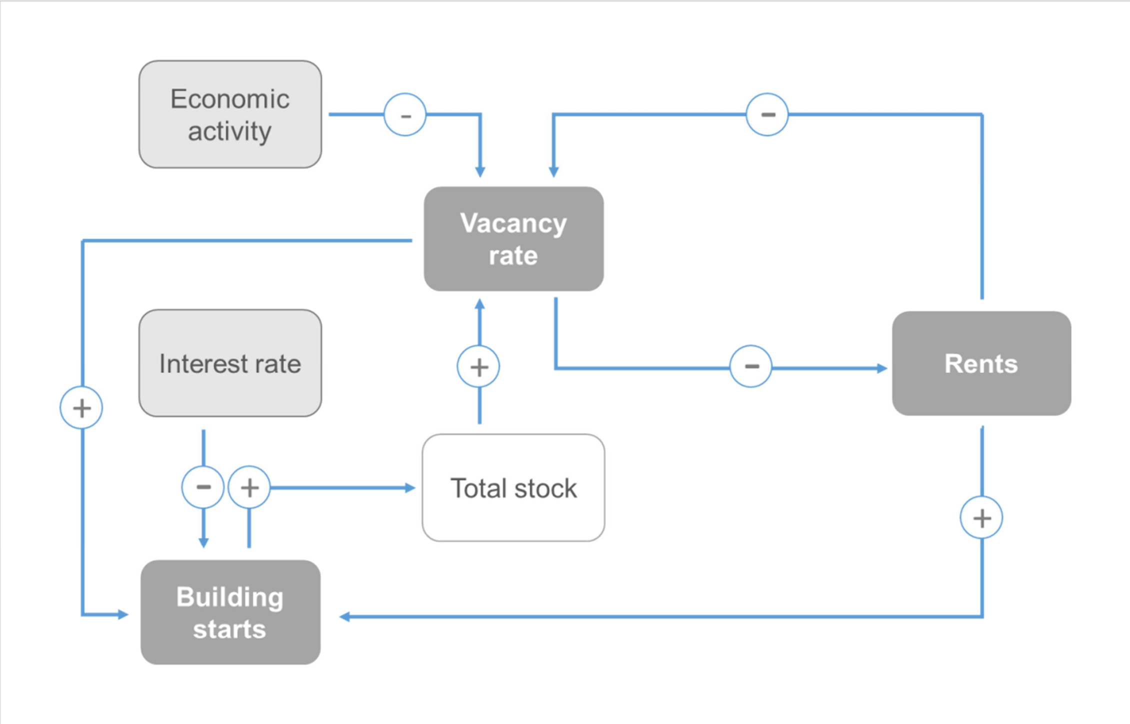
Businesses demand office space from landlords aiming to obtain the maximum return on their investment. According to BNP Paribas Real Estate[17] the general market practice is that 80% of transactions to offices leases, 5% pre-lets and the rest for sale. It is therefore a reasonable assumption in most empirical studies (including this one) that the owners are limited to rent space (never sell) and end users to exclusively let spaces (never purchase). This will facilitate the analysis and focus on the dynamics of rents, side-stepping selling prices, as they are balanced in the investment market.

Office stock is the market supply and has the characteristics of a capital asset subject to depreciation (destruction and change of use) as well as accumulation through new construction and refurbishments. There will be new stock added when property prices charged by developers exceeds construction costs (interest rate, land, construction, materials, etc.). In other words, once the shortage in the stock is transferred to rental increases in the user market, and finally to the selling market, developers will begin construction of new buildings to benefit of the higher prices of the property. Developments cease at the moment in which the stock available caters all demand, causing prices of the property to fall back to the level of replacement costs. In this sense, the office promotion market can be considered as an imbalance phenomenon (Ball et al., 1998). Once such imbalance is observed in the user market, new stock is added in the next period, thus forming a real estate cycle. Figure 1 presents a conceptual framework that helps to explain the key relationships of an office market, that may be employed to any other empirical modelling of a non-residential market (Brooks and Tsolacos, 2010).

[17] BNP Paribas Real Estate Spain, 2011. Madrid and Barcelona office market, second quarter.

Figure 2.1. An analytical model for the property market

Light grey background variables are exogenous; Dark grey background variables are endogenous and white background is a variable that may be determined by calculation.



The direction of the arrows indicates whether a variable affects or is affected by other(s). Only two variables of the scheme are not affected but only affect: The level of economic activity and the interest rate. Therefore, those will be considered as exogenous to the model, specifying its nature of partial equilibrium. The sign accompanying the arrow corresponds to the effect of a positive change in the origin variable on the target variable. As an example an increase in level of economic activity will decrease the vacancy rate. The endogenous variables, therefore, are vacancy rate, building starts and rent levels. In the following sections we specify the equations derived of this scheme.

Developers will construct new buildings according the balance of the asset price and their replacement costs. That is, office supply responds positively to higher property prices and negatively to the production costs and financing, which in this work are assumed exogenous. Meanwhile, property prices are higher the scarcer the available stock is (once exhausted the reduction of space per employee), that is, the lower the vacancy rate, which is the ratio between the total available floor area and stock, the higher the rental values. In turn, this shortage is greater in periods of increased economic activity. In summary, the office market depends positively on the real business cycle and employment. The high correlation between activity variables (production, economic sentiment, etc.) and employment, as well as the correlation between national and local employment allow for obtaining similar adjustments in the commercial real estate models. According to Brounen and Jennen (2009) no significant differences are obtained. Nevertheless, we have tested our models both for national activity variables and local activity variables. I.e. we modelled the Madrid's office market both with Spain's GDP and Madrid's service sector employment level[18]. Both give similar results, confirming Brounen et al. (2009) findings.

2.3. Modelling

Following Englund et al. (2008, hereafter EGHS) and Hendershott et al. (2013, hereafter HJM) we use a cointegration approach which employs a single long term equation between rents, economic activity and stock as ECM in the three expressions of the adjustments of rents, vacancy and stock. Therefore, our approach specifies the short run dynamics as a system of three equations to be solved simultaneously.

Businesses' office demand is a function of their activity level and the new contract's rent level

$$_t = \gamma_0 \; {}^{\gamma_{1G}}_{tG} \; {}^{\gamma_2}_{tG} \qquad \text{(Eq. 2.1)}$$

Where $\gamma_{1G}$and $\gamma G$are the (negative) price and (positive) income elasticities for the logarithmic expression of (1). The equilibrium rent is reached when vacancy rate is at its long term (constant) level and demand is equal to the total supply (S_t) minus the natural vacancy level

$$_{tG} \;\; _t,G_t \;\; = \;\; 1 - V^{*G} \;\; _{tG} \qquad \text{(Eq. 2.2)}$$

Equating (Eq. 2.1) and (Eq. 2.2) we obtain:

$$_t = \gamma_0'^{G} \; {}^{\gamma_1}_{tG} \;\; 1 - V^{*G} \;\; {}_{tG}{}^{\gamma_{2G}} \qquad \text{(Eq. 2.3)}$$

Which corresponds to our expression of long run rent that in logs may be expressed as:

$$\ln G_t = \ln\gamma_0' + \gamma_1 \ln G_t + \gamma \;\ln\; 1 - V^{*G} + \gamma \;\ln G_{yG} \qquad \text{(Eq. 2.4)}$$

(Eq. 2.4) may be re-expressed taking into account that $\ln G V^{*G} = v^{*G}$is a constant value

$$\ln G_t = \gamma_0 + \gamma_1 \; nG_t + \gamma \;\ln G_{yG} \qquad \text{(Eq. 2.5)}$$

Where $\gamma_{0G}= \ln\gamma_0'^{G}+ \gamma G\ln G 1 - v^{*G}$. Note that because $\ln\gamma_0'^{G}$is unknown, the natural vacancy rate may not be found in this expression (HJM). Nevertheless, we may derive such value from the short run expressions.

The short run expressions for our modelling are standard for the dynamics under ECM:

$$\Delta\ln G_t = \alpha_0 + \sum_{i=0G}^{n_1} \alpha_{1,i} \Delta\ln G_{t-iG} + \sum_{i=0G}^{n_2} \alpha_{\;,i} \Delta\ln G_{t-iG} + \sum_{i=0G}^{n_3} \alpha_{3,i} \Delta\ln G_{t-iG} + \sum_{i=0G}^{n_4} \alpha_{4,i} \mathrm{v}_{t-i-1G} + \sum_{i=0G}^{n_5} \alpha_{5,i} \varepsilon_{t-i-1G} \qquad \text{(Eq. 2.6)}$$

In (Eq. 2.6) the adjustment term for the vacancy rate do not have the long term level for vacancy rate as it is constant and is embedded in the constant term. Actually departing from such constant term, we can estimate to the long term (or natural) level of the vacancy rate knowing that $\alpha_0 = -v^* \sum_{i=0G}^{n_4} \alpha_{4,iG}$ therefore: $v^* = -\alpha_0 / \sum_{i=0G}^{n_4} \alpha_{4,iG}$

Taking (Eq. 2.6) as reference we can specify the short run dynamics for the vacancy rate

[18] Although a clear definition for office employment exists, no such statistical series is found for the period and frequency used in this work (2001 Q1 – 2015 Q2).

$$\Delta v_t = \beta_0 + \sum_{i=0}^{m_1} \beta_{1,i} \Delta v_{t-i} + \sum_{i=0}^{m_2} \beta_{,i} \Delta \ln G_{t-i} + \sum_{i=0}^{m_3} \beta_{3,i} \Delta \ln G_{t-i} + \sum_{i=0}^{m_4} \beta_{4,i} v_{t-i-1} + \sum_{i=0}^{m_5} \beta_{5,i} \varepsilon_{t-i-1} \quad \text{(Eq. 2.7)}$$

From (7) it is also possible to estimate the natural value of the vacancy rate with $\beta_0 = -v^* \sum_{i=0}^{n_4} \beta_{4,i}$ so $v^* = -\beta_0 / \sum_{i=0}^{n_4} \beta_{4,i}$

The short run adjustment of the stock level is estimated by means of the gap existent between the natural vacancy rate and the actual vacancy rate. The rational of this, comes from idea that the higher the gap the higher the rent. At the same time, HJM assert that the present value of future rents is the value of new stock investment, or change in office stock which is actually our third short term equation. This is a useful specification for our work as we lack series of new deliveries and stock destruction or depreciation. The adjustment of stock is therefore as follows:

$$\Delta G_t = \delta_0 + \sum_{i=0}^{l_1} \delta_{1,i} \Delta G_{t-i} + \sum_{i=0}^{l_2} \delta_{,i} v_{t-i-1} + \sum_{i=0}^{l_3} \delta_{3,i} \varepsilon_{t-i-1} \quad \text{(Eq. 2.8)}$$

Where again $-\delta_0 / \sum_{i=0}^{l_2} \delta G_{,i}$ is an estimation of the long run vacancy rate.

For equations 2.6 to 2.8 the expected sings for the ECM estimated coefficient is negative; it is expected that variables return to equilibrium when rents and vacancy are above long term value.

2.4. Database and variables description

Office market variable's database for this work was provided by BNP Paribas Real Estate and contains quarterly observations from 2001Q1 to 2015Q2. Exogenous economic activity variables are available in the Spanish National Statistics Office (INE) web site. The geographical scope corresponds to the offices within metropolitan Madrid area, plus municipalities of Las Rozas de Madrid, Pozuelo de Alarcón, Alcobendas and San Sebastian de los Reyes. The database conveniently comprises two cycles for the Spanish economy: the aftermath of the dot-com crisis, the Great Crisis 2007-2013 and the most recent recovery phase (2014-2015). As presented in the modelling section, the system integrates one economic activity variable. There is certain flexibility when choosing the economic drive for the model due to the high correlation between activity variables (production, economic sentiment, etc.) and employment, as well as the correlation between national and local employment. This allows for obtaining similar adjustments in the commercial real estate models. According to Brounen and Jennen (2009) no significant differences are obtained. Using this framework, we have estimated two sets of models: one using Spain's GDP as economic activity variable and other using Madrid's service sector employment, to test the best fitted model and also obtain information on the exposure of Madrid's business environment (office market) to national macroeconomic indicators (Spanish GDP). Table 2.1 presents key statistics of the variables used in this work.

Table 2.1. Main variables used in the empirical analysis

	Real rent (RENT)	*Vacancy rate* (VACR)	*Office Stock* (STOCK)	*Spanish GDP* (GDP)	*Service sector employment* (SEMP)	*Occupied Space* (OS)	*Vacant space* (VAC)
Unit of measure	€/m ^{2}/month	%	m ^{2}	Index 2010=100	000 persons	m ^{2}	m ^{2}
Mean	18.3	10.4%	10,845,798	95.8	2,259.4	9,688,903	1,156,896
Median	18.0	9.5%	11,163,405	97.5	2,343.5	9,998,857	993,293
Max	29.7 (2001Q2)	16.3% (2015 Q1)	11,885,563 (2013 Q1)	104.4 (2008 Q2)	2,515.0 (2008 Q4)	10,332,478 (2008 Q1)	1,933,485 (2015 Q1)

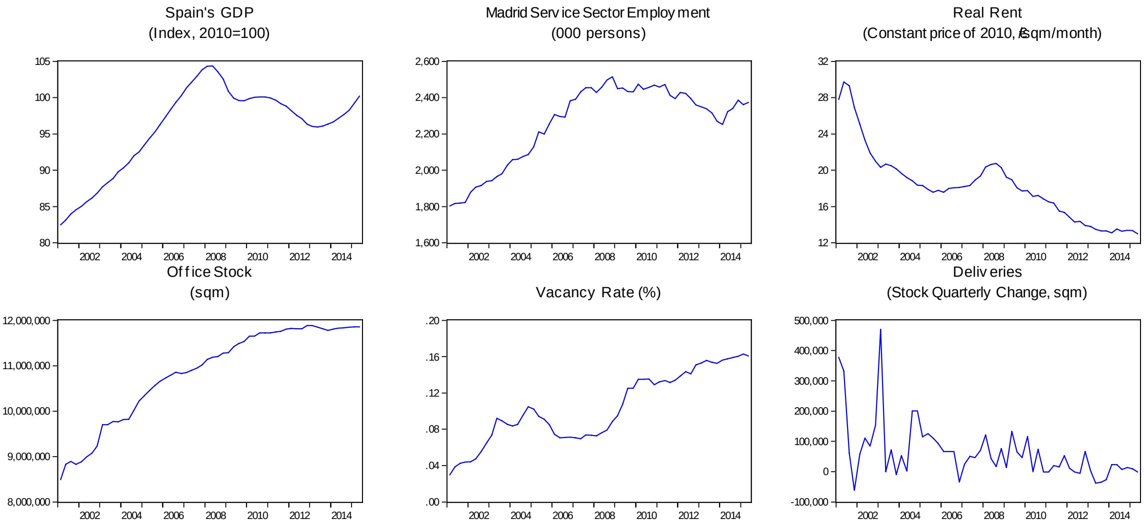
Min	13.0 (2013 Q2)	3.0% (2001 Q1)	8,493,109 (2001 Q1)	82.5 (2001 Q1)	1,802.0 (2001 Q1)	8,240,115 (2001 Q1)	252,994 (2001 Q1)
Std. Deviation	3.9	3.9%	1,035,969	6.0	216.7	636,415	504,466
Observations	58	58	58	58	58	58	58

Note: (Names in parenthesis are those used in the econometric specification). Real rent has been deflated with Spanish GDP deflator at constant price of 2010. From the left, the first three variables comprise our endogenous variables, GDP and employment comprise, separately, the exogenous ones. The last two are calculated variables extracted from vacancy rate and office stock.

Real rent in table 1 corresponds to Madrid's quarterly new letting contracts average headline rent. It is measured in €/m²/month and is expressed in real terms at 2010 prices, using GDP deflator. Values in parentheses show the periods where extreme observations are obtained. Maximum values are to be seen in 2008 for GDP, service sector employment and occupied space, reflecting the highest point of expansion of Spanish and Madrid's economy and real estate markets. After the explosion of the Bubble, economic activity dropped, causing reduction in rents as well as in occupancy. It is in Q1 2015 when rents get to their minimum point and vacancy rate and vacant space to their maximum. Figure 2.2 gives a clearer picture of the recent property cycle in Madrid.

Figure 2.2. Trends in the main variables used to model Madrid office market

Time series span: 2001Q1 – 2015Q2



The maximum levels of service sector employment and GDP are observed in the second half of 2008, coinciding with the maximum historical levels in occupied space and a local minimum (after 2005Q1) in vacancy rate. From that moment onwards, occupancy started to fall and vacancy rate increased in a swiftly fashion. Just before the last crisis hit Spanish and Madrid's economies, deliveries were constantly increasing the stock at an average pace of nearly 60.000 sqm per quarter but demand activity managed to generate positive net absorption and decreases in vacancy rate (7%, 2007Q2). After 2008Q2, with the economy shrinking, new contract's rents started a continuous descent until 2015. With low expectations on returns, developers hastily halted new building starts. Nevertheless, deliveries of new schemes did not stop as construction process lasts for at least 18 months, giving some momentum to variation in stock. In the period 2009-2010 such variation was nearly 55.000 sqm per quarter (construction inertia) while in the period 2011-2015 such variation was of 7.500 sqm per quarter. In figure 2 it is clear the shared trend among rents, vacancy rates (inversed) economic activity and stock variation. Such trend is indicating a likely common long-term growth which, in other words, signals the possible existence of cointegration among those series. The co-movements of the series have been traced through their correlations and collected in Table 2.2.

Table 2.2. Correlation analysis

Sample: 2001Q1 2015Q2

Included observations: 58

Correlation *p-value*	Rent	Vacancy rate	Office stock	Spanish GDP	Service sector employment	Occupied space	Vacant space	Variation in stock
Rent	1.0000 ----							
Vacancy rate	-0.8992 *0.0000*	1.0000 ----						
Office stock	-0.8624 *0.0000*	0.8535 *0.0000*	1.0000 ----					
Spanish GDP	-0.5942 *0.0000*	0.4707 *0.0002*	0.8413 *0.0000*	1.0000 ---				
Service sector employment	-0.6712 *0.0000*	0.5914 *0.0000*	0.9112 *0.0000*	0.9705 *0.0000*	1.0000 ----			
Occupied space	-0.6947 *0.0000*	0.599702 *0.0000*	0.9281 *0.0000*	0.9635 *0.0000*	0.9810 *0.0000*	1.0000 ----		
Vacant space	-0.8946 *0.0000*	0.9962 *0.0000*	0.8828 *0.0000*	0.5122 *0.0000*	0.63356 *0.0000*	0.6443 *0.0000*	1.0000 ----	
Variation in stock	0.4856 *0.0001*	-0.4492 *0.0004*	-0.4979 *0.0001*	-0.4221 *0.0010*	-0.4541 *0.0003*	-0.4458 *0.0005*	-0.4601 *0.0003*	1.0000 ----

Correlation of rent, vacancy rate and stock with the economic activity proxies (Spanish GDP and Madrid's service sector employment) is strong (exception made for vacancy rate and GDP), supporting their role as main drivers and to be confirmed with cointegration tests. It also is an indicator that the series are not stationary[19]. The correlation of -0.9 between average real rent and vacancy rate (p-value of zero) sets the strong interplay of the real estate variables. Although such correlation is high, it does not equal one due to the existence of rigidities in the space markets. These rigidities come mainly in the form of lease contracts (Torto et al., 1997 and HJM), making businesses to be off their optimal space demand when they receive activity shocks. Another part may be played by structural vacancy which is composed by office stock that does not have quality, location and access apt to compete within the market (Remøy, 2010).

New deliveries have no strong correlation with the selected variables. The high volatility of the series reduces their correlation with the other fundamentals.

2.5. Econometric Specification

In order to implement our cointegrating regression analysis we have tested stationarity for the variables participating in the ECM. Table 2.3 summarises the results.

Table 2.3. Tests of integration

Augmented Dickey-Fuller Test (null hypothesis: series has unit root)					
	lag (AIC)	Model	t-statistic	Critical value (5%)	Critical value (1%)
RENT	5	Constant	-1.2838	-2.9126	-3.5482
ΔRENT***	3	Constant	-3.8438	-2.9126	-3.5482
STOCK	6	Constant + Trend	-1.3657	-3.4892	-4.1242
ΔSTOCK***	5	Constant + Trend	-6.8113	-3.4892	-4.1242
GDP	9	Constant	-1.7763	-2.9126	-3.5482
ΔGDP	8	Constant	-1.5002	-2.9126	-3.5482
SEMP	0	Constant	-2.5791	-2.9126	-3.5482
ΔSEMP***	0	Constant	-6.6930	-2.9126	-3.5482

Perron test with structural break (null hypothesis: series has unit root with a structural break)						
	lag	Model	t-statistic	Critical value (5%)	Critical value (1%)	Date of structural break
GDP	4	Constant	-4.3343	-5.23	-5.92	NA
ΔGDP***	3	Constant	-6.1336	-5.23	-5.92	Q4 2007

*** denotes significance at 1% level of confidence. ADF gives strong evidence for first order of integration for rent, stock and Madrid's service sector employment. Evidence on first degree of stationarity for GDP is given by the Perron test, with a structural break in Q4 2007.

[19] It is because if series were stationary correlation should be around 50% which is the mere correlation given by 'the flip of a coin'.

All variables participating in the cointegrating equation have unit root. Nevertheless, ADF fails to reject the hypothesis of first degree of integration for GDP. The reason is that the last crisis linked several quarters of negative variations. Yet, we resorted to test stationary with structural break using Perron (1995) test. As expected, we reject the null hypothesis for the difference of GDP, so we may conclude that the level of GDP has a unit root when a structural break is accounted for in 2007Q4, period when Spanish crisis started. We may have opted to include such structural break in our modelling by means of a dummy variable, taking a value of zero before 2007Q4 and one since such date. Nevertheless, from theory we know the long term the relationship between local markets office rents and national GDP, especially for capital cities, such as the Madrid's case. Using this framework, we do not include such dummy and maintain a simpler modelling of long term equations.

Having stated the order of integration of the variables to participate in the cointegrating equation we tested for cointegration among them.

Using both the Johansen (1991) procedure and Engle and Granger (1987) single equation cointegration test we identified at least one cointegrating relationship i.e. one long term equilibrium relationship among our non-stationary variables RENT, STOCK, GDP or RENT, STOCK, SEMP (see table 2.4).

Table 2.4. Cointegration tests results

Johansen Cointegration test among Rent, GDP and Stock– P-values for the cointegration rank test
Cointegrating regression tested with a constant term and 1 to 4 lags interval

		Null hypothesis of:		
		No Cointegrating equations	One cointegrating Equation	Two cointegrating Equations
Cointegration test using	Trace	0.0000***	0.0789*	0.4713
	Maximum eigenvalue	0.0000***	0.649*	0.4713

Both the Trace and Maximum eigenvalue tests reject the existence of two cointegrating relationships at 5% of confidence level. This supports the existence of one cointegrating relationship.

Johansen Cointegration test among Rent, SEMP and Stock – P-values for the cointegration rank test
Cointegrating regression tested with a constant term and 1 to 4 lags interval

		Null hypothesis of:		
		No Cointegrating equations	One cointegrating Equation	Two cointegrating Equations
Cointegration test using	Trace	0.0000***	0.0664*	0.7882
	Maximum eigenvalue	0.0000***	0.0288**	0.7882

The trace test rejects the existence of a two cointegrating relationships at 5% of confidence level. The maximum eigenvalue test rejects the existence of three cointegrating relationships at 5% of confidence level. This supports the existence of one or two cointegrating relationships.

Engle-Granger Cointegration test among Rent, GDP and Stock
P-values for the cointegration test. Null hypothesis of no-cointegration
Cointegrating regression tested with a constant term and seven lags

		RENT	GDP	STOCK
Cointegration test using	Engle-Granger tau-statistic	0.7106	0.5898	0.5788
	Normalised autocorrelation coefficient	0.0014**	0.5569	0.0000***

Although the Engle-Granger tau statistic fails to reject the hypothesis of no cointegration, the normalised autocorrelation coefficient test signals some degree of cointegration among the series.

Engle-Granger Cointegration test among Rent, SEMP and Stock
P-values for the cointegration test. Null hypothesis of no-cointegration
Cointegrating regression tested with a constant term and one lag

		RENT	SEMP	STOCK
Cointegration test using	Engle-Granger tau-statistic	0.0509**	0.0492**	0.0301**
	Normalised autocorrelation coefficient	0.0900*	0.0018***	0.0022***

Both the Engle-Granger tau statistic and the normalised autocorrelation coefficient test reject the null hypothesis of non-existence of cointegration at a 5% of confidence level. ***Denotes significance at 1% of confidence level, ** denotes significance at 5% of confidence level and * denotes significance at 10% of confidence level. All variables tested in logs.

All the tests indicate the existence of la long term relationship between office rents, gross domestic product and office stock or between office rents, Madrid service sector employment and office stock at the traditional confidence levels. It is worth to mention that the Engle-Granger test for Rent, GDP and STOCK was the less indicative of existence of cointegration, whether using or not a dummy variable representing the sock of the crisis of 2007. Yet, the Johansen test for the same variables effectively supports the existence of cointegration.

2.6. Error correction models

Under the light of non-stationarity of the variables, we have chosen two methods for estimating error correction models. One is the classical Engle and Granger two step method and the other is the single equation error correction mechanism (SEECM, Banerjee, 1993). With these methods, the standard assumptions of the asymptotic analysis are valid in the presence of first-order non-stationary and cointegrated series. The inference on the estimated coefficients is possible because the t-statistic and f-distributions behave optimally. In this sense, a structural modelling in a multivariate system is performed using seemingly unrelated regressions (SUR) as residual terms may be correlated. The system of equations estimated correspond to equations (2.6) to (2.8).

a. Two-step methodology estimates

Recognizing a long term relationship in our variables, we estimated the long run equation for rents by fully modified least squares (FMLS) proposed by Phillips and Hansen (1990), as long as OLS estimates yield biased estimated coefficients. The results of estimating equation 2.5 are presented in table 2.5.

Table 2.5. Cointegrating equations

Long run models. Endogenous variable: Logarithm of Real Average Office Rent - LOG(RENT)					
	Coefficient	t-Statistic		Coefficient	t-Statistic
LOG(GDP)	2.3636	5.6574***	LOG(SEMP)	2.4657	7.0437***
LOG(STOCK)	-3.1597	-13.0565***	LOG(STOCK)	-4.1233	-11.7047***
INTERCEPT	4.2766	14.1425***	INTERCEPT	50.6283	15.0515***
Adjusted R-squared	*0.8372*			*0.8642*	
Durbin-Watson stat	*0.2490*			*0.5637*	
Jarque Bera (p-value)	*0.1949*			*0.4427*	

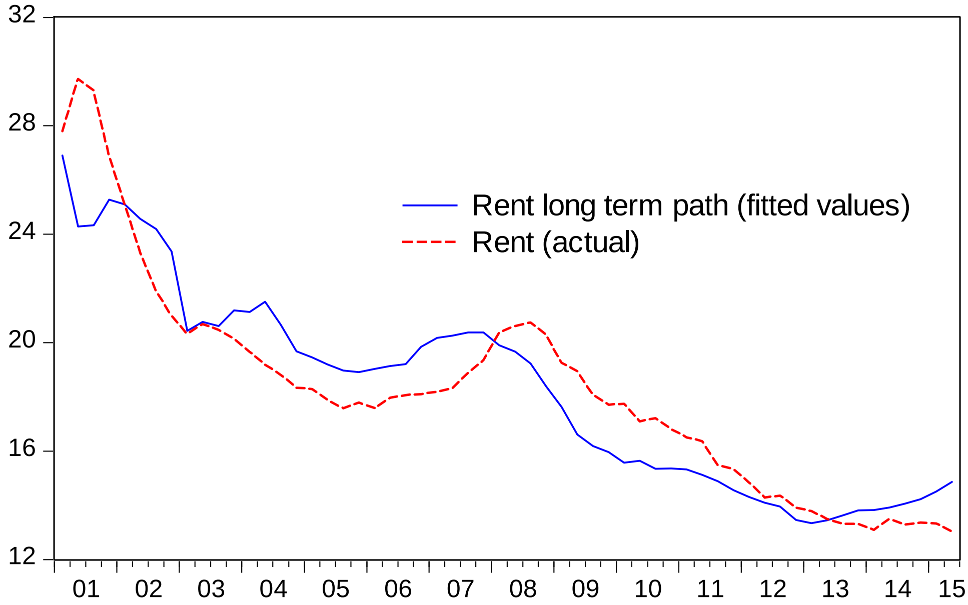
Note: Cointegrating equation estimated by FMLS, using Spanish GDP and Service sector employment (SEMP) as regressors for the long term expression for average rents.
*** Denotes significance at 1% confidence level; Sample 2001Q1 – 2015Q2; Included observations: 58

Both expressions explain similarly the long term path for rents with positive GDP and SEMP elasticities. On the other hand, long term elasticity for STOCK is negative in the two equations. The adjusted R-squared is high as expected in regressions with variables in levels with a time trend.

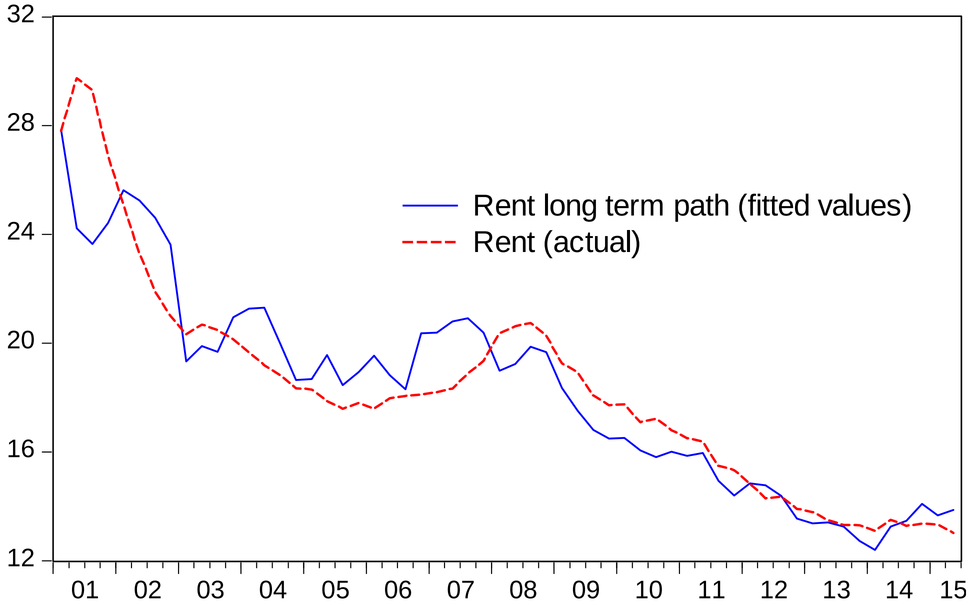
One advantage of estimating long term expression for prices is the possibility to check periods of under and over valuation. In the figure 3 we have represented the actual rental prices versus the estimated long term rent values. In both cases actual rents present some 5 year periods of under and over valuation. After the Dot-Com bubble bust rents were above their equilibrium. Since 2002 rents decreased and remained below their long term level until 2007, coinciding with the end of the expansion period of the Spanish economy. After the beginning of the last crisis, fundamentals set lower levels of equilibrium rents. In the period 2013-2015 rents are below long term path.

Figure 2.2. Long run rent estimation using cointegrating equations of table 5

Rent gap between estimated equilibrim rent and actual rent (€/m²/month)
(Cointegrating equation estimated with GDP and Stock as regressors)



Rent gap between estimated equilibrim rent and actual rent (€/m²/month)
(Cointegrating equation estimated with
Madrid's service sector employment and Stock as regressors)



The relationship between the long term rent and the actual values is similar among the two models estimated. However, the levels are different; In the case of the model using Spanish GDP as regressor for the cointegrating equation the average over pricing is 8% whilst using SEMP as regressor, the average overprice is 6%. Under pricing periods with both GDP and SEMP approaches have an average deviation of 6%. The estimated error correction model is presented in table 2.6.

Table 2.6. ECM estimates (2SECM)

Short run models.
Spanish GDP as Demand Proxy
Estimation method: Seemingly Unrelated Least Squares (SUR)

	Coefficient	t-Statistic	P-value
Rent - DLOG(RENT)			
INTERCEPT	-0.0423	-2.0963	0.0377
DLOG(RENT(t-1))	0.4973	5.7388	0.0000
DLOG(RENT(t-6))	-0.1974	-2.2241	0.0276
DLOG(STOCK(-6))	-0.6107	-1.9435	0.0537
LOG(VACR(t-1))	-0.0166	-1.9634	0.0514
ECM_{REnt}(t-1)	-0.1545	-4.1598	0.0001
Adjusted R-squared	0.5453		
Durbin-Watson stat	1.9334		
Vacancy - DLOG(VACR)			
INTERCEPT	-0.1618	-4.5187	0.0000
DLOG(VACR(-1))	0.4078	3.7170	0.0003
DLOG(STOCK(-1))	2.4988	3.3974	0.0009
DLOG(STOCK(-2))	-2.4176	-3.0834	0.0024
DLOG(GDP(-1))	-4.4476	-2.9856	0.0033
LOG(VACR(-1))	-0.0824	-5.0580	0.0000
ECM_{RENT}(t-1)	-0.1634	-1.6546	0.1000
Adjusted R-squared	0.6656		
Durbin-Watson stat	2.2416		
Stock - DLOG(STOCK)			
INTERCEPT	-0.0123	-2.7036	0.0076
DLOG(STOCK(-7))	0.3286	3.4618	0.0007
VACR(-4)	-0.0064	-3.4130	0.0008
ECM_{RENT}(t-2)	-0.0428	-4.0800	0.0001
Adjusted R-squared	0.5344		
Durbin-Watson stat	1.8802		

Short run models.
Madrid Service Sector Employment as Demand Proxy
Estimation method: Seemingly Unrelated Least Squares (SUR)

	Coefficient	t-Statistic	P-value
Rent - DLOG(RENT)			
INTERCEPT	-0.0681	-2.9805	0.0033
DLOG(RENT(t-1))	0.4511	5.1460	0.0000
DLOG(RENT(t-6))	-0.3589	-3.9902	0.0001
DLOG(STOCK(t-4))	-0.9690	-2.9812	0.0033
VACR(t-1)	-0.0280	-2.8327	0.0052
ECM_{REnt}(t-1)	-0.1116	-3.1189	0.0022

Adjusted R-squared	0.6352		
Durbin-Watson stat	2.1227		
Vacancy - DLOG(VACR)			
INTERCEPT	-0.1214	-2.8351	0.0052
DLOG(VACR(-1))	0.3331	3.0314	0.0028
DLOG(STOCK(-1))	2.1559	2.5695	0.0111
DLOG(SEMP(-1))	-1.5584	-2.9630	0.0035
DLOG(SEMP(-4))	-1.0053	-1.8813	0.0618
LOG(VACR(-1))	-0.0592	-3.0681	0.0025
ECM_{RENT}(t-1)	-0.2429	-2.5986	0.0103
Adjusted R-squared	0.5309		
Durbin-Watson stat	1.7279		
Stock - DLOG(STOCK)			
INTERCEPT	-0.0106	-2.4348	0.0160
DLOG(STOCK(-7))	0.2764	2.9641	0.0035
LOG(VACR(t-4))	-0.0058	-3.2384	0.0015
ECM_{RENT}(t-2)	-0.0492	-5.0439	0.0000
Adjusted R-squared	0.5825		
Durbin-Watson stat	2.1158		

Sample 2001 Q1 – 2015 Q2; Included observations: 58; Total System observations: 174

We have estimated two systems of short run equations for average office rent. One uses Spanish GDP in the cointegrating equation and short term dynamics, the other uses Madrid's service sector employment (SEMP hereafter) instead. Using a database of quarterly observations, we have restricted the model to a maximum of 8 laggards, as in real estate literature it is common to include two years in order to capture construction dynamics which take such time to deliver new buildings to the market[20]. To get to the final models we present in table 6 we have used a backward procedure, which progressively omits all insignificant estimators from a general specification (Hendry et al. 1999).

Adjusted R-squared values range from 53% to 66%. The lowest values are obtained in the estimations for change in vacancy rate and stock when using SEMP as activity proxy (53% in both cases). The equations of variation of vacancy rate with GDP as an activity proxy obtains the highest value (66%). Adjustments mechanisms (rent ECM and vacancy rate's) obtain the expected negative sign however the speed of adjustment is not the same. Serial correlation among residuals don't seem to pose a problem as Durbin-Watson statistic falls all the time in the acceptable range of 1.5-2.5. In order to test higher degree of autocorrelation we have tested the Portmanteau test (Ljung and Box, 1978). As our modelling takes into account up to eight lags we have test residual serial correlation up to that lag plus other four periods. The results of the Portmanteau Q-statistic (H0: no serial correlation) reject the null hypothesis for lags tested (please see details in the tables of annex of chapter 2).

[20] We also ran a lag structure test using a simple VAR model. Most of the criteria used with the GDP specification pointed to a lag structure of 8 lags while the SEMP specification had a less homogeneous structure with two criteria pointing to 8 lags structure, one to 7 lags and two to two lags. Please see Lag Order Selection title in the annex of chapter 2 to see all tests with the 5 criteria used.

Rental dynamics: When GDP is the selected as activity variable, rent ECM results higher than when SEMP is used. In particular, rent deviations from long term equilibrium are corrected 15% each quarter when one models with GDP and 11% each quarter when using SEMP. All other factors being equal, rent deviations are offset in 6.5 quarters (19 months) when modelling with GDP and in 9 quarters (36 months) when using SEMP. Vacancy rate happens to adjust faster when SEMP is used but the coefficients have similar magnitudes: 2.6% each quarter when modelling with job market figures and 1.6% each quarter when using national output. Rent variations also negatively depend on lags of stock variation and rents themselves in both specifications. At the same time, GDP variations or SEMP variation resulted significant for rent dynamics and its main impact is derived from the ECM.

Vacancy rate dynamics: Both approaches respond similarly to their own first lag as well as strongly positively to the first lag of stock variation. The variation in economic activity negatively impacts vacancy rates variations and it is important to stress the values of such elasticities: GDP modelling yields a strong impact of GDP on vacancy rate dynamics of around -2,41 points. On the other hand, SEMP variations impact with the first (-1.5) and fourth lags (-1.0). The log-level of vacancy rate has higher impact when GDP is used in the model (8%) than when SEMP is used (6%). When checking the rent ECM on vacancy rate variation we obtain higher speed of adjustment with the SEMP model (22% each quarter) than when GDP is used (16% each quarter).

Stock dynamics: Supply equations are the most parsimonious of the system and the main components are vacancy rate and rent gap mechanisms. For both cases (GDP and SEMP) the seventh lag of the stock variation plays an important role, with estimated coefficients of 0.33 and 0.28 for GDP and SEMP cases, respectively. The correction mechanisms from rent and vacancy participate with the second and fourth lags respectively. This means that stock growth, which is a proxy of new deliveries, is affected by disequilibria observed in vacancy rate one year ago and is rents two quarters ago. This is in line with EJM who argue that longer lags of the regressors affect the stock dynamics due to the time it takes developers to deliver new buildings to the market. Yet, for them, the lag of the ECM is two years.

b. Single equation methodology estimates

We proceed now to estimate equations (6) to (8) with the single equation error correction modelling (SEECM). With this framework we construct a system of equations which may be estimated by SUR in spite the presence of non-stationary and co-integrated variables. This is thanks to the fact that dependent variables of the system are in differences and therefore the estimation of spurious regressions are omitted (De Boef et al., 2004). Table 7 presents the results of the SEECM for the GDP and SEMP cases with the SUR estimation method.

Table 2.1. ECM estimates (SEECM)

Short run models ***Spanish GDP as Demand Proxy*** *Estimation method: Seemingly Unrelated Least Squares (SUR)*			
	Coefficient	t-Statistic	P-value
Long term coefficients			
LOG(GDP)	1.7250	-3.6755	0.0003
LOG(STOCK)	-2.3085	4.4910	0.0000
Rent DLOG(RENT)			
INTERCEPT	6.4366	3.6970	0.0003
DLOG(RENT(t-1))	0.4825	5.3085	0.0000
DLOG(STOCK(-6))	-0.7894	-2.5205	0.0127
LOG(VACR(t-1))	-0.0299	-1.7986	0.0740
ECM_{REnt}(t-1)	-0.2008	-5.4959	0.0000
Adjusted R-squared	0.5352		
Durbin-Watson stat	2.1423		
Vacancy DLOG(VACR)			
INTERCEPT	10.4803	2.7843	0.0060
DLOG(VACR(-1))	0.2529	2.4458	0.0155
DLOG(GDP(-1))	-6.7335	-5.1831	0.0000
DLOG(STOCK(-2))	2.7215	3.5562	0.0005
LOG(VACR(-1))	-0.1167	-4.1107	0.0001
ECM_{RENT}(t-1)	-0.3311	-3.6296	0.0004
Adjusted R-squared	0.6163		
Durbin-Watson stat	1.6916		
Stock DLOG(STOCK)			
INTERCEPT	1.8383	3.7080	0.0003
DLOG(STOCK(-7))	-0.0146	-3.1777	0.0018
VACR(-4)	-0.0576	-4.0472	0.0001
ECM_{RENT}(t-2)	1.8383	3.7080	0.0003
Adjusted R-squared	0.3879		
Durbin-Watson stat	1.6111		

Short run models ***Madrid Service Sector Employment as Demand Proxy*** *Estimation method: Seemingly Unrelated Least Squares (SUR)*			
	Coefficient	t-Statistic	P-value
Long term coefficients			
LOG(GDP)	-2.2901	-6.7315	0.0000
LOG(STOCK)	4.0787	7.9061	0.0000
Rent DLOG(RENT)			
INTERCEPT	6.6601	3.1260	0.0021
DLOG(RENT(t-1))	0.5441	5.9162	0.0000
DLOG(STOCK(-1))	0.1452	1.7972	0.0742
DLOG(SEMP(t-6))	-0.5179	-2.1386	0.0340

LOG(VACR(t-1))	-0.0293	-1.6851	0.0940
ECM_{REnt}(t-1)	-0.1772	-4.9275	0.0000
Adjusted R-squared	0.5283		
Durbin-Watson stat	2.3225		
Vacancy DLOG(VACR)			
INTERCEPT	13.7737	2.7133	0.0074
DLOG(VACR(-1))	0.3092	2.7517	0.0066
DLOG(SEMP(-1))	-1.7100	-3.2035	0.0016
DLOG(SEMP(-2))	-0.9775	-1.7835	0.0765
DLOG(STOCK(-1))	2.0910	2.5027	0.0134
LOG(VACR(-1))	-0.0529	-2.1710	0.0314
ECM_{RENT}(t-1)	-0.2707	-2.8148	0.0055
Adjusted R-squared	0.5077		
Durbin-Watson stat	1.5832		
Stock DLOG(STOCK)			
INTERCEPT	3.4493	7.0339	0.0000
VACR(-8)	-0.0069	-1.7633	0.0798
ECM_{RENT}(t-2)	-0.0675	-4.9681	0.0000
Adjusted R-squared	0.5528		
Durbin-Watson stat	2.1251		

Sample 2001 Q1 – 2015 Q2; Included observations: 58; Total System observations: 174

The estimated SEECM for GDP and SEMP behave, to the greatest extent, similarly. Nevertheless, the R-squared values are less than with the 2SECM. This is explained in part from the fact that coefficients of the long term deviations are simultaneously estimated, decreasing degrees of freedom. It also may be derived from the fact that each long term coefficient actually is estimated in each variation equation. The adjusted R-squared values now range between 37% and 61%, lower than what we obtained with the 2SECM. Nevertheless, adjusted R-squared values were uniform for the three equations with the SEMP approach, and ranging from 50% to 55%.

Rental dynamics: When modelled with GDP rent variation depends on its one quarter lagged value as well as the first lag of stock variation. This coefficient holds a negative value. The coefficients of the correction mechanisms for rents and vacancy have also negative value. On the other hand, when using SEMP as demand proxy the same variables resulted significant for the model but the change in the exogenous economic driver (SEMP) appeared with its sixth lag. Regarding the correction mechanisms, that derived from rent gap suggests a speed of adjustment of 20% each quarter when using GDP, pointing to a complete correction, *ceteris paribus*, of 15 months. When SEMP is employed as exogenous demand driver, the speed of correction is 17% per quarter, which means rent adjustment takes place in around 18 months. For the vacancy rate gap, rents are offset by vacancy in 3% each quarter for both GDP and SEMP approaches.

Vacancy rate dynamics: Vacancy rate change depends on its first lag, also negatively depends first lag of GDP or SEMP and the positively from the second lag of stock. (Both specifications resulted quite similar). The correction mechanism from rent indicates a quick adjustment of vacancy (close to 30% in both economic variables). The vacancy rate gap is also similar with estimated values of 5% in both cases.

Stock dynamics: Stock growth rate depends on the seventh lag of stock when GDP is used. When SEMP is employed, stock only depends on the vacancy rate gap in its fourth lag and the rent ECM in its second. The same happened with GDP modelling, but in this case the observation of rent two years ago determines the current variation of stock. It is the second lag of the rent ECM that affects current deliveries.

Long run vacancies

As commented in the modelling section, a different definition of the long run vacancy rate is embedded in each of the short run equations. We used the estimated values to retrieve the long run vacancy rate for each equation estimated in the 2SECM but not with the estimations, as long as the information embedded in the constant term also includes the constant of the cointegrating relationship times the adjustment coefficient. Table 2.8 presents the results.

Table 2.8. Estimated values for the long run vacancy rate

Estimated long run vacancy rates

	Equation to retrieve vacancy rate	Growth equation	GDP as demand proxy (%)	SEMP as demand proxy (%)
Two step ECM	$v^* = -\alpha_0/\sum_{i=0}^{n_4}\alpha_{4,i}$	Rent	12.8	11.4
	$v^* = -\beta_0/\sum_{i=0}^{n_4}\beta_{4,i}$	Vacancy rate	7.1	7.8
	$v^* = -\gamma_0/\sum_{i=0}^{l_2}\gamma_{\ ,i}$	*Stock*	6.8	6.1

Note: Values retrieved as $expGv^*$

Although the results are similar across GDP and SEMP modelling, they differ among growth equations. They are closer for the vacancy rate and stock equations ranging between 6.1% and 7.8%. As per the equations of rents, the long run values for rents are 11.4% and 12.8% indicating high values of stationary vacancy rate of the Madrid office market. The estimates from vacancy rate and stock equations seem more reasonable and in line with the research of Hendershott *et al*. (2013).

In order to summarize the findings of the estimation we present the results of the obtained error mechanisms in all methods utilized. Table 2.9 contains the values of rent ECM and vacancy rate gaps.

Table 2.9. Summary of rent and vacancy ECM

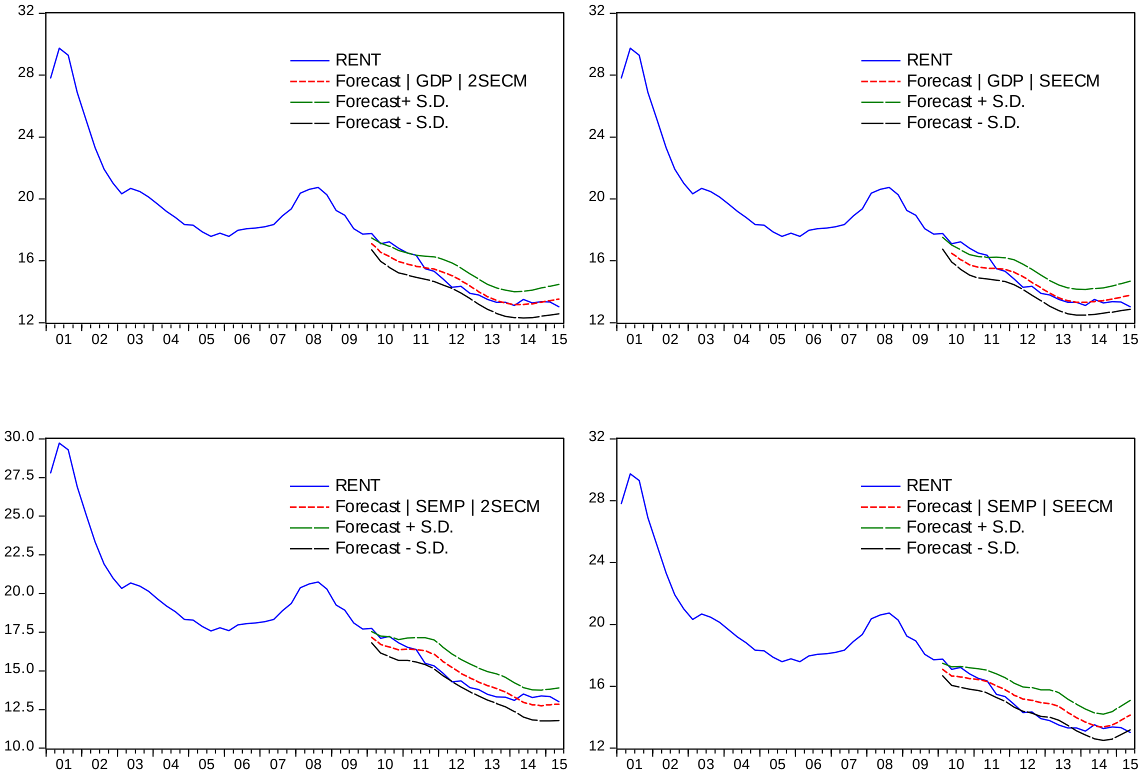
	Estimated coefficients of error correction mechanisms (Lags of the correction mechanism in parenthesis)			
		Growth equation	GDP as demand proxy	SEMP as demand proxy
Two step ECM	Rent ECM	Rent	-0.1545 (t-1)	-0.1116 (t-1)
		Vacancy rate	-0.1634 (t-1)	-0.2429 (t-1)
		Stock	-0.0428 (t-2)	-0.0492 (t-2)
	Vacancy gap	Rent	-0.0824 (t-1)	-0.0280 (t-1)
		Vacancy rate	-0.0824 (t-1)	-0.0592 (t-1)
		Stock	-0.0064 (t-4)	-0.0058 (t-4)
Single Equation ECM	Rent ECM	Rent	-0.2008 (t-1)	-0.1772 (t-1)
		Vacancy rate	-0.3311 (t-1)	-0.2707 (t-1)
		Stock	-0.0576 (t-2)	-0.0675 (t-2)
	Vacancy gap	Rent	-0.0299 (t-1)	-0.0293 (t-1)
		Vacancy rate	-0.1167 (t-1)	-0.0529 (t-1)
		Stock	-0.0146 (t-4)	-0.0069 (t-2)

Note: All values are significant at a 5% of confidence level.

2.7. Forecast performance comparison

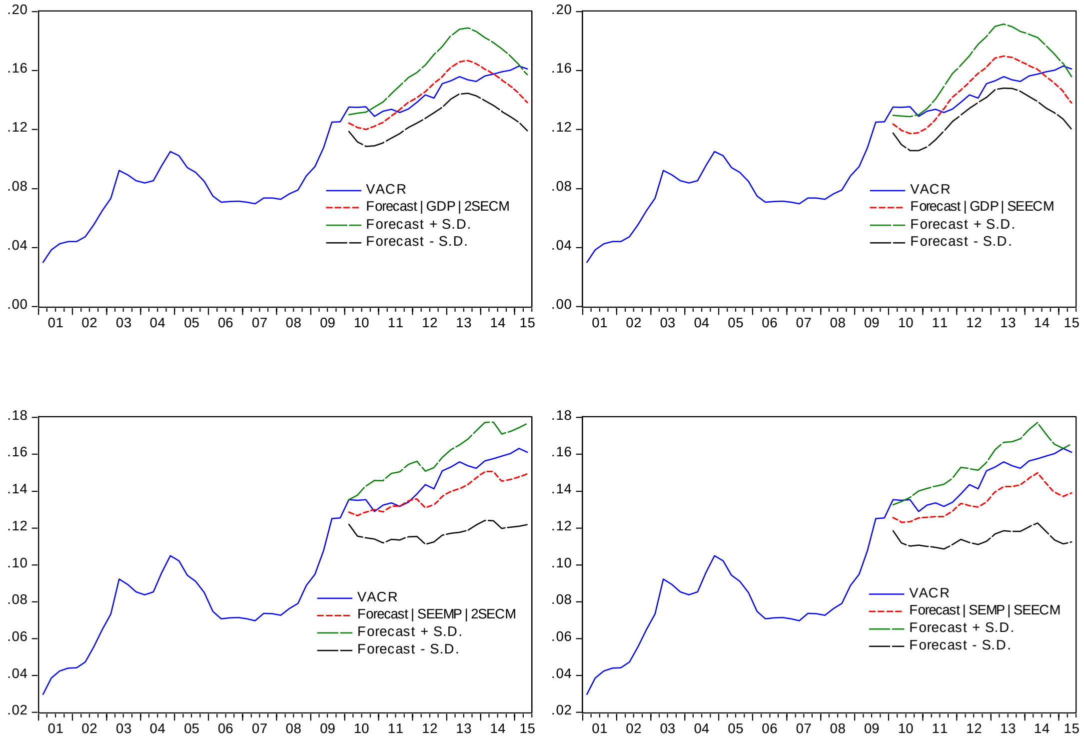
To initially illustrate the differences in forecasting performance of our four models we present the charts of the dynamic forecast in the period 2010 Q1 to 2015 Q2.

Figure 2.4. Rent dynamic forecast



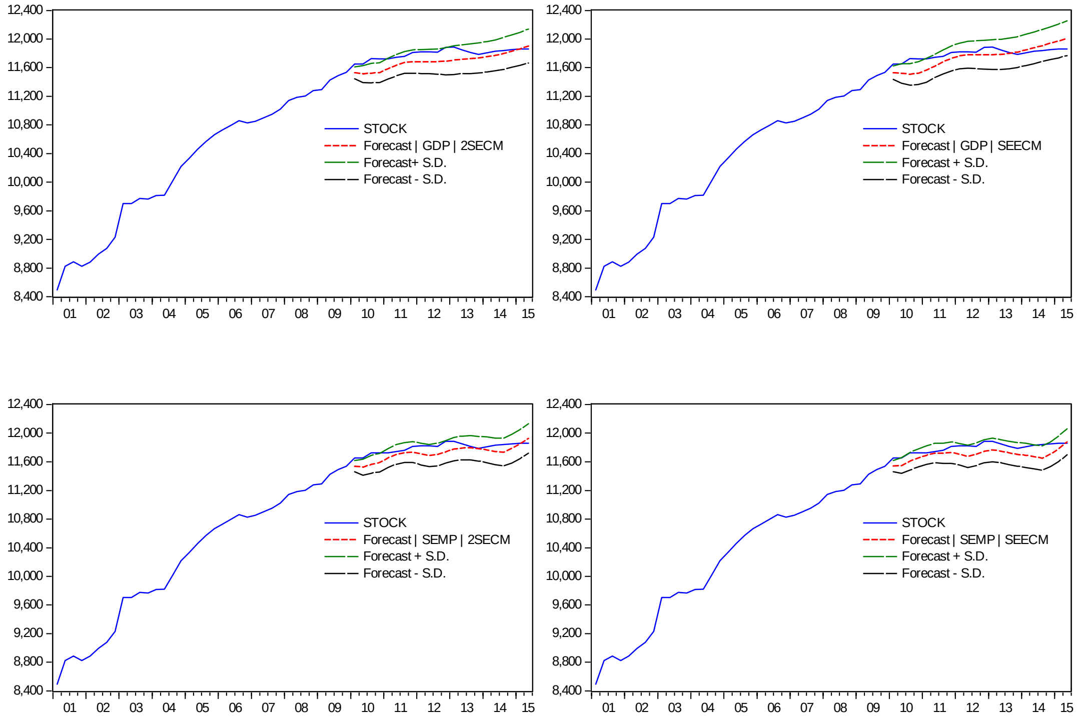
Rent dynamic forecast with the four approaches employed. Sample 2001 Q1 – 2015 Q2; Included observations: 58; Total System observations: 174.

Figure 2.5. Vacancy Rate dynamic



Vacancy Rate dynamic forecast with the four approaches employed. Sample 2001 Q1 – 2015 Q2; Included observations: 58; Total System observations: 174.

Figure 2.6. Stock dynamic forecast



Stock dynamic forecast with the four approaches employed. Sample 2001 Q1 – 2015 Q2; Included observations: 58; Total System observations: 174. The less biased forecasts are those modelled with Spanish GDP.

As a general trait, the models predict a market recovery since the pick-up of the Spanish economy in H1 2014. In particular, rents are forecasted to increase in 2015, as well as stock. Also, vacancy rate should be falling during 2015. The goodness of fit seems higher in rent and stock equations, but less in vacancy rate. We have computed the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE) and the Theil Coefficient (THEIL) for all the forecasts produced to have a quantitative assessment of the forecast performance. Table 10 contains the results obtained as well a scoring value, that will help to aggregate the information of goodness of fitness of each forecast produced in a single figure.

Table 2.10. Results for forecast performance evaluation

	Rent forecast				*Vacancy rate forecast*				*Stock forecast*			
Model	GDP 2SECM	GDP SEECM	SEMP 2SECM	SEMP SEECM	GDP 2SECM	GDP SEECM	SEMP 2SECM	SEMP SEECM	GDP 2SECM	GDP SEECM	SEMP 2SECM	SEMP SEECM
*RMSQ**	0.50	0.53	0.55	0.70	1.03	1.30	0.92	1.34	129.29	104.14	88.12	104.37
	0.00	*0.12*	*0.21*	*1.00*	*0.25*	*0.91*	*0.00*	*1.00*	*1.00*	*0.39*	*0.00*	*0.39*
*MAE**	0.40	0.39	0.49	0.60	0.86	1.19	0.79	1.20	114.75	89.96	77.39	95.83
	0.03	*0.00*	*0.46*	*1.00*	*0.15*	*0.97*	*0.00*	*1.00*	*1.00*	*0.34*	*0.00*	*0.49*
MAPE	2.58	2.53	3.34	4.22	5.83	8.11	5.27	8.01	0.97	0.76	0.66	0.81
	0.03	*0.00*	*0.48*	*1.00*	*0.20*	*1.00*	*0.00*	*0.97*	*1.00*	*0.34*	*0.00*	*0.49*
THEIL	0.02	0.02	0.02	0.02	0.04	0.04	0.03	0.05	0.01	0.00	0.00	0.00
	0.00	*0.13*	*0.20*	*1.00*	*0.19*	*0.78*	*0.00*	*1.00*	*1.00*	*0.38*	*0.00*	*0.40*

Note: Low scores indicate better performance. The black numbers correspond to the obtained performance statistic for each variable forecasted, for each modelling approach and for each exogenous variable used. The grey numbers bellow correspond to the scoring system employed to aggregated and rate the forecasting performance statistics. *In €/sqm/month for Rent forecast; In % for Vacancy Rate forecast; In 000 sqm for Stock forecast.

As we have gathered 48 indicators of forecasting performance we have designed a normalized scoring system that allows to discern the best modelling techniques. Apart of ranking the scores, we made a measure of relative distance among each statistic computing the following formula:

$$G_i = \frac{G_i - G_{minG}}{G_{maxG} - G_{minG}};\ 0 \leq G_i \leq 1$$

In this ratio the maximum performance statistic G_{maxG} takes value of one $G_i = 1$ and the minimum performance statistic G_{minG} take value of zero $G_i = 0$. The intermediate performance statistics G_{iG} obtains a value of the relative distance between the maximum and the minimum values. This allows taking into account similar forecast performance among statistics. In other words, we weight the performance statistics as a function of their relative situation to avoid the homogenous weighting derived from a simple ranking. To aggregate the performance comparison of the individual performance measures we simply sum normalized scores and select that with the lowest result. The next table presents the main results of the aggregation of the normalized performance statistics:

Table 2.11. Results of the standardised forecast performance statistics

		Variable forecasted			
		Rent	Vacancy	Stock	Overall score
Exogenous variable and employed methodology	GDP | 2SECM	**0.05**	**1.27**	5.1	6.4
	GDP | SEECM	0.24	4.66	2.4	7.4
	SEMP | 2SECM	1.34	**1.27**	**0.9**	**3.5**
	SEMP | SEECM	4.00	4.96	2.8	11.7

We have aggregated the results of each performance statistic for each variable in order to obtain the best approach to make predictions. 2SECM performs better than SEECM in the partial 'equation-specific scores' as well as in the overall score, except for the Stock equation when GDP is used as exogenous variable.

Several readings can be made with Table 11. If we set the objective for comparing modelling techniques, we should compare row 1 against row 2 and row 3 against row 4. Doing this we may conclude that the 2SECM yields lower scores, therefore does it better than SEECM. The only exception is made in stock equation when using GDP as exogenous demand driver. Notice, however, that is a worse predictor when using SEMP as exogenous variable.

The second reading can be made to assess performance when using GDP or SEMP; we have to compare now row 1 against row 3 and row 2 against row 4. In this case, the results are mixed. In the partial particular-equation assessment, GDP modelling does it better when using SEECM in all equations. Nonetheless, the forecast performance when using 2SECM is mixed, and depends the equation one is focusing on. Checking on the overall score (last column to the right in Table 11) it is lower, therefore better, when using GDP and SEECM than when using SEMP and SEECM. On the other hand, the general score is lower when using SEMP and 2SECM than when using GDP and 2SECM. So when comparing forecast performance from the perspective of exogenous variables we may say that there is draw and researcher criterion is important to decide which model to use.

Checking single variable performance forecasts, the scoring system indicates that using jointly GDP and 2SECM is the best approach to predict rents and vacancy rate. The most fitted approach to make stock forecast is using regional service sector employment (SEMP) and two stage error correction mechanism (2SECM). This methodology also does it well to forecast vacancy rate.

Finally, comparing only overall scores leaves us with the simple task possible and it is to select the approach with the least overall score. The good performance in stock and vacancy rate forecasting allows the two stage error correction mechanism (2SECM) using Madrid's service sector employment (SEMP) to be the best approach to forecasts rents, vacancy and stock in a single system.

2.8. Concluding remarks

We have modelled Madrid's office market with a system of equations for stock variation, vacancy rate variation and rental prices (average real rent) variation, within an error correction mechanism framework. This framework allows for capturing long term development paths and, therefore, analyse short term deviations from the long term track. Having rejected the hypothesis of non-existence of first degree stationarity of the variables participating in the model (i.e. rents, vacancy rate, stock, GDP and Madrid's service sector employment) we have failed to reject the hypothesis of non-existence of cointegration, setting a solid ground for co-integration estimation techniques. We have used two approaches for estimation of error correction mechanisms: The two stage error correction mechanism (2SECM) and the single equation error correction mechanism (SEECM). The latter approach is innovative in the context of commercial real estate, as the 2SECM is classical in property research literature and, to the best of our knowledge, the SEECM has not been used in real estate papers so far.

Both techniques were tested using two different exogenous variables proxying economic activity: Spanish GDP and Madrid's service sector employment. As a consequence, we have fitted and compared four models. Our results suggest quite similar explanatory capabilities of these two exogenous economic variables. When modelling the short run, we produce a robust structure with the high degree of significance of regressors as well as high goodness of fitness for the four models estimated. For the case of rents dynamics, the economic driver gives feedback through the long term expression. They also rely their lagged value and changes in the stock level. Vacancy rates actually depend on their lagged values as well as on the dynamics of economic driver (GDP or service sector employment). Stock tends to be the most rigid expression and depending only on its lagged values and the error correction mechanism of vacancy rate and rents.

The speed of adjustment to long term rent gaps and long term vacancy rate gaps have the expected –negative– sign and magnitudes in all estimated equations systems. Although there is variation among models, we may say that Madrid office rents adjust each quarter around 15% of their deviation from long term rent equilibrium. Rents' average adjustment speed to long term vacancy rate gaps is around 4% in each quarter. The quarterly adjustments of vacancy rates to long term rent gaps and long term vacancy rate gaps are 25% and 7.5%, respectively. Regarding stock, the speed of adjustment is the lowest and is around 5% in the case of rent gap while less than 1% for the case of vacancy rate gaps.

Recurring to the properties of our theoretical equations (6 to 8 equations above), we have derived the long term values of vacancy rate or natural vacancy rate. When using the rent dynamics expression, we obtain values around 12%. Nevertheless, when using vacancy rate and stock's short term equations to solve for the long term vacancy, we obtain values around 7% which are more in accordance to related literature (EGHS and HJM). The full sample average vacancy rate is 10.4% and using it as a benchmark, we see more realistic the long term value derived from vacancy rate and stock equations. Also from the perspective of the authors this value is more in line with the sound level for an office market.

We test our models to dynamically forecast a period of five years. As a general trait, rents and stock forecasts have the lowest levels of error. This means that the endeavour to forecast vacancy rates is more challenging. Nevertheless, the forecasts of the four models estimated present low levels of error and fit well to the actual values of the endogenous variables (please refer to table 2.10).

Finally, we have designed a comparative scoring system to aggregate the results of 4 different forecast performance indices. Using this technique, we posit that the best model to forecast rent is the two stages error correction mechanism (2SECM) using GDP as exogenous economic variable. It is therefore important to mention the feedback of an aggregated variable, such as GDP, on local businesses decisions is strong and worth to analyse. This combination also holds for vacancy rate forecasting. Yet, when stock is forecasted it is better to use Madrid service sector employment

(SEMP) as exogenous demand proxy, but maintaining the usage of the 2SECM. This last combination remains as the best approach to estimate the system of three equations, as its forecasts of vacancy rates are as good as with GDP and 2SECM and its forecast error is low for the rents case.

Although the introduction of the single equation error correction mechanism is innovative, it did not yield consistently better results than the more classical 2SECM. Nor did the SEECM allowed to inform about long term vacancy rate as the constant term of the long run equation (whether significant or not) is embedded in the short run expression.

As research paths opened with this research we suggest testing of asymmetrical shocks as well as impulse-response analysis. Other line of investigation may come from panel data modelling, pooling market data from European capital cities and extract also fixed effects of each market apart of the classical elasticities.

References

Alberts, William W. 1962. Business Cycles, Residential Construction Cycles, and the Mortgage Market. The Journal of Political Economy, 263–81.

Ball, Michael, Colin Lizieri, and Bryan D. MacGregor. 1998. The Economics of Commercial Property Markets. Psychology Press. https://books.google.es/books?hl=en&lr=&id=PxIqjANgCI4C&oi=fnd&pg=PR12&dq=Ball,+Lizieri,+%26+MacGregor,+1998&ots=UAn0V6AotN&sig=KC_O-ofRNXa7R-yL5oaN1rEkF6g.

Banerjee, Anindya, Juan J. Dolado, John W. Galbraith, and David Hendry. 1993. Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data. OUP Catalogue. Oxford University Press. https://ideas.repec.org/b/oxp/obooks/9780198288107.html.

Blank, David M., and Louis Winnick. 1953. The Structure of the Housing Market. The Quarterly Journal of Economics 67 (2): 181–208. doi:10.2307/1885333.

Brooks, Chris, and Sotiris Tsolacos. 2010. Real Estate Modelling and Forecasting. Cambridge University Press. https://books.google.es/books?hl=en&lr=&id=qSBn9qMYp1oC&oi=fnd&pg=PR11&dq=Brooks+%26+Tsolacos,+2010&ots=i6bXjXgxAo&sig=PQU9mUhmElPgXJe9I-m7nbJIRvs.

Brounen, Dirk, and Maarten Jennen. 2009. Asymmetric Properties of Office Rent Adjustment. The Journal of Real Estate Finance and Economics 39 (3): 336–58. doi:10.1007/s11146-009-9188-9.

Engle, Robert F., and C. W. J. Granger. 1987. Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica 55 (2): 251–76. doi:10.2307/1913236.

Englund, Peter, Åke Gunnelin, Patric H. Hendershott, and Bo Söderberg. 2008. Adjustment in Property Space Markets: Taking Long-Term Leases and Transaction Costs Seriously. Real Estate Economics 36 (1): 81–109. doi:10.1111/j.1540-6229.2008.00208.x.

Ferri, Michael G. 1977. An Application of Hedonic Indexing Methods to Monthly Changes in Housing Prices: 1965–1975*. Real Estate Economics 5 (4): 455–62.

Franz Fuerst, and Patrick McAllister. 2010. Supply Elasticities and Developers' Expectations: A Study of European Office Markets. Journal of European Real Estate Research 3 (1): 5–23. doi:10.1108/17539261011040514.

Hendershott, Patric H., Maarten Jennen, and Bryan D. MacGregor. 2013. Modeling Space Market Dynamics: An Illustration Using Panel Data for US Retail. The Journal of Real Estate Finance and Economics 47 (4): 659–87. doi:10.1007/s11146-013-9426-z.

Hendry, David F., and Hans-Martin Krolzig. 1999. Improving on'Data Mining Reconsidered'by KD Hoover and SJ Perez. The Econometrics Journal, 202–19.

Johansen, Søren. 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. Econometrica 59 (6): 1551–80. doi:10.2307/2938278.

Keele, Luke, and Suzanna De Boef. 2004. Not Just for Cointegration: Error Correction Models with Stationary Data. Documento de Trabajo. Departamento de Política Y Relaciones Internacionales, Nuffield College Y Oxford University. http://www.nuffield.ox.ac.uk/Research/Politics%20Group/Working%20papers/Documents/Working%20papers/2005/Keele%20DeBoef%20ECM%20041213.pdf.

Krolzig, Hans-Martin, and David F. Hendry. 2001. Computer Automation of General-to-Specific Model Selection Procedures. Journal of Economic Dynamics and Control 25 (6): 831–66.

Ng, Serena, and Pierre Perron. 1995. Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag. Journal of the American Statistical Association 90 (429): 268–81.

Pritchett, Clayton P. 1977. The Effect of Regional Growth Characteristics on Regional Housing Prices. Real Estate Economics 5 (2): 189–208.

Remøy, Hilde Therese. 2010. Out of Office: A Study on the Cause of Office Vacancy and Transformation as a Means to Cope and Prevent. IOS Press. https://books.google.es/books?hl=en&lr=&id=_Toa4ag6rEcC&oi=fnd&pg=PR10&dq=Rem%C3%B8y,+2010&ots=Hx1dDeXqJ_&sig=T9HIRwp8htP7ocxpnmWs5EXy-to.

Rosen, Kenneth T. 1984. Toward a Model of the Office Building Sector. Real Estate Economics 12 (3): 261–69.

Wheaton, William C. 1987. The Cyclic Behavior of the National Office Market. Real Estate Economics 15 (4): 281–99.

Wheaton, William C., Raymond G. Torto, and Peter Evans. 1997. The Cyclic Behavior of the Greater London Office Market. The Journal of Real Estate Finance and Economics 15 (1): 77–92.

HEDONICAL OFFICE RENTS WITH SPATIAL ECONOMETRICS[21][22]

[21] This work has been presented in the main sessions of the European Real Estate Society Congress 2015 in Istanbul, Turkey.

[22] The results of this research have been submitted to the Journal of Real Estate Finance and Economics (JREFE) in February 2016. As of April 2016 it is under review of for publication. The JREFE is a double-blind refereed academic journal edited by Springer with an 2014 Thomson Reuters' Impact Factor of 0.657 SNIP of 1.105 and a SC Imago Journal Rank of 0.838.

3.1. Introduction

Being property prices the baseline for real estate research, analysts have to choose from several definitions of prices. Research findings have to be clearly delimited for the sake of interpretation and decision taking as well as policy making. Since the 1990s data availability allowed new non-residential analytical papers on the evolution and cyclical behaviour of the office, retail and logistics occupier markets, mainly in London and New York/Manhattan areas. The techniques used swiftly revealed the necessity of utilization of well performing rental indexes in order to avoid 'imputed noise' to the estimations and extracting conclusions on non-market rent drives. In this sense 'controlled' experiments on rental behaviour were born and econometric modelling took a focus point once occupied by several definitions of average rents. These types of experiments are still new to markets outside UK and USA and there is still room to prove theoretical developments on real estate economics in alternative markets. In this sense we have explored the estimation of a hedonical rent index for the office market in Madrid, taking advantage of new data availability and new estimation tools quite fitted to the real estate analysis (*i.e.* spatial econometrics). In this paper we estimate a spatial lag rent model by maximum likelihood and calculate the rent for an 'average office' the latter defined as that with average characteristics as well as an average location. The data for our study has two main sources: a) a list of new letting contracts provided by BNP Paribas Real Estate[23] with a semi-annual structure and information on each transaction such as headline rent, occupier company, space let and address, and b) a database published by the Spanish Land Registry (Cadastre) comprising a list of all the properties registered as an office in the city of Madrid and the cities in its catchment area. Each line of this database comprises hedonical characteristics such as size of the building housing a particular office, age of construction (and date of refurbishment), an index of technical quality of the building and geographical coordinates (Universal Transverse Mercator -UTM- coordinates) for geo-localization of the buildings. Added to the provided information we have calculated other hedonical characteristics such as distance to the closest metro station entrance and the type of company signing a new contract (if it was a multinational corporate, which is linked to the type of commitment of the contract signed). By crossing these two databases we pooled a full set of hedonical characteristics as well as the required geographical data to estimate hedonical spatial rent models. Estimations by standard Ordinary Least Squares (OLS) have been obtained by a researcher-driven General-to-Specific (GETS) modelling. This model is compared with the spatial-lag model to test for possible improvements derived from the geographical approach. We do detect improvements in the estimations as our spatial technique takes into account the existence of interaction between the rent level of a particular contract and the rent level of its neighbours, fact that we call spatial dependence. In other words the rents of an office reach certain level, partly, because the rent of its neighbours and partly for its hedonic characteristics. If this interaction is not taken into account (not using spatial modelling) biased estimated parameters are obtained and wrong conclusions will be drawn (Ward and Gleditsch, 2007).

Our estimated rent index is more accurate to capture the rental cycle in the Madrid office market than that based on (weighted) average rents. Compared to OLS estimations, spatial regressions have better performance in terms of explanatory

[23] BNP Paribas Real Estate at its turn exchanges its own transactions with the most important market players in order to develop market research, reaching 95% of the transactions in Madrid office market.

capacity and estimated coefficients stability distributions. This, in turn, gives a better understanding of the actual market evolution as depicts as less biased rent path than the benchmark OLS model.

The rest of the paper is structured as follows. In Section 3.2, we describe the mainstream in hedonical rental modelling and spatial econometrics, Section 3.3 is devoted to explain the spatial econometrics modelling. Section 3.4 presents the market and the data used in this research, discussing our dataset's advantages and limitations. In Section 3.5, we present our empirical exercise. Finally, in Section 3.6 we offer some concluding remarks. The annexes include detailed test and estimation outputs, software procedures and complementary charts and tables.

3.2. Literature review

Hedonical estimation has been used in several real estate fields both to analyse impacts of hedonical characteristics (e.g. externalities) as well as to obtain non-biased and well specified price or rents estimations since for more than 30 years (see Kain and Quigley, 1970; Straszheim, 1974). This research started with the housing market, as data sources are more bountiful than for commercial real estate. Clapp (1980) estimates a hedonical office rent model for the city of Los Angeles in his quest for explaining businesses rationale for choosing a particular location in that city. Wheaton and Torto (1994) explore for the first time the construction of a hedonical rent index for the commercial property market with a twofold objective: a) estimate the rental value of an archetypical office for the several office markets in metropolitan areas of the USA and b) compare the resulting index dynamics with the evolution of vacancy rate in order to check their opposite co-movement. Actually they mirror existing exercises of this type already explored for the residential market (Rosen, 1974; Case and Shiller, 1987). Southard et al. (1997) take the same approach with a smaller set of regressors but keeping the aim of estimating the rent value of the typical office in terms of lease, location and building characteristics. As hedonic techniques for price index studies proved their superiority over weighted average definitions and other indexes (Hill and Melser, 2008) panel data estimation techniques were adopted adding time components to regions/zones analysis and controlling for unobservable property characteristics[24] such as fixed and random effects. One more time, the majority of the hedonic estimation with panel data literature studied housing markets and, also once again, dwell mostly on the marginal effects estimation than on the out-of-the sample rent/price estimation (see Quigley, 1995; Gao and Wang, 2007; Hansen, 2009 and Osland, 2013).

In more recent iterations, researchers have explored hedonical methods misspecification issues when estimations are made with panel data (Kuminoff et al., 2010). These techniques try to capture spatial fixed effects with different levels for the constant between groups. However, as noted by Osland (2013), fixed effects estimation normally results in spurious solutions for unobserved spatial feedback. Osland (2010) develops a hedonic modelling based on spatial econometrics as alternative to fixed effects. In this paper, Osland specifies the different techniques developed for spatial econometrics, especially the Lag model, the error model and the Durbin model, all three to be explained in the next chapter.

[24] Such as city submarkets or neighbourhood attributes

3.3. Spatial feedback and its econometric modelling

Anselin (1988) and LeSage and Pace (2009) give an important caveat for researchers on the real estate field: in the presence of spatial feedback OLS modelling yields biased and inconsistent coefficient estimators.

LeSage and Pace (2009) and Anselin (2013) comment that spatial econometrics is worth to implement when moving across the Cartesian plane one finds a) spatial interdependence among data and b) spatial heterogeneity in model parameters. a) implies that explanatory variables are not fixed in repeated sampling exercises and therefore error terms tend to be correlated (Can, 1990). On the other hand b) implies that the assumption of a single linear relationship does not hold between different sub-samples of the observations composing the sample, yielding homoscedasticity violation issues, as explained by LeSage and Pace (2009). Furthermore, as much of property markets data are gathered and analysed with reference to their physical location in space, it is important to integrate the locational dimension to the economic modelling. As stated by LeSage and Pace (2009) the fundamental theorem of regional science is *distance matters.* In other words, at a geographical level, prices of properties close to each other are more related than to properties more distant in space. We have then set the necessary conditions to bring forth our chosen methodology to study spatial feedback.

a. Spatial Lag Model

Office rents, as other prices attached to property markets are bounded to location, case in point, near office rent levels tend to be more related to closer than to distant offices (Chasco and Sanchez 2015). If this is the case, there exist spatial autocorrelation or dependency. This in turn, dwindles as distance among offices increases (positive autocorrelation). The rationale for this is twofold. Asymmetries of information make economic agents to overcome it by referencing the price of the transacted asset to its peers' comparable transactions, relying on the principle of regional science. On the other side, there is a spill over effect whereby externalities are generated when agents make decisions (e.g. refurbishing) on their own properties. The hedonic model that controls for spatial dependency is formulated as:

$$\log G \quad = r = \rho G\ r + \alpha I_n + X\beta + G\,\gamma + \varepsilon,\ \ \varepsilon \sim G0_{n\times1}, \sigma\ \ I_{nG} \qquad \text{(Eq.3.1)}$$

Where is a vector of rents with dimension $n \times 1$. $\rho G$is a scalar parameter to be estimated and indicates spatial dependence intensity or, in other words, captures the average impact of neighbouring observations on the rents vector. As put by LeSage and Pace (2009), the existence of ρ in the regression exercise enables that some part of the total variation [in the dependant variable] across the spatial sample would be explained by each observation's dependence on its neighbours. The parameter ρG would reflect this in the typical sense of regression.

is the $n \times n$ weight matrix of spatial interconnexions. This matrix has the following characteristics:

- Each element comprises a measure of the distance between the *i*-th column observation and the *j*-th row observation
- This measure is the inverse of the distance, representing the increasing autocorrelation of closer offices letting rent levels

- It is row standardized, dividing each row element by the sum of the row, yielding unity sums in each row
- Diagonal set to zero, saying the distance to the proper office is non existent
- In this paper, distances are calculated with an Euclidean formula[25]

The matrix product *Wr* is then the spatial lag term, which takes into account the aforementioned spatial interaction of the endogenous variable. It is our view that this particular spatial model is the most intuitively fitted for office market analysis. This is because it straightforward captures the feedback between neighbouring rents, rather than associating such interdependencies to other parts of the specification, such as the error term[26] or the set of regressors[27], models to be discussed in the following sections.

X is a $n \times k$ matrix containing the independent variables' observations, β is a $k \times 1G$ matrix of the coefficients of the regressors, is a $n \times tG$matrix comprising time dummy variables, γ is a $t \times 1$ matrix of the coefficients of the time dummies and $\varepsilon G$is a vector of disturbances that are iid.

The spatial lag model seems familiar to the classical hedonic methodology and actually the latter is a special case of (1) when ρ equals 0. Nevertheless, the marginal effects of the regressors have special interpretations as implied by

$$r = GI_n - \rho G \quad ^{-1G} X\beta G + GI_n - \rho G \quad ^{-1\,G}\; \gamma G + GI_n - \rho\, G \quad ^{-1G} \varepsilon \qquad \text{(Eq.3.2)}$$

when solving for r in (Eq.3.1). Not only X is determining r in (Eq.3.2*)* but also all the spatial interdependencies captured by $I_n - \rho G \quad ^{-1}$, the so called *spatial multiplier* by Anselin (2003). From (Eq.3.2) we can obtain for a particular regressor X_{jG}or a particular period $_{kG}$

$$\frac{\delta rG}{\delta X_{jG}} = GI_n - \rho\, G \quad ^{-1}\left(I_n \beta_j\right) \qquad \text{(Eq.3.3)}$$

$$\frac{\delta rG}{\delta T_{kG}} = GI_n - \rho\, G \quad ^{-1G} I_n \gamma_k \qquad \text{(Eq.3.4)}$$

(Eq.3.3) and (Eq.3.4) are the marginal impacts of the regressors contrasting with the marginal impacts of the OLS regression $\beta_{jG}$and γ_k. It is worth to note that (Eq.3.3) and (Eq.3.4) are not scalars such as in the OLS case. They are $n \times n$ matrixes and therefore their interpretation is not straightforward. As noted by Fernandez-Aviles et al., (2012) the marginal information conveyed by (Eq.3.3) and (Eq.3.4) can by broken down by what it is in:

- The diagonal of the matrix: the marginal effect of the X_{iG}in a particular row (*i*) on the dependent variable observation *i*
- The off-diagonal of the matrix: the marginal effect of X_{jG}in a particular row (*i*) on the dependent variable observation *i*

[25] LeSage and Pace (2009) present a comprehensive collection of methodologies to build spatial weight matrixes
[26] Modelled with the Spatial Autoregressive Error model
[27] Modelled with the Spatial Durbin Model

- The sum of the row *i* yields the marginal impact of the of all (n) observations on the *i* observation of the dependent variable
- The sum of the column *i* yields the marginal impact of the X_j variable on all endogenous observations by the amount of the j-th observation

As commented, the OLS estimation in the presence of spatial feedback (e.g. equation 1) yields biased and inconsistent estimates. LeSage and Pace (2009) give an exit to this case proposing a classical maximum likelihood estimation method for estimating (Eq.3.1) and (Eq.3.2), which is the method followed in this research.

b. Spatial Autoregressive Error Model

It is also possible to assign a spatial structure to the error term of the model. Anselin (2003) propose such a model where the spatial error structure can be associated with a moving average process across the space. For the particular case of rent modelling the spatial autoregressive error model (SAEM) expression is:

$$\boldsymbol{r} = \boldsymbol{X\beta} + \boldsymbol{T\gamma} + \boldsymbol{\varepsilon} \qquad \text{(Eq.3.5)}$$

$$\boldsymbol{\varepsilon} = \boldsymbol{\lambda W \varepsilon} + \boldsymbol{u}, \quad \boldsymbol{u} \sim \boldsymbol{N}\left(\boldsymbol{0}_{\boldsymbol{n}\times\boldsymbol{1}}, \boldsymbol{\sigma^2 I_n}\right) \qquad \text{(Eq.3.6)}$$

In the SAEM $\boldsymbol{\lambda}$ is the spatial autoregressive coefficient (Osland, 2010). W is the weight matrix and $\boldsymbol{\varepsilon}$ and $\boldsymbol{u}$ are assumed uncorrelated. Solving 6 for $\boldsymbol{\varepsilon}$ and replacing it in (Eq.3.5) we get to the expression

$$\boldsymbol{r} = \boldsymbol{X\beta} + \boldsymbol{T\gamma} + \left(\boldsymbol{I_n} - \boldsymbol{\lambda W}\right)^{-1}\boldsymbol{u} \qquad \text{(Eq.3.7)}$$

In (Eq.3.7) the location of each office is affecting the stochastic error and therefore the level of each location's rental value. Osland (2010) stresses that when compared to the OLS regression the hedonic coefficients don't have to change substantially otherwise spatially correlated residuals are rather signalling omitted variables issues. When the spatial model correctly represents the market structure it is acknowledged that some degree of non-explained spatial feedback exists and it is in consequence exogenous. Such spatial shocks can be for example redefinition of business clusters derived from new thoroughfares or changes in physical infrastructure. Of course minor omitted variables may not alter a correct spatial error model structure.

Alternative structures for spatial modelling of the error term have been developed in literature. One of their main exponents is the Spatial Moving Average model

$$\boldsymbol{r} = \boldsymbol{X\beta} + \boldsymbol{T\gamma} + \boldsymbol{\varepsilon} \qquad \text{(Eq.3.8)}$$

$$\boldsymbol{\varepsilon} = \boldsymbol{\theta W u} + \boldsymbol{u}, \quad \boldsymbol{u} \sim \boldsymbol{N}\left(\boldsymbol{0}_{\boldsymbol{n}\times\boldsymbol{1}}, \boldsymbol{\sigma^2 I_n}\right) \qquad \text{(Eq.3.9)}$$

Where $\boldsymbol{\theta}$ is a spatial moving average coefficient and *u* and is an uncorrelated perturbation term (Moreno and Vayá Valcarcel, 2002).

c. Spatial Durbin Model

Adding restrictions to the parameters of equation (Eq.3.1) it is possible to get a model that combines spatial elements both on the endogenous variables as well as the exogenous. This model is time-series equivalent and is a spinoff of the spatial error model and/or the spatial lag model (see Anselin, 2003; Bivand, 1984) . We can denote the spatial Durbin model as:

$$r = \rho W r + X\beta + T\gamma + \rho W X \beta' + \varepsilon G \quad \text{(Eq.3.10)}$$

The similarity with the spatial lag model is straightforward as the expression $\rho \quad rG$ recognizes the impact of rent levels of the neighbours on each office. The added factor is $\rho G \; X\beta'^{G}$which captures the impact of neighbouring building characteristics on each office rent level.

The selection among the three models is suggested by Florax et al., (2003) in the following terms, using statistics tests, explained below:

Table 3.1. Florax model selection procedure

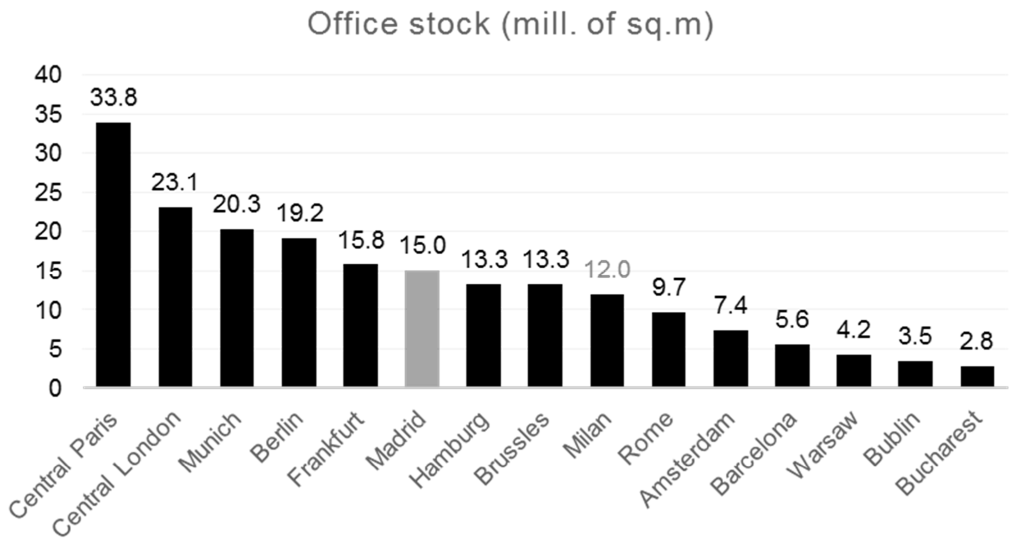
Step	Statistic	H0	H1	Action if H0 is rejected
1	Moran's I	Residuals with no spatial effects	Unspecified spatial process in residuals	Use Lagrange multiplier
2	Lagrange Multipliers error (LM-error)	No spatial autocorrelation in error structure (θ=0 and assumption of ρ=0)	Spatial autocorrelation (θ≠0)	Estimate Spatial error model
3	LM-lag	No spatial autocorrelation in endogenous variable spatial lag structure (ρ=0 and assumption of θ=0)	Spatial autocorrelation (ρ≠0)	Estimate Spatial lag model
If both H0 of steps 2 and 3 are rejected, the procedures is to select the model with the highest LM statistic				

3.4. Region description and dataset

a. The Madrid Office Market

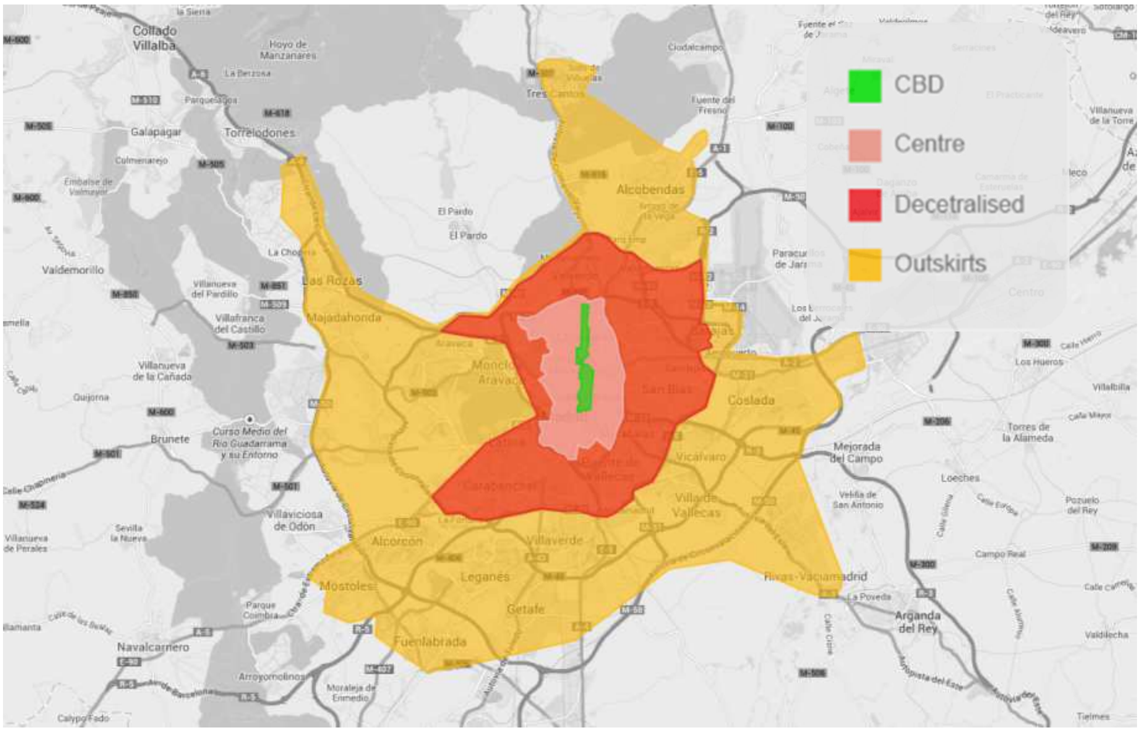
In the European context, the Madrid office market is average sized (Figure 3.1). With a stock of nearly 15 million sqm, the city ranks 6th among the main office hubs in Western Europe.

Figure 3.1. Office stock of the main cities of Western and Central Europe (2014)
Source: BNP Paribas Real Estate



With respect to market zones it is a general practice to divide the market in four sub-regions which are CBD, Centre, Decentralized (Dec) and Outskirts (Out) as seen on Figure 3.2.

Figure 3.2. Business districts of the Madrid office market
They may be also referred to as office zones or sub-markets. Source of the base map: Google Maps

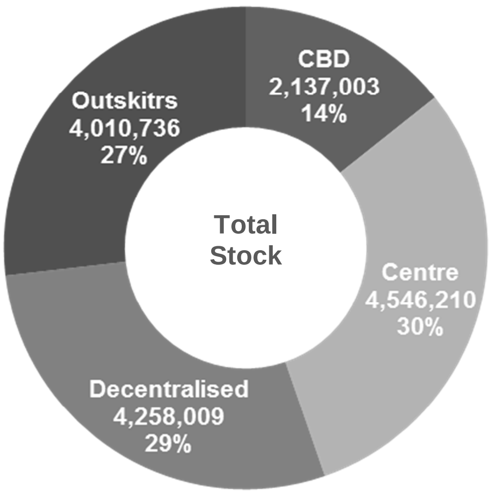


The physical conformation of the office market is outlined by:

- Four sub-markets that are concentric and delimited by, more less, clear physical boundaries comprised by three semi-circular thoroughfares (M-30, M-40 and M-50 motorways).
- The ‘backbone’ of the office activity is the Catellana-Recoletos axis in the centre if the city, conforming the CBD.
- All office buildings outside the CBD but inside the first ring of the M-30 motorway are considered part of the Centre zone
- Office buildings in between the two rings of M-30 and M-40 motorways are considered Decentralized
- Offices located outside the M-40 belong to the Outskirts zone

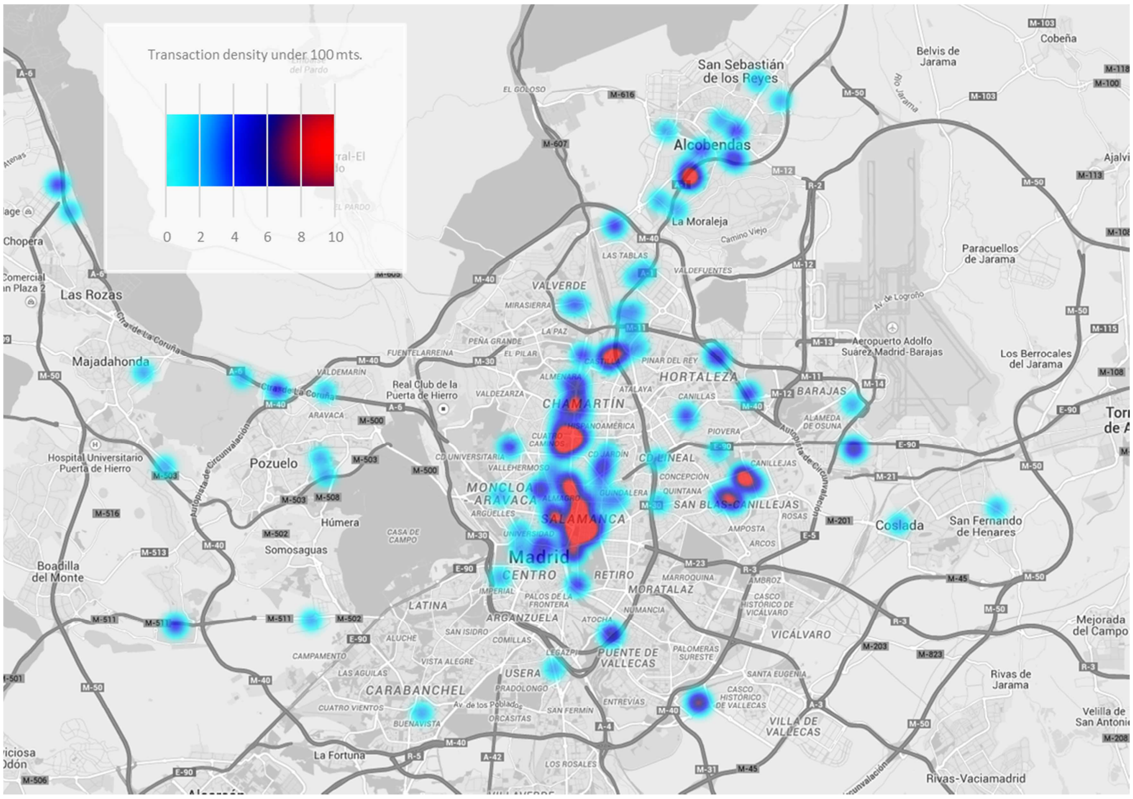
Figure 3.2 gives insight on the composition of the stock. Nonetheless, it is worth to mention the importance of the CBD as a ‘networking hub’ attracting most of the office take-up activity as seen in Figure 3.3.

Figure 3.3. Office stock composition of the office market in Madrid
Based on data of the Spanish Cadastre (units: sqm and percentage)



Although a lesser proportion of the supply is located on the CBD (Castellana-Recoletos axis), much of the transaction activity is done there. Figure 3.4 is a heat map that changes colours from light to dark, as there are more new leases closer to each other.

Figure 3.4. Market intensity of the Madrid office market 2014
Spatially close lease transactions are red coloured. More distant leases are coloured light blue. Source of the base map: Google Maps



One of the main impressions is that a significant proportion of transaction activity is drawn by the CBD. This graphical representation is then a first step towards proving the existence of spatial feedback in the Madrid office market.

3.5. Database

We have two main sources of information regarding our dataset. One is a list of new office lease transactions provided by BNP Paribas Real Estate, with the following characteristics:

- Semi-annual structure
- From 2003:1 to 2014:1
- Number of transactions: 3,912
- Tenant name and business sector
- Specific address of the building
- Surface area or deal size
- Transaction zonification within the four of the business districts
- Reference rent is the headline rent of the contract[28]

The other one is a comprehensive list of office buildings for the Madrid market and its surrounding areas, where transaction activity is present. This database is the source of

[28] Deflated by the GDP deflator at constant prices of 2010

most hedonic characteristics used in the econometric exercise. The source is Spain's Land Registry or Cadastre. The components of this database are the following:

- List of the 16,420 properties containing the total built office stock of the city
- Specific address of the building
- Geographic coordinates of each particular building under the UTM format[29]
- Date of construction and refurbishment, if applicable
- Number of floors
- Total surface area of the building
- Office surface area of the building
- Building technical quality index (0=best; 9=worst)
- Registered surface area for office usage
- Number of landlords inside the property,

among other properties not used in this exercise[30]

The two databases have been carefully cross-referenced using a homogenized address format as link between the two of them.

Complementary variables have been created to improve the list of regressors. One of them was distance from the leased office to the closest metro entrance (called Metro), also calculated with a Euclidean formula. The procedure was calculating the minimum linear distance from a matrix of distances between all the lease transactions and all the metro entrance points. The other calculated dummy was a variable called Corporate which identified all new tenants with international presence. The aim for controlling the nature of the tenant was our view that international corporates tend to pay an extra for their offices in response to their willingness to pay for high quality premises.

The main limitations of our dataset may come from the lack of information on lease incentives such as free rent periods or staggered rent arrangements that separate the headline rent (used in this work) from the effective rent. Also, we lack information on the specifics of the lease contract such as binding term and break options that are hedonic characteristics themselves and their impacts are to linger on the residual structure of the econometric exercise.

Tables 3.1 to 3.3 contain the main descriptive statistics for the variables collected in both sources. We have divided the descriptive tables in three given the different nature of the data. Table 1 comprises the main hedonic variables with data measured in a continuum. Table 3.2 offers information on hedonic dummy variables and Table 3.3 contains the time dummy variables.

[29] Specifically, this format is the UTM projection in the ED-50 system

[30] Including: registry number, underground and above ground surface, Municipality and province of where the building is located

Table 3.2. Descriptive statistics of the main hedonic variables
Variables collected BNPPRE transaction list and the Land Registry of Spain

	Real rent (€ per sqm per month)	Lease surface area (sqm)	Age (years)	Floors	Bld. Surface (sqm)	Technical quality index (0=best; 9=worst)	Closest metro entrance distance (m)
Average	17.7	809.2	24.2	7.3	23,023.7	3.0	844.0
Median	16.6	363.4	16.0	6.0	12,052.0	3.0	336.4
Std. Dev.	6.2	1,608.1	24.8	6.7	32,192.7	1.0	2,036.8
Min	4.0	20	0.0	-5.0	48	0.0	4.6
Max	49.3	30,600.0	295.6	56.0	299,433.0	8.4	14,820.6
Count	3,912	3,912	3,912	3,912	3,912	3,912	3,912

Most of the variables have reasonable standard deviations when compared to the average and relatively close median and average values. Surface measures are the most volatile, such as Lease and Building surface areas. Yet, these variables were left out of the regression analysis as they were not significant. The variable Age presents a possible outlier in its maximum value. Nevertheless its impact is restricted in the regression analysis with the dummy variable Stately, that takes value of 1 when the building's age is over 60 years. Regarding Metro Entrance Distance its maximum value also appears to be extreme, but it is an actual distance from the Madrid's subway network, as office buildings in the surrounding villages of Madrid have no metro coverage.

Table 3.3. Descriptive statistics of dummy variables

	CBD	Centre	Dec	Out	Stately	Exclusive	*Corporate*
Obs. with value = 1	1,070	1,004	877	961	580	2,392	873
Proportion in sample	*27%*	*26%*	*22%*	*25%*	*15%*	*61%*	*22%*
Obs. with value = 0	2,842	2,908	3,035	2,951	3,332	1,520	3,039
Proportion in sample	*73%*	*74%*	*78%*	*75%*	*85%*	*39%*	*78%*
Count	3,912	3,912	3,912	3,912	3,912	3,912	3,912

Some office characteristics were captured with dummy variables. Table 3.3 presents the seven of them. The first four make reference to the business district the leased office is located in. Among them, the sum of observations with unity value is the size of the sample (n=3,912). The higher value of transaction happens to be in the CBD, which at the same time is the smallest of the four districts, both in terms of surface area (Figure 1) and share in stock (Figure 2), giving evidence of the business hub Madrid's CBD actually is. The least proportion of office buildings has the condition of stately and representative asset. Only 15% percent of the sample has this characteristic however, this low share is a matter of scarcity in supply than lack of demand. The dummy Exclusive, which takes value of 1 if the leased office is housed in a building used only for office activities, has a dominant share (61%) implying that the Madrid's office market has a strong skewness on office buildings that have no mixed use, for example with residential, hotel and/or retail sales activities. Lastly, the dummy for Corporate informs whether the tenant is an international large-scale company. The share in the sample is most probably related to the actual composition of the service business sector in Madrid, which has a great share of small and medium enterprises.

Table 3.4. Descriptive statistics of time dummy variables

	Obs. with value = 1	Proportion in sample		Obs. with value = 1	Proportion in sample
H1 2003	136	3.5%	H1 2009	114	2.9%
H2 2003	172	4.4%	H2 2009	136	3.5%
H1 2004	270	6.9%	H1 2010	143	3.7%
H2 2005	207	5.3%	H2 2010	122	3.1%
H1 2005	267	6.8%	H1 2011	115	2.9%
H2 2005	232	5.9%	H2 2011	132	3.4%
H1 2006	250	6.4%	H1 2012	128	3.3%
H2 2006	217	5.5%	H2 2012	126	3.2%
H1 2007	247	6.3%	H1 2013	99	2.5%
H2 2007	223	5.7%	H2 2013	108	2.8%
H1 2008	213	5.4%	H1 2014	103	2.6%
H2 2008	152	3.9%			

The size and share in the sample of each time dummy variable is given in Table 3.4. These figures give an idea of the exposition of the office leasing activity to the business cycle. In the upturn phase, the number of transactions increased steadily from 136 in the first half of 2003 to 247 in H1 2007. In the same relationship, during the downward phase of the cycle, the number of transaction fell to 99 in the first half of 2013 to slightly recover in the next two half year periods.

Table 3.5. Correlation coefficient of the hedonic variables

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Lease surface area	1.0000					
Building surface	0.2213	1.0000				
Quality index	-0.0403	-0.2072	1.0000			
Metro entrance distance	0.0127	0.0599	0.0560	1.0000		
Building floors	-0.0152	0.2957	-0.3219	-0.1826	1.0000	
Building age	-0.1498	-0.2727	0.0076	-0.2049	0.0702	1.0000

Table 3.5 gives a solid ground for econometric estimation as hedonic variables hold low levels of correlation. The highest degree of co-movement is given by the correlation coefficient among the Quality index and Building Floors of around 32% which is still away of indicating slight covariance. Therefore, the variables selected for the regression exercise have no issues of endogeneity and no instruments are needed for them to participate in the modelling exercise. It is worth to mention that lease surface area and building surface did not participate in the final regression because did not produced significant coefficient estimators. Yet, intuition and former research in hedonic literature use similar metrics with negative impact from Lease Surface Area - derived from bargain capacity - and positive impact from building size - derived from recognition, amenities and services of large office buildings - (Caduff, 2013; Costello, 2012; Franklin and Waddell, 2003; Limehouse and McCormick, 2011; Ustaoğlu, 2003).

3.6. Hedonical rent estimation

When trying to find spatial feedback among a set of variables the classical starting point is to find evidence of spatial effects in a particular model. Osland (2010) states that such effects are to be found in the residual structure of an OLS model. In this sense the spatial effects tests are run over the residuals of a well specified OLS model. If the baseline OLS model is not well specified, spatial effects may be intertwined with omitted variables issues (i.e. non normal residuals distribution and inconsistent estimated regression coefficients). Nevertheless, theory is not comprehensive regarding hedonic modelling structure and researchers have to test their hedonic model from different perspectives to avoid as much as possible misspecification issues before proceeding to test for spatial feedback. They also have to have strong reasoning on spatial effects to complement econometrical tests. Even after controlling for missing variables, the correct structure to control for spatial effects has to be carefully selected to avoid false interpretations.

3.7. Empirical Modelling[31]

a. OLS Hedonic Model

The baseline model used here has been formed from seminal hedonic research on the office market such as Torto and Southard with Wheaton William C. (1997) and Wheaton and Torto (1994), and more recent studies on the residential side such as Caduff (2013) and Osland (2013). The initial hedonic model to be estimated is the following:

$$r = \alpha + X\beta + G\gamma + \varepsilon G \qquad \text{(Eq. 3.11)}$$

In which *X* represent the matrix of hedonic characteristics and *T* represents a matrix of time dummies. The particular expression estimated is the following:

$$r = G$$
$$\alpha + \beta_{1G} \quad + \beta G \quad + \beta_{3G} \quad + \beta_{4G} \quad + \beta_{5G} \quad + \beta_{6G} XG \quad IVG + G$$
$$\beta_7 QG \quad IG \quad + \beta_8 MG \quad + \beta_{9G} \quad PG \quad + \gamma \sum G_{003h1G}^{014h1} H \quad + \varepsilon G$$

(Eq. 3.12)

Where r represents the natural logarithm of the real rent (at constant prices of 2010 using the GDP deflator) and the right hand variables are office buildings characteristics. *CBD*, *CENTRE* and *DEC* are dummy variables indicating the business district which the office belongs to. We have excluded the Outskirts district; therefore, our model embeds the price of the Outskirts offices in the intercept. As offices located in the CBD are the most expensive, and prices decrease gradually until the Outskirts area, the district dummies estimators are expected to be positive and respecting that $\beta_1 > \beta G > \beta_3 > \alpha$. *AGE* is a variable that measures the years from the construction to the year of the transaction[32]. Its impact is expected to be negative; yet, older buildings that become classical architectonical pieces tend to be more expensive. This led us to introduce *AGE* to the power of two but resulted non-significant for the specification implemented. Our second approach was creating a dummy variable which gets the value of 1 when the age of the building was over 60 years. This variable was called *STATELY* and has got positive sign as expected. The variable *FLOORS* aims to capture the impact of the height of the building on rents. As the Madrid's office market is split between exclusive office buildings and other of mixed use, the variable *EXCLUSIVE* was introduced to asses if actually office-only buildings have some actual differential with mixed buildings. The variable *QUALITY* is an important one. Each building the Cadastre database has assigned a quality index that ranges from 0 to 9, where 0 is the best technical quality and 9 is the worst. Therefore the expected sign is negative. As mentioned before, we estimated the linear distance to closest metro entrance with the UTM coordinates of the transactions and the UTM coordinates of the metro entrances *(METRO)*, extracted from a GPS' Points of Interest file. The expected sign is therefore negative, as the less irrigation of the office by transportation services the lesser tends to be its rent. The extent of our database was limited to hedonic characteristics; therefore we did not have specific information on the contract structure behind each lease contract such as term, break options, etc. However, we recurred to the database to approximate some contract information. We wanted to test

---

[31] Econometric estimations have been performed with Stata software.

[32] We have replaced the year of construction for the year of complete refurbishment when such information was available in the Land Registry database.

if the commitment of multinational corporates had some effect on the final rent agreed, as the practice points that these companies tend to pay a premium for entering some flagship buildings. In the sense, *CORPORATE* is a dummy variable identifying tenants that were multinational companies with an estimator sign expected to be positive. The last set of regressors corresponded to dummy variables identifying each half year period in the database. The reference period in the regression model was the first of the database (2003 H1). These were expected to yield positive estimators in periods of economic expansion and negative in the crisis periods. At the same time, the values should tend to be more negative in the most severe crisis periods and more positive in the booming periods. We employed constant estimators' values for the full sample period indicating our assumption of constant technology, which makes sense for the service sector in the period 2003-2014. Table 3.6 contains the results of the regression exercise.

Table 3.7. OLS estimation output

	Estimator	p-value		Estimator	p-value		Estimator	p-value
CONSTANT	2.7608	0.0000						
CBD	0.5687	0.0000	H12004	-0.1071	0.0000	H12010	-0.1877	0.0000
CENTRE	0.3760	0.0000	H22004	-0.1246	0.0000	H22010	-0.2523	0.0000
DEC	0.1688	0.0000	H12005	-0.1306	0.0000	H12011	-0.2523	0.0000
AGE	-0.0012	0.0000	H22005	-0.0982	0.0000	H22011	-0.3103	0.0000
FLOORS	0.0019	0.0000	H12006	-0.0688	0.0000	H12012	-0.3408	0.0000
EXCLUSIVE	0.0760	0.0000	H22006	-0.0600	0.0010	H22012	-0.4081	0.0000
QUALITY	-0.0499	0.0000	H12008	0.0606	0.0010	H12013	-0.4471	0.0000
METRO	-0.00001	0.0000	H12009	-0.0915	0.0000	H22013	-0.4417	0.0000
CORPORATE	0.0924	0.0000	H22009	-0.1476	0.0000	H12014	-0.4745	0.0000

Notes: n=3,912; R^2=0.59; Root MSE= 0.227

All variables are highly significant, but some commentaries are to be made on the estimators. The intercept is the highest of the estimators with a value of 2.76, and is the reference to estimate any in or out of the sample office value. The second estimator in importance is the *CBD.* It reports higher lease values in this district than any other district of the city. Its importance is well documented (see for example Wheaton and Torto, 1994; Osland, 2010) and denotes not only a strategical position in geographical terms, but also associated values such as business networking and showcase offices. The estimators for the *CENTRE* and *DEC* variables are also positive, as these areas are more expensive than the Outskirts district, but Decentralized cheaper than Centre. Building characteristics such as number of plants, age, quality and metro distance, give the expected sign.

Regarding the estimators of the time variables it is worth to mention that the following periods have non-significant results and as a result do not appear in Table 5: H2 2003, H1 2007, H2 2007 and H2 2008. This is due to the similarity of values in 2003 and the commented periods. The negative values since 2009 represent the constant discount in real rents in response to the property crisis in Spain. As a matter of fact, the coefficient estimators in recent periods are more negative than those close to 2009. The forecasting capacity of the model is acceptable, with an R^2 of 0.591. More details on performance will be given on the Performance Comparison chapter.

3.8. Evidence of Spatial Effects

Obtaining OLS results is the basic to step for testing spatial feedback. Specifically, the spatial dependence tests asses the null hypothesis of random distribution of a variable across the space, against the alternative of significant association of values also across the Cartesian plane. The Moran's *I* test (Moran, 1948) on the residuals is the common practice in the spatial econometrics literature. It is a global test with which spatial feedback is tested among all observations, but not among regions. The latter case is assessed with local spatial dependence tests such as the Local Moran's *I* and New-*G* tests (Moreno and Vayá Valcarcel, 2002). However, the scope of this paper is already regional therefore the global test is apt for its task. We test the following expression for the Moran's *I*:

$$I = G\frac{NG}{S_{0G}} \times \frac{\sum_{ijG}^{N} W_{ijG} e_i - \bar{e}\ (e_j - \bar{e})}{S_i^{NG} e_i - \bar{e}G^{2G}},\ i \neq jG \qquad \text{(Eq. 3.13)}$$

where N is the sample size, ${}_{0}$ is the sum all elements of the weigh matrix and acts as a standardizing factor; ${}_{ijG}$ is a particular observation of the weight matrix, e_{iG} and e_{jG} are particular residuals from the OLS estimation; $\bar{e}G$ is the average of the residuals (Anselin, 1988). The calculation of the weigh matrix in this research has been a traditional one based on Anselin (1988). In this framework the calculated weight matrix includes the standardized inverse distance of each transaction and considers as 'neighbour' all transactions in the first 15.5 kms to ensure all transactions have at least one neighbour.

The Moran's *I* test has an alternative hypothesis of spatial effects of unspecified kind. This leaves room for testing such kind of spatial error distribution under the alternative hypothesis. As a response the Lagrange Multipliers (LM) are commonly employed. There are two main types of these tests: Those for measuring spatial lag structure and those for spatial error structure. Each of them has a version for global spatial distribution and local spatial distribution. As this paper focuses on an already local market it also focuses on global LM tests. The null hypothesis of both *LM-lag* and *LM-error* tests is the non-existence of spatial effects, which is tested with the expressions:

$$M - errGr = G\frac{\frac{e'WeG}{S^{2G}}}{T_1}, Gi \neq j, G\ {}_{1} = trG\ {}'^{G}\ + G \qquad \text{(Eq. 3.14)}$$

$$M - \ aG = G\frac{\frac{e'WrG}{S^{2G}}}{RJ_{\rho-\beta G}},\ i \neq j,\ \ J_{\rho-\beta G} = \left[\ {}_{1} + \frac{WXBG'M\ WXBG}{S^{2G}}\right]; MG = I - X\ X'X\ ^{-1}X'G$$

(Eq. 3.15)

where r in (14) and (15) refers to the endogenous variable. Both LM tests are asymptotically distributed as $\chi\ \ 1G$. The alternative hypothesis is that actually there exists spatial autocorrelation of the type of the test ran. If both tests yield confirmation of spatial correlation the common practice is to select the specification with highest LM statistic (Florax and De Graaff, 2004). We will not follow this recommendation as our knowledge of the market is that spatial lag is the actual model for the local practice when letting an office. The results of the tests indicate the existence of spatial feedback in the residuals of the OLS estimation (equation 3.6). Table 3.6 presents the results:

Table 3.7. Spatial effects tests

	Statistic	p value
Moran's *I*	60.861	0.0000
LM-error	2696.527	0.0000
LM-lag	863.762	0.0000
Weight matrix: Row-standardized, with a distance band of 15.5 KMS		

It is clear that the tests reject the hypothesis of normal distribution of the residuals across the Cartesian plane giving solid ground to estimate the spatial econometrics model.

3.9. Estimated Spatial Lag Model

The Spatial model used in this study has been the Lag model. It is because in a market with asymmetries in information, such as the office market, agents rely heavily on comparable transactions to assess the rent of the office of their interest, making them spatially correlated. In other words, this approach makes more sense than the spatial-error or the Durbin Models from a market practice perspective. The model estimated is the following:

$$r = \rho G\ r + \alpha I_n + \beta_{1G} \quad + \beta G \quad + \beta_{3G} \quad + \beta_{4G} \quad + \beta_{5G} \quad + G$$
$$\beta_{6G} \quad + \beta_7\ XG \quad IV \quad + \beta_8 QG \quad IG \quad + \beta_{9G} \quad PG \quad + \quad \gamma + \varepsilon G$$

(Eq. 3.16)

Equation (3.16) models the natural logarithm of rents as a function of asset of variables where the business districts of the city are *CBD*, *CENTRE* and *DEC*, in the form of dummy variables. The variable *AGE* is the number of years since the constructions of the building and the date of the lease transaction. *STATELY* is a dummy variable taking value of one if the building has more than 60 years of construction. *FLOORS* is the number of stories of the building; *EXCLUSIVE* is a dummy variable that takes value of one if the transacted office is located in a building with 100% office usage. *QUALITY* is a technical quality index, which gives value of cero to the best buildings and 9 to the worst. Finally, *CORPORATE* is a dummy variable that take value of one if the tenant is a multinational corporate. ρ is the key parameter to estimate in the Spatial Lag model. Its significance proves the spatial feedback is well captured by this particular model. On the other hand, its value measures the strength of such spatial feedback. W is the weight matrix and is the same used for the Moran's ***I*** test.

When modelling a variable with its own spatial lag, OLS return skewed and inconsistent estimators. This holds true even if there is no spatial correlation in the error term (Moreno and Vayá Valcarcel, 2004). Therefore, in spatial econometrics the classical approach is the estimation by maximum likelihood (ML) and it is the one used in this paper. Nevertheless, the estimation method in spatial econometrics is not limited to the ML but other methods. These include instrumental variables (Anselin, 1995) and the generalized method of moments (Kelejian and Prucha, 1999). Table 3.8 shows the estimated coefficients of (Eq. 3.16) by ML:

Table 3.8. Spatial Lag Model Estimation Output

	Estimator	p-value		Estimator	p-value		Estimator	p-value
Constant	0.4134	0.0000	H22003	-0.0307	0.1070	H12010	-0.2077	0.0000
$\hat{\rho}$	0.8863	0.0000	H12004	-0.1326	0.0000	H22010	-0.2625	0.0000
CBD	0.2751	0.0000	H22004	-0.1551	0.0000	H12011	-0.2667	0.0000
CENTRE	0.1243	0.0000	H12005	-0.1564	0.0000	H22011	-0.3291	0.0000
DEC	0.0268	0.0330	H22005	-0.1217	0.0000	H12012	-0.3662	0.0000
AGE	-0.0017	0.0000	H12006	-0.0916	0.0000	H22012	-0.4259	0.0000
STATELY	0.0273	0.0400	H22006	-0.0803	0.0000	H12013	-0.4661	0.0000
FLOORS	0.0026	0.0000	H12007	-0.0474	0.0050	H22013	-0.4681	0.0000
EXCLUSIVE	0.0804	0.0000	H12008	0.0448	0.0110	H12014	-0.4846	0.0000
QUALITY	-0.0464	0.0000	H12009	-0.1090	0.0000			
CORPORATE	0.0877	0.0000	H22009	-0.1688	0.0000			

Notes: n=3,912; Squared correlation: 0.58; Root MSE=0.215

As in the OLS estimation, all regressors are highly significant and with the expected sign. It is worthwhile to comment particularities:

- The values of the estimated coefficients for the intercept and the office districts have significantly reduced with respect to the OLS exercise, but proper-building characteristics estimators keep similar values. The reason is that the estimation of the spatial effects coefficient (ρ) is pulling spatial effects out of variables subject to spatial feedback. Therefore, the number of floors or the age of the building do not impact prices because of the location, but to their implicit value. However, being in a different district actually affects rents due to localization; with the spatial regression that effect is translated onto the spatial-feedback strength coefficient (ρ). In the spatial model the effect of the office districts on rents is a 'skimmed' version of the OLS version, isolating spatial effects and keeping other effects such as prestige, ease of networking or access to business clusters in the marginal effect on rent.

- The variable *METRO*, which is the distance in linear meters to the closest metro entrance, does not appear in the spatial regression due to its loss of significance.

- The variable *STATELY*, is present in the spatial model, but not in the OLS approach because in the latter, it was not significant. As a matter of fact, the sign of its estimated coefficient is positive, confirming that there is a lease premium to pay for classical office buildings, as expected.

- Estimators of H2 2007 and H2 2008 resulted non-significant as they have similar values as those of H1 2003.

- The squared correlation[33] is slightly less than the R^2 of the OLS regression (0.58<0.59) but both values signal similar and relatively high explanatory capacity of both models[34].

- The estimated factor of intensity of spatial effects (ρ) was also highly significant and reports a high degree of spatial feedback as its value is close to its upper boundary (0.96)[35].

3.10. Marginal effects

OLS hedonical exercises are a decomposition of the endogenous variable among its components and an error term. As a result, the interpretation of the estimated coefficients is straightforward. In the present exercise the estimated coefficients are the semi elasticities of the office rents. Yet, in the spatial regression the marginal effects are given by matrixes of equations (3) and (4) and not by the simple estimated coefficients (Kim et al., 2003; Mobley et al., 2008). In the particular case of this study each marginal effects matrix has a dimension of 3,912 x 3,912 being 3,912 the size of the sample. This implies that for each of the 29 regressors of equation (16) we obtain a matrix of around 15.3 million elements. For the sake of simplicity we describe only the marginal effects matrix of the CBD regressor36, as this paper is more focused on the out-of-the-sample rent estimation. When calculating (Eq. 3.16), we are also able to compute:

$$\sum r\ \ w\ \ i\ \ |G\frac{\delta rG}{\delta CBDG} = 2{,}41\ G \qquad \text{(Eq. 3.17)}$$

which can be interpreted as the increase in the logarithm of the rent derived from leasing an office in the CBD and not in the Outskirts district, given the particular rent level of the office of row i and the indirect marginal effects (of column j ≠i) across the Cartesian plane. Therefore to obtain a normalized impact and not a particular one for office i, we proceed to divide (Eq. 3.17) by the average of log(R) = ($\bar{r}G$:

$$\frac{\sum row\ \ i\ \ |G\frac{\delta rG}{\delta CBDG}}{\bar{r}G} = G\frac{.410G}{.809G} = 0.858G \qquad \text{(Eq. 3.18)}$$

$$\exp\ \ 0.858G = 2.358G$$

which may be conceived as the marginal impact of the regressor CBD on the average rent. In other words, hiring an office located in the CBD adds €2.36 sqm/month to the

---

[33] No R-squared produced by ML estimation

[34] There can be no direct comparison between R^2 and squared correlation in the ML procedure, as their results are not identical (Spanos, 1989)

[35] The 95% confidence interval for Rho is 0.808 < rho < 0.964

[36] The results of the other marginal effects matrixes are available upon request to the corresponding author

office rent of one located in the Outskirts zone when rent is estimated by the spatial approach, all other factors being equal. The OLS regression same marginal impact is estimated in €1.76 sqm/month euros[37].

Nevertheless, the marginal impact may also be derived from a simpler calculation, applying the fact that each row of the spatial weigh matrix has a sum of 1/(1-ρ) (Kim et al., 2003). This term is known as the Global Spatial Multiplier (in this case with value of 8.795) and under the light of (3) the marginal effects of the spatial specification may be calculated by means of:

$$\frac{\delta rG}{\delta X_j} = \left({}^{1}/_{1-\rho}\right) \beta_{jG} \qquad \text{(Eq. 3.19)}$$

Table 3.9 shows the estimated results of (Eq. 3.19) for all the hedonic regressors normalized by the average rent in comparison with the OLS estimators:

Table 3.9. Marginal effects comparative results of OLS and Spatial regression

Regressor	$\boldsymbol{\beta_{OLSG}}$	$\boldsymbol{\beta_{SPATIALG}}$	Global effect $\boldsymbol{\beta_{SPATIALG}} \times \boldsymbol{Global\ multiplier}$	OLS monetary impact (€sqm/month) $\boldsymbol{e^{\beta_{OLS}}}$	Spatial monetary impact (€sqm/month) $\boldsymbol{e^{\left(\frac{Global\ efect}{\log\ Avrg.\ rent}\right)}}$
Constant	2.7595	0.4134	3.6359	15.7919	3.6487
CBD	0.5658	0.2751	2.4195	1.7609	2.3663
CENTRE	0.3753	0.1243	1.0932	1.4554	1.4758
DEC	0.169	0.0268	0.2357	1.1841	1.0875
AGE	-0.0012	-0.0017	-0.0150	0.9988	0.9947
STATELY	NA	0.0273	0.2401	NA	1.0892
FLOORS	0.0021	0.0026	0.0229	1.0021	1.0082
METRO	-0.00001	NA	NA	-0.9999	NA
EXCLUSIVE	0.0752	0.0804	0.7071	1.0781	1.2863
QUALITY	-0.05	-0.0464	-0.4081	0.9512	0.8648
CORPORATE	0.0939	0.0877	0.7713	1.0984	1.3160

Global multiplier [1/(1-ρ)]: 8.7950; Log average real rent: 2.809

The results of Table 8 give a straightforward comparison of the monetary impacts of the OLS and spatial exercises. In general the monetary breakdown of letting rents among hedonic characteristics is similar in the two econometric approaches.

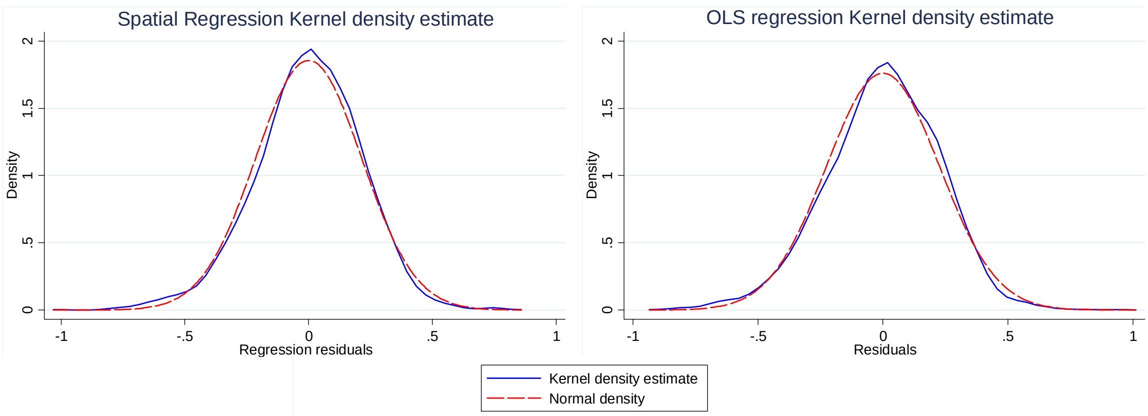
[37] Bear in mind that the value of the constant for the OLS regression is 2.76 which translated onto impact in euros is exp (constant_OLS) = €15.81, while in the spatial lag regression that same parameter is exp (constant_SPATIAL) = €1.59 so the final impact is much higher in the OLS approach than in the spatial approach.

Nevertheless, it is necessary to comment some differences. The most important is the intercept. In the OLS approach this parameter is more than four times the size of the spatial-lag regression. As the rest of the hedonical regressors have closer values that differ at the most in 25%, is its clear that the final hedonical estimation will be biased upwards in the OLS approach as documented by Mobley et al. (2008) and Anselin (2003), because the spatial feedback among rent levels is not accounted for in the least squared regression, producing biased estimates of marginal effects and misrepresentative standard errors. In other words, OLS presents specification issues as letting rent determinants such as autocorrelation of the explanatory variable and spatial spill over effects are ignored. The spatial lag approach actually corrects such defects giving unbiased marginal and global effects of hedonic characteristics on rent levels.

3.11. Performance comparison between OLS and Spatial models

The results of the models are now compared by means of the distribution of their residuals. As a starting point we comment that the classical tests of normality distribution of the residuals such as the Jarque-Bera and Shapiro-Wilk tests report non-normally distributed residuals (see annex of chapter 3). Nevertheless, these tests do not perform well under large sample sizes (more than a few hundred observations). In those cases the null hypothesis of normal distribution is systematically rejected as slight variations of the empirical distributions are reported as non-normality. Rather, we used the Kernel Density Estimation, using the Epanechnikov kernel (Epanechnikov, 1969). This approach is a non-parametric specification by which we measure the 'distance' between the residuals' distribution to the normal distribution. Such distance is actually the size a of tension factor, or bandwidth, in the function that minimizes the average squared residuals of the current distribution from the Gaussian distribution (Silverman, 1986). As seen in Figure 3.5, the distribution of the residuals of both OLS and spatial regressions are quite well distributed over the zero average.

Figure 3.5. Kernel Density Estimation



A simple graphical comparison indicates that the left-hand chart (spatial regression) has a more smoothed distribution than that of the OLS regression. This is confirmed by the commented tension factor, as presented in Table 3.10.

Table 3.11. Bandwidth used in the kernel density estimation

	Spatial regression residuals	OLS regression residuals
Bandwidth	0.035	0.038
Kernel used: Epanechnikov		

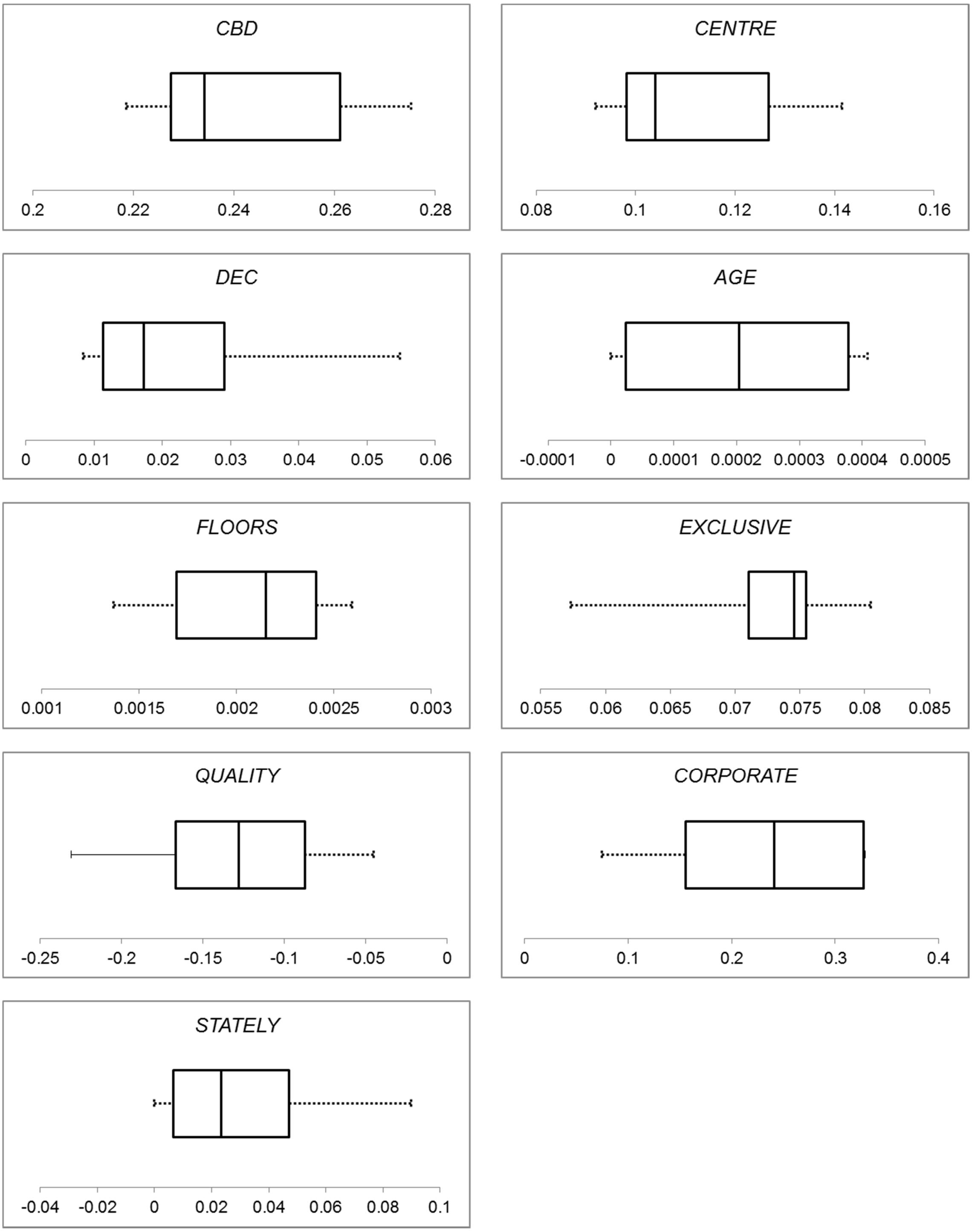
In kernel density estimation, getting higher smoothness, and therefore closer empirical distribution to the Gaussian distribution, one needs to use a higher bandwidth. Consequently, as the bandwidth of the spatial regression is lesser than the OLS regression it is possible to assert that the residuals of the spatial approach are closer to the Gaussian distribution than those of the OLS regression.

Another way to compare performance of the two approaches is using a goodness of fit measure such as the root mean squared error (RMSE). For the case of spatial regression it is obtained an RMSE of 0.215 while the forecast of the OLS regression yields an RMSE of 0.226. Once more, the spatial approach outperforms the classical approach due to its higher accuracy in the inner sample forecast.

3.12. Stability tests of the spatial regression

Returning to the spatial model, we test the stability of the regressors using different sample sizes. The purpose is to check if there is a major change in the magnitude of the regressors in different points in time. We proceeded to graphically analyse from a chart the homogeneity of the main hedonic regressors of the estimation to see if they drastically changed when restricting the data to different periods (sample sizes). Figure 6 presents the results for this assessment, restricting the sample by half year periods since H12014 to H12007, which is equivalent of 15 different sample sizes (See annex 6).

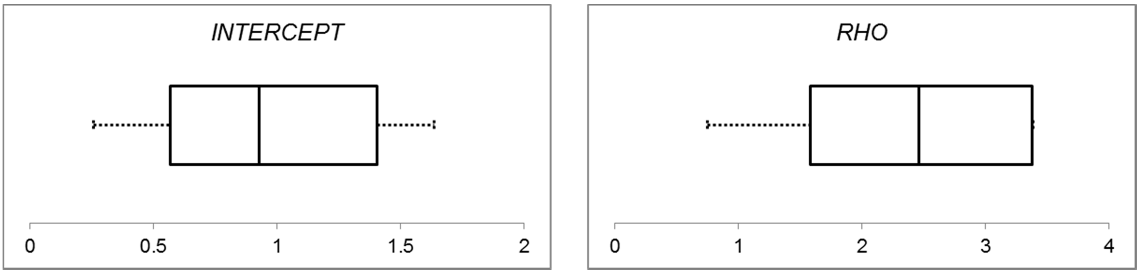
Figure 3.6. Box plots of the estimated coefficients of the spatial regression



The charts in Figure 6 present the distribution of the estimated coefficients of the hedonic variables used in the spatial model, as well as the intercept and the spatial feedback strength coefficient (ρ). The general trend is most of the coefficients' maximum and minimum values lay inside 1.5 times the interquartile range (span of percentiles 25^{th} and 75^{th}) and therefore the estimators are stable regardless of the span of the data. Nonetheless, it is worth to mention a possible outlier for the coefficient of the *EXCLUSIVE* dummy variable that identifies when an office is leased in a building used only for offices and not mixed used with residential, retail or industrial activities. A possible extreme value is registered when the regression is restricted to 2003 H1 to

2007 H2 data. However, with those sample sizes, both estimators render marginally insignificant at 1% of significance with p-values of 0.77 and 0.045.

Figure 3.7. Box plots of the estimated constant parameters



The parameters corresponding to the Intercept and the spatial strength coefficient are also quite stable with no extreme values when restricting the sample by half year periods between 2007 H1 and (full sample) 2014 H1 (Figure 3.8).

Figure 3.8. Size of hedonic regressors across time in the spatial regression

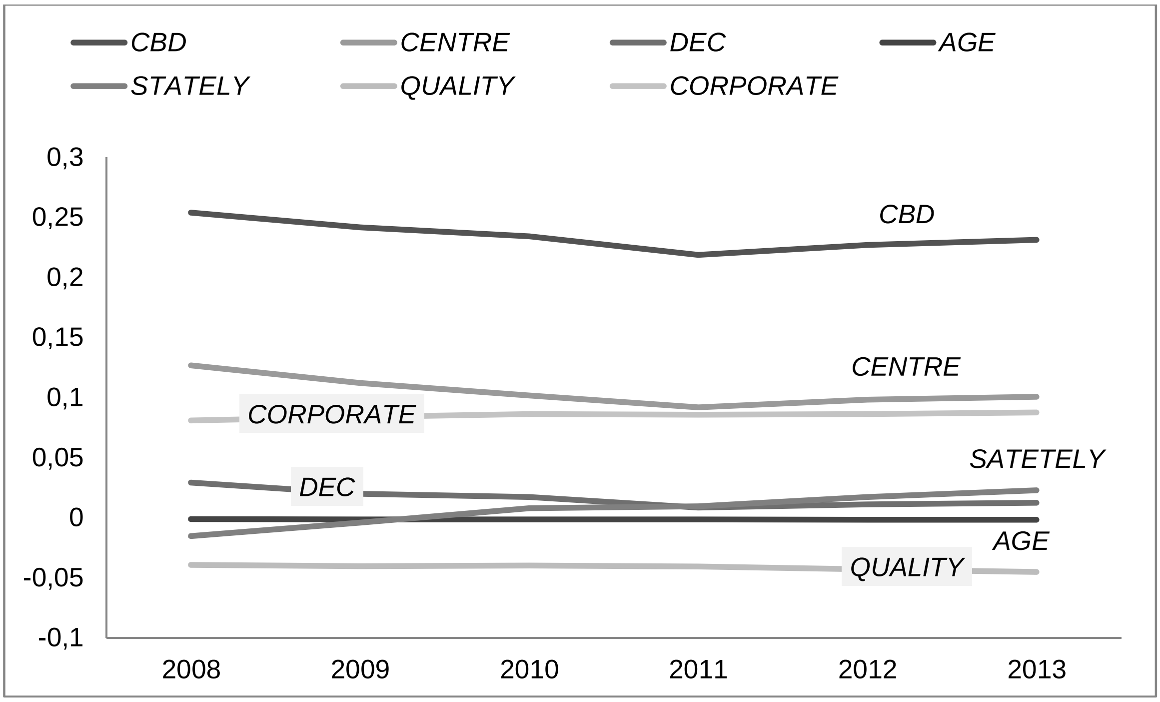


Figure 3.8 shows the evolution over time of the size of hedonic regressors in the spatial regression. The main characteristic of this figure is the relative flatness of the lines drawn which gives a clear idea of the homogeneous impact of the regressors when restricting the sample size of the spatial regression. In other words the rent levels decomposition for the Madrid office market is robust among its hedonical components regardless of the sample taken. The less stable seems to be that for Metro Distance (which captures the marginal impact of the linear distance of the closest metro entrance on the office lease price). As the sample is less restricted *METRO* reduces its value to converge to zero since H1 2012 to the end of the sample in H1 2014. That is the reason it did not participate in the final spatial regression (H1 2003-H1 2014).

3.13. The archetype office and the out of the sample estimation

We focus again on one of the main targets of this study. This is the rent estimation of the average office in Madrid. Consequently, the main subject of this chapter is to define such an average office. Bearing in mind the database and estimated models structure, we will estimate four archetypical offices each of them belonging to one of the four business districts presented in Figure 1. The two models (spatial and OLS) are used to estimate the rent of such four typical offices, but with each of them with the average hedonic characteristics of at the intra-district level. Table 3.11 presents the characteristics used for this purpose.

Table 3.11. Hedonic characteristics to be used in the out of the sample estimation

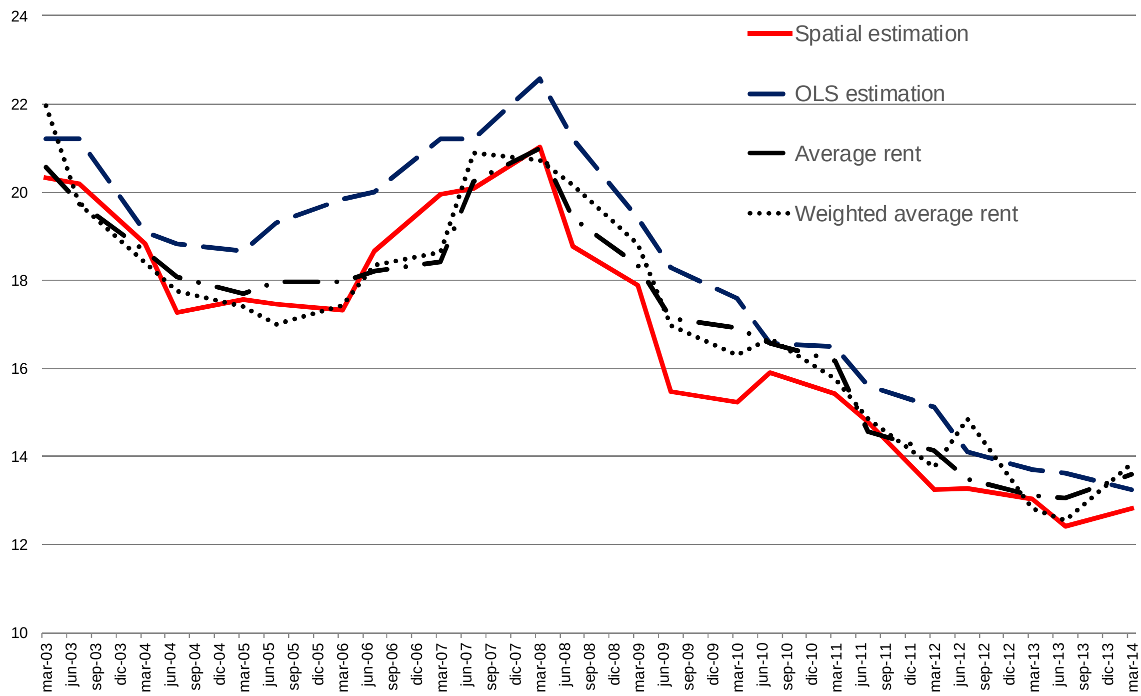
	X Coord	Y Coord	*AGE* (Years)	*STATELY*	*FLOORS*	*EXCLUSIVE*	*QUALITY*	METRO (meters)	*CORPORATE*
CBD	441,666.68	4,476,074.24	41	1	12.9	1	2.7	183	1
CENTRE	442,425.28	4,477,113.92	34	1	7.3	1	2.6	186	1
DEC	446,734.16	4,475,984.32	16	1	6.6	1	3.2	562	1
OUT	450,738.92	4,477,704.32	19	1	2.4	1	2.9	1,540	1

For the variables *STATELY, EXCLUSIVE* and *CORPORATE*, the number 1 means that this hedonic characteristic is selected to estimate the office's rent level.

The values of the hedonic characteristics have been selected as of the first half of 2014 for the variables *AGE, FLOORS, QUALITY* and *METRO*. The assumption is over an office located in an exclusive and stately office building, with corporate tenant.

After evaluating the OLS and spatial equations specified in (Eq. 3.12) and (Eq. 3.16) with the hedonic characteristics of Table 3.11 combined with each time dummy, we build-up the time series for each office district. We take the average of the four district's rent series to obtain the representative rent of the Madrid's office market. Figure 3.9 offers a graphical representation of the estimated results.

Figure 3.9. Estimation of the typical office rent (€/sqm/month)



Both econometrically estimated series share trend and dates of the maximum values. Minimum values are far-off just one period among themselves. At the same time, the OLS estimation is smoother than the Spatial. Nevertheless taking as a reference the average rent and the deal-size weighted average[38] to assess the relative distance of the OLS and spatial estimation, the over estimation of the rent is obvious for the OLS case, as commented in Section 3.10. On the other hand, the spatial estimation shares levels with both types of averages, except for the crisis period, between the second half of 2008 and the first half of 2013. This apparent downward bias has an explanation, and actually is a remarkable property of the spatial model. Thanks to the crisis period new leases rents constantly adjusted. However, given the existence of contracts, companies had to wait until the extinction of the contract or activation of break option to leave their office without paying compensations to the landlord, and have access to lesser rent levels. This most probably made business to encompass such relatively higher rent levels to their cost structure during the crisis. Reaching a break date for the contract most of the companies decided to maintain such cost structure but to move to better locations. Therefore, companies took advantage of now affordable rents in better locations, increasing the share of relatively more expensive offices in the new-leases-average rent calculation, dragging it upwards. This is what we call a composition effect, when estimating market rents. Therefore, for the Spanish crisis period it is reasonable to see lesser level of rents once one controls for this composition effect. Actually this is what the econometric estimation does when estimating out of the sample rent levels for each of the four office districts. As commented, in this paper we estimate rents for an archetypical office for each of the four office districts of Madrid and with those four estimations we proceed to calculate the average. This, apart of using the full power of the sample and controlling for hedonic characteristics, completely isolates the composition effect. It happens that

[38] Deal size refers to the amount of squared meters hired with the new lease contract

when also controlling for spatial feedback, the estimated rent ends up below the OLS estimated rent. The average OLS-rent is 8% higher than the Spatial-rent.

Finally, notice that the turning point of the market since H1 2013 is well captured by all the methods, except for the OLS one. This is other evidence of the better suitability of the spatial approach, as it is not as stiff as the OLS benchmark.

3.14. Concluding remarks

1. The general purpose of this paper has been to contribute to the empirical literature on office rent modelling. After reviewing the literature, there is no lack of discussion in terms of hedonic estimation of real estate prices. Yet, research is not extensive in terms of spatial econometrics and most of it is related to housing prices. To the best of our knowledge, spatial-hedonical references for the commercial property markets are quite restricted. So, this paper is a bold initiative both for the commercial real estate and the Spanish office market.

2. An OLS hedonical estimation was selected as a benchmark against which we compare the spatial lag approach. This OLS benchmark model was selected with a GETS methodology, as we had a large set of hedonic characteristics that acted as candidate regressors to explain office rents. The spatial model was also selected by GETS obtaining a similar model to the OLS but not an exact one. Although there are more techniques to control for spatial feedback such as the spatial error model or the Durbin model, we maintained the spatial lag approach as the most adjusted to the actual market practice of using comparable transactions to obtain references when negotiating rents in lease contracts.

3. The results of both approaches were correct in terms of expected sign and both models perform well regarding explanatory capacity (R^2), goodness of fit (RMSE) and error distribution. Though, on the relative performance field, we identified better results in the case of the spatial model. This improvement is not unexpected, as controlling for spatial feedback, *ceteris paribus*, improves performance in the commented metrics. Therefore estimations are to be interpreted as more accurate than in the OLS approach and suggest a significant role of location in this market.

4. Both the OLS and spatial methodologies are well adjusted to solve sample composition effects. This means that when average rent calculation have skewness towards expensive zones due to large numbers of transactions, estimating four hedonical rents for each of the four business districts of the city solves the issue without any special weight methodology to be developed.

5. Also, the common issue with deal size weighted average is solved, limiting the effect of dragging the average towards the rent of large letting deals.

6. Market insights can be obtained from the estimation of a measure of spatial feedback strength (ρ). The spatial lag regression estimates such factor and its size is indicator of the intensity of spatial effects. In the case of this study, the estimated upper bound to rho is the unity and the lower bound is -2. As its estimated value is 0.8863 we can assert that the spatial feedback for the Madrid office market is strong, and such relationship among rent levels has to be considered whenever possible.

7. Spatial estimated coefficients interpretation differs substantially from the OLS regression. In the latter case, the marginal effects are straightforward associated with the estimated coefficients. In the spatial exercise the marginal effects are stored in an effects matrix which elements are spatially weighted. This means that interpretation of its elements is not simple therefore is to be made with caution. In this exercise we estimated the marginal impact of the hedonic variables as the sum of row-*i* of the marginal impact matrix (all rows have the same sum value and are associated to the particular rent level of the deal) and divided by the average log-rent to get a normalized marginal effect and comparable to that of the OLS regression.

8. The global effect value [1/(1-ρ)] is the translator from spatial coefficients to marginal impacts, as it is the factor that takes into account all spill over effects contained in the spatial weight matrix. In the case of this paper, the marginal effects of the OLS regression (coefficients) are similar to the global effects except for the intercept. The OLS intercept has a value four times that of the spatial lag model. Therefore, the endogenous variable estimations are biased upwards.

9. The estimated spatial coefficients are quite stable when changing the sample size. We restricted the sample size by forming subsets of observations of each half year since H1 2008 to H1 2014. The stability test produces quite centred values of the hedonic estimators meaning that rent estimations and forecasts remain valid with different sub-samples thanks to homogeneity in the data underlying the estimations.

10. The spatial lag estimation gives a better idea of the trend of the office market rents in Madrid than the OLS estimate, the simple average and the weighted average. This can be stated thanks to the soundness of the estimation when compared to the OLS exercise, the non-bias of the estimators, their homogeneity and the more accurate value of the out-of-the-sample estimated rent. This increased accuracy is based in the nonexistence of systematic positive deviations as presented by the OLS estimate. Additionally, the results clearly correct the concentration of transactions in the crisis period (2008-2013) around the central areas that causes inflated rent averages thanks to a pulling-up effect of composition of the sample rather than a market phenomenon derived from scarcity. In this period the spatial estimation is the lower among OLS estimated, average rent and weighted average rent, but keeps similar levels to the averages in the rest of the periods.

11. The use of spatial lag estimation taking account submarkets has allowed us to detect valuable insights on the office market. In view of the encouraging results of the present study, some optimism about the benefits from implementing this analysis seems justified.

References

Anselin, L., 1988. A test for spatial autocorrelation in seemingly unrelated regressions. Economics Letters 28, 335–341.

Anselin, L., 1995. Local Indicators of Spatial Association—LISA. Geographical Analysis 27, 93–115. doi:10.1111/j.1538-4632.1995.tb00338.x

Anselin, L., 2003. Spatial externalities, spatial multipliers, and spatial econometrics. International Regional Science Review 26, 153–166.

Anselin, L., 2013. Spatial econometrics: Methods and models. Springer Science & Business Media.

Bivand, R.S., 1984. Regression modeling with spatial dependence: an application of some class selection and estimation methods. Geographical Analysis 16, 25–37.

Caduff, L., 2013. Spatial hedonic regression modelling of commercial and office space prices in Singapore. Singapore ETH Centre, ETH Zurich, Singapore, Zurich.

Can, A., 1990. The measurement of neighborhood dynamics in urban house prices. Economic Geography 66, 254–272.

Case, K. E., Shiller, R., 1987. Prices of single-family homes since 1970: New indexesfor four cities.New England Economic Review Sept/Oct: 45-56.

Chasco C and B Sánchez (2015) Valuation of environmental pollution in the city of Madrid: An application with hedonic models and spatial quantile regression, Journal of Regional and Urban Economics/Revue d'Économie Régionale & Urbaine (RERU) 2-3/2015, 343-370

Clapp, J. M., 1980. The intrametropolitan location of office activities. Journal of Regional Science 20, 387–399

Costello, G., 2012. Building age, depreciation and real option value – an Australian case study, in Proceedings of the 18th Annual Pacific-Rim Real Estate Society Conference.

Epanechnikov, V., 1969. Non-parametric estimation of a multivariate probability density. Theory of Probability and Its Applications 14, 153–158. doi:10.1137/1114019

Fernandez-Aviles, G., Minguez, R., Montero, J.-M., 2012. Geostatistical air pollution indexes in spatial hedonic models: The case of Madrid, Spain. Journal of Real Estate Research 34, 243–274.

Florax, R.J., De Graaff, T., 2004. The performance of diagnostic tests for spatial dependence in linear regression models: a meta-analysis of simulation studies, in: Advances in Spatial Econometrics. Springer, pp. 29–65.

Florax, R.J., Folmer, H., Rey, S.J., 2003. Specification searches in spatial econometrics: the relevance of Hendry's methodology. Regional Science and Urban Economics 33, 557–579.

Franklin, J.P., Waddell, P., 2003. A hedonic regression of home prices in King County, Washington, using activity-specific accessibility measures, in: Proceedings of the Transportation Research Board 82nd Annual Meeting, Washington, DC.

Gao, A., Wang, G., 2007. Multiple Transactions Model: A Panel Data Approach to Estimate Housing Market Indices. Journal of Real Estate Research 29, 241–266. doi:10.5555/rees.29.3.uuj80773h7274360

Hansen, J., 2009. Australian House Prices: A comparison of hedonic and repeat-sales measures (SSRN Scholarly Paper No. ID 1403578). Social Science Research Network, Rochester, NY.

Hill, R.J., Melser, D. 2008. Hedonic imputation and the price index problem: An application to housing. Economic Enquiry 46, 593-609.

Kain, J.F., Quigley, J.M., 1970. Measuring the Value of Housing Quality. Journal of the American Statistical Association 65, 532–548. doi:http://amstat.tandfonline.com/loi/uasa20

Kelejian, H.H., Prucha, I.R., 1999. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. International Economic Review 40, 509–533.

Kim, C.W., Phipps, T.T., Anselin, L., 2003. Measuring the benefits of air quality improvement: a spatial hedonic approach. Journal of Environmental Economics and Management 45, 24–39.

Kuminoff, N.V., Parmeter, C.F., Pope, J.C., 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? Journal of Environmental Economics and Management 60, 145–160. doi:10.1016/j.jeem.2010.06.001

LeSage, J.P., Pace, R.K., 2009. Introduction to Spatial Econometrics. Chapman & Hall/CRC.

Limehouse, F., McCormick, R.E., 2011. Impacts of Central Business District location: A hedonic analysis of legal service establishments. US Census Bureau Center for Economic Studies Working Paper No. CES 11–21.

Mobley, L.R., Frech, T.E., Anselin, L., 2008. Spatial interaction, spatial multipliers, and hospital competition. Working Paper 07-08. Economics Department, University of California at Santa Barbara

Moran, P.A.P., 1948. The interpretation of statistical maps. Journal of the Royal Statistical Society. Series B (Methodological) 10, 243–251.

Moreno, R., Vayá Valcarcel, E., 2002. Space Econometry: new techniques for the regional analysis. An application to the European regions. Investigaciones Regionales - Journal of Regional Research 1, 83–106.

Osland, L., 2010. An application of spatial econometrics in relation to hedonic house price modeling. Journal of Real Estate Research 32, 289–320. doi:10.5555/rees.32.3.d4713v80614728x1

Osland, L. 2013. The importance of unobserved attributes in hedonic house price models. International Journal of Housing Markets and Analysis 6, 63–78. doi:10.1108/17538271311306020

Quigley, J.M., 1995. A simple hybrid model for estimating real estate price indexes. Journal of Housing Economics 4, 1–12. doi:10.1006/jhec.1995.1001

Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. Journal of Political Economy 82: 34-55.

Spanos, A., 1989. Statistical foundations of econometric modelling. Cambridge University Press

Silverman, B.W., 1986. Density estimation for statistics and data analysis. CRC press.

Straszheim, M.R., 1974. Hedonic estimation of housing market prices: A further comment. Review of Economics and Statistics 56, 404–406. doi:http://www.mitpressjournals.org/loi/rest

Ustaoğlu, E., 2003. Hedonic price analysis of office rents: A case study of the office market in Ankara. Middle East Technical University.

Ward, M. C., Gleditsch, K. S., 2007. An introduction to Spatial Regression Models in the Social Sciences. Available at http://dces.wisc.edu/wp-content/uploads/sites/30/2013/08/W4_W7_WardGleditsch.pdf

Wheaton, W.C., Torto, R.G., 1994. Office rent indices and their behavior over time. Journal of Urban Economics 35, 121–139. doi:http://www.sciencedirect.com/science/journal/00941190

Wheaton William C., Torto, R.G., Southard, J.A., 1997. The CB Commercial/Torto Wheaton Database. Journal of Real Estate Literature 5, 59–66. doi:http://ares.metapress.com/content/121542/

SUMMARY, CONCLUSIONS AND FURTHER RESEARCH

Summary

The Spanish property market is an interesting case study due to the collapse after the overshooting from a long-term price increase of Spanish real estate prices. Indeed, house prices in Spain showed one of the biggest cumulative growth rates among the OECD during the 1990s, being supported by rapid economic expansion, strong employment growth, an immigration boom, and low real interest rates. With the abrupt drying up of funding since mid-2007, these factors have eroded quickly.

This thesis has attempted to provide a re-assessment of the evolution of the residential property market in Spain by exploring the role played by various factors affecting home prices such as population growth, demand and supply of land, investors' expectations, general economic conditions, cultural factors and economic fundamentals and by applying several complementary econometric techniques.

Hicks (1981, p.232) contented that economic models are rays of light, which illuminating a part of a whole, living the rest in dark. Therefore, it is sensible to have different economic models and econometric techniques to analyse a given topic, so that their conclusions can be compared and further insights can be gained. Un this thesis we have adopted an eclectic approach and have employed data-based methods for establishing the relevant determinants of the price set in the Spanish commercial and residential real estate market.

Main findings

In Chapter 1, we developed a structural model for residential prices in Spain based on a cointegrating and error correction mechanism framework and both a broad theoretical approach to select the fundamental variables that govern house price dynamics and an combination of different specification econometric techniques. A long-run relationship was obtained and an appropriated error-correction model for the short-run dynamics was also found, informing us on how any deviation from the long-run equilibrium is feed-back in order to force the movement towards the long-run equilibrium.

Chapter 2 examined Madrid's office market using with a system of equations for stock variation, vacancy rate variation and rental prices (average real rent) variation, within an error correction mechanism framework. This framework allowed us to capture long term development paths and, therefore, analyse short term deviations from the long term track.

In Chapter 3, we contribute to the empirical literature on office rent modelling by providing an OLS hedonical model and comparing its results with those obtaining from the spatial lag approach, concluding that the spatial lag estimation gives a better idea of the trend of the office market rents in Madrid than the OLS estimate, the simple average and the weighted average.

Future lines of research

There are a number of directions that extensions from the present research might take. Six avenues that seem worthy of further research are:

i) The modelling strategy proposed in Chapter 1 could be used to analyse the determinant of commercial property prices in other Spanish (e. g., Barcelona) and European (e. g., London and Paris) cities to detect similarities and discrepancies and to relate them to the institutional and/or the legal framework where these markets operate.

ii) The use of the RETINA automatic predictive modelling (see Pérez-Amaral et al, 2004) could extend the analysis made in Chapter 1. RETINA is designed to embody flexibility (using nonlinear transformations of the predictors of interest), selective search within the range of possible models, control of collinearity, out-of-sample forecasting ability, and computational simplicity. We can the performance of RETINA with both GASIC and PcGets, a well-known automatic modelling method proposed by Hendry and Krolzig (2004).

iii) The auto selection technique involved in GASIC, RETINA and PcGETS may be easily improved by assisting the algorithm with 'guided' process, in which the researcher has previously grouped variables by similarities in nature, for example, by theoretical proximity.

iv) The use of MIxed DAta Sampling (MIDAS) models could also bear fruit in modelling residential prices. These models provide parsimonious specifications based on distributed lag polynomials, which flexibly deal with data sampled at different frequencies (see, e. g., Ghysels et al. 2004, and Clements and Galvao, 2008), allowing us to combine indicators with different sampling frequency (daily, weekly, monthly, quarterly and yearly) and further extend the analysis made in Chapter 1.

v) The ECM approach adopted in Chapter 2 could be applied to the study of the commercial property market in other Spanish (e. g., Barcelona) and European (e. g., London and Paris) cities to explore the possible differences between rents, economic activity, vacancy and stock and to examine their behaviour during the successive cycles experience by this type of property.

vi) The hedonical rent estimation techniques expanded with spatial econometrics used in Chapter 3 could be applied to other property markets in Spain, both metropolitan cities like Barcelona and Seville, and touristic cities like Malaga, and Granada.

In view of the encouraging results of the present thesis, some optimism about the benefits from implementing these extensions seems justified.

References

Clements, M. P., Galvao, A. B., 2008. Macroeconomic forecasting with mixed frequency data: Forecasting US output growth, Journal of Business and Economic Statistics 26, 546–554.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2004. The MIDAS touch: Mixed data sampling regressions, Discussion paper UNC and UCLA.

Hendry, D. F., Krolzig, H. M., 2004. Sub-sample model selection procedures in general-to-specific modelling. In R Becker and S Hurn (eds.), Contemporary Issues in Economics and Econometrics: Theory and Applications, pp. 53-75. Edward Elgar.

Hicks, J. R. (1981). Wealth and Welfare. Collected Essays on Economic Theory, Vol. I (Oxford: Basil Blackwell).

Perez-Amaral, T., Gallo, G. M., White, H., 2003. A flexible tool for model building: The Relevant Transformation of the Inputs Network Approach (RETINA). Oxford Bulletin of Economics and Statistics 65, 821-838.

Annex to chapter 1

1. List of referenced papers used in our eclectic approach.

It was intended to select the variables that conformed both our structural model and the GUM used for the model selection algorithm.

Tag	Source Type	Author Last Name	First & Middle	Title	Year	Journal Name	Vol.	Issue	Pages
Bla1953	Journal Article	Blanck	D. M.	The structure of the housing Market	1953	Quarterly journal of economics			181-203
		Winnick	L.	The structure of the housing Market	1953	Quarterly journal of economics			181-203
Alb1962	Journal Article	Alberts	W. W.	Business cycles, residential construction cycles and the mortgage market	1962	The journal of political economy	LXX		263-281
Pri1977	Journal Article	Pritchett	Clayton P	The Effect of Regional Growth Characteristics on Regional Housing Prices	1977	American Real Estate and Urban Economics Association Journal	5	2	189-208
Fer1977	Journal Article	Ferri	Michael G.	An Application of Hedonic Indexing Methods to Monthly Changes in Housin	1977	American Real Estate and Urban Economics Association Journal. Winter	5	2	455-462

				g Prices: 1965-1975					
Ros1984	Journal Article	Rosen	Kenneth T.	Toward a Model of the Office Building Sector	1984	Real Estate Economics, fall 1984			261
Whe1987	Journal Article	Wheaton	W.C.	The Cyclic Behavior Of The National Office Market	1987	Real Estate Economics	15	4	221-281
Eng1987	Journal Article	Engle	R..F.	Cointegration and errror correction: representation, estimation and testing	1987	Econométrica (55-2)			251-276
		Granger	C. W.	Cointegration and errror correction: representation, estimation and testing	1987	Econométrica (55-2)			251-276
Phi1990	Journal Article	Phillips	Peter C.	Statistical inference in instrumental variables regressions with I(1) processes	1990	The review of economic studies			99-125

		Hansen	Bruce E.	Statistical inference in instrumental variables regressions with I(1) processes	1990	The review of economic studies			99-125
Joh1991	Journal Article	Johansen	Soren	Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models	1991	Econometrica			1551-1580
Han1992a	Journal Article	Hansen	Bruce E.	Efficient estimation and testing of coitegrating vector in the presence of deterministic trends	1992a	Journal of econometrics			87-121
Han1992b	Journal Article	Hansen	Bruce E.	Tests for parameter instability in regressions with I(1) processes	1992b	Journal of business and economic statistics			321-335
DiP1995	Book	DiPascuale	D.	Urban Economics and Real Estate Market	1995				

				s					
		Wheaton	W.	Urban Economics and Real Estate Markets	1995				
		Johansen	Soren	Likelihood-based inference in cointegrated vector autoregressive models	1996				
Whe1997	Journal Article	Wheaton	William C.	The Cyclic Behavior of the Greater London Office Market	1997	Journal of Real Estate Finance and Economics, vol. 15, iss. 1			77-92
		Torto	Raymond G.	The Cyclic Behavior of the Greater London Office Market	1997	Journal of Real Estate Finance and Economics, vol. 15, iss. 1			77-92
		Evans	Peter	The Cyclic Behavior of the Greater London Office Market	1997	Journal of Real Estate Finance and Economics, vol. 15, iss. 1			77-92
Bal1998	Book	Ball	Michael	The economics of commercial property markets	1998				

		Lizieri	Colin	The economics of commercial property markets	1998				
		MacGregor	Brian D.	The economics of commercial property markets	1998				
Bro2009	Journal Article	Brounen	Dirk	Local office rent dynamics. A tale of ten cities	2009	Journal of real estate finance and economics			385-402
		Jennen	Maarten	Local office rent dynamics. A tale of ten cities	2009	Journal of real estate finance and economics			385-402
Fue2010	Journal Article	Fuerst	Franz	Supply elasticities and developers' expectations: a study of European office markets	2010	Journal of European Real Estate Research, vol. 3, No. 1			5-23
Bro2010	Book	Brooks	Chris	Real Estate Modelling and Forecasting	2010				
		Tsolacos	Sotiris	Real Estate Modelling and Forecasting	2010				

BNP2011	Report	BNP Paribas Real Estate España		El mercado de oficinas en Madrid y Barcelona, segundo trimestre	2011				
Hue2008	Journal Article	Huerta	Ramón	A housing-demographic multilayered nonlinear model to test regulation strategies	2008	Working Paper			
		Corbacho	Fernando	A housing-demographic multilayered nonlinear model to test regulation strategies	2008	Working Paper			
		Lago-Fernández	Luis F.	A housing-demographic multilayered nonlinear model to test regulation strategies	2008	Working Paper			

Car2001	Journal Article	Caridad y Ocerin	JM	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	2001	Estudios de economía inmobiliaria	18		68-71
		Ceular Villamandos	N	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	2001	Estudios de economía inmobiliaria	18		68-71
Qui1999	Journal Article	Quigley	John M.	Real Estate Prices and Economic Cycles	1999	International Real Estate Review	2	1	1-20
Qua1991	Journal Article	Quan	Daniel C.	Price formation and the Appraisal Function in Real Estate Markets	1991	Journal of Real Estate Finance and Economics	4		127-146
		Quigley	John M.	Price formation and the Appraisal Function in Real Estate Markets	1991	Journal of Real Estate Finance and Economics	4		127-146

Mac2011	WorkingPaper	Mack	Adrienne	A Cross-Country Database of Real House Prices: A Methodological Note	2011				
		Martínez-García	Enrique						
Kos2004	Journal Article	Tsatsaronis	Kostas	What drives housing prices dynamics: cross country evidence	2004	BIS quarterly review	March		65-78
		Zhu	Haibin	What drives housing prices dynamics: cross country evidence	2004	BIS quarterly review	March		65-78
Omb2011	Journal Article	Omboi	Bernard Messah	Factors influencing Real Estate Prices - A Survey of Real Estates in Meru Municipaliy, Kenya	2011	Journal of Economics and Sustainable Development	2	4	34-53
		Kigige	Anderson M	Factors influencing Real Estate Prices - A Survey of Real Estates in Meru Municipaliy, Kenya	2011	Journal of Economics and Sustainable Development	2	4	34-53

Fav2012	Journal Article	Favikukis	Jack	International Capital Flows and House Prices: Theory and Evidence	2012	Nber Working Paper Series	17751		
		Kohn	David	International Capital Flows and House Prices: Theory and Evidence	2012	Nber Working Paper Series	17751		
		Ludvigson	Sydney C.	International Capital Flows and House Prices: Theory and Evidence	2012	Nber Working Paper Series	17751		
		Van Nieuwerburgh	Stijn	International Capital Flows and House Prices: Theory and Evidence	2012	Nber Working Paper Series	17751		
Lam2012	Journal Article	Lambertini	Lisa	Expectations Driven Cycles in the Housing Market	2012	Bank of Finland Research Discussion Papers	2		
		Mendicino	Caterina	Expectations Driven Cycles in the Housing Market	2012	Bank of Finland Research Discussion Papers	2		

		Punzi	María T.	Expectations Driven Cycles in the Housing Market	2012	Bank of Finland Research Discussion Papers	2		
Fer2012	WorkingPaper	Fernandez Durán	Laura	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	2011				
		Llorca	Alicia	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	2012				
		Ruiz	Nancy	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	2013				

		Valero	Soledad	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	2014				
		Botti	Vicente	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	2015				
Cra2011	WorkingPaper	Craig	Sean R	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	2011			Nov	
		Changchun	Hua	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	2011			Nov	

Anj2011	WorkingPaper	Rosen	Anja	On the predictive content of nonlinear transformations of lagged autoregression residuals and time series observations	2011		Paper 113		
Cop2011	WorkingPaper	Coporale	Guglielmo María	Are Stock and Housing Returns Complements or Substitutes? Evidence from OCDE countries	2011		NIPE WP 33		
		Sousa	Ricardo M.	Are Stock and Housing Returns Complements or Substitutes? Evidence from OCDE countries	2011		NIPE WP 33		
Aco2010	WorkingPaper	Acosta-González	Eduardo	On factors explaining the 2008 financial crisis	2010				
		Fernández-Rodríguez	Fernando	On factors explaining the 2008 financia	2010				

				l crisis					
		Sosvilla-Rivero	Simón	On factors explaining the 2008 financial crisis	2010				
Whe1992	Journal Article	Wheaton	William C.	Office Rent Indices and Their Behavior over Time	1992	Journal of Urban Economics	2 - march 1994	35	121-139
		Torto	Raymond	Office Rent Indices and Their Behavior over Time	1992	Journal of Urban Economics	2 - march 1994	35	121-139
mal1998	Journal Article	Malpenzzi	Stephen	A simple error correction model of house prices		Journal of Housing Economics		8	27-62

2. *List of theoretical variables collected from real estate economics literature*

Source Tag	Article name	Theoretical Variable	Proxy	Expected effect	Left-side/Right-side	Residential/Commercial
Hue2008	A housing-demographic multilayered nonlinear model to test regulation strategies	House occupancy	Number of occupied housing units	N/A	Left-side	Residential
Hue2008	A housing-demographic multilayered nonlinear model to test regulation strategies	House stock	Total number of housing units	N/A	Left-side	Residential
Hue2008	A housing-demographic	Housing	Ratio	N/A	Left-side	Residential

	multilayered nonlinear model to test regulation strategies	Vacancy rate				
Hue2008	A housing-demographic multilayered nonlinear model to test regulation strategies	Families that can enter vacant sites	Number of families	Positive	Right-side	Residential
Hue2008	A housing-demographic multilayered nonlinear model to test regulation strategies	House price	Layers or Price housing bands	N/A	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	House price	Market price (pesetas)	N/A	Left-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Surface Area (usable sqm)	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Rooms	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Wardrobes	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Bathrooms	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Air Conditioning	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales	Hedonic feature	Sink	Positive	Right-side	Residential

	artificiales					
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Pantry	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Cellar	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Kitchen quality	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Entrance hall quality	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Garage access	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Lift	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Floor	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Pool	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Parabolic antenna	Positive	Right-side	Residential
Car2001	Un análisis del mercado de la vivienda a	Hedonic feature	Neighborhood income	Positive	Right-side	Residential

	través de redes neuronales artificiales					
Car2001	Un análisis del mercado de la vivienda a través de redes neuronales artificiales	Hedonic feature	Neighborhood quality	Positive	Right-side	Residential
Qui1999	Real Estate Prices and Economic Cycles	House price	Price of owner occupied housing	N/A	Left-side	Residential
Qui1999	Real Estate Prices and Economic Cycles	Population	Total population	N/A	Right-side	Residential
Qui1999	Real Estate Prices and Economic Cycles	Income	Income	Positive	Right-side	Residential
Qui1999	Real Estate Prices and Economic Cycles	Employment	Aggregate employment	Positive	Right-side	Residential
Qui1999	Real Estate Prices and Economic Cycles	Construction	Construction permits	Negative	Right-side	Residential
Qui1999	Real Estate Prices and Economic Cycles	Vacancy	Vacancy rate	Negative	Right-side	Residential
Qui1999	Real Estate Prices and Economic Cycles	Lagged prices	Lagged prices	Unknown	Right-side	Residential
Qua1991	Price formation and the Appraisal Function in Real Estate Markets	Price of a class on similar properties	Price of any property	N/A	Left-side	Residential
Qua1991	Price formation and the Appraisal Function in Real Estate Markets	Estimated price of a class on similar properties	Price of any property plus an error term	N/A	Left-side	Res/Comm
Qua1991	Price formation and the Appraisal Function in Real Estate Markets	Information set	Error term of estimated price	Unknown	Right-side	Res/Comm

Mac2011	A Cross-Country Database of Real House Prices: A Methodological Note	House price	Index of houses	N/A	Left-side	Residential
Mac2011	A Cross-Country Database of Real House Prices: A Methodological Note	Income	Private disposable income	N/A	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Long term housing demand	Long term housing demand	N/A	Left-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Income growth	Growth in household disposable income	Positive	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Population age	Relative size of older and younger inhabitants	Unknown	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Fiscal incentives	Tax rates	Unknown	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Interest rate	Loan rates	Negative	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Long run inflation	Inflation rate	Unknown	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Housing supply	Hosing stock	N/A	Left-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Land stock	Land stock	Unknown	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country	Land price	Land price	Unknown	Right-side	Residential

	evidence					
Kos2004	What drives housing prices dynamics: cross country evidence	Construction costs	Construction costs permits	Positive	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Housing prices	House price index	N/A	Left-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Business cycle	GDP	Positive	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Household income	GDP	Positive	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Inflation	Change in consumer index	Negative	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Interest rate	Real short term interest rate	Negative	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Term spread	Bond minus interest rate	Negative	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Credit availability	Growth rate in inflation-adjusted bank credit	Positive	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Adjustment of Interest rate	Dummy (1if mortgage interest rate is fixed, 0 if variable)	Unknown	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Mortgage equity withdrawal	Dummy (1if used, 0 if not)	Unknown	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	LTV ratio	Max LTV ratio	Unknown	Right-side	Residential

Kos2004	What drives housing prices dynamics: cross country evidence	Valuation method	Dummy (1if Open market value, 0 if Mortgage lending value)	Unknown	Right-side	Residential
Kos2004	What drives housing prices dynamics: cross country evidence	Securitization (Mortgage backed)	Dummy (1if used, 0 if not)	Unknown	Right-side	Residential
Omb2011	Factors influencing Real Estate Prices - A Survey of Real Estates in Meru Municipaliy, Kenya	Price of real estate	Price of houses	N/A	Left-side	Residential
Omb2011	Factors influencing Real Estate Prices - A Survey of Real Estates in Meru Municipaliy, Kenya	Demand	Incomes of real estate investors	Positive	Right-side	Residential
Omb2011	Factors influencing Real Estate Prices - A Survey of Real Estates in Meru Municipaliy, Kenya	Location	Address	Positive	Right-side	Residential
Omb2011	Factors influencing Real Estate Prices - A Survey of Real Estates in Meru Municipaliy, Kenya	Intermediation	Deals closed with realtors	Positive	Right-side	Residential
Fav2012	International Capital Flows and House Prices: Theory and Evidence	House price	Real house price growth	N/A	Left-side	Residential
Fav2012	International Capital Flows and House Prices: Theory and Evidence	Credit availability	% Banks relaxing credit standard for mortgages loans	Positive	Right-side	Residential
Fav2012	International Capital Flows and House Prices: Theory	Capital flow	Current account deficit/GDP	Positive	Right-side	Residential

	and Evidence					
Fav2012	International Capital Flows and House Prices: Theory and Evidence	Capital flow	Current account deficit/GDP*CS	Positive	Right-side	Residential
Fav2012	International Capital Flows and House Prices: Theory and Evidence	Capital flow	Net foreign holdings of total securities to GDP	Unknown	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	House price	House price	N/A	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	GDP	GDP	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Consumption	Real consumption	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Investment	Real business investment	N/A	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Housing investment	Real housing investment	N/A	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Wage	Real wages (Consumption and housing)	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Loans	Real Loans	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Hours worked	Hours worked	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Consumers' belief of favorable buying conditions		Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	House price	Log change in real house prices	N/A	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing	Interest rate	Short term real interest rate (Difference in	Positive	Left-side	Residential

	Market		3-month-treasury-bill-rate and the GDP deflator)			
Lam2012	Expectations Driven Cycles in the Housing Market	GDP	GDP	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Population	Civilian non-institutional population	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	News of changes in business conditions	Perception of the current state of the economy (University of Michigan survey of consumers)	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Consumer optimism	Index of Consumers' Sentiment (ICS)	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	News of changes in business conditions	Perception of the current state of the economy	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	News of changes in business conditions	Expectations of rising housing prices	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	News of changes in business conditions	Expectations of tightening future credit	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Household indebtness	Households debt	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	News on productivity shocks	IDEM	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	News on monetary policy shocks	IDEM	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Expectations of policy rate (nominal variable)	IDEM	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Expectations of inflation (nominal variable)	IDEM	Positive	Left-side	Residential

Lam2012	Expectations Driven Cycles in the Housing Market	Loan-to-value ratios	IDEM	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Utility	IDEM	Positive	Right-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Consumption	Private consumption	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Housing services	Housing services	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Hours worked in the good-sector	Housing services	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Hours worked in the construction	Housing services	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Elasticity of substitution of sectors in work	Estimated parameter	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Frish inverse elasticity of labor supply	Estimated parameter	Unknown	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Relative weigh in utility of housing services	Estimated parameter	Unknown	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Lending interest rate	IDEM	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Depreciation of capital	IDEM	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Depreciation of houses	IDEM	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Land	Land stock	Unknown	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing	Land price	Prince index	Negative	Left-side	Residential

	Market					
Lam2012	Expectations Driven Cycles in the Housing Market	Houses price	Prince index	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Capital utilization rates of transforming potential capital to effective capital	IDEM	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Capital utilization rates of transforming potential capital to effective houses	IDEM	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Lump-sum profits paid to households	IDEM	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Marginal cost of producing consumption-good-sector specific capital	Estimated parameter	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Trend growth rate of real consumption	Variation in consumption	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Trend growth rate of capital	Gross capital formation	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Trend growth rate of housing capital	Gross capital formation	Negative	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Scaling factor of marginal utility of consumption	Estimated parameter	Unknown	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Mark-up on the wages paid in the capital sector	Estimated parameter	Positive	Left-side	Residential
Lam2012	Expectations Driven Cycles in the Housing Market	Mark-up on the wages paid in the house sector	Estimated parameter	Positive	Left-side	Residential
Fer2012	The impact on location on housing prices: Applying the	House price	Offer prices per square meter	N/A	Left-side	Residential

	Artificial Neural Network Model as analytical tool					
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Distance to CBD	Distance to CBD	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Square footage	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	House Height	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Penthouse and similar	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Age	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Condition	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Parking space	IDEM	Positive	Right-side	Residential

Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Number of bedrooms	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Number of bathrooms	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Views aspect	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Lift	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Natural gas installation	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Central heating	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Green zones	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model	Swimming pool	IDEM	Positive	Right-side	Residential

	as analytical tool					
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Sports facilities	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Playground	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Street width	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Pavement width	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Quality of urban	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Proximity to metro/train station	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Proximity to motorways	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices:	Education centres	IDEM	Positive	Right-side	Residential

	Applying the Artificial Neural Network Model as analytical tool					
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Health centres	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Cultural centres	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Sport centres	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Parks	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Traffic density	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Special buildings	IDEM	Positive	Right-side	Residential
Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Socioeconomic status	IDEM	Positive	Right-side	Residential

Fer2012	The impact on location on housing prices: Applying the Artificial Neural Network Model as analytical tool	Immigration population	IDEM	Unknown	Right-side	Residential
Cra2011	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	Equilibrium house price	IDEM	N/A	Left-side	Residential
Cra2011	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	Land supply	Sqm sold at land auctions	Negative	Right-side	Residential
Cra2011	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	Building Costs	Index of material and labor costs	Positive	Right-side	Residential
Cra2011	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	GDPpc	Household income	Positive	Right-side	Residential
Cra2011	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	Interest rate	Prime rate	Negative	Right-side	Residential
Cra2011	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	Domestic credit	Domestic credit	Positive	Right-side	Residential
Cra2011	Determinants of Property Prices in Hong Kong SAR: Implications for Policy	Domestic credit	Mortgage credit	Positive	Right-side	Residential
Anj2011	On the predictive content of nonlinear transformations of lagged autoregression residuals and time series observations	Pool of economic indicators			Right-side	Non Real Estate

Aco2010	On factors explaining the 2008 financial crisis	Severity of the crisis	4 severity stages	N/A	Left-side	Non Real Estate
Aco2010	On factors explaining the 2008 financial crisis	Pool of 60 variables		N/A	Right-side	Non Real Estate
Whe1992	Office Rent Indices and Their Behavior over Time	Rent	Consideration rent: Average gross payment per sqf to be paid over the full term the lease. Includes movements over time in the base rent as well as free rent periods.	N/A	Left-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	Surface of the lease	Sqm	Unknown	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	Length of lease	Years	Positive	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	1 if 5+ stories; 0 otherwise	IDEM	Positive	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	1 if new building; 0 otherwise	IDEM	Positive	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	1 if turn-key; 0 otherwise	IDEM	Positive	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	1 if lease in gross rent; 0 otherwise	IDEM	Positive	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	1 if lease in gross rent with taxes passed through; 0 otherwise	IDEM	Positive	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	Date of lease firm	Dummy for each year	Unknown	Right-side	Commercial
Whe1992	Office Rent Indices and Their Behavior over Time	Location	Dummy for submarkets	Unknown	Right-side	Commercial

3. *List of collected indicators that actually fed both our structural and automatic selected models*

Variable name	Unit	Definition	Source	Original Frequency	Transformation method	Acronym	Integration order
Current account deficit	Thousands of euros of 2008	Accumulated values at the end of the year	Spanish Central Bank	Monthly	Deflated by GDP Deflator (GDPD) and monthly data aggregated at the end of each quarter	CURRENT_ACCOUNT	FDS
EC's consumer confidence index	Percentage (net balance)	Weighted average of confidence survey undertaken by the EC in Spain	European Commission	Quarterly	None	CONSUMER_CONFIDENCE	FDS
Spain CPI	Index (2011=100)	Price index of a bundle of goods representative of the average household in Spain	National Statistics Institute	Monthly	3 month average for each quarter	CPI	SDS
Durable goods credit	Thousands of euros of 2008	Net position of loans of credit entities	Spanish Central Bank	Quarterly	Deflated by GDP Deflator (GDPD)	CREDIT_CONSUMPTION	FDS
House acquisition credit	Thousands of euros of 2008	Net position of housing acquisition loans of credit entities	Spanish Central Bank	Quarterly	Deflated by GDP Deflator (GDPD)	CREDIT_HOUSE	SDS
Total household credit	Thousands of euros of 2008	Net position of household loans of credit entities	Spanish Central Bank	Quarterly	Deflated by GDP Deflator (GDPD)	CREDIT_HOUSEHOLDS	SDS
Productive credit	Thousands of euros of 2008	Net position of corporate loans of credit entities	Spanish Central Bank	Quarterly	Deflated by GDP Deflator (GDPD)	CREDIT_PRODUCTIVE	SDS
Total non-public credit	Thousands of euros of 2008	Net position of non-corporate non-household loans of credit entities	Spanish Central Bank	Quarterly	Deflated by GDP Deflator (GDPD)	CREDIT_TOTAL	SDS
Doubtful credit ratio	Percentage	Total amount of house acquisition credit not attended in at least 3 months	Spanish Central Bank	Quarterly	None	DOUBT_CREDIT	SDS
Household credit for house	Percentage	Share of the net position of	Spanish Central Bank	Quarterly	None	HOUSE_CREDIT_TO_GDP	SDS

acquisition		housing acquisition loans of credit entities to GDP				
Credit subject to house purchase saving plans	Thousands of euros of 2008	Outstanding credit for housing purchase conditioned to saving plans supported by the Spanish government	Spanish Central Bank	Quarterly	Deflated by GDP Deflator (GDPD)	LOANS_ST_HOUSE_PURCH	FDS
Total mortgage backed credit (% GDP)	Percentage	Total outstanding credit with a mortgage collateral dedicated to housing purchase to GDP	Spanish Central Bank	Quarterly	None	MORTGAGE_TO_GDP	SDS
Employees	Thousands of persons	People with a current employment (Self or third party employed)	National Statistics Institute	Quarterly	None	EMPLOYEES	SDS
Households	Thousands of units	Total number of households	National Statistics Institute	Quarterly	None	HOUSEHOLDS	SDS
Total population	Persons	Total of persons that have residence in Spain	National Statistics Institute	Semi-annual	Interpolation of missing quarterly data using the cardinal spline algorithm, that uses a non-linear pattern and the two previous and two next available observations plus a tension parameter (curviness) as parameters	POP	SDS
Population less than 15	Persons		National Statistics Institute	Semi-annual	Interpolation of missing quarterly data using the cardinal spline algorithm, that uses a non-linear pattern and the two previous and two next available observations plus a tension	POP_LESS 15	Stationary

				parameter (curviness) as parameters		
Population 15-19	Persons	National Statistics Institute	Semi-annual	Interpolation of missing quarterly data using the cardinal spline algorithm, that uses a non-linear pattern and the two previous and two next available observations plus a tension parameter (curviness) as parameters	POP1519	FDS
Population 20-24	Persons	National Statistics Institute	Semi-annual	Interpolation of missing quarterly data using the cardinal spline algorithm, that uses a non-linear pattern and the two previous and two next available observations plus a tension parameter (curviness) as parameters	POP2024	FDS
Population 25-54	Persons	National Statistics Institute	Semi-annual	Interpolation of missing quarterly data using the cardinal spline algorithm, that uses a non-linear pattern and the two previous and two next available observations plus a tension parameter (curviness) as parameters	POP2454	SDS
25 yo+ population	Persons	National Statistics Institute	Semi-annual	Interpolation of missing quarterly	POP25MORE	SDS

					data using the cardinal spline algorithm, that uses a non-linear pattern and the two previous and two next available observations plus a tension parameter (curviness) as parameters		
Population 55 and more	Persons		Spanish Central Bank	Quarterly	None	POP55MORE	FDS
Youth unemployment rate (20-29 yo)	Percentage	Total employed persons to active persons	National Statistics Institute	Quarterly	None	RATE_UNEMP_2029	SDS
Theoretical annual effort without tax deduction	Percentage	Total annual value of mortgage credit payments of a median income household to its total annual disposable income	Spanish Central Bank	Quarterly	None	EFFORT_DEDUCT	FDS
Theoretical annual effort with tax deduction	Percentage	Total annual value of mortgage credit payments of a median income household to its total annual disposable income adding the fiscal deduction for housing acquisition	Spanish Central Bank	Quarterly	None	EFFORT_NO_DEDUCT	FDS
Real GDP, 2008	Millions of euros of 2008	Real gross domestic product in prices of 2008 deflated GDP deflator	National Statistics Institute	Quarterly	Deflated using GDP Deflator	GDP_2008	SDS
DGP volume index	Index (2008 = 100)	Index of GDP	National Statistics Institute	Quarterly	Linked series joining different independent series	GDP_VOL_INDEX_2008	SDS
GDP Per Capita real 2008	Euros of 2008	Real GDP, 2008 to total population	National Statistics Institute	Quarterly	Own calculation using GDP_2008 and POP	GDPPC	SDS

Net national disposable income PC	Euros of 2008	National disposable income to total population	EUROSTAT/INE	Quarterly	Deseasoned, deflated with GDP deflator Net National Disposable Income divided by POP	incomepc_d_real	FDS
Gross capital formation - dwellings	Thousands of euros of 2008	Gross Capital formation in houses	National Statistics Institute	Quarterly	Deflated using GDP Deflator. Estimated by OLS for the period 1995Q1 To 1999Q4 using GKF in construction that was available for the whole sample period (1995Q1-2012Q4)	GCF_DWELL	SDS
Real estate foreign direct investment	Thousands of euros of 2008	Foreign direct investment dedicated to real estate	National Statistics Institute	Quarterly	Deflated using GDP Deflator	RE_FDI	FDS
Real estate foreign direct investment to GDP	Percentage (ratio)	Foreign direct investment dedicated to real estate to GDP	National Statistics Institute	Quarterly	None	RE_FDI_TO_GDP	FDS
Ibex-35	Index (1989=3000)	Spanish stock market index of the 35 most liquid companies in the Madrid Stock Exchange	National Statistics Institute	Monthly	Quarterly average	IBEX	FDS
Weighted average of more than 3 years credit	Percentage (rate)	Average of interest rate for housing purchase of the credit entities	Spanish Central Bank	Quarterly	None	MORTG_RATE	FDS
Residential yield (last 12 months)	Percentage	Estimated residential yield	Spanish Central Bank	Quarterly	None	YIELD_HOUSING	Stationary
Average residential price (<=2 years)	Euros of 2008 per SQM	Valuation based house price average collected from valuation companies	Ministry of Public Works	Quarterly	Deflated using GDP Deflator	HOUSE_PRICE_2LESS	SDS
Average residential price (>2 years)	Euros of 2008 per SQM	Valuation based house price average collected from valuation	Ministry of Public Works	Quarterly	Deflated using GDP Deflator	HOUSE_PRICE_2MORE	SDS

		companies					
Average residential price	Euros of 2008 per SQM	Valuation based house price average collected from valuation companies	Ministry of Public Works	Quarterly	Deflated using GDP Deflator	HOUSE_PRICE_M2	SDS
Residential price index in Netherlands	Index (1996 = 100)	Transaction based house price index of houses in the Netherlands	EUROSTAT	Quarterly	Deflated using GDP Deflator	HOUSE_PRICE_NL	FDS
Average residential price (coast line and Islands)	Euros of 2008 per SQM	Valuation based house price average collected from valuation companies	Ministry of Public Works	Quarterly	Deflated using GDP Deflator	RES_RPRICE_INX_COAST	SDS
Average residential price (Madrid and Barcelona)	Euros of 2008 per SQM	Valuation based house price average collected from valuation companies	Ministry of Public Works	Quarterly	Deflated using GDP Deflator	RES_RPRICE_INX_MADBAR	SDS
Average residential price (Rest of provinces)	Euros of 2008 per SQM	Valuation based house price average collected from valuation companies	Ministry of Public Works	Quarterly	Deflated using GDP Deflator	RES_RPRICE_INX_REST	SDS
Residential price index	Index (1996 = 100)	Transaction based house price index	National Statistics Institute	Quarterly	None	RPRICEINDEX	SDS
Started households	Residential units	Estimated building starts as a function of new development permits	Ministry of Public Works	Quarterly	None	BULD_STRT	FDS
Started free households	Residential units	Estimated building starts as a function of new development permits	Ministry of Public Works	Quarterly	None	BULD_STRT_FREE	FDS
Started protected households	Residential units	Estimated building starts as a function of new development permits	Ministry of Public Works	Quarterly	None	BULD_STRT_PROTECT	FDS
Household permits	Residential units	Construction permits	Ministry of Public Works	Quarterly	None	HOUSE_PERMITS	FDS
House stock	Residential units	Total number of households calculated form the decennial census plus the flow of	Ministry of Public Works	Quarterly	None	STOCK	SDS

		new deliveries					
House stock as of permits of Colegio de Arquitectos	Residential units	Total number of households calculated form the decennial census plus the flow of new deliveries	Ministry of Public Works	Quarterly	None	STOCK_COLEGIOAR	SDS
Apparent Concrete Consumption	Thousands of tons	Number of metric tons of concrete produced reported by the national concrete producers plus concrete imports	Ministry of Public Works	Quarterly	None	CONCRETE_CONSUM	SDS
Concrete Output	Thousands of tons	Number of metric tons of concrete in stock reported by the national concrete producers plus concrete imports	Ministry of Public Works	Quarterly	None	CONCRETE_OUTPUT	SDS
Net financial wealth	Thousands of euros of 2008	Total financial assets less total liabilities in hands of Spanish households	Spanish Central Bank	Quarterly	Deflated using GDP Deflator	NET_FINANC_WEALTH	FDS
Real estate household wealth	Millions of euros of 2008	Total real estate assets less total real estate liabilities in hands of Spanish households	Spanish Central Bank	Quarterly	Deflated using GDP Deflator	RE_HOUSEH_WEALTH	SDS
Wealth to GDP	Percentage (ratio)	Total financial assets less total liabilities in hands of Spanish households to GDP	Spanish Central Bank	Quarterly	None	WEALTH_TO_GDP	FDS

4. Calculation of the series of DGP deflator

As of the moment of the elaboration of this study we did not have available a complete and homogeneous quarterly series of DGP deflator to deflate monetary series. Therefore we proceeded to calculate such series, as explained below:

1. Recollect the series of Nominal GDP and Chained-linked GDP Volume available in INE:

	Methodology (year)	Base
1995-2011	2000	2000
2000-2013	2008	2008

2. Calculate two series of GDP Deflator with the following formula:

$$MIG \quad \_G \quad P_{t,iG} / G \quad P\_VG \quad \_IG \quad X_{t.iG}$$

   Where *t* is a particular quarter and *i* an INE GDP estimation methodology.

3. As we wanted the deflator to have Base=2008 the series 2000-2013 was already calculated in step 2. However we had to re-base the DGP Deflator (GDPD) series base 2000. To do so we applied the quarterly (backward) variation of the GDPD_2000 to GDPD_2008 to have the full series 1995Q1 to 2012Q4 in the same base.

GDP Deflator base 2008

(Chained series using nominal and volume index GDP series

with 2000 and 2008 bases)

Mar-95	63.09468964
Jun-95	63.81949225
Sep-95	64.4140921
Dec-95	65.0804248
Mar-96	65.76244799
Jun-96	66.10931187
Sep-96	66.49497172
Dec-96	66.90483152
Mar-97	67.16405213
Jun-97	67.66377168
Sep-97	68.21792944
Dec-97	68.53801417
Mar-98	67.37130369
Jun-98	69.68365458
Sep-98	70.26262266
Dec-98	70.9703301
Mar-99	70.50958178
Jun-99	71.20113326
Sep-99	71.66885666

Dec-99	72.24068835
Mar-00	72.80480793
Jun-00	73.27905326
Sep-00	74.15136517
Dec-00	75.13018333
Mar-01	75.71548981
Jun-01	76.63617718
Sep-01	77.40684319
Dec-01	78.00003743
Mar-02	79.01882521
Jun-02	79.84167352
Sep-02	80.74212846
Dec-02	81.55435983
Mar-03	82.47636061
Jun-03	83.18435073
Sep-03	83.98602585
Dec-03	84.87843926
Mar-04	85.52009584
Jun-04	86.59615442
Sep-04	87.53221144
Dec-04	88.39494397
Mar-05	89.38157524
Jun-05	90.31062274
Sep-05	91.18497376
Dec-05	92.2774162
Mar-06	93.26386218
Jun-06	94.10858948
Sep-06	95.15402182
Dec-06	95.67323232
Mar-07	96.59963103
Jun-07	97.32709324
Sep-07	97.83950421
Dec-07	98.80725241
Mar-08	99.31887031
Jun-08	99.90509038
Sep-08	100.3815094
Dec-08	100.2801869
Mar-09	100.2518446
Jun-09	100.0851757
Sep-09	99.87668567
Dec-09	99.96320098
Mar-10	100.0428465
Jun-10	100.0683068
Sep-10	100.1744283
Dec-10	100.213143
Mar-11	100.1179778
Jun-11	100.2168745
Sep-11	100.0606235
Dec-11	100.1728557
Mar-12	100.0511361

Jun-12	100.0895095
Sep-12	100.2179142
Dec-12	100.1523398
Mar-13	100.9471298
Jun-13	100.7562369

5. Confirmatory analysis of order of integration

The output of the 530 tests made - 53 variables times 3 (number of models in the ADF test) times 2 (levels and first difference) plus 53 variables times 2 (number of models in KPSS test) times 2 (levels and first difference) - can be sent upon request to the corresponding author.

Model estimated	**Variable name**	**Null hypothesis**	**Test result**	**Order of integration**
ADF (Constant and Trend)	BULD_STRT	BULD_STRT has a unit root	Non stationary	FDS
ADF (Constant and Trend)	BULD_STRT	D(BULD_STRT) has a unit root	Stationary	
ADF (Constant)	BULD_STRT	BULD_STRT has a unit root	Non stationary	
ADF (Constant)	BULD_STRT	D(BULD_STRT) has a unit root	Non stationary	
ADF (no exogenous)	BULD_STRT	BULD_STRT has a unit root	Non stationary	
ADF (no exogenous)	BULD_STRT	D(BULD_STRT) has a unit root	Stationary	
KPSS (Constant and Trend)	BULD_STRT	BULD_STRT is stationary	Non stationary	
KPSS (Constant and Trend)	BULD_STRT	D(BULD_STRT) is stationary	Stationary	
KPSS (Constant)	BULD_STRT	BULD_STRT is stationary	Stationary	
KPSS (Constant)	BULD_STRT	D(BULD_STRT) is stationary	Non stationary	
ADF (Constant and Trend)	BULD_STRT_FREE	BULD_STRT_FREE has a unit root	Non stationary	FDS
ADF (Constant and Trend)	BULD_STRT_FREE	D(BULD_STRT_FREE) has a unit root	Stationary	
ADF (Constant)	BULD_STRT_FREE	BULD_STRT_FREE has a unit root	Non stationary	
ADF (Constant)	BULD_STRT_FREE	D(BULD_STRT_FREE) has a unit root	Stationary	
ADF (no exogenous)	BULD_STRT_FREE	BULD_STRT_FREE has a unit root	Non stationary	
ADF (no exogenous)	BULD_STRT_FREE	D(BULD_STRT_FREE) has a unit root	Stationary	
KPSS (Constant and Trend)	BULD_STRT_FREE	BULD_STRT_FREE is stationary	Non stationary	
KPSS (Constant and Trend)	BULD_STRT_FREE	D(BULD_STRT_FREE) is stationary	Stationary	
KPSS (Constant)	BULD_STRT_FREE	BULD_STRT_FREE is stationary	Stationary	
KPSS (Constant)	BULD_STRT_FREE	D(BULD_STRT_FREE) is stationary	Stationary	
ADF (Constant and Trend)	BULD_STRT_PROTECT	BULD_STRT_PROTECT has a unit root	Non stationary	FDS
ADF (Constant and Trend)	BULD_STRT_PROTECT	D(BULD_STRT_PROTECT) has a unit root	Stationary	
ADF (Constant)	BULD_STRT_PROTECT	BULD_STRT_PROTECT has a unit root	Non stationary	
ADF (Constant)	BULD_STRT_PROTECT	D(BULD_STRT_PROTECT) has a unit root	Stationary	
ADF (no exogenous)	BULD_STRT_PROTECT	BULD_STRT_PROTECT has a unit root	Non stationary	
ADF (no exogenous)	BULD_STRT_PROTECT	D(BULD_STRT_PROTECT) has a unit root	Stationary	
KPSS (Constant and Trend)	BULD_STRT_PROTECT	BULD_STRT_PROTECT is stationary	Non stationary	
KPSS (Constant and Trend)	BULD_STRT_PROTECT	D(BULD_STRT_PROTECT) is stationary	Stationary	
KPSS (Constant)	BULD_STRT_PROTECT	BULD_STRT_PROTECT is stationary	Stationary	
KPSS (Constant)	BULD_STRT_PROTECT	D(BULD_STRT_PROTECT) is stationary	Stationary	

ADF (Constant and Trend)	CONCRETE_CONSUM	CONCRETE_CONSUM has a unit root	Non stationary	SDS
ADF (Constant and Trend)	CONCRETE_CONSUM	D(CONCRETE_CONSUM) has a unit root	Non stationary	
ADF (Constant)	CONCRETE_CONSUM	CONCRETE_CONSUM has a unit root	Non stationary	
ADF (Constant)	CONCRETE_CONSUM	D(CONCRETE_CONSUM) has a unit root	Non stationary	
ADF (no exogenous)	CONCRETE_CONSUM	CONCRETE_CONSUM has a unit root	Non stationary	
ADF (no exogenous)	CONCRETE_CONSUM	D(CONCRETE_CONSUM) has a unit root	Non stationary	
KPSS (Constant and Trend)	CONCRETE_CONSUM	CONCRETE_CONSUM is stationary	Non stationary	
KPSS (Constant and Trend)	CONCRETE_CONSUM -	D(CONCRETE_CONSUM) is stationary	Stationary	
KPSS (Constant)	CONCRETE_CONSUM	CONCRETE_CONSUM is stationary	Stationary	
KPSS (Constant)	CONCRETE_CONSUM -	D(CONCRETE_CONSUM) is stationary	Non stationary	
ADF (Constant and Trend)	CONCRETE_OUTPUT	CONCRETE_OUTPUT has a unit root	Non stationary	SDS
ADF (Constant and Trend)	CONCRETE_OUTPUT -	D(CONCRETE_OUTPUT) has a unit root	Non stationary	
ADF (Constant)	CONCRETE_OUTPUT	CONCRETE_OUTPUT has a unit root	Non stationary	
ADF (Constant)	CONCRETE_OUTPUT -	D(CONCRETE_OUTPUT) has a unit root	Non stationary	
ADF (no exogenous)	CONCRETE_OUTPUT	CONCRETE_OUTPUT has a unit root	Non stationary	
ADF (no exogenous)	CONCRETE_OUTPUT -	D(CONCRETE_OUTPUT) has a unit root	Non stationary	
KPSS (Constant and Trend)	CONCRETE_OUTPUT	CONCRETE_OUTPUT is stationary	Non stationary	
KPSS (Constant and Trend)	CONCRETE_OUTPUT	D(CONCRETE_OUTPUT) is stationary	Non stationary	
KPSS (Constant)	CONCRETE_OUTPUT	CONCRETE_OUTPUT is stationary	Stationary	
KPSS (Constant)	CONCRETE_OUTPUT	D(CONCRETE_OUTPUT) is stationary	Non stationary	
ADF (Constant and Trend)	CONSUMER_CONFIDENCE	CONSUMER_CONFIDENCE has a unit root	Non stationary	FDS
ADF (Constant and Trend)	CONSUMER_CONFIDENCE	D(CONSUMER_CONFIDENCE) has a unit root	Stationary	
ADF (Constant)	CONSUMER_CONFIDENCE	CONSUMER_CONFIDENCE has a unit root	Non stationary	
ADF (Constant)	CONSUMER_CONFIDENCE	D(CONSUMER_CONFIDENCE) has a unit root	Stationary	
ADF (no exogenous)	CONSUMER_CONFIDENCE	CONSUMER_CONFIDENCE has a unit root	Non stationary	
ADF (no exogenous)	CONSUMER_CONFIDENCE	D(CONSUMER_CONFIDENCE) has a unit root	Stationary	
KPSS (Constant and Trend)	CONSUMER_CONFIDENCE	CONSUMER_CONFIDENCE is stationary	Stationary	
KPSS (Constant and Trend)	CONSUMER_CONFIDENCE	D(CONSUMER_CONFIDENCE) is stationary	Stationary	
KPSS (Constant)	CONSUMER_CONFIDENCE	CONSUMER_CONFIDENCE is stationary	Non stationary	
KPSS (Constant)	CONSUMER_CONFIDENCE	D(CONSUMER_CONFIDENCE) is stationary	Stationary	
ADF (Constant and Trend)	CPI -	CPI has a unit root	Non stationary	SDS
ADF (Constant and Trend)	CPI -	D(CPI) has a unit root	Stationary	
ADF (Constant)	CPI -	CPI has a unit root	Non stationary	
ADF (Constant)	CPI -	D(CPI) has a unit root	Stationary	
ADF (no exogenous)	CPI -	CPI has a unit root	Non stationary	
ADF (no exogenous)	CPI -	D(CPI) has a unit root	Non stationary	
KPSS (Constant and Trend)	CPI -	CPI is stationary	Stationary	
KPSS (Constant and Trend)	CPI -	D(CPI) is stationary	Non stationary	
KPSS (Constant)	CPI -	CPI is stationary	Non stationary	
KPSS (Constant)	CPI -	D(CPI) is stationary	Non stationary	
ADF (Constant and Trend)	CREDIT_CONSUMPTION	CREDIT_CONSUMPTION has a unit root	Non stationary	FDS

ADF (Constant and Trend)	CREDIT_CONSUMPTION	D(CREDIT_CONSUMPTION) has a unit root	Stationary	
ADF (Constant)	CREDIT_CONSUMPTION	CREDIT_CONSUMPTION has a unit root	Non stationary	
ADF (Constant)	CREDIT_CONSUMPTION	D(CREDIT_CONSUMPTION) has a unit root	Stationary	
ADF (no exogenous)	CREDIT_CONSUMPTION	CREDIT_CONSUMPTION has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_CONSUMPTION	D(CREDIT_CONSUMPTION) has a unit root	Stationary	
KPSS (Constant and Trend)	CREDIT_CONSUMPTION	CREDIT_CONSUMPTION is stationary	Non stationary	
KPSS (Constant and Trend)	CREDIT_CONSUMPTION	D(CREDIT_CONSUMPTION) is stationary	Non stationary	
KPSS (Constant)	CREDIT_CONSUMPTION	CREDIT_CONSUMPTION is stationary	Non stationary	
KPSS (Constant)	CREDIT_CONSUMPTION	D(CREDIT_CONSUMPTION) is stationary	Non stationary	
ADF (Constant and Trend)	CREDIT_HOUSE	CREDIT_HOUSE has a unit root	Non stationary	SDS
ADF (Constant and Trend)	CREDIT_HOUSE	D(CREDIT_HOUSE) has a unit root	Non stationary	
ADF (Constant)	CREDIT_HOUSE	CREDIT_HOUSE has a unit root	Non stationary	
ADF (Constant)	CREDIT_HOUSE	D(CREDIT_HOUSE) has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_HOUSE	CREDIT_HOUSE has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_HOUSE	D(CREDIT_HOUSE) has a unit root	Non stationary	
KPSS (Constant and Trend)	CREDIT_HOUSE	CREDIT_HOUSE is stationary	Stationary	
KPSS (Constant and Trend)	CREDIT_HOUSE	D(CREDIT_HOUSE) is stationary	Non stationary	
KPSS (Constant)	CREDIT_HOUSE	CREDIT_HOUSE is stationary	Non stationary	
KPSS (Constant)	CREDIT_HOUSE	D(CREDIT_HOUSE) is stationary	Stationary	
ADF (Constant and Trend)	CREDIT_HOUSEHOLDS	CREDIT_HOUSEHOLDS has a unit root	Non stationary	SDS
ADF (Constant and Trend)	CREDIT_HOUSEHOLDS	D(CREDIT_HOUSEHOLDS) has a unit root	Non stationary	
ADF (Constant)	CREDIT_HOUSEHOLDS	CREDIT_HOUSEHOLDS has a unit root	Non stationary	
ADF (Constant)	CREDIT_HOUSEHOLDS	D(CREDIT_HOUSEHOLDS) has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_HOUSEHOLDS	CREDIT_HOUSEHOLDS has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_HOUSEHOLDS	D(CREDIT_HOUSEHOLDS) has a unit root	Non stationary	
KPSS (Constant and Trend)	CREDIT_HOUSEHOLDS	CREDIT_HOUSEHOLDS is stationary	Non stationary	
KPSS (Constant and Trend)	CREDIT_HOUSEHOLDS	D(CREDIT_HOUSEHOLDS) is stationary	Non stationary	
KPSS (Constant)	CREDIT_HOUSEHOLDS	CREDIT_HOUSEHOLDS is stationary	Non stationary	
KPSS (Constant)	CREDIT_HOUSEHOLDS	D(CREDIT_HOUSEHOLDS) is stationary	Stationary	
ADF (Constant and Trend)	CREDIT_PRODUCTIVE	CREDIT_PRODUCTIVE has a unit root	Non stationary	SDS
ADF (Constant and Trend)	CREDIT_PRODUCTIVE	D(CREDIT_PRODUCTIVE) has a unit root	Non stationary	
ADF (Constant)	CREDIT_PRODUCTIVE	CREDIT_PRODUCTIVE has a unit root	Non stationary	
ADF (Constant)	CREDIT_PRODUCTIVE	D(CREDIT_PRODUCTIVE) has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_PRODUCTIVE	CREDIT_PRODUCTIVE has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_PRODUCTIVE	D(CREDIT_PRODUCTIVE) has a unit root	Non stationary	
KPSS (Constant and Trend)	CREDIT_PRODUCTIVE	CREDIT_PRODUCTIVE is stationary	Stationary	
KPSS (Constant and Trend)	CREDIT_PRODUCTIVE	D(CREDIT_PRODUCTIVE) is stationary	Non stationary	
KPSS (Constant)	CREDIT_PRODUCTIVE	CREDIT_PRODUCTIVE is stationary	Non stationary	
KPSS (Constant)	CREDIT_PRODUCTIVE	D(CREDIT_PRODUCTIVE) is stationary	Stationary	
ADF (Constant and Trend)	CREDIT_TOTAL	CREDIT_TOTAL has a unit root	Non stationary	SDS
ADF (Constant and Trend)	CREDIT_TOTAL	D(CREDIT_TOTAL) has a unit root	Non stationary	

ADF (Constant)	CREDIT_TOTAL	CREDIT_TOTAL has a unit root	Non stationary	
ADF (Constant)	CREDIT_TOTAL	D(CREDIT_TOTAL) has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_TOTAL	CREDIT_TOTAL has a unit root	Non stationary	
ADF (no exogenous)	CREDIT_TOTAL	D(CREDIT_TOTAL) has a unit root	Non stationary	
KPSS (Constant and Trend)	CREDIT_TOTAL	CREDIT_TOTAL is stationary	Stationary	
KPSS (Constant and Trend)	CREDIT_TOTAL	D(CREDIT_TOTAL) is stationary	Non stationary	
KPSS (Constant)	CREDIT_TOTAL	CREDIT_TOTAL is stationary	Non stationary	
KPSS (Constant)	CREDIT_TOTAL	D(CREDIT_TOTAL) is stationary	Stationary	
ADF (Constant and Trend)	CURRENT_ACCOUNT	CURRENT_ACCOUNT has a unit root	Non stationary	FDS
ADF (Constant and Trend)	CURRENT_ACCOUNT	D(CURRENT_ACCOUNT) has a unit root	Stationary	
ADF (Constant)	CURRENT_ACCOUNT	CURRENT_ACCOUNT has a unit root	Non stationary	
ADF (Constant)	CURRENT_ACCOUNT	D(CURRENT_ACCOUNT) has a unit root	Stationary	
ADF (no exogenous)	CURRENT_ACCOUNT	CURRENT_ACCOUNT has a unit root	Non stationary	
ADF (no exogenous)	CURRENT_ACCOUNT	D(CURRENT_ACCOUNT) has a unit root	Stationary	
KPSS (Constant and Trend)	CURRENT_ACCOUNT	CURRENT_ACCOUNT is stationary	Non stationary	
KPSS (Constant and Trend)	CURRENT_ACCOUNT	D(CURRENT_ACCOUNT) is stationary	Non stationary	
KPSS (Constant)	CURRENT_ACCOUNT	CURRENT_ACCOUNT is stationary	Stationary	
KPSS (Constant)	CURRENT_ACCOUNT	D(CURRENT_ACCOUNT) is stationary	Stationary	
ADF (Constant and Trend)	DOUBT_CREDIT	DOUBT_CREDIT has a unit root	Non stationary	SDS
ADF (Constant and Trend)	DOUBT_CREDIT	D(DOUBT_CREDIT) has a unit root	Non stationary	
ADF (Constant)	DOUBT_CREDIT	DOUBT_CREDIT has a unit root	Non stationary	
ADF (Constant)	DOUBT_CREDIT	D(DOUBT_CREDIT) has a unit root	Non stationary	
ADF (no exogenous)	DOUBT_CREDIT	DOUBT_CREDIT has a unit root	Non stationary	
ADF (no exogenous)	DOUBT_CREDIT	D(DOUBT_CREDIT) has a unit root	Non stationary	
KPSS (Constant and Trend)	DOUBT_CREDIT	DOUBT_CREDIT is stationary	Non stationary	
KPSS (Constant and Trend)	DOUBT_CREDIT	D(DOUBT_CREDIT) is stationary	Stationary	
KPSS (Constant)	DOUBT_CREDIT	DOUBT_CREDIT is stationary	Stationary	
KPSS (Constant)	DOUBT_CREDIT	D(DOUBT_CREDIT) is stationary	Non stationary	
ADF (Constant and Trend)	EFFORT_DEDUCT	EFFORT_DEDUCT has a unit root	Non stationary	FDS
ADF (Constant and Trend)	EFFORT_DEDUCT	D(EFFORT_DEDUCT) has a unit root	Stationary	
ADF (Constant)	EFFORT_DEDUCT	EFFORT_DEDUCT has a unit root	Non stationary	
ADF (Constant)	EFFORT_DEDUCT	D(EFFORT_DEDUCT) has a unit root	Stationary	
ADF (no exogenous)	EFFORT_DEDUCT	EFFORT_DEDUCT has a unit root	Non stationary	
ADF (no exogenous)	EFFORT_DEDUCT	D(EFFORT_DEDUCT) has a unit root	Stationary	
KPSS (Constant and Trend)	EFFORT_DEDUCT	EFFORT_DEDUCT is stationary	Stationary	
KPSS (Constant and Trend)	EFFORT_DEDUCT	D(EFFORT_DEDUCT) is stationary	Non stationary	
KPSS (Constant)	EFFORT_DEDUCT	EFFORT_DEDUCT is stationary	Stationary	
KPSS (Constant)	EFFORT_DEDUCT	D(EFFORT_DEDUCT) is stationary	Stationary	
ADF (Constant and Trend)	EFFORT_NO_DEDUCT	EFFORT_NO_DEDUCT has a unit root	Non stationary	FDS
ADF (Constant and Trend)	EFFORT_NO_DEDUCT	D(EFFORT_NO_DEDUCT) has a unit root	Stationary	
ADF (Constant)	EFFORT_NO_DEDUCT	EFFORT_NO_DEDUCT has a unit root	Non stationary	
ADF (Constant)	EFFORT_NO_DEDUCT	D(EFFORT_NO_DEDUCT) has a unit root	Stationary	
ADF (no exogenous)	EFFORT_NO_DEDUCT	EFFORT_NO_DEDUCT has a unit root	Non stationary	
ADF (no exogenous)	EFFORT_NO_DEDUCT	D(EFFORT_NO_DEDUCT) has a unit root	Stationary	
KPSS (Constant and Trend)	EFFORT_NO_DEDUCT	EFFORT_NO_DEDUCT is stationary	Stationary	

KPSS (Constant and Trend)	EFFORT_NO_DEDUCT	D(EFFORT_NO_DEDUCT) is stationary	Stationary	
KPSS (Constant)	EFFORT_NO_DEDUCT	EFFORT_NO_DEDUCT is stationary	Stationary	
KPSS (Constant)	EFFORT_NO_DEDUCT	D(EFFORT_NO_DEDUCT) is stationary	Stationary	
ADF (Constant and Trend)	EMPLOYEES	EMPLOYEES has a unit root	Non stationary	SDS
ADF (Constant and Trend)	EMPLOYEES	D(EMPLOYEES) has a unit root	Non stationary	
ADF (Constant)	EMPLOYEES	EMPLOYEES has a unit root	Non stationary	
ADF (Constant)	EMPLOYEES	D(EMPLOYEES) has a unit root	Non stationary	
ADF (no exogenous)	EMPLOYEES	EMPLOYEES has a unit root	Non stationary	
ADF (no exogenous)	EMPLOYEES	D(EMPLOYEES) has a unit root	Non stationary	
KPSS (Constant and Trend)	EMPLOYEES	EMPLOYEES is stationary	Non stationary	
KPSS (Constant and Trend)	EMPLOYEES	D(EMPLOYEES) is stationary	Non stationary	
KPSS (Constant)	EMPLOYEES	EMPLOYEES is stationary	Non stationary	
KPSS (Constant)	EMPLOYEES	D(EMPLOYEES) is stationary	Non stationary	
ADF (Constant and Trend)	GCF_DWELL	GCF_DWELL has a unit root	Non stationary	SDS
ADF (Constant and Trend)	GCF_DWELL	D(GCF_DWELL) has a unit root	Non stationary	
ADF (Constant)	GCF_DWELL	GCF_DWELL has a unit root	Non stationary	
ADF (Constant)	GCF_DWELL	D(GCF_DWELL) has a unit root	Non stationary	
ADF (no exogenous)	GCF_DWELL	GCF_DWELL has a unit root	Non stationary	
ADF (no exogenous)	GCF_DWELL	D(GCF_DWELL) has a unit root	Stationary	
KPSS (Constant and Trend)	GCF_DWELL	GCF_DWELL is stationary	Non stationary	
KPSS (Constant and Trend)	GCF_DWELL	D(GCF_DWELL) is stationary	Non stationary	
KPSS (Constant)	GCF_DWELL	GCF_DWELL is stationary	Stationary	
KPSS (Constant)	GCF_DWELL	D(GCF_DWELL) is stationary	Stationary	
ADF (Constant and Trend)	GDP_2008	GDP_2008 has a unit root	Non stationary	SDS
ADF (Constant and Trend)	GDP_2008	D(GDP_2008) has a unit root	Stationary	
ADF (Constant)	GDP_2008	GDP_2008 has a unit root	Non stationary	
ADF (Constant)	GDP_2008	D(GDP_2008) has a unit root	Non stationary	
ADF (no exogenous)	GDP_2008	GDP_2008 has a unit root	Non stationary	
ADF (no exogenous)	GDP_2008	D(GDP_2008) has a unit root	Non stationary	
KPSS (Constant and Trend)	GDP_2008	GDP_2008 is stationary	Non stationary	
KPSS (Constant and Trend)	GDP_2008	D(GDP_2008) is stationary	Non stationary	
KPSS (Constant)	GDP_2008	GDP_2008 is stationary	Non stationary	
KPSS (Constant)	GDP_2008	D(GDP_2008) is stationary	Non stationary	
ADF (Constant and Trend)	GDP_VOL_INDEX_2008	GDP_VOL_INDEX_2008 has a unit root	Non stationary	SDS
ADF (Constant and Trend)	GDP_VOL_INDEX_2008	D(GDP_VOL_INDEX_2008) has a unit root	Stationary	
ADF (Constant)	GDP_VOL_INDEX_2008	GDP_VOL_INDEX_2008 has a unit root	Non stationary	
ADF (Constant)	GDP_VOL_INDEX_2008	D(GDP_VOL_INDEX_2008) has a unit root	Non stationary	
ADF (no exogenous)	GDP_VOL_INDEX_2008	GDP_VOL_INDEX_2008 has a unit root	Non stationary	
ADF (no exogenous)	GDP_VOL_INDEX_2008	D(GDP_VOL_INDEX_2008) has a unit root	Non stationary	
KPSS (Constant and Trend)	GDP_VOL_INDEX_2008	GDP_VOL_INDEX_2008 is stationary	Non stationary	
KPSS (Constant and Trend)	GDP_VOL_INDEX_2008	D(GDP_VOL_INDEX_2008) is stationary	Non stationary	
KPSS (Constant)	GDP_VOL_INDEX_2008	GDP_VOL_INDEX_2008 is stationary	Non stationary	
KPSS (Constant)	GDP_VOL_INDEX_2008	D(GDP_VOL_INDEX_2008) is stationary	Non stationary	
ADF (Constant and Trend)	GDPPC	GDPPC has a unit root	Non stationary	SDS

ADF (Constant and Trend)	GDPPC	D(GDPPC) has a unit root	Non stationary	
ADF (Constant)	GDPPC	GDPPC has a unit root	Non stationary	
ADF (Constant)	GDPPC	D(GDPPC) has a unit root	Non stationary	
ADF (no exogenous)	GDPPC	GDPPC has a unit root	Non stationary	
ADF (no exogenous)	GDPPC	D(GDPPC) has a unit root	Stationary	
KPSS (Constant and Trend)	GDPPC	GDPPC is stationary	Non stationary	
KPSS (Constant and Trend)	GDPPC	D(GDPPC) is stationary	Stationary	
KPSS (Constant)	GDPPC	GDPPC is stationary	Non stationary	
KPSS (Constant)	GDPPC	D(GDPPC) is stationary	Non stationary	
ADF (Constant and Trend)	HOUSE_CREDIT_TO_GDP	HOUSE_CREDIT_TO_GDP has a unit root	Non stationary	SDS
ADF (Constant and Trend)	HOUSE_CREDIT_TO_GDP	D(HOUSE_CREDIT_TO_GDP) has a unit root	Non stationary	
ADF (Constant)	HOUSE_CREDIT_TO_GDP	HOUSE_CREDIT_TO_GDP has a unit root	Non stationary	
ADF (Constant)	HOUSE_CREDIT_TO_GDP	D(HOUSE_CREDIT_TO_GDP) has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_CREDIT_TO_GDP	HOUSE_CREDIT_TO_GDP has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_CREDIT_TO_GDP	D(HOUSE_CREDIT_TO_GDP) has a unit root	Non stationary	
KPSS (Constant and Trend)	HOUSE_CREDIT_TO_GDP	HOUSE_CREDIT_TO_GDP is stationary	Non stationary	
KPSS (Constant and Trend)	HOUSE_CREDIT_TO_GDP	D(HOUSE_CREDIT_TO_GDP) is stationary	Non stationary	
KPSS (Constant)	HOUSE_CREDIT_TO_GDP	HOUSE_CREDIT_TO_GDP is stationary	Non stationary	
KPSS (Constant)	HOUSE_CREDIT_TO_GDP	D(HOUSE_CREDIT_TO_GDP) is stationary	Stationary	
ADF (Constant and Trend)	HOUSE_PERMITS	HOUSE_PERMITS has a unit root	Non stationary	FDS
ADF (Constant and Trend)	HOUSE_PERMITS	D(HOUSE_PERMITS) has a unit root	Stationary	
ADF (Constant)	HOUSE_PERMITS	HOUSE_PERMITS has a unit root	Non stationary	
ADF (Constant)	HOUSE_PERMITS	D(HOUSE_PERMITS) has a unit root	Stationary	
ADF (no exogenous)	HOUSE_PERMITS	HOUSE_PERMITS has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PERMITS	D(HOUSE_PERMITS) has a unit root	Stationary	
KPSS (Constant and Trend)	HOUSE_PERMITS	HOUSE_PERMITS is stationary	Non stationary	
KPSS (Constant and Trend)	HOUSE_PERMITS	D(HOUSE_PERMITS) is stationary	Stationary	
KPSS (Constant)	HOUSE_PERMITS	HOUSE_PERMITS is stationary	Stationary	
KPSS (Constant)	HOUSE_PERMITS	D(HOUSE_PERMITS) is stationary	Stationary	
ADF (Constant and Trend)	HOUSE_PRICE_2LESS	HOUSE_PRICE_2LESS has a unit root	Non stationary	SDS
ADF (Constant and Trend)	HOUSE_PRICE_2LESS	D(HOUSE_PRICE_2LESS) has a unit root	Non stationary	
ADF (Constant)	HOUSE_PRICE_2LESS	HOUSE_PRICE_2LESS has a unit root	Non stationary	
ADF (Constant)	HOUSE_PRICE_2LESS	D(HOUSE_PRICE_2LESS) has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PRICE_2LESS	HOUSE_PRICE_2LESS has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PRICE_2LESS	D(HOUSE_PRICE_2LESS) has a unit root	Non stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_2LESS	HOUSE_PRICE_2LESS is stationary	Non stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_2LESS	D(HOUSE_PRICE_2LESS) is stationary	Non stationary	
KPSS (Constant)	HOUSE_PRICE_2LESS	HOUSE_PRICE_2LESS is stationary	Non stationary	
KPSS (Constant)	HOUSE_PRICE_2LESS	D(HOUSE_PRICE_2LESS) is stationary	Non stationary	
ADF (Constant and Trend)	HOUSE_PRICE_2MORE	HOUSE_PRICE_2MORE has a unit root	Non stationary	SDS
ADF (Constant and Trend)	HOUSE_PRICE_2MORE	D(HOUSE_PRICE_2MORE) has a unit root	Non stationary	
ADF (Constant)	HOUSE_PRICE_2MORE	HOUSE_PRICE_2MORE has a unit root	Stationary	

ADF (Constant)	HOUSE_PRICE_2MORE	D(HOUSE_PRICE_2MORE) has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PRICE_2MORE	HOUSE_PRICE_2MORE has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PRICE_2MORE	D(HOUSE_PRICE_2MORE) has a unit root	Non stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_2MORE	HOUSE_PRICE_2MORE is stationary	Non stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_2MORE	D(HOUSE_PRICE_2MORE) is stationary	Non stationary	
KPSS (Constant)	HOUSE_PRICE_2MORE	HOUSE_PRICE_2MORE is stationary	Non stationary	
KPSS (Constant)	HOUSE_PRICE_2MORE	D(HOUSE_PRICE_2MORE) is stationary	Non stationary	
ADF (Constant and Trend)	HOUSE_PRICE_M2	HOUSE_PRICE_M2 has a unit root	Non stationary	SDS
ADF (Constant and Trend)	HOUSE_PRICE_M2	D(HOUSE_PRICE_M2) has a unit root	Non stationary	
ADF (Constant)	HOUSE_PRICE_M2	HOUSE_PRICE_M2 has a unit root	Non stationary	
ADF (Constant)	HOUSE_PRICE_M2	D(HOUSE_PRICE_M2) has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PRICE_M2	HOUSE_PRICE_M2 has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PRICE_M2	D(HOUSE_PRICE_M2) has a unit root	Non stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_M2	HOUSE_PRICE_M2 is stationary	Non stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_M2	D(HOUSE_PRICE_M2) is stationary	Non stationary	
KPSS (Constant)	HOUSE_PRICE_M2	HOUSE_PRICE_M2 is stationary	Non stationary	
KPSS (Constant)	HOUSE_PRICE_M2	D(HOUSE_PRICE_M2) is stationary	Non stationary	
ADF (Constant and Trend)	HOUSE_PRICE_NL	HOUSE_PRICE_NL has a unit root	Non stationary	FDS
ADF (Constant and Trend)	HOUSE_PRICE_NL	D(HOUSE_PRICE_NL) has a unit root	Stationary	
ADF (Constant)	HOUSE_PRICE_NL	HOUSE_PRICE_NL has a unit root	Non stationary	
ADF (Constant)	HOUSE_PRICE_NL	D(HOUSE_PRICE_NL) has a unit root	Stationary	
ADF (no exogenous)	HOUSE_PRICE_NL	HOUSE_PRICE_NL has a unit root	Non stationary	
ADF (no exogenous)	HOUSE_PRICE_NL	D(HOUSE_PRICE_NL) has a unit root	Stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_NL	HOUSE_PRICE_NL is stationary	Non stationary	
KPSS (Constant and Trend)	HOUSE_PRICE_NL	D(HOUSE_PRICE_NL) is stationary	Stationary	
KPSS (Constant)	HOUSE_PRICE_NL	HOUSE_PRICE_NL is stationary	Non stationary	
KPSS (Constant)	HOUSE_PRICE_NL	D(HOUSE_PRICE_NL) is stationary	Non stationary	
ADF (Constant and Trend)	HOUSEHOLDS	HOUSEHOLDS has a unit root	Non stationary	SDS
ADF (Constant and Trend)	HOUSEHOLDS	D(HOUSEHOLDS) has a unit root	Non stationary	
ADF (Constant)	HOUSEHOLDS	HOUSEHOLDS has a unit root	Non stationary	
ADF (Constant)	HOUSEHOLDS	D(HOUSEHOLDS) has a unit root	Non stationary	
ADF (no exogenous)	HOUSEHOLDS	HOUSEHOLDS has a unit root	Non stationary	
ADF (no exogenous)	HOUSEHOLDS	D(HOUSEHOLDS) has a unit root	Non stationary	
KPSS (Constant and Trend)	HOUSEHOLDS	HOUSEHOLDS is stationary	Non stationary	
KPSS (Constant and Trend)	HOUSEHOLDS	D(HOUSEHOLDS) is stationary	Non stationary	
KPSS (Constant)	HOUSEHOLDS	HOUSEHOLDS is stationary	Non stationary	
KPSS (Constant)	HOUSEHOLDS	D(HOUSEHOLDS) is stationary	Stationary	
ADF (Constant and Trend)	IBEX	IBEX has a unit root	Non stationary	FDS
ADF (Constant and Trend)	IBEX	D(IBEX) has a unit root	Stationary	
ADF (Constant)	IBEX	IBEX has a unit root	Non stationary	
ADF (Constant)	IBEX	D(IBEX) has a unit root	Stationary	
ADF (no exogenous)	IBEX	IBEX has a unit root	Non stationary	
ADF (no exogenous)	IBEX	D(IBEX) has a unit root	Stationary	
KPSS (Constant and Trend)	IBEX	IBEX is stationary	Stationary	
KPSS (Constant and Trend)	IBEX	D(IBEX) is stationary	Stationary	
KPSS (Constant)	IBEX	IBEX is stationary	Stationary	

KPSS (Constant)	IBEX	D(IBEX) is stationary	Stationary	
ADF (Constant and Trend)	INCOMEPC_D_REAL	INCOMEPC_D_REAL has a unit root	Non stationary	FDS
ADF (Constant and Trend)	INCOMEPC_D_REAL	D(INCOMEPC_D_REAL) has a unit root	Stationary	
ADF (Constant)	INCOMEPC_D_REAL	INCOMEPC_D_REAL has a unit root	Stationary	
ADF (Constant)	INCOMEPC_D_REAL	D(INCOMEPC_D_REAL) has a unit root	Stationary	
ADF (no exogenous)	INCOMEPC_D_REAL	INCOMEPC_D_REAL has a unit root	Non stationary	
ADF (no exogenous)	INCOMEPC_D_REAL	D(INCOMEPC_D_REAL) has a unit root	Stationary	
KPSS (Constant and Trend)	INCOMEPC_D_REAL	INCOMEPC_D_REAL is stationary	Non stationary	
KPSS (Constant and Trend)	INCOMEPC_D_REAL	D(INCOMEPC_D_REAL) is stationary	Non stationary	
KPSS (Constant)	INCOMEPC_D_REAL	INCOMEPC_D_REAL is stationary	Stationary	
KPSS (Constant)	INCOMEPC_D_REAL	D(INCOMEPC_D_REAL) is stationary	Non stationary	
ADF (Constant and Trend)	LOANS_ST_HOUSE_PURCH	LOANS_ST_HOUSE_PURCH has a unit root	Non stationary	FDS
ADF (Constant and Trend)	LOANS_ST_HOUSE_PURCH	D(LOANS_ST_HOUSE_PURCH) has a unit root	Stationary	
ADF (Constant)	LOANS_ST_HOUSE_PURCH	LOANS_ST_HOUSE_PURCH has a unit root	Non stationary	
ADF (Constant)	LOANS_ST_HOUSE_PURCH	D(LOANS_ST_HOUSE_PURCH) has a unit root	Stationary	
ADF (no exogenous)	LOANS_ST_HOUSE_PURCH	LOANS_ST_HOUSE_PURCH has a unit root	Non stationary	
ADF (no exogenous)	LOANS_ST_HOUSE_PURCH	D(LOANS_ST_HOUSE_PURCH) has a unit root	Stationary	
KPSS (Constant and Trend)	LOANS_ST_HOUSE_PURCH	LOANS_ST_HOUSE_PURCH is stationary	Non stationary	
KPSS (Constant and Trend)	LOANS_ST_HOUSE_PURCH	D(LOANS_ST_HOUSE_PURCH) is stationary	Non stationary	
KPSS (Constant)	LOANS_ST_HOUSE_PURCH	LOANS_ST_HOUSE_PURCH is stationary	Non stationary	
KPSS (Constant)	LOANS_ST_HOUSE_PURCH	D(LOANS_ST_HOUSE_PURCH) is stationary	Stationary	
ADF (Constant and Trend)	MORTG_RATE	MORTG_RATE has a unit root	Non stationary	FDS
ADF (Constant and Trend)	MORTG_RATE	D(MORTG_RATE) has a unit root	Stationary	
ADF (Constant)	MORTG_RATE	MORTG_RATE has a unit root	Stationary	
ADF (Constant)	MORTG_RATE	D(MORTG_RATE) has a unit root	Stationary	
ADF (no exogenous)	MORTG_RATE	MORTG_RATE has a unit root	Stationary	
ADF (no exogenous)	MORTG_RATE	D(MORTG_RATE) has a unit root	Stationary	
KPSS (Constant and Trend)	MORTG_RATE	MORTG_RATE is stationary	Non stationary	
KPSS (Constant and Trend)	MORTG_RATE	D(MORTG_RATE) is stationary	Stationary	
KPSS (Constant)	MORTG_RATE	MORTG_RATE is stationary	Non stationary	
KPSS (Constant)	MORTG_RATE	D(MORTG_RATE) is stationary	Stationary	
ADF (Constant and Trend)	MORTGAGE_TO_GDP	MORTGAGE_TO_GDP has a unit root	Non stationary	SDS
ADF (Constant and Trend)	MORTGAGE_TO_GDP	D(MORTGAGE_TO_GDP) has a unit root	Non stationary	
ADF (Constant)	MORTGAGE_TO_GDP	MORTGAGE_TO_GDP has a unit root	Non stationary	
ADF (Constant)	MORTGAGE_TO_GDP	D(MORTGAGE_TO_GDP) has a unit root	Non stationary	
ADF (no exogenous)	MORTGAGE_TO_GDP	MORTGAGE_TO_GDP has a unit root	Non stationary	
ADF (no exogenous)	MORTGAGE_TO_GDP	D(MORTGAGE_TO_GDP) has a unit root	Non stationary	
KPSS (Constant and Trend)	MORTGAGE_TO_GDP	MORTGAGE_TO_GDP is stationary	Non stationary	
KPSS (Constant and Trend)	MORTGAGE_TO_GDP	D(MORTGAGE_TO_GDP) is stationary	Non stationary	
KPSS (Constant)	MORTGAGE_TO_GDP	MORTGAGE_TO_GDP is stationary	Non stationary	
KPSS (Constant)	MORTGAGE_TO_GDP	D(MORTGAGE_TO_GDP) is stationary	Stationary	
ADF (Constant and Trend)	NET_FINANC_WEALTH	NET_FINANC_WEALTH has a unit	Non stationary	FDS

		root		
ADF (Constant and Trend)	NET_FINANC_WEALTH	D(NET_FINANC_WEALTH) has a unit root	Stationary	
ADF (Constant)	NET_FINANC_WEALTH	NET_FINANC_WEALTH has a unit root	Non stationary	
ADF (Constant)	NET_FINANC_WEALTH	D(NET_FINANC_WEALTH) has a unit root	Stationary	
ADF (no exogenous)	NET_FINANC_WEALTH	NET_FINANC_WEALTH has a unit root	Non stationary	
ADF (no exogenous)	NET_FINANC_WEALTH	D(NET_FINANC_WEALTH) has a unit root	Stationary	
KPSS (Constant and Trend)	NET_FINANC_WEALTH	NET_FINANC_WEALTH is stationary	Non stationary	
KPSS (Constant and Trend)	NET_FINANC_WEALTH	D(NET_FINANC_WEALTH) is stationary	Stationary	
KPSS (Constant)	NET_FINANC_WEALTH	NET_FINANC_WEALTH is stationary	Stationary	
KPSS (Constant)	NET_FINANC_WEALTH	D(NET_FINANC_WEALTH) is stationary	Stationary	
ADF (Constant and Trend)	POP	POP has a unit root	Non stationary	SDS
ADF (Constant and Trend)	POP	D(POP) has a unit root	Non stationary	
ADF (Constant)	POP	POP has a unit root	Non stationary	
ADF (Constant)	POP	D(POP) has a unit root	Non stationary	
ADF (no exogenous)	POP	POP has a unit root	Non stationary	
ADF (no exogenous)	POP	D(POP) has a unit root	Non stationary	
KPSS (Constant and Trend)	POP	POP is stationary	Non stationary	
KPSS (Constant and Trend)	POP	D(POP) is stationary	Non stationary	
KPSS (Constant)	POP	POP is stationary	Non stationary	
KPSS (Constant)	POP	D(POP) is stationary	Stationary	
ADF (Constant and Trend)	POP1519	POP1519 has a unit root	Non stationary	FDS
ADF (Constant and Trend)	POP1519	D(POP1519) has a unit root	Stationary	
ADF (Constant)	POP1519	POP1519 has a unit root	Non stationary	
ADF (Constant)	POP1519	D(POP1519) has a unit root	Stationary	
ADF (no exogenous)	POP1519	POP1519 has a unit root	Non stationary	
ADF (no exogenous)	POP1519	D(POP1519) has a unit root	Stationary	
KPSS (Constant and Trend)	POP1519	POP1519 is stationary	Stationary	
KPSS (Constant and Trend)	POP1519	D(POP1519) is stationary	Stationary	
KPSS (Constant)	POP1519	POP1519 is stationary	Non stationary	
KPSS (Constant)	POP1519	D(POP1519) is stationary	Stationary	
ADF (Constant and Trend)	POP2024	POP2024 has a unit root	Non stationary	FDS
ADF (Constant and Trend)	POP2024	D(POP2024) has a unit root	Stationary	
ADF (Constant)	POP2024	POP2024 has a unit root	Non stationary	
ADF (Constant)	POP2024	D(POP2024) has a unit root	Stationary	
ADF (no exogenous)	POP2024	POP2024 has a unit root	Non stationary	
ADF (no exogenous)	POP2024	D(POP2024) has a unit root	Stationary	
KPSS (Constant and Trend)	POP2024	POP2024 is stationary	Stationary	
KPSS (Constant and Trend)	POP2024	D(POP2024) is stationary	Stationary	
KPSS (Constant)	POP2024	POP2024 is stationary	Non stationary	
KPSS (Constant)	POP2024	D(POP2024) is stationary	Stationary	
ADF (Constant and Trend)	POP2454	POP2454 has a unit root	Non stationary	SDS
ADF (Constant and Trend)	POP2454	D(POP2454) has a unit root	Non stationary	
ADF (Constant)	POP2454	POP2454 has a unit root	Non stationary	
ADF (Constant)	POP2454	D(POP2454) has a unit root	Non stationary	
ADF (no exogenous)	POP2454	POP2454 has a unit root	Non stationary	

ADF (no exogenous)	POP2454	D(POP2454) has a unit root	Non stationary	
KPSS (Constant and Trend)	POP2454	POP2454 is stationary	Non stationary	
KPSS (Constant and Trend)	POP2454	D(POP2454) is stationary	Stationary	
KPSS (Constant)	POP2454	POP2454 is stationary	Stationary	
KPSS (Constant)	POP2454	D(POP2454) is stationary	Stationary	
ADF (Constant and Trend)	POP25MORE	POP25MORE has a unit root	Non stationary	SDS
ADF (Constant and Trend)	POP25MORE	D(POP25MORE) has a unit root	Non stationary	
ADF (Constant)	POP25MORE	POP25MORE has a unit root	Non stationary	
ADF (Constant)	POP25MORE	D(POP25MORE) has a unit root	Non stationary	
ADF (no exogenous)	POP25MORE	POP25MORE has a unit root	Non stationary	
ADF (no exogenous)	POP25MORE	D(POP25MORE) has a unit root	Non stationary	
KPSS (Constant and Trend)	POP25MORE	POP25MORE is stationary	Non stationary	
KPSS (Constant and Trend)	POP25MORE	D(POP25MORE) is stationary	Non stationary	
KPSS (Constant)	POP25MORE	POP25MORE is stationary	Non stationary	
KPSS (Constant)	POP25MORE	D(POP25MORE) is stationary	Stationary	
ADF (Constant and Trend)	POP55MORE	POP55MORE has a unit root	Non stationary	FDS
ADF (Constant and Trend)	POP55MORE	D(POP55MORE) has a unit root	Stationary	
ADF (Constant)	POP55MORE	POP55MORE has a unit root	Non stationary	
ADF (Constant)	POP55MORE	D(POP55MORE) has a unit root	Stationary	
ADF (no exogenous)	POP55MORE	POP55MORE has a unit root	Non stationary	
ADF (no exogenous)	POP55MORE	D(POP55MORE) has a unit root	Stationary	
KPSS (Constant and Trend)	POP55MORE	POP55MORE is stationary	Stationary	
KPSS (Constant and Trend)	POP55MORE	D(POP55MORE) is stationary	Stationary	
KPSS (Constant)	POP55MORE	POP55MORE is stationary	Stationary	
KPSS (Constant)	POP55MORE	D(POP55MORE) is stationary	Stationary	
ADF (Constant and Trend)	POP_LESS15	POP_LESS15 has a unit root	Non stationary	Stationary
ADF (Constant and Trend)	POP_LESS15	D(POP_LESS15) has a unit root	Non stationary	
ADF (Constant)	POP_LESS15	POP_LESS15 has a unit root	Stationary	
ADF (Constant)	POP_LESS15	D(POP_LESS15) has a unit root	Stationary	
ADF (no exogenous)	POP_LESS15	POP_LESS15 has a unit root	Non stationary	
ADF (no exogenous)	POP_LESS15	D(POP_LESS15) has a unit root	Stationary	
KPSS (Constant and Trend)	POP_LESS15	POP_LESS15 is stationary	Stationary	
KPSS (Constant and Trend)	POP_LESS15	D(POP_LESS15) is stationary	Stationary	
KPSS (Constant)	POP_LESS15	POP_LESS15 is stationary	Stationary	
KPSS (Constant)	POP_LESS15	D(POP_LESS15) is stationary	Stationary	
ADF (Constant and Trend)	RATE_UNEMP_2029	RATE_UNEMP_2029 has a unit root	Non stationary	SDS
ADF (Constant and Trend)	RATE_UNEMP_2029	D(RATE_UNEMP_2029) has a unit root	Non stationary	
ADF (Constant)	RATE_UNEMP_2029	RATE_UNEMP_2029 has a unit root	Non stationary	
ADF (Constant)	RATE_UNEMP_2029	D(RATE_UNEMP_2029) has a unit root	Non stationary	
ADF (no exogenous)	RATE_UNEMP_2029	RATE_UNEMP_2029 has a unit root	Non stationary	
ADF (no exogenous)	RATE_UNEMP_2029	D(RATE_UNEMP_2029) has a unit root	Stationary	
KPSS (Constant and Trend)	RATE_UNEMP_2029	RATE_UNEMP_2029 is stationary	Non stationary	
KPSS (Constant and Trend)	RATE_UNEMP_2029	D(RATE_UNEMP_2029) is stationary	Stationary	
KPSS (Constant)	RATE_UNEMP_2029	RATE_UNEMP_2029 is stationary	Stationary	
KPSS (Constant)	RATE_UNEMP_2029	D(RATE_UNEMP_2029) is stationary	Non stationary	
ADF (Constant and Trend)	RE_FDI	RE_FDI has a unit root	Non stationary	FDS

ADF (Constant and Trend)	RE_FDI	D(RE_FDI) has a unit root	Non stationary	
ADF (Constant)	RE_FDI	RE_FDI has a unit root	Non stationary	
ADF (Constant)	RE_FDI	D(RE_FDI) has a unit root	Stationary	
ADF (no exogenous)	RE_FDI	RE_FDI has a unit root	Non stationary	
ADF (no exogenous)	RE_FDI	D(RE_FDI) has a unit root	Stationary	
KPSS (Constant and Trend)	RE_FDI	RE_FDI is stationary	Non stationary	
KPSS (Constant and Trend)	RE_FDI	D(RE_FDI) is stationary	Stationary	
KPSS (Constant)	RE_FDI	RE_FDI is stationary	Stationary	
KPSS (Constant)	RE_FDI	D(RE_FDI) is stationary	Stationary	
ADF (Constant and Trend)	RE_FDI_TO_GDP	RE_FDI_TO_GDP has a unit root	Non stationary	FDS
ADF (Constant and Trend)	RE_FDI_TO_GDP	D(RE_FDI_TO_GDP) has a unit root	Non stationary	
ADF (Constant)	RE_FDI_TO_GDP	RE_FDI_TO_GDP has a unit root	Non stationary	
ADF (Constant)	RE_FDI_TO_GDP	D(RE_FDI_TO_GDP) has a unit root	Non stationary	
ADF (no exogenous)	RE_FDI_TO_GDP	RE_FDI_TO_GDP has a unit root	Non stationary	
ADF (no exogenous)	RE_FDI_TO_GDP	D(RE_FDI_TO_GDP) has a unit root	Stationary	
KPSS (Constant and Trend)	RE_FDI_TO_GDP	RE_FDI_TO_GDP is stationary	Non stationary	
KPSS (Constant and Trend)	RE_FDI_TO_GDP	D(RE_FDI_TO_GDP) is stationary	Stationary	
KPSS (Constant)	RE_FDI_TO_GDP	RE_FDI_TO_GDP is stationary	Stationary	
KPSS (Constant)	RE_FDI_TO_GDP	D(RE_FDI_TO_GDP) is stationary	Stationary	
ADF (Constant and Trend)	RE_HOUSEH_WEALTH	RE_HOUSEH_WEALTH has a unit root	Non stationary	SDS
ADF (Constant and Trend)	RE_HOUSEH_WEALTH	D(RE_HOUSEH_WEALTH) has a unit root	Non stationary	
ADF (Constant)	RE_HOUSEH_WEALTH	RE_HOUSEH_WEALTH has a unit root	Stationary	
ADF (Constant)	RE_HOUSEH_WEALTH	D(RE_HOUSEH_WEALTH) has a unit root	Non stationary	
ADF (no exogenous)	RE_HOUSEH_WEALTH	RE_HOUSEH_WEALTH has a unit root	Non stationary	
ADF (no exogenous)	RE_HOUSEH_WEALTH	D(RE_HOUSEH_WEALTH) has a unit root	Non stationary	
KPSS (Constant and Trend)	RE_HOUSEH_WEALTH	RE_HOUSEH_WEALTH is stationary	Non stationary	
KPSS (Constant and Trend)	RE_HOUSEH_WEALTH	D(RE_HOUSEH_WEALTH) is stationary	Non stationary	
KPSS (Constant)	RE_HOUSEH_WEALTH	RE_HOUSEH_WEALTH is stationary	Non stationary	
KPSS (Constant)	RE_HOUSEH_WEALTH	D(RE_HOUSEH_WEALTH) is stationary	Stationary	
ADF (Constant and Trend)	RES_RPRICE_INX_COAST	RES_RPRICE_INX_COAST has a unit root	Non stationary	SDS
ADF (Constant and Trend)	RES_RPRICE_INX_COAST	D(RES_RPRICE_INX_COAST) has a unit root	Non stationary	
ADF (Constant)	RES_RPRICE_INX_COAST	RES_RPRICE_INX_COAST has a unit root	Non stationary	
ADF (Constant)	RES_RPRICE_INX_COAST	D(RES_RPRICE_INX_COAST) has a unit root	Non stationary	
ADF (no exogenous)	RES_RPRICE_INX_COAST	RES_RPRICE_INX_COAST has a unit root	Non stationary	
ADF (no exogenous)	RES_RPRICE_INX_COAST	D(RES_RPRICE_INX_COAST) has a unit root	Non stationary	
KPSS (Constant and Trend)	RES_RPRICE_INX_COAST	RES_RPRICE_INX_COAST is stationary	Non stationary	
KPSS (Constant and Trend)	RES_RPRICE_INX_COAST	D(RES_RPRICE_INX_COAST) is stationary	Non stationary	
KPSS (Constant)	RES_RPRICE_INX_COAST	RES_RPRICE_INX_COAST is stationary	Non stationary	
KPSS (Constant)	RES_RPRICE_INX_COAST	D(RES_RPRICE_INX_COAST) is stationary	Non stationary	
ADF (Constant and Trend)	RES_RPRICE_INX_MADBAR	RES_RPRICE_INX_MADBAR has a unit root	Non stationary	SDS
ADF (Constant and Trend)	RES_RPRICE_INX_MADBAR	D(RES_RPRICE_INX_MADBAR) has a unit root	Non stationary	

ADF (Constant)	RES_RPRICE_INX_MADBAR	RES_RPRICE_INX_MADBAR has a unit root	Non stationary	
ADF (Constant)	RES_RPRICE_INX_MADBAR	D(RES_RPRICE_INX_MADBAR) has a unit root	Non stationary	
ADF (no exogenous)	RES_RPRICE_INX_MADBAR	RES_RPRICE_INX_MADBAR has a unit root	Non stationary	
ADF (no exogenous)	RES_RPRICE_INX_MADBAR	D(RES_RPRICE_INX_MADBAR) has a unit root	Non stationary	
KPSS (Constant and Trend)	RES_RPRICE_INX_MADBAR	RES_RPRICE_INX_MADBAR is stationary	Non stationary	
KPSS (Constant and Trend)	RES_RPRICE_INX_MADBAR	D(RES_RPRICE_INX_MADBAR) is stationary	Non stationary	
KPSS (Constant)	RES_RPRICE_INX_MADBAR	RES_RPRICE_INX_MADBAR is stationary	Non stationary	
KPSS (Constant)	RES_RPRICE_INX_MADBAR	D(RES_RPRICE_INX_MADBAR) is stationary	Non stationary	
ADF (Constant and Trend)	RES_RPRICE_INX_REST	RES_RPRICE_INX_REST has a unit root	Non stationary	SDS
ADF (Constant and Trend)	RES_RPRICE_INX_REST	D(RES_RPRICE_INX_REST) has a unit root	Non stationary	
ADF (Constant)	RES_RPRICE_INX_REST	RES_RPRICE_INX_REST has a unit root	Non stationary	
ADF (Constant)	RES_RPRICE_INX_REST	D(RES_RPRICE_INX_REST) has a unit root	Non stationary	
ADF (no exogenous)	RES_RPRICE_INX_REST	RES_RPRICE_INX_REST has a unit root	Non stationary	
ADF (no exogenous)	RES_RPRICE_INX_REST	D(RES_RPRICE_INX_REST) has a unit root	Non stationary	
KPSS (Constant and Trend)	RES_RPRICE_INX_REST	RES_RPRICE_INX_REST is stationary	Non stationary	
KPSS (Constant and Trend)	RES_RPRICE_INX_REST	D(RES_RPRICE_INX_REST) is stationary	Non stationary	
KPSS (Constant)	RES_RPRICE_INX_REST	RES_RPRICE_INX_REST is stationary	Non stationary	
KPSS (Constant)	RES_RPRICE_INX_REST	D(RES_RPRICE_INX_REST) is stationary	Non stationary	
ADF (Constant and Trend)	RPRICEINDEX	RPRICEINDEX has a unit root	Non stationary	SDS
ADF (Constant and Trend)	RPRICEINDEX	D(RPRICEINDEX) has a unit root	Non stationary	
ADF (Constant)	RPRICEINDEX	RPRICEINDEX has a unit root	Non stationary	
ADF (Constant)	RPRICEINDEX	D(RPRICEINDEX) has a unit root	Non stationary	
ADF (no exogenous)	RPRICEINDEX	RPRICEINDEX has a unit root	Non stationary	
ADF (no exogenous)	RPRICEINDEX	D(RPRICEINDEX) has a unit root	Non stationary	
KPSS (Constant and Trend)	RPRICEINDEX	RPRICEINDEX is stationary	Non stationary	
KPSS (Constant and Trend)	RPRICEINDEX	D(RPRICEINDEX) is stationary	Non stationary	
KPSS (Constant)	RPRICEINDEX	RPRICEINDEX is stationary	Non stationary	
KPSS (Constant)	RPRICEINDEX	D(RPRICEINDEX) is stationary	Stationary	
ADF (Constant and Trend)	STOCK	STOCK has a unit root	Non stationary	SDS
ADF (Constant and Trend)	STOCK	D(STOCK) has a unit root	Non stationary	
ADF (Constant)	STOCK	STOCK has a unit root	Non stationary	
ADF (Constant)	STOCK	D(STOCK) has a unit root	Non stationary	
ADF (no exogenous)	STOCK	STOCK has a unit root	Non stationary	
ADF (no exogenous)	STOCK	D(STOCK) has a unit root	Non stationary	
KPSS (Constant and Trend)	STOCK	STOCK is stationary	Non stationary	
KPSS (Constant and Trend)	STOCK	D(STOCK) is stationary	Non stationary	
KPSS (Constant)	STOCK	STOCK is stationary	Non stationary	
KPSS (Constant)	STOCK	D(STOCK) is stationary	Stationary	
ADF (Constant and Trend)	STOCK_COLEGIOAR	STOCK_COLEGIOAR has a unit root	Non stationary	SDS
ADF (Constant and Trend)	STOCK_COLEGIOAR	D(STOCK_COLEGIOAR) has a unit root	Non stationary	
ADF (Constant)	STOCK_COLEGIOAR	STOCK_COLEGIOAR has a unit root	Non stationary	
ADF (Constant)	STOCK_COLEGIOAR	D(STOCK_COLEGIOAR) has a unit root	Non stationary	

ADF (no exogenous)	STOCK_COLEGIOAR	STOCK_COLEGIOAR has a unit root	Non stationary	
ADF (no exogenous)	STOCK_COLEGIOAR	D(STOCK_COLEGIOAR) has a unit root	Non stationary	
KPSS (Constant and Trend)	STOCK_COLEGIOAR	STOCK_COLEGIOAR is stationary	Stationary	
KPSS (Constant and Trend)	STOCK_COLEGIOAR	D(STOCK_COLEGIOAR) is stationary	Non stationary	
KPSS (Constant)	STOCK_COLEGIOAR	STOCK_COLEGIOAR is stationary	Non stationary	
KPSS (Constant)	STOCK_COLEGIOAR	D(STOCK_COLEGIOAR) is stationary	Stationary	
ADF (Constant and Trend)	WEALTH_TO_GDP	WEALTH_TO_GDP has a unit root	Non stationary	FDS
ADF (Constant and Trend)	WEALTH_TO_GDP	D(WEALTH_TO_GDP) has a unit root	Stationary	
ADF (Constant)	WEALTH_TO_GDP	WEALTH_TO_GDP has a unit root	Non stationary	
ADF (Constant)	WEALTH_TO_GDP	D(WEALTH_TO_GDP) has a unit root	Stationary	
ADF (no exogenous)	WEALTH_TO_GDP	WEALTH_TO_GDP has a unit root	Non stationary	
ADF (no exogenous)	WEALTH_TO_GDP	D(WEALTH_TO_GDP) has a unit root	Stationary	
KPSS (Constant and Trend)	WEALTH_TO_GDP	WEALTH_TO_GDP is stationary	Stationary	
KPSS (Constant and Trend)	WEALTH_TO_GDP	D(WEALTH_TO_GDP) is stationary	Stationary	
KPSS (Constant)	WEALTH_TO_GDP	WEALTH_TO_GDP is stationary	Non stationary	
KPSS (Constant)	WEALTH_TO_GDP	D(WEALTH_TO_GDP) is stationary	Stationary	
ADF (Constant and Trend)	YIELD_HOUSING	YIELD_HOUSING has a unit root	Non stationary	Stationary
ADF (Constant and Trend)	YIELD_HOUSING	D(YIELD_HOUSING) has a unit root	Non stationary	
ADF (Constant)	YIELD_HOUSING	YIELD_HOUSING has a unit root	Stationary	
ADF (Constant)	YIELD_HOUSING	D(YIELD_HOUSING) has a unit root	Non stationary	
ADF (no exogenous)	YIELD_HOUSING	YIELD_HOUSING has a unit root	Non stationary	
ADF (no exogenous)	YIELD_HOUSING	D(YIELD_HOUSING) has a unit root	Non stationary	
KPSS (Constant and Trend)	YIELD_HOUSING	YIELD_HOUSING is stationary	Non stationary	
KPSS (Constant and Trend)	YIELD_HOUSING	D(YIELD_HOUSING) is stationary	Non stationary	
KPSS (Constant)	YIELD_HOUSING	YIELD_HOUSING is stationary	Non stationary	
KPSS (Constant)	YIELD_HOUSING	D(YIELD_HOUSING) is stationary	Stationary	

6. *Unit root tests with structural break*

H0: Variable has a unit root

Conventions to read the following tables:
R: Rejects H0 at a 95% of level of confidence
R*: Rejects H0 at a 90% of level of confidence
A: Does not rejects H0

Variable: First difference of Real average house price (£sqm), at prices of 2008

Perron unit root test with structural break

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	A			A			A		
1	A			A			A		
2	A			A			A		
3	R	3	2008 Q1	R*	3	1998 Q4	R	3	2008 Q1
4	R	3	2008 Q1	R*	3	1998 Q4	R	3	2008 Q1
5	R	3	2008 Q1	R*	3	1998 Q4	R	3	2008 Q1
6	R	3	2008 Q1	R*	3	1998 Q4	R	3	2008 Q1
7	R	3	2008 Q1	A			R	7	2008 Q1
8	R*	7	2008 Q1	A			A		
9	A			A			A		
10	A			A			A		

Zivot-Andrews Test

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	A			A			A		
1	A			A			A		
2	A			A			A		
3	R	3	2008 Q2	A			R	3	2008 Q2
4	A			A			R	4	2008 Q2
5	A			A			R	4	2008 Q2
6	A			A			R	4	2008 Q2
7	R	7	2008 Q2	A			R	7	2008 Q2
8	A			A			A		
9	A			A			A		

10	A			A			A		

Variable: First difference of Real Gross Domestic Product per Capita (₺, at prices of 2008

Perron unit root test with structural break

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date
0	A			A	0		A	0	
1	A			A	1		A	0	
2	A			A	0		A	0	
3	R	3	2008 Q1	R	3	2008 Q1	R*	3	1998 Q4
4	R	3	2008 Q1	R	3	2008 Q1	R*	3	1998 Q4
5	R	3	2008 Q1	R	3	2008 Q1	R*	3	1998 Q4
6	R	3	2008 Q1	R	3	2008 Q1	R*	3	1998 Q4
7	R	3	2008 Q1	R	7	2008 Q1	A	7	
8	R	7	2008 Q1	A	8		A	8	
9	A			A	8		A	8	
10	A			A	8		A	8	

Zivot-Andrews Test

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	A			A			A		
1	A			A			A		
2	A			A			A		
3	R	3	2008 Q2	A			R	3	2008 Q2
4	A			A			R	4	2008 Q2
5	A			A			R	4	2008 Q2
6	A			A			R	4	2008 Q2
7	R	7	2008 Q2	A			R	7	2008 Q2
8	A			A			A		
9	A			A			A		
10	A			A			A		

Variable: First difference of Gross Capital Formation in real estate (₺, at prices of 2008

Perron unit root test with structural break

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	R*	0		A			A		
1	R*	0		A			A		
2	R*	0		A			A		
3	R*	0		A			A		
4	R*	0		A			A		
5	R*	0		A			A		
6	R*	0		A			A		
7	R*	0		A			A		
8	R*	0		A			A		
9	R*	0		A			A		
10	R*	0		A			A		

Zivot-Andrews Test

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	A			A			A		
1	A			A			A		
2	A			A			A		
3	A			A			R	0	2008 Q1
4	R	0	2007 Q1	A			R	0	2008 Q1
5	R	0	2007 Q1	A			R	0	2008 Q1
6	A			A			A		
7	R	0	2007 Q1	A			R	0	2008 Q1
8	R	0	2007 Q1	A			R	0	2008 Q1
9	R	0	2007 Q1	A			R	0	2008 Q1
10	R	0	2007 Q1	A			R	0	2008 Q1

Variable: First difference of Real Gross Domestic (£, at prices of 2008

Perron unit root test with structural break

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	A			A			A		
1	A			A			A		
2	A			A			A		
3	R	3	2008 Q1	A			R	3	2008 Q1

4	R	3	2008 Q1	A			R	3	2008 Q1
5	R	4	2008 Q1	A			R	4	2008 Q1
6	R	3	2008 Q1	A			R	3	2008 Q1
7	R	7	2008 Q2	A			R	7	2008 Q2
8	R	8	2008 Q1	A			R	8	2008 Q1
9	R	8	2008 Q1	A			R	8	2008 Q2
10	R	8	2008 Q1	A			R	8	2008 Q2

Zivot-Andrews Test

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	A			A			A		
1	A			A			A		
2	A			A			A		
3	R	3	2008 Q2	R	3	2006 Q1	R	3	2008 Q2
4	R	4	2008 Q2	A			R	4	2008 Q2
5	R	5	2008 Q2	A			R	5	2008 Q2
6	R	4	2008 Q2	A			R	4	2008 Q2
7	R	7	2008 Q2	A			R	7	2008 Q2
8	R	8	2008 Q2	A			R	8	2008 Q2
9	R	8	2008 Q2	A			R	8	2008 Q2
10	R	8	2008 Q2	A			R	8	2008 Q2

Variable: First difference of Apparent concrete consumption (000 metric tons)

Perron unit root test with structural break

Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0	R	0	2006 Q2	R	0	1998 Q3	R	0	2007 Q2
1	R	1	2007 Q4	R	1	2000 Q4	R	1	2007 Q4
2	R	1	2007 Q4	R	1	2000 Q4	R	1	2007 Q4
3	R	3	2007 Q3	A			A		
4	R	3	2007 Q3	A			A		
5	R	0	2006 Q2	R	0	1998 Q3	R	0	2007 Q2
6	R	6	2007 Q3	A			R	6	2007 Q4
7	R	0	2006 Q2	R	0	1998 Q3	R	0	2007 Q2
8	R	0	2006 Q2	R	0	1998 Q3	R	0	2007 Q2
9	R	0	2006 Q2	R	0	1998 Q3	R	0	2007 Q2
10	A			A			A		

Zivot-Andrews Test									
Maximum lags used in test:	Shock in: Intercept	Lag chosen	Shock date	Shock in: Trend	Lag chosen	Shock date	Shock in: Intercept and trend	Lag chosen	Shock date
0									
1	R	1	2007 Q4	R	1	2001 Q3	R	1	2008 Q1
2	R	2	2007 Q4	R	2	1998 Q4	R	2	2008 Q1
3	A			A			A		
4	A			A			A		
5	A			A			A		
6	R	6	2008 Q1	A			R	6	2008 Q1
7	A			A			A		
8	R	6	2008 Q1	A			R	6	2008 Q1
9	A			A			A		
10	A			A			A		

7. Cointegration Tests

Structural Modelling

Cointegration Test - Engle-Granger

Equation: PRICE_LR

Specification: LOG(HOUSE_PRICE_M2) LOG(GDPPC(-0))
LOG(MORTG_RATE(-0)) LOG(BULD_STRT_FREE(-0))
LOG(GCF_DWELL(-0)) C DUMMY

Cointegrating equation deterministics: C DUMMY

Null hypothesis: Series are not cointegrated

Automatic lag specification (lag=0 based on Schwarz Info Criterion, maxlag=11)

	Value	Prob.*
Engle-Granger tau-statistic	-5.307134	0.0093
Engle-Granger z-statistic	-39.78611	0.0091

*MacKinnon (1996) p-values.

Warning: p-values do not account for user-specified deterministic regressors.

Intermediate Results:

Rho - 1	-0.568373
Rho S.E.	0.107096
Residual variance	0.000386
Long-run residual variance	0.000386
Number of lags	0
Number of observations	70
Number of stochastic trends**	5

**Number of stochastic trends in asymptotic distribution.

Engle-Granger Test Equation:

Dependent Variable: D(RESID)

Method: Least Squares

Sample (adjusted): 1995Q2 2012Q3

Included observations: 70 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID(-1)	-0.568373	0.107096	-5.307134	0.0000

R-squared	0.289715	Mean dependent var	0.000344
Adjusted R-squared	0.289715	S.D. dependent var	0.023313
S.E. of regression	0.019648	Akaike info criterion	-5.007540
Sum squared resid	0.026636	Schwarz criterion	-4.975419
Log likelihood	176.2639	Hannan-Quinn criter.	-4.994781
Durbin-Watson stat	1.907508		

Automatic selected model with GASIC

Cointegration Test - Engle-Granger

Equation: LR_PRICE_GASIC

Specification: LOG(PRICE) LOG(EFFORT_NO_DEDUCT)

LOG(CONCRETE_CONSUM) LOG(MORTG_RATE) LOG(GDP_2008) C

Cointegrating equation deterministics: C

Null hypothesis: Series are not cointegrated

Automatic lag specification (lag=0 based on Schwarz Info Criterion,

maxlag=11)

	Value	Prob.*
Engle-Granger tau-statistic	-4.760394	0.0360
Engle-Granger z-statistic	-34.72545	0.0305

*MacKinnon (1996) p-values.

Intermediate Results:

Rho - 1	-0.489091
Rho S.E.	0.102742
Residual variance	0.000272
Long-run residual variance	0.000272
Number of lags	0
Number of observations	71
Number of stochastic trends**	5

**Number of stochastic trends in asymptotic distribution.

Engle-Granger Test Equation:

Dependent Variable: D(RESID)

Method: Least Squares

Sample (adjusted): 1995Q2 2012Q4

Included observations: 71 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID(-1)	-0.489091	0.102742	-4.760394	0.0000

R-squared	0.244558	Mean dependent var	-3.89E-05
Adjusted R-squared	0.244558	S.D. dependent var	0.018967
S.E. of regression	0.016485	Akaike info criterion	-5.358732
Sum squared resid	0.019023	Schwarz criterion	-5.326864
Log likelihood	191.2350	Hannan-Quinn criter.	-5.346059
Durbin-Watson stat	2.031471		

8. Long run estimation output - Structural modelling

Dependent Variable: LOG(HOUSE_PRICE_M2)

Method: Fully Modified Least Squares (FMOLS)

Sample (adjusted): 1995Q2 2012Q3

Included observations: 70 after adjustments

Cointegrating equation deterministics: C DUMMY

Long-run covariance estimate (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(GDPPC)	1.56439	0.077833	20.39477	0.0000
LOG(MORTG_RATE)	-0.13980	0.016371	-7.234887	0.0000
LOG(BULD_STRT_FREE)	-0.07348	0.008005	-10.78238	0.0000
LOG(GCF_DWELL)	0.50291	0.029623	16.99713	0.0000
C	-10.17855	0.679104	-15.14194	0.0000
DUMMY	0.07240	0.016133	5.227061	0.0000

R-squared	0.992640	Mean dependent var	7.314782
Adjusted R-squared	0.992066	S.D. dependent var	0.260815
S.E. of regression	0.023232	Sum squared resid	0.034543
Durbin-Watson stat	1.080153	Long-run variance	0.000784

9. Short run estimation output - Structural modelling

Dependent Variable: DLOG(HOUSE_PRICE_M2)

Method: Least Squares

Sample (adjusted): 1996Q2 2012Q4

Included observations: 67 after adjustments

Variable	Coefficient	Std. Error	t-Statistic
LOG(HOUSE_PRICE_M2(-1))-1.56446145578*LOG(GDPPC(-1))+0.139817496438*LOG(MORTG_RATE(-1))+0.0735931696716*LOG(BULD_STRT_FREE(-1))-0.502995059157*LOG(GCF_DWELL(-1))+10.1785842463-0.072496921077*DUMMY	-0.287858	0.059037	-4.875900
DLOG(HOUSE_PRICE_M2(-4))	0.627044	0.064479	9.724857
DLOG(MORTG_RATE(-2))	-0.040475	0.017333	-2.335226
DLOG(BULD_STRT_FREE(-4))	0.018552	0.008745	2.121359
DLOG(BULD_STRT_FREE(-2))	0.018253	0.008372	2.180214

R-squared	0.829447	Mean dependent var	
Adjusted R-squared	0.818443	S.D. dependent var	
S.E. of regression	0.009897	Akaike info criterion	
Sum squared resid	0.006073	Schwarz criterion	
Log likelihood	216.7713	Hannan-Quinn criter.	
Durbin-Watson stat	1.907415		

10. Long run estimation output - GETS modelling

Dependent Variable: LOG(PRICE)

Method: Fully Modified Least Squares (FMOLS)

Date: 05/31/14 Time: 12:29

Sample (adjusted): 1995Q2 2012Q4

Included observations: 71 after adjustments

Cointegrating equation deterministics: C

Long-run covariance estimate (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(EFFORT_NO_DEDUCT)	0.639301	0.029548	21.63627	0.0000
LOG(CONCRETE_CONSUM)	0.160411	0.009314	17.22280	0.0000
LOG(MORTG_RATE)	-0.332009	0.022267	-14.91011	0.0000
LOG(GDP_2008)	0.693200	0.060472	11.46309	0.0000
C	-4.298316	0.667614	-6.438330	0.0000

R-squared	0.994379	Mean dependent var	7.315029
Adjusted R-squared	0.994038	S.D. dependent var	0.258954
S.E. of regression	0.019995	Sum squared resid	0.026387
Durbin-Watson stat	1.025831	Long-run variance	0.000740

11. Long run estimation output - GETS modelling with a dummy variable recognizing structural break

Dependent Variable: LOG(PRICE)

Method: Fully Modified Least Squares (FMOLS)

Date: 07/26/14 Time: 13:03

Sample (adjusted): 1995Q2 2012Q4

Included observations: 71 after adjustments

Cointegrating equation deterministics: C DUMMY

Long-run covariance estimate (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(EFFORT_NO_DEDUCT)	0.598760	0.041985	14.26141	0.0000
LOG(CONCRETE_CONSUM)	0.161496	0.009423	17.13758	0.0000
LOG(MORTG_RATE)	-0.309612	0.026452	-11.70448	0.0000
LOG(GDP_2008)	0.681667	0.061668	11.05378	0.0000
C	-4.076790	0.700487	-5.819933	0.0000
DUMMY	0.021420	0.018754	1.142145	0.2576

R-squared	0.994737	Mean dependent var	7.315029
Adjusted R-squared	0.994332	S.D. dependent var	0.258954
S.E. of regression	0.019495	Sum squared resid	0.024704
Durbin-Watson stat	1.034213	Long-run variance	0.000727

12. Short run estimation output - GETS Modelling with Error Correction Mechanism

Dependent Variable: DLOG(PRICE)

Method: Least Squares

Date: 09/09/14 Time: 16:48

Sample (adjusted): 1996Q2 2012Q4

Included observations: 67 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(PRICE(-1))-0.639301292294*LOG(EFFORT_NO_DEDUCT(-1))-0.160410998866*LOG(CONCRETE_CONSUM(-1))+0.332008749696*LOG(MORTG_RATE(-1))-0.693200026756*LOG(GDP_2008(-1))+4.2983159666	-0.096204	0.072995	-1.317948	0.1925
DLOG(PRICE(-4))	0.683577	0.083669	8.170039	0.0000
DLOG(PRICE(-3))	0.166395	0.073555	2.262187	0.0273
DLOG(CONCRETE_CONSUM(-4))	0.049039	0.016365	2.996613	0.0040
DLOG(MORTG_RATE(-2))	-0.051025	0.017925	-2.846522	0.0060
DLOG(GDP_2008(-3))	-0.800466	0.351669	-2.276190	0.0264
DLOG(GDP_2008(-1))	0.915952	0.342165	2.676931	0.0096

R-squared	0.787273	Mean dependent var	0.005706
Adjusted R-squared	0.766000	S.D. dependent var	0.023227
S.E. of regression	0.011236	Akaike info criterion	-6.040868
Sum squared resid	0.007574	Schwarz criterion	-5.810527
Log likelihood	209.3691	Hannan-Quinn criter.	-5.949722
Durbin-Watson stat	1.814639		

13. Short run estimation output - GETS Modelling without Error Correction Mechanism

Dependent Variable: DLOG(PRICE)

Method: Least Squares

Date: 09/25/14 Time: 19:35

Sample (adjusted): 1996Q2 2012Q4

Included observations: 67 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOG(PRICE(-4))	0.676891	0.084018	8.056514	0.0000
DLOG(PRICE(-3))	0.165995	0.073997	2.243262	0.0285
DLOG(CONCRETE_CONSUM(-4))	0.045999	0.016299	2.822171	0.0064
DLOG(MORTG_RATE(-2))	-0.055645	0.017685	-3.146402	0.0026
DLOG(GDP_2008(-3))	-0.785412	0.353600	-2.221185	0.0301
DLOG(GDP_2008(-1))	0.913845	0.344222	2.654811	0.0101

R-squared	0.781115	Mean dependent var	0.005706
Adjusted R-squared	0.763173	S.D. dependent var	0.023227
S.E. of regression	0.011303	Akaike info criterion	-6.042180
Sum squared resid	0.007793	Schwarz criterion	-5.844745
Log likelihood	208.4130	Hannan-Quinn criter.	-5.964055
Durbin-Watson stat	1.892734		

Annex to chapter 2

System residual autocorrelations

Portmanteau autocorrelation test

Null hypothesis: No residual autocorrelations up to lag h

Sample: 2001Q1 – 2015Q2

Included observations: 58

GDP as exogenous variable

Estimation method: 2SECM

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	9.001786	0.4371	9.159712	0.4227	9
2	19.40110	0.3675	19.93043	0.3368	18
3	28.70528	0.3753	29.74212	0.3259	27
4	37.27710	0.4101	38.94888	0.3385	36
5	57.13987	0.1059	60.68549	0.0592	45
6	60.88383	0.2420	64.86145	0.1479	54
7	71.93890	0.2060	77.43389	0.1043	63
8	86.39331	0.1185	94.20100	0.0407	72
9	92.24112	0.1848	101.1229	0.0645	81
10	99.16224	0.2387	109.4859	0.0796	90
11	108.7259	0.2367	121.2879	0.0636	99
12	118.7584	0.2254	133.9375	0.0460	108

SEMP as exogenous variable

Estimation method: 2SECM

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	7.689030	0.5658	7.823925	0.5520	9
2	18.13213	0.4470	18.63999	0.4143	18
3	29.97563	0.3152	31.12950	0.2659	27
4	39.22844	0.3272	41.06770	0.2581	36
5	58.26080	0.0887	61.89557	0.0479	45

6	64.91047	0.1469	69.31251	0.0783	54
7	72.54845	0.1922	77.99884	0.0966	63
8	84.77159	0.1441	92.17768	0.0548	72
9	95.99469	0.1222	105.4622	0.0353	81
10	100.5885	0.2091	111.0130	0.0658	90
11	113.7219	0.1479	127.2202	0.0295	99
12	125.0850	0.1248	141.5476	0.0168	108

GDP as exogenous variable

Estimation method: SEECM

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	9.715002	0.3740	9.885440	0.3598	9
2	21.99419	0.2322	22.60317	0.2063	18
3	32.00111	0.2320	33.15593	0.1920	27
4	39.30304	0.3242	40.99874	0.2605	36
5	57.18503	0.1051	60.56771	0.0604	45
6	66.70903	0.1148	71.19063	0.0584	54
7	81.42039	0.0592	87.92120	0.0208	63
8	91.63444	0.0592	99.76950	0.0169	72
9	95.50253	0.1294	104.3480	0.0414	81
10	105.6801	0.1238	116.6460	0.0310	90
11	117.1223	0.1032	130.7661	0.0179	99
12	130.0781	0.0728	147.1017	0.0074	108

SEMP as exogenous variable

Estimation method: SEECM

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	13.11258	0.1576	13.34262	0.1477	9
2	22.41522	0.2141	22.97750	0.1915	18
3	27.86585	0.4179	28.72544	0.3743	27
4	45.54019	0.1324	47.70899	0.0917	36
5	58.58720	0.0841	61.98685	0.0472	45
6	69.75698	0.0732	74.44545	0.0340	54
7	77.79577	0.0993	83.58761	0.0424	63
8	86.40273	0.1184	93.57167	0.0447	72
9	91.60669	0.1973	99.73147	0.0774	81
10	102.5244	0.1729	112.9237	0.0515	90
11	116.8509	0.1063	130.6032	0.0183	99
12	122.9501	0.1542	138.2936	0.0262	108

Lag order selection

Exogenous variable: Spanish GDP

Variables: LOG(RENT) LOG(GDP) LOG(STOCK)

Exogenous variables: C

Sample: 2001Q1 2015Q2

Included observations: 58

Lag	LogL	LR	FPE	AIC	SC	HQ
0	233.6850	NA	7.05e-08	-7.954654	-7.848079	-7.913141
1	560.5647	608.6726	1.22e-12	-18.91602	-18.48973	-18.74997
2	618.7488	102.3238	2.25e-13	-20.61203	-19.86601*	-20.32144
3	627.0284	13.70404	2.32e-13	-20.58718	-19.52144	-20.17206
4	636.8266	15.20409	2.29e-13	-20.61471	-19.22924	-20.07504
5	648.2656	16.56687	2.15e-13	-20.69881	-18.99362	-20.03461
6	671.6561	31.45624	1.35e-13	-21.19504	-19.17012	-20.40629
7	688.6786	21.13130	1.07e-13	-21.47167	-19.12703	-20.55839
8	706.7204	20.53037*	8.33e-14*	-21.78346*	-19.11910	-20.74564*

* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Exogeneous variable: Madrid’s service sector employment

Variables: LOG(RENT) LOG(STOCK) LOG(SEMP)

Exogenous variables: C

Sample: 2001Q1 2015Q2

Included observations: 58

Lag	LogL	LR	FPE	AIC	SC	HQ

0	230.2857	NA	7.92e-08	-7.837438	-7.730864	-7.795925
1	498.5961	499.6124	1.04e-11	-16.77918	-16.35288	-16.61312
2	528.7028	52.94621	5.02e-12	-17.50699	-16.76097*	-17.21640*
3	538.9220	16.91456	4.85e-12	-17.54903	-16.48329	-17.13390
4	545.7848	10.64924	5.29e-12	-17.47534	-16.08987	-16.93567
5	555.0545	13.42506	5.35e-12	-17.48464	-15.77944	-16.82043
6	570.1167	20.25608	4.48e-12	-17.69368	-15.66876	-16.90493
7	589.3215	23.84041*	3.29e-12	-18.04557	-15.70093	-17.13228
8	602.0431	14.47636	3.08e-12*	-18.17390*	-15.50954	-17.13608

* indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Annex to chapter 3

1. Automatic General to Specific linear regression output

Number of observations = 3,912

F (27, 3884) = 207.01

Prob > F = 0.0000

R-squared = 0.5915

Root MSE = 0.22737

Robust estimation: yes

Dependant variable: Logarithm of real rent = log (RENT) = rent

	Coefficient	Std. Error	t-statistic	p-value	Conf. interval (95%)	
CONSTANT	2.760828	0.0202591	136.28	0	2.721109	2.800548
CBD	0.5687084	0.0142138	40.01	0	0.5408412	0.5965756
CENTRE	0.3760973	0.012919	29.11	0	0.3507686	0.401426
DEC	0.1688657	0.0115259	14.65	0	0.1462684	0.1914631
AGE	-0.0012464	0.0002061	-6.05	0	-0.0016504	-0.0008423
FLOORS	0.0019627	0.0005471	3.59	0	0.0008902	0.0030353
EXCLUSIVE	0.0760644	0.0087212	8.72	0	0.0589657	0.093163
QUALITY	-0.0499036	0.0041166	-12.12	0	-0.0579745	-0.0418327
METRO	-0.0000101	2.01E-06	-5	0	-0.000014	-6.11E-06
CORPORATE	0.0924265	0.0082755	11.17	0	0.0762017	0.1086513
H1 2004	-0.1071558	0.0161099	-6.65	0	-0.1387404	-0.0755711
H2 2004	-0.1246149	0.016228	-7.68	0	-0.156431	-0.0927987
H1 2005	-0.1306223	0.0150875	-8.66	0	-0.1602025	-0.1010421
H2 2005	-0.0982072	0.0161611	-6.08	0	-0.1298923	-0.0665221
H1 2006	-0.0688543	0.0147592	-4.67	0	-0.0977907	-0.0399178
H2 2006	-0.0600422	0.0174457	-3.44	0.001	-0.0942459	-0.0258385
H1 2008	0.0606914	0.0185287	3.28	0.001	0.0243646	0.0970183
H1 2009	-0.091502	0.0236832	-3.86	0	-0.1379348	-0.0450693
H2 2009	-0.147671	0.0208115	-7.1	0	-0.1884735	-0.1068685

H1 2010	-0.1877481	0.0179521	-10.46	0	-0.2229444	-0.1525517
H2 2010	-0.2463606	0.0217394	-11.33	0	-0.2889823	-0.2037388
H1 2011	-0.2523337	0.0217579	-11.6	0	-0.2949918	-0.2096757
H2 2011	-0.3103149	0.0231312	-13.42	0	-0.3556654	-0.2649645
H1 2012	-0.3408938	0.0205219	-16.61	0	-0.3811285	-0.300659
H2 2012	-0.4081387	0.02218	-18.4	0	-0.4516243	-0.3646532
H1 2013	-0.4471181	0.0253825	-17.62	0	-0.4968824	-0.3973538
H2 2013	-0.4417805	0.0263821	-16.75	0	-0.4935046	-0.3900563
H1 2014	-0.4745045	0.0270087	-17.57	0	-0.5274571	-0.4215519

2. Spatial lag regression output

Number of observations = 3,912

Wald Chi-squared (29) = 3,091.155

Prob > Chi-squared = 0.000

Variance ratio = 0.665

Squared correlation = 0.588

Sigma = 0.22

Log likelihood = 451.22263

Dependant variable: Logarithm of real rent = log (RENT) = rent

	Coefficient	Std. Error	t-statistic	p-value	Conf. interval (95%)	
CONSTANT	0.4133777	0.1066221	3.88	0	0.2044021	0.6223532
ρ-hat	0.8862797	0.039695	22.33	0	0.8084788	0.9640805
CBD	0.275142	0.0188166	14.62	0	0.2382621	0.3120219
CENTRE	0.1243125	0.0168027	7.4	0	0.0913798	0.1572451
DEC	0.0268392	0.0125994	2.13	0.033	0.0021449	0.0515335
AGE	-0.0017148	0.0002166	-7.92	0	-0.0021394	-0.0012902
FLOORS	0.0025937	0.000639	4.06	0	0.0013413	0.0038462
EXCLUSIVE	0.0804345	0.0079265	10.15	0	0.0648988	0.0959702

QUALITY	-0.0464124	0.0037195	-12.48	0	-0.0537024	-0.0391224
STATELY	0.0272887	0.0132616	2.06	0.04	0.0012965	0.053281
CORPORATE	0.0877032	0.0084579	10.37	0	0.0711261	0.1042803
H2 2003	-0.0306982	0.0190277	-1.61	0.107	-0.0679919	0.0065955
H1 2014	-0.1325876	0.0162635	-8.15	0	-0.1644635	-0.1007117
H2 2004	-0.1551145	0.0177917	-8.72	0	-0.1899857	-0.1202434
H1 2005	-0.1564056	0.0162797	-9.61	0	-0.1883134	-0.1244979
H2 2005	-0.1217061	0.0170901	-7.12	0	-0.1552021	-0.0882101
H1 2006	-0.0915591	0.01665	-5.5	0	-0.1241925	-0.0589256
H2 2006	-0.0802829	0.0175088	-4.59	0	-0.1145996	-0.0459663
H1 2007	-0.0474218	0.0167209	-2.84	0.005	-0.0801942	-0.0146495
H1 2008	0.0447924	0.0175797	2.55	0.011	0.0103369	0.079248
H1 2009	-0.1089712	0.0223275	-4.88	0	-0.1527323	-0.0652101
H2 2009	-0.1688286	0.0208119	-8.11	0	-0.2096191	-0.128038
H1 2010	-0.2076707	0.0204165	-10.17	0	-0.2476862	-0.1676551
H2 2010	-0.26246	0.021733	-12.08	0	-0.3050558	-0.2198642
H1 2011	-0.2666867	0.022256	-11.98	0	-0.3103076	-0.2230657
H2 2011	-0.3290702	0.0210357	-15.64	0	-0.3702993	-0.287841
H1 2012	-0.3661958	0.0213011	-17.19	0	-0.4079452	-0.3244464
H2 2012	-0.4259155	0.0214511	-19.86	0	-0.4679588	-0.3838722
H1 2013	-0.4661288	0.0236676	-19.69	0	-0.5125165	-0.4197411
H2 2013	-0.4681388	0.0228572	-20.48	0	-0.512938	-0.4233396
H1 2014	-0.4845845	0.0232982	-20.8	0	-0.5302481	-0.4389209

Wald test of rho=0: Chi-squared (1) = 498.505 (0.000)

Likelihood ratio test of rho=0: Chi-squared (1) = 401.216 (0.000)

3. Spatial error regression output

Number of observations = 3,912

Wald Chi-squared (29) = 3,256.775

Prob > Chi-squared = 0.000

Variance ratio = 0.445

Squared correlation = 0.573

Sigma = 0.22

Log likelihood = 448.53737

Dependant variable: Logarithm of real rent = log (RENT) = rent

	Coefficient	Std. Error	t-statistic	p-value	Conf. interval (95%)	
λ-hat	0.9916425	0.0082181	120.67	0	0.9755354	1.00775
CBD	0.4492984	0.022416	20.04	0	0.4053639	0.4932329
CENTRE	0.2747866	0.0208695	13.17	0	0.2338832	0.3156901
DEC	0.1116968	0.0182638	6.12	0	0.0759004	0.1474931
AGE	-0.001693	0.0002195	-7.71	0	-0.0021231	-0.0012629
FLOORS	0.0211618	0.014243	1.49	0.137	-0.0067538	0.0490775
EXCLUSIVE	0.0032225	0.0006788	4.75	0	0.001892	0.004553
QUALITY	0.074848	0.0078913	9.48	0	0.0593813	0.0903147
STATELY	-0.0464086	0.0039385	-11.78	0	-0.0541279	-0.0386893
CORPORATE	0.0862867	0.0084597	10.2	0	0.069706	0.1028674
H1 2014	-0.123212	0.0155206	-7.94	0	-0.1536317	-0.0927922
H2 2004	-0.1443814	0.0170891	-8.45	0	-0.1778755	-0.1108874
H1 2005	-0.1488796	0.0154991	-9.61	0	-0.1792573	-0.118502
H2 2005	-0.1146997	0.0163409	-7.02	0	-0.1467274	-0.0826721
H1 2006	-0.0862971	0.0158776	-5.44	0	-0.1174165	-0.0551776
H2 2006	-0.0755207	0.0168521	-4.48	0	-0.1085502	-0.0424911
H1 2007	-0.0452595	0.0159941	-2.83	0.005	-0.0766074	-0.0139117
H1 2008	0.0483749	0.0168997	2.86	0.004	0.015252	0.0814978
H1 2009	-0.0996844	0.021807	-4.57	0	-0.1424253	-0.0569435
H2 2009	-0.1671023	0.0201788	-8.28	0	-0.2066521	-0.1275525
H1 2010	-0.1951967	0.0198414	-9.84	0	-0.2340851	-0.1563082
H2 2010	-0.2546432	0.021171	-12.03	0	-0.2961377	-0.2131488

H1 2011	-0.2552293	0.0217199	-11.75	0	-0.2977995	-0.2126591
H2 2011	-0.328258	0.0204303	-16.07	0	-0.3683005	-0.2882154
H1 2012	-0.3521208	0.0207256	-16.99	0	-0.3927422	-0.3114994
H2 2012	-0.4177566	0.0208528	-20.03	0	-0.4586273	-0.3768858
H1 2013	-0.4611991	0.0231292	-19.94	0	-0.5065315	-0.4158668
H2 2013	-0.4641555	0.0223341	-20.78	0	-0.5079295	-0.4203815
H1 2014	-0.4801344	0.0227655	-21.09	0	-0.524754	-0.4355147

Wald test of lambda=0: Chi-squared (1) = 1.5e+04 (0.000)

Likelihood ratio test of lambda=0: Chi-squared (1) = 398.423 (0.000)

4. Durbin model regression output

Number of observation = 3912

Wald Chi-squared (56) = 3,575.961

Prob > Chi-squared = 0.000

Variance ratio = 0.637

Squared correlation = 0.620

Sigma = 0.21

Log likelihood = 555.34371

Dependant variable: Logarithm of real rent = log (RENT) = rent

	Coefficient	Std. Error	t-statistic	p-value	Conf. interval (95%)	
CONSTANT	-0.29531	0.203381	-1.45	0.147	-0.69393	0.103311
ρ-hat	0.8862797	0.039695	22.33	0	0.8084788	0.9640805
CBD	0.32961	0.0291	11.33	0	0.272576	0.386645
CENTRE	0.201211	0.026769	7.52	0	0.148745	0.253677
DEC	0.065048	0.024781	2.62	0.009	0.016477	0.113618
AGE	-0.00179	0.000219	-8.17	0	-0.00222	-0.00136
STATELY	0.027925	0.014405	1.94	0.053	-0.00031	0.056159
FLOORS	0.002869	0.000694	4.13	0	0.001508	0.00423
EXCLYUSIVE	0.084385	0.008003	10.54	0	0.068701	0.10007
QUALITY	-0.04265	0.004003	-10.66	0	-0.0505	-0.03481
CORPORATE	0.082521	0.008274	9.97	0	0.066305	0.098737
H12004	-0.12855	0.015188	-8.46	0	-0.15832	-0.09878
H22004	-0.1508	0.016782	-8.99	0	-0.18369	-0.11791
H12005	-0.14553	0.015233	-9.55	0	-0.17539	-0.11567
H22005	-0.11891	0.016084	-7.39	0	-0.15044	-0.08739
H12006	-0.08466	0.015678	-5.4	0	-0.11539	-0.05394
H22006	-0.07992	0.016555	-4.83	0	-0.11236	-0.04747
H12007	-0.04175	0.015719	-2.66	0.008	-0.07256	-0.01094
H12008	0.054124	0.016562	3.27	0.001	0.021663	0.086586

H12009	-0.09393	0.021338	-4.4	0	-0.13576	-0.05211
H22009	-0.16221	0.019848	-8.17	0	-0.20111	-0.1233
H12010	-0.20255	0.019384	-10.45	0	-0.24054	-0.16456
H22010	-0.2562	0.020694	-12.38	0	-0.29675	-0.21564
H12011	-0.26649	0.021257	-12.54	0	-0.30816	-0.22483
H22011	-0.32643	0.020068	-16.27	0	-0.36576	-0.2871
H12012	-0.36293	0.020337	-17.85	0	-0.40279	-0.32307
H22012	-0.42003	0.020444	-20.55	0	-0.4601	-0.37996
H12013	-0.45871	0.022618	-20.28	0	-0.50304	-0.41438
H22013	-0.4667	0.021859	-21.35	0	-0.50954	-0.42386
H12014	-0.47097	0.022305	-21.11	0	-0.51469	-0.42725
wx_cbd	0.049695	0.107706	0.46	0.645	-0.1614	0.260794
wx_centre	-0.23064	0.099798	-2.31	0.021	-0.42624	-0.03504
wx_dec	0.047483	0.072079	0.66	0.51	-0.09379	0.188755
wx_age	0.00944	0.002613	3.61	0	0.004319	0.01456
wx_stately	-0.5555	0.114173	-4.87	0	-0.77927	-0.33172
wx_floors	-0.0167	0.004569	-3.66	0	-0.02565	-0.00774
wx_exclusive	0.216205	0.086491	2.5	0.012	0.046685	0.385724
wx_qual_adj	-0.07738	0.029299	-2.64	0.008	-0.13481	-0.01996
wx_corporate	0.660792	0.172241	3.84	0	0.323206	0.998379
wx_H12004	0.049826	0.317412	0.16	0.875	-0.57229	0.671942
wx_H22004	-0.31569	0.341393	-0.92	0.355	-0.98481	0.353422
wx_H12005	0.765149	0.34961	2.19	0.029	0.079925	1.450373
wx_H22005	1.120909	0.296286	3.78	0	0.540199	1.701619
wx_H12006	0.295691	0.378155	0.78	0.434	-0.44548	1.036861
wx_H22006	-0.02276	0.313029	-0.07	0.942	-0.63628	0.590768
wx_H12007	1.038493	0.285017	3.64	0	0.479869	1.597116
wx_H12008	-0.1741	0.32957	-0.53	0.597	-0.82004	0.471849
wx_H12009	0.383698	0.497597	0.77	0.441	-0.59157	1.358969
wx_H22009	1.328228	0.357283	3.72	0	0.627966	2.028489
wx_H12010	-0.91739	0.528763	-1.73	0.083	-1.95374	0.118971

wx_H22010	1.10442	0.517133	2.14	0.033	0.090858	2.117982
wx_H12011	-2.9958	0.629613	-4.76	0	-4.22982	-1.76178
wx_H22011	2.081854	0.428879	4.85	0	1.241266	2.922442
wx_H12012	-0.10754	0.517444	-0.21	0.835	-1.12171	0.906628
wx_H22012	0.328896	0.428386	0.77	0.443	-0.51072	1.168516
wx_H12013	1.853857	0.567075	3.27	0.001	0.742411	2.965304
wx_H22013	3.467443	0.529172	6.55	0	2.430285	4.504601
wx_H12014	-0.72101	0.336726	-2.14	0.032	-1.38098	-0.06103

Wald test of rho=0:Chi-squared (1) = 2020.612 (0.000)

Wald test for coefficients on lags of X's =0: Chi-squared (56) = 220.215 (0.000)

Likelihood ratio test of SDM vs. OLS: Chi-squared (29) = 173.359 (0.000)

5. Normality test for Spatial Lag regression residuals

				Joint test	
Variable	Observations	Pr. (Skewness)*	Pr. (Kurtosis)**	Adj. Chi-squared (2)	Prob>Chi-squared***
Spatial residuals vector	3912	0.0000	0.0000	66.31	0.0000

***H0: Skewness =0**

**** H0: Kurtosis=3**

*****H0: Skewness =0 and Kurtosis=3**

Skewness/Kurtosis normality tests:

Shapiro-Wilk test:

Variable	Observations	W	V	z	Prob>z*
Spatial residuals vector	3912	0.99407	12.905	6.659	0.0000

***H0: Residuals normally distributed**

6. Stability tests numerical results

Data for box-plot charts: Hedonic characteristic's estimators with different sample sizes.

Sample from H1 2003 to	*CBD*	*CENTRE*	*DEC*	*AGE*	*STATELY*	*FLOORS*
2007 H1	0.2708482	0.1414372	0.0548402	-0.0013779	-0.0089434	0.0021497
2007 H2	0.2691923	0.1388602	0.0498262	-0.001346	-0.0095255	0.0019367
2008 H1	0.2610081	0.1336603	0.0338852	-0.0013158	-0.0098505	0.0017138
2008 H2	0.254013	0.1267517	0.0290736	-0.001315	-0.015428	0.001388
2009 H1	0.2460786	0.1151013	0.0261734	-0.0013393	-0.0086769	0.0013673
2009 H2	0.2417469	0.1119935	0.0199615	-0.0014516	-0.0042415	0.0015379
2010 H1	0.2320529	0.103963	0.0163174	-0.0014828	0.0031885	0.0016907
2010 H2	0.234145	0.1016793	0.0172672	-0.001526	0.0079272	0.0019681
2011 H1	0.2310983	0.1013381	0.0162237	-0.0015193	0.009038	0.0022102
2011 H2	0.2186129	0.091908	0.0084196	-0.0015372	0.0094775	0.0023257
2012 H1	0.2237426	0.0963686	0.0086729	-0.0016481	0.0115734	0.002257
2012 H2	0.2270197	0.098172	0.0109565	-0.0017239	0.0173023	0.002409
2013 H1	0.2274304	0.0971763	0.0112858	-0.0017003	0.0157149	0.0024373
2013 H2	0.2313111	0.1007008	0.0123979	-0.0017058	0.0226646	0.0024375
2014 H1	0.275142	0.1243125	0.0268392	-0.0017148	0.0272887	0.0025937
MIN	0.2186129	0.091908	0.0084196	-0.0017239	-0.015428	0.0013673
Q1	0.2274304	0.098172	0.0112858	-0.0017003	-0.0089434	0.0016907
MEDIAN	0.234145	0.103963	0.0172672	-0.0015193	0.0079272	0.0021497
Q3	0.2610081	0.1267517	0.0290736	-0.001346	0.0157149	0.002409
MAX	0.275142	0.1414372	0.0548402	-0.001315	0.0272887	0.0025937
IQ range	0.0335777	0.0285797	0.0177878	0.0003543	0.0246583	0.0007183
1.5 IQ range	0.05036655	0.04286955	0.0266817	0.00053145	0.03698745	0.00107745

Annex 6 (Cont). Stability tests numerical results

Sample from H1 2003 to	*EXCLUSIVE*	*QUALITY*	*CORPORATE*	*CONSTANT*	*RHO*
2007 H1	0.0572839	-0.0323395	0.0747125	0.7101883	0.7477164
2007 H2	0.06093	-0.0344085	0.0801013	0.6346807	0.7725057
2008 H1	0.071071	-0.0367066	0.0795126	0.5501386	0.8011143
2008 H2	0.0700708	-0.0394135	0.0809432	0.4756867	0.8312232
2009 H1	0.0733453	-0.0407493	0.0850513	0.4161184	0.8541687
2009 H2	0.0726551	-0.0404415	0.0839657	0.4125976	0.8582347
2010 H1	0.0745221	-0.040091	0.0826146	0.3498288	0.8821702
2010 H2	0.0754081	-0.03991	0.0862045	0.3398824	0.8860161
2011 H1	0.0754472	-0.0403486	0.0852545	0.3125527	0.8970549
2011 H2	0.0758506	-0.0405963	0.0857382	0.2561656	0.9218906
2012 H1	0.0739245	-0.04193	0.0862595	0.2665524	0.9229763
2012 H2	0.0745443	-0.0430045	0.0863168	0.3133247	0.9103171
2013 H1	0.0753031	-0.0435206	0.0860297	0.3221593	0.9104563
2013 H2	0.0777248	-0.0451853	0.0874872	0.3584828	0.9003657
2014 H1	0.0804345		0.0877032	0.4133777	0.8862797
MIN	0.0572839	-0.0451853	0.0747125	0.2561656	0.7477164
Q1	0.071071	-0.04219863	0.0809432	0.3133247	0.8312232
MEDIAN	0.0745221	-0.04039505	0.0852545	0.3584828	0.8860161
Q3	0.0754472	-0.03873678	0.0862595	0.4756867	0.9103171
MAX	0.0804345	-0.0323395	0.0877032	0.7101883	0.9229763
IQ range	0.0043762	0.00346185	0.0053163	0.162362	0.0790939
1.5 IQ range	0.0065643	0.00519278	0.00797445	0.243543	0.11864085

7. Numerical results of out-of-the-sample rent estimation

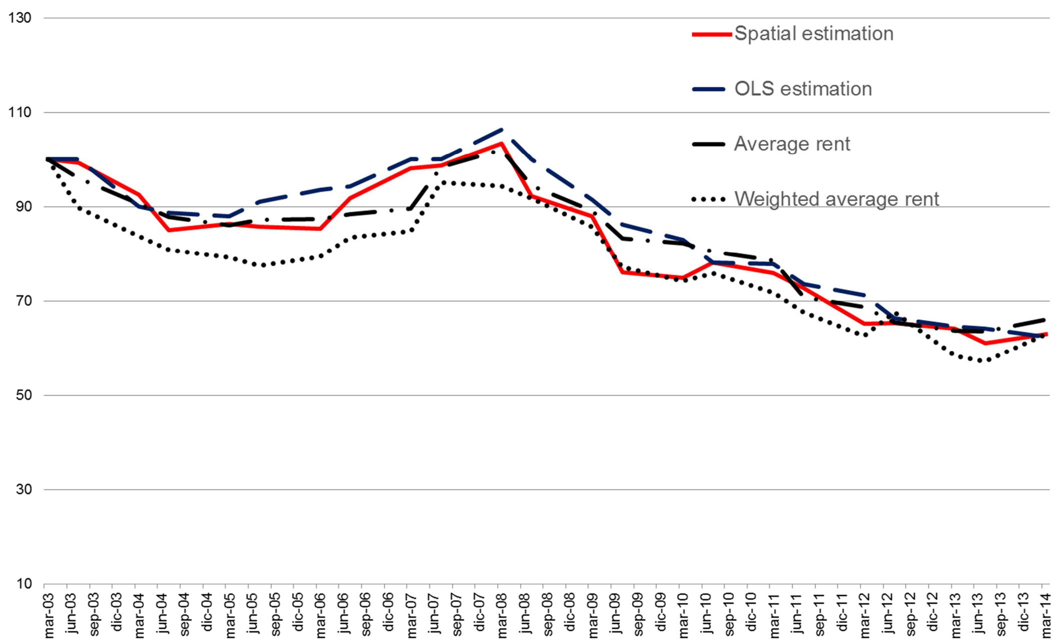
Results are obtained using assumptions of table 10 and expressed in €/sqm/month.

	Spatial estimation	OLS estimation	Average rent	Weighted average rent
H1 2003	20.3245395	21.2124447	20.5703152	21.9645946
H2 2003	20.1838343	21.2124447	19.7730494	19.7418502
H1 2004	18.823823	19.0978302	18.6500233	18.4058008
H2 2004	17.2794857	18.824376	18.0619931	17.7618888
H1 2005	17.5591951	18.6720532	17.695359	17.40818
H2 2005	17.4490531	19.3159258	17.9571775	17.0050426
H1 2006	17.3325632	19.835773	17.9778988	17.4307555
H2 2006	18.6747622	20.0041995	18.2047828	18.333184
H1 2007	19.957859	21.2124447	18.4317788	18.6320821
H2 2007	20.0724956	21.2124447	20.2528023	20.8918875
H1 2008	21.0138659	22.5738606	20.994552	20.7371228
H2 2008	18.7629419	21.2124447	19.4257116	20.1537109
H1 2009	17.8812558	19.3999856	18.3360597	18.8188719
H2 2009	15.4622257	18.2897213	17.1441021	16.9631112
H1 20010	15.2170651	17.5948303	16.9210815	16.2920359
H2 2010	15.8881651	16.5805467	16.5644379	16.6645866
H1 2011	15.4228743	16.5030997	16.161173	15.7607085
H2 2011	14.7745374	15.6159281	14.5484302	14.8618588
H1 2012	13.2441871	15.1224278	14.1219717	13.7608156
H2 2012	13.2821529	14.0902888	13.4847205	14.8475703
H1 2013	13.0325598	13.7110027	13.1007934	12.8055182
H2 2013	12.4086679	13.6241728	13.0702935	12.5564597
H1 2014	12.8052085	13.2404233	13.5847031	13.8266693

8. A rent index with the numerical results of out-of-the-sample rent estimation

Results are obtained using assumptions of table 10.

	Spatial estimation	OLS estimation	Average rent	Weighted average rent
H1 2003	100	100	100	100
H2 2003	99.3077075	100	96.1241926	89.8803305
H1 2004	92.6162337	90.0312548	90.6647424	83.7975896
H2 2004	85.0178459	88.7421333	87.806108	80.865999
H1 2005	86.3940611	88.0240512	86.0237623	79.2556399
H2 2005	85.8521448	91.0594041	87.29656	77.4202437
H1 2006	85.2789959	93.5100751	87.3972938	79.3584214
H2 2006	91.8828307	94.3040735	88.5002617	83.4669811
H1 2007	98.1958728	100	89.6037744	84.8277988
H2 2007	98.7599034	100	98.4564512	95.116199
H1 2008	103.391596	106.418005	102.062374	94.411589
H2 2008	92.3166888	100	94.4356536	91.7554421
H1 2009	87.9786513	91.4556802	89.1384479	85.678212
H2 2009	76.0766349	86.2216569	83.3438959	77.2293389
H1 20010	74.870405	82.945792	82.2597094	74.1740797
H2 2010	78.1723251	78.1642425	80.5259317	75.8702219
H1 2011	75.8830194	77.7991407	78.56551	71.7550623
H2 2011	72.6930976	73.616824	70.7253637	67.6627961
H1 2012	65.1635287	71.2903579	68.6521891	62.6499867
H2 2012	65.3503265	66.4246343	65.5542729	67.5977435
H1 2013	64.122288	64.6365982	63.6878594	58.3007267
H2 2013	61.0526396	64.2272635	63.5395882	57.1668175
H1 2014	63.0036831	62.4181867	66.0403254	62.9498045



9. Stata software procedure

1. Weight matrix calculation

```
spwmatrix gecon x_coord y_coord , cart rowstand wname(wght) eignvar(eigen) wtype(inv) dband(0 10500)
```

2. Moran's I and LM test calculation

```
spatdiag, weights(wght)
```

3. Spatial lag model estimation

```
spmlreg lrrent cbd centre dec age stately floors exclusive quality corporate H22003 H12004 H22004 H12005 H22005 H12006 H22006 H12007 H12008 H12009 H22009 H12010 H22010 H12011 H22011 H12012 H22012 H12013 H22013 H12014, weights(wght) wfrom(Stata) eignvar(eigen) model(lag) sr2
```