

The influence of macroeconomic factors on Portuguese Real Estate: a statistical approach

João Fragoso Januário

Mestrado em Engenharia Civil

Abstract

The Real Estate market is of utmost importance in today's modern economy and in every citizen's life. Hence, in order to improve urban planning and management as well as Real Estate investment, it is fundamental to study and understand this market's dynamics. Through this essay we aim to understand the Portuguese Real Estate market and its evolution since 1990. A variety of macroeconomic data will be used as input, in addition to a comprehensive study of the policies applied on this sector. We will be using Linear Regression methods on both explanatory and predictive approaches by analyzing the Real Estate market, its dynamics and its most important factors. Based on DiPasquale and Weathon's (1992) model for Real Estate Assets and Space, along with the Supply and Demand model, we attempted to understand the relationship between Real Estate inputs (such as building costs and access to finance) and its outputs (construction). In order to implement and access the models, we used data from 2001 to 2016. We found a great similarity between average value (€ per square meter) throughout the country, although there were some regional specificities worth noting. Moreover, we demonstrated that Linear Multiple Regression models are well fitted for an explanatory analysis of the data analyzed, however perform poorly when used as predictive models.

1. Introduction

The real estate sector has had a strong influence in the Portuguese economy (Cushman & Wakefield 2014). In 1995, this sector represented 7% of the GVA (Gross Value Added) to the Portuguese economy (according to PORDATA) and its weight has been increasing ever since. Currently, the real estate and construction sectors account for 17% of GVA. Thus, a comprehensive study of the real estate market would provide a better understanding of one of the major drivers of economy (Case 2000), (Leung 2004). The increasing capital gains on

Portuguese realty, especially on luxury real estate, has been noted by local and foreign investors. Although there is no formal definition of what constitutes a luxury real estate property, it can be assumed that properties valued over 500.000€ are considered luxury ones, seeing as this is the value set as the threshold for the *Golden Visa* or Residence Permit for Investment Activities (ARI). Both private investors and REITs (Real Estate Investment Funds) have been investing in the Portuguese commercial real estate (as rental properties) and residential segments (as second residences or rental properties).

The literature on the relationship between macroeconomic factors and real estate pricing is arguably less than expected (Leung 2004), especially considering that real estate is often used as collateral asset on credit and risk analysis in numerous investment projects (Mera e Renaud 2000). However, some studies have shown the relationship between real estate prices and economic factors such as Foreign Direct Investment (FDI) for instance Tsatsaronis and Zhu (2004), Fereidouni et al. (2010), Moshirian and Pham (2000) and Silva (2006).

The Markets for Real Estate Assets and Space: A Conceptual Framework, by DiPasquale and Wheaton (1992) is one of the most comprehensive studies on real estate market dynamics, and as a result has served as the basis for the development of this essay. Fischer, DiPasquale and Wheaton have studied the relationship between housing stock, rent, price and new construction, presenting the Fischer-DiPasquale-Wheaton model (*FDW-model*) for real estate markets.

Throughout this essay we have used the average bank valuation as a decision variable (Lourenço and Paulo 2014) for both national and regional real estate models while trying to understand the relationship between macroeconomic factors and the Portuguese real estate markets.

2. Fundamentals of Real Estate markets

The real estate market as an “economic market” follows the economic demand and supply model. The value of an asset is determined not only based on its physical and geographic characteristics but also on its demand and on availability of similar assets. This relationship of demand and supply can be simplified through the following figure based on the theory presented on *Wealth of Nations* (Smith 1776). The price is therefore given by the intersection between demand (D) and supply (S).

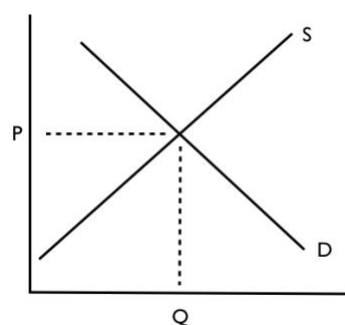


Figure 1 - Demand and Supply

While demand can be influenced by such factors as the average household income, supply can be influenced by the stock of real estate assets as well as the vacancy rates. Both demand and supply, are influenced by various factors, therefore this equilibrium is dynamic serving as an guideline for all market stakeholders.

Denise DiPasquale and William Wheaton (1992) further studied this relationship and presented an iterative model relating real estate rents, price, stock and construction. This model can be represented as per figure below.

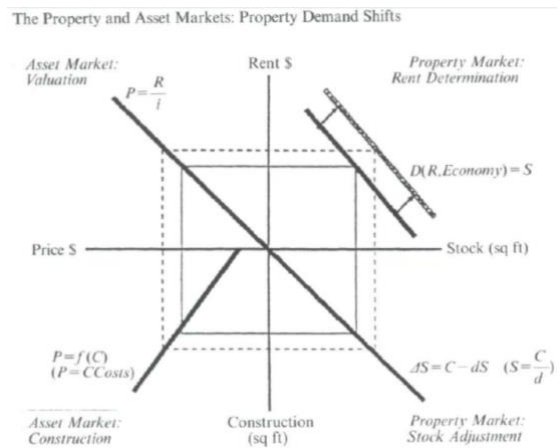


Figure 2 - DiPasquale-Wheaton (DW) model

The relationship between the real estate market and the financial market can be found on the second quadrant (Colwell 2002). However, according to the same author, it is important to note that this model does not account for vacancy rates (percentage of occupied residences).

The relationship between the property and the asset markets can also be expressed by the price-to-rent ratio (Lourenço and Paulo 2014). A high price-to-rent might lead to the decline of housing prices, since new residents will be more likely to rent than to buy, whereas a low price-to-rent ratio might lead to an increase in market prices, seeing as new residents will be more likely to buy and become homeowners (Pagourtzi et al. 2003). The *price-to-income ratio* can also be used as a tendency indicator on the future development of the real estate property market. In case of a high price-to-income ratio people will be less likely to buy, which should lead to a decline in real estate prices. On the other hand, a decline on this ratio may lead to an increase on the likelihood to buy properties.

This dynamic can often be represented as a real estate cycle (Gottlieb 1976). According to Case (Case 2000), there are several factors that might influence these dynamics such as vacancy rates, construction of new proprieties, as well as investors' speculations, which can potentially lead to a real estate market crash. Real estate properties can be overpriced due to investors'

speculations, leading to a crash in case this is not supported by a change in demographics (Mankiw e Weil 1989).

Furthermore, speculation can also negatively impact construction, leading construction companies and players to overbuild and subsequently cause a housing bubble (Hong, Scheinkman and Xiong 2006) (Glaeser, Gyourko and Saiz 2008).

3. The Portuguese Real Estate market

The Portuguese real estate market can be defined by its high percentage of homeownership (Braga 2013). There is a strict link between the Portuguese financial sector and the real estate market. After the liberalization of commercial banking, which began in 1984 (Santos, Teles and Serra 2014), easy access to credit led to an increased demand for homeownership. Numerous families were able to buy residences for the first time. This higher demand led to new construction and to an over-indebtedness. This excessive demand was noted worldwide and the Portuguese market was referred to an example of good capital gains on real estate investment (Goetzmann and Rouwenhorst 2000). From 2000 to 2007 the number of permits issued for new construction started to decline however, the average price per square meter has remained roughly constant. This collapsed after the 2007 financial crisis with an average decline of 12% in the housing prices from 2007 to 2013. Several government sponsored programs have offered reduced interest rates to young couples and affordable housing (*Porta 65* program).

The increasing influx of tourists has also influenced the Portuguese real estate markets, especially in the Lisbon area, due to increased demand for short term rent (through *AirBnB* and other real estate rental platforms) and homeownership by foreign private investors.

4. National and regional models

While analyzing state of the art real estate models, it has been found that majority of models are created based on two different techniques: Artificial Neural Networks (ANN) (Nguyen e Cripps 2001) (Teixeira 2012) and Linear Multiple Regression (Zurada, Guan e Levitan 2011).

Moreover, models traditionally based on Artificial Neural Networks perform better when using large datasets with a large number of independent variables and observations. However, Linear Multiple Regression is also an established method for modeling real estate by both academics and professionals (Zurada, Guan e Levitan 2011). During the development of this study we have tested both approaches and achieved better results when using Linear Multiple Regression models. This could be a result of a limited number of observation (56 observations).

We have used Python programming language in order to develop this study and model both national and regional markets.

5. Model Mathematical Formulation

Linear Multiple Regression models form a relationship between a dependent variable y_i and several independent variables x_{ik} . The coefficients β_1 , defined as constants, express the impact on the dependent variable of each unit change on each independent variable (Hair et al. 2010). Additionally, it is considered a noise coefficient given by e_i . The equation can be written as follows:

$$y_i = \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + \dots + x_{ik}\beta_k + e \quad i = 1,2, \dots, n \quad [1]$$

In order to estimate the coefficients we used the Ordinary Least Squares (OLS) method (Gustafsson & Wogenius, 2014), in which the difference between the predicted and observed value is deduced as follows:

$$\sum_{i=1}^n \hat{e}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad [2]$$

In order to validate the model to the national average bank valuation we have chosen the Goodness to Fit (R^2) and the Mean Absolute Percentage Error (MAPE) (Nguyen e Cripps 2001). Other methods have been used, as per table 1.

Table 1 - Adapted from (Zurada, Guan and Levitan 2011); pi- predicted value; ai- observed value; n – number of observations; i = 1,...,n

Performance Criteria	Formula
Root mean-squared Error (RMSE)	$\sqrt{\frac{\sum_{i=1}^n (p_i - a_i)^2}{n}}$
Mean Absolute Error (MAE)	$\frac{\sum_i^n p_i - a_i }{n}$
Mean Absolute Percentage Error (MAPE)	$\frac{\sum_i^n \left \frac{p_i - a_i}{a_i} \right }{n}$
Goodness to Fit	$\left(\frac{S_{PA}}{\sqrt{S_p S_a}} \right)^2$ with $S_{PA} = \frac{\sum_i^n (p_i - \bar{p})(a_i - \bar{a})}{n-1}$ $S_p = \frac{\sum_i^n (p_i - \bar{p})^2}{n-1}$; $S_a = \frac{\sum_i^n (a_i - \bar{a})^2}{n-1}$; $\bar{p} = \frac{\sum_i^n p_i}{n}$

6. Decision Variable

As a decision (dependent) variable we have used the average bank valuation per square metre (Pagourtzi et al. 2003). This is a widely used metric and it is a good indicator of markets overall tendency (Chandler e Disney 2014), (Gloude-mans e Miller 1976).

7. Independent variables

This essay has included a variety of independent variables related to the macroeconomic tendencies, seeing as several studies have proven the importance of including several indicators as inputs to a real estate model (Tsatsaronis e Zhu 2004) (Ross 2011) (Winkler e Jud 2002). The independent variables considered in this study can be found on the following table:

Table 2 – Independent variables

#	Simbol	Variable	Description	Unit	Source
1	Tdes	Unemployment Rate (Quarterly)	Unemployment rate of active population (between 15 and 74 years old)	percentage	INE
2	PIB	Gross Domestic Product	Total gross domestic product	million euros	INE
3	Cpriv	Private Consumption	total private consumption; annual rate of change	percentage	BdP
4	Cpub	Public Consumption	total public consumption; annual rate of change	percentage	BdP
5	FBCF	Gross Fixed Capital Formation	Equipment acquired over 500€; annual rate of change	percentage	BdP
6	Print	Domestic Demand	Total consumption and investment by residents; annual rate of change	percentage	BdP
7	Exp	Exports	Total exports	euros	INE
8	Imp	Imports	Total imports	euros	INE
9	RMM	Net Average Household Income	Net average household income of employees	euros	BdP
10	IHPC	Harmonised Index of Consumer Prices	Average increase in consumer prices; annual rate of change	percentage	BdP
11	ISE	Economic Sentiment Indicator	Public opinion indicator	index	EC
12	Tax_Hab	Mortgages Average Interest Rate	Average interest rates offered by banks to residents	percentage	BdP

Table 3 - Independent variables

13	E3M	Euribor 3 Months	3 Months average Euro Interbank Offered Rate	percentage	EBF
14	E6M	Euribor 6 Months	6 Months average Euro Interbank Offered Rate	percentage	EBF
15	OT10	Treasury Bonds	Average return on 10-year Treasury Bonds	percentage	REUTERS
16	PSI20	Index PSI20	Euronext Lisbon stock index (20 largest companies)	percentage	NYSE Euronext
17	EUR_USD	EUR/USD Exchange rate	Average EUR/USD rexchange rate (Quarterly)	<i>Euro/USD</i>	ECB
18	Res	Total Resident Population	Total Resident Population	millions	INE
19	Act	Total Active Population	Number of citizens between 15 and 74 years old	millions	INE
20	IDE	Foreign Direct Investment	Total investment in Portugal by foreign investors	million euros	BdP
21	DEB	Gross Foreign Debt	Total debt to foreign entities	million euros	BdP
22	C_Contr	Construction Costs Index	Construction costs index; Based on year 2000 = 100	Percentage	INE

The majority of data was gathered from the *Instituto Nacional de Estatística* (INE) statistical reports. However, we have also used Bank of Portugal (*Banco de Portugal* – BdP), Reuters, New York Stock Exchange (NYSE), the European Central Bank (ECB) and the European Banking Federation (EBF) as sources for the development of this project.

8. Case Study

During this essay we have studied the impact of macroeconomics factors on Portuguese real estate through the assessment of the average bank valuation. After a first analysis on the real estate panorama we have gathered and pre-processed all variables. Then, we have modeled the selected variable. First we used the national average bank valuation as a decision variable. Then, on a second phase we have used several regional data to model real estate markets on the top tem cities per yearly transactions, namely the Lisbon Metropolitan Area and Oporto Metropolitan Area as well as the cities of Almada, Amadora, Cascais, Lisbon, Oeiras, Sintra, Matosinhos, Oporto, Vila Nova de Gaia and Braga.

9. Input Data

We started by analyzing the correlation between variables through the below cluster heatmap. This map uses color to identify strong positive correlations (in red) and strong negative correlations (in blue) and organizes the independent variables in clusters using Euclidean distance as criteria.

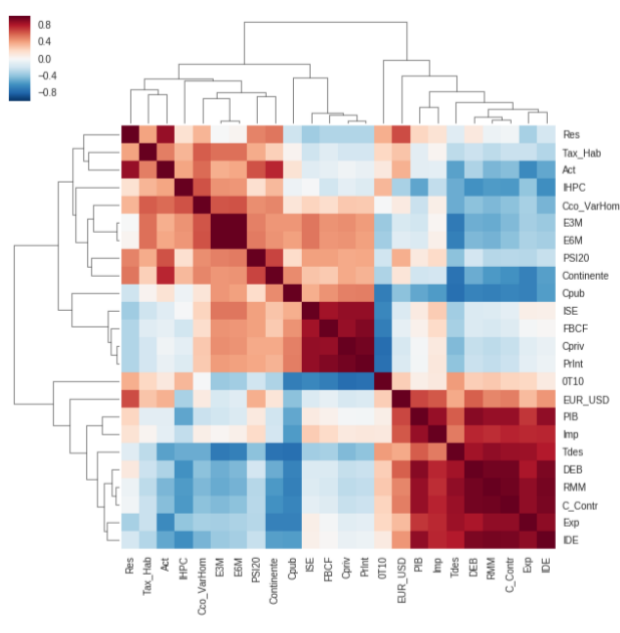


Figure 3 - Cluster heatmap of independent variables

We have also analyzed the correlation between regional data and observed a high correlation between all selected regional areas. However, it is important to note that Lisbon as its distinct dynamics as per above heatmap.

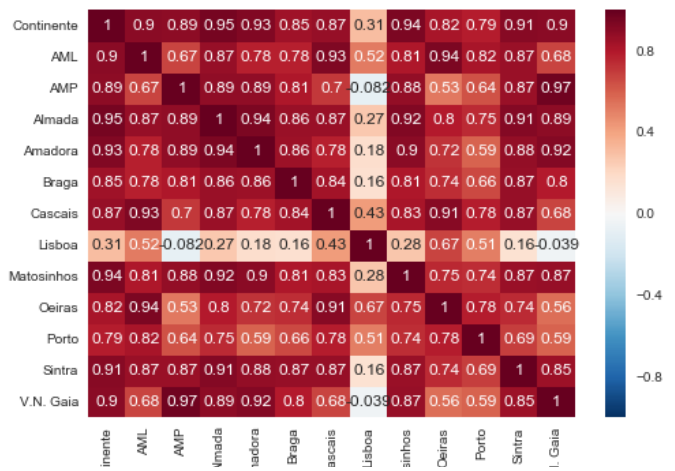


Figure 4 - correlation matrix heatmap of regional data

10. National results

On a first regression model we have used the standard Linear Multiple Regression taking all independent variables as input. However, results have shown a p -value of over 10% on the majority of independent variables. Therefore, we have tried a second approach using Stepwise elimination as a feature selection technique. Backwards Stepwise Regression have shown to produce the best results. The equation obtained can be written as per below:

$$Cont \left(\frac{\text{€}}{m^2} \right) = 1175,58 + 93,04PIB + 54,11RMM + 18,54ISE - 21,51E3M - 16,55OT10 + 50,03Res - 98,27DEB - 94,43C_Constr \quad [1]$$

The results met our expectations on the model. It shows that an increase in GDP (PIB), Net Average Household Income (RMM) and number of residents have a positive influence on the

increase of average bank valuation. A positive public opinion, given by the Economic Sentiment Index (ISE) will also increase the likelihood of investing in real estate thus increasing the average bank valuation. On the other hand, an increase in interest rates and an increase in Treasury Bonds yields should lead to a decrease on bank valuation. The goodness to fit (R^2) stands at 92,2%, which means a high percentage of bank valuation variance can be explained by the above model. The Durbin-Watson residuals equal to 0,952 which may suggest a high correlation between variables which can be explained by the high correlation shown amongst macroeconomic variables.

11. Regional results

On a regional level we came to conclude that each city was influence by a different set of macroeconomic factors. However, we have noted that the Gross Domestic Product (PIB) as well as Euribor interest rates are part of the majority of the equations found. The results table can be found on table 4.

Table 4 - Regional models after Backwards Stepwise elimination

Region	Model	
Lisbon Metropolitan Area	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1354,92 + 215,39\text{PIB} + 46,24\text{Cpub} - 70,38\text{Imp} + 73,64\text{IHPC} + 50,80\text{ISE} - 47,90\text{E6M} - 127,18\text{IDE}$	[2]
Almada	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1380,49 + 175,86\text{PIB} + 49,52\text{PrInt} - 63,73\text{Imp} + 38,71\text{IHPC} - 31,51\text{E3M} + 80,48\text{Res} - 198,05\text{DEB}$	[3]
Amadora	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1354,70 + 241,25\text{PIB} - 71,08\text{FBCF} + 55,23\text{PrInt} - 409,41\text{E3M} + 364,85\text{E6M} + 37,13\text{EUR_USD} - 201,16\text{DEB} - 157,57\text{C_Contr}$	[4]
Cascais	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1633,21 - 135,14\text{PIB} + 54,44\text{Cpriv} + 74,11\text{Cpub} + 148,68\text{RMM} + 71,11\text{IHPC} + 60,56\text{PSI20} + 105,91\text{Res}$	[5]
Lisbon	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1847,82 - 148,68\text{Tdes} - 79,29\text{Imp} + 61,16\text{IHPC} - 43,33\text{Tax_Hab} + 732,18\text{E3M} - 722,03\text{E6M} + 59,22\text{PSI20} + 68,83\text{Res} + 296,54\text{C_Contr}$	[6]
Oeiras	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1601,34 + 231,18\text{PIB} + 34,89\text{Cpriv} - 86,04\text{Imp} + 60,64\text{IHPC} + 41,11\text{PSI20} - 129,75\text{IDE}$	[7]
Sintra	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1157,86 - 56,13\text{Tdes} - 50,35\text{Imp} + 99,17\text{RMM} + 60,52\text{ISE} + 102,53\text{Res} - 125,54\text{DEB}$	[8]

Oporto Metropolitan Area	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1111,85 + 95,87Tdes + 154,55PIB + 21,60PrInt$ $+ 30,46Tax_Hab - 444,11E3M + 425,90E6M$ $- 24,43OT10 - 40,87PSI20 + 41,76EUR_USD$ $- 353,76C_Contr$	[9]
Matosinhos	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1240,02 + 69,79PIB + 41,46Cpub + 70,38Exp$ $+ 48,22ISE + 24,24Tax_Hab - 250,67E3M$ $+ 233,01E6M - 30,69PSI20 + 163,84Res$ $- 84,08Act - 220,45C_Contr$	[10]
Oporto	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1370,39 + 49,12Cpub + 31,25FBCF + 32,46IHPC$ $- 21,31E6M + 62,47Res$	[11]
Vila Nova de Gaia	$y\left(\frac{\text{€}}{\text{m}^2}\right) = 1056,92 + 96,33Tdes + 161,22PIB - 428,12E3M$ $+ 438,30E6M - 27,73OT10 - 39,24PSI20$ $+ 43,33EUR_USD - 346,77C_Contr$	[12]

12. Conclusions

The models presented previously offered an interesting overview on the macroeconomic factors that influence the Portuguese real estate markets. There is evidence that the Gross Domestic Product, the Net Average Household Income as well as Imports and Exports play an important role in this market dynamics. Moreover, interest rates showed evidence of a negative impact on the increase of bank valuation. In regard to regional markets, Lisbon shows a distinct trend while the majority of the remaining analyzed cities show a similar dynamic and a high correlation.

During the development of this study we understood that several regression techniques can be applied to real estate models namely Artificial Neural Networks (ANN) as well as Linear Multiple Regression (LMR). While ANN have shown better results when a large number of observations is available, LMR have shown fair results when applied to smaller datasets. In this case, we have achieved better results using Backwards Stepwise Elimination with a goodness of fit of 92,2% on the national average bank valuation model.

Additionally, for future studies, we would suggest Autoregressive Conditional Heteroscedasticity (ARCH) models (Hamilton 1994) as well as the analysis of Granger causality (Darrat e Glascock 1993). We would also suggest using a larger dataset, with further household characteristics, allowing a segment analysis and modelling of real estate assets (Case and Shiller 1994).

References

- Braga, F. M. 2013. "Mercado Imobiliário Em Portugal." *Jurismat: Revista Jurídica*, no. 2: 199–216.
- Case, K. E. 2000. "Real Estate and the Macroeconomy." *Brookings Papers on Economic Activity* 31 (2): 119–62. doi:10.1353/eca.2000.0011.
- Case, K. E., and R. J. Shiller. 1994. "A Decade of Boom and Bust in the Price of Single-Family Homes: Boston and Los Angeles, 1983-1993." *New England Economic Review*, no. March: 40–51.
- Chandler, D. and . Disney. 2014. "Measuring House Prices: A Comparison of Different Indices." *Institute for Fiscal Studies*.
- Colwell, P. F. 2002. "Tweaking the DiPasquale-Wheaton Model." *Journal of Housing Economics* 11 (1): 24–39. doi:10.1006/jhec.2001.0301.
- Cushman & Wakefield. 2014. "Marketbeat: Portugal Outono 2014."
- Darrat, All E, and J. L Glascock. 1993. "On the Real Estate Market Efficiency." *Journal of Real Estate Finance and Economics* 7: 55–72. doi:10.1007/BF01096936.
- Glaeser, E. L, J. Gyourko, and A. Saiz. 2008. "Housing Supply and Housing Bubbles." *Journal of Urban Economics* 64 (2): 198–217. doi:10.3386/w14193.
- Gloudemans, R. J., and D. W. Miller. 1976. "Multiple Regression Analysis Applied To Residential Properties." *Decision Sciences* 7 (2): 294–304. doi:10.1111/j.1540-5915.1976.tb00676.x.
- Goetzmann, W. N, and K Geert Rouwenhorst. 2000. "Global Real Estate Markets - Cycles and Fundamentals." Yale ICF Working Paper No. 99-03. doi:10.3386/w7566.
- Gottlieb, M. 1976. "The Behavior of the Real Estate Market in Long Local Cycles." In *Long Swings in Urban Development*, National B, 83–112. National Bureau of Economic Research.
- Hair, Joseph F, William C Black, Barry J Babin, and Rolph E Anderson. 2010. *Multivariate Data Analysis. Vectors*. Pearson Education (US). doi:10.1016/j.ijpharm.2011.02.019.
- Hamilton, J. D. 1994. *Time Series Analysis*. Princeton University Press.
- Hong, H, J. Scheinkman, and W. Xiong. 2006. "Asset Float and Speculative Bubbles ." *The Journal of Finance* LXI (3). doi:10.3386/w11367.
- Leung, C.. 2004. "Macroeconomics and Housing: A Review of the Literature." *Journal of Housing Economics* 13 (4 SPEC.ISS.): 249–67. doi:10.1016/j.jhe.2004.09.002.
- Lourenço, R., and M M Paulo. 2014. "The Dynamics and Contrast of House Prices in Portugal and Spain." *Economic Bulletin*. Lisboa.
- Mankiw, N. Gregory, and D. N. Weil. 1989. "The Baby Boom, The Baby Bust, and The Housing Market." *Regional Science and Urban Economics* 19 (2): 235–58. doi:10.3386/w2794.
- Mera, Kōichi, and B. Renaud. 2000. *Asia's Financial Crisis and the Role of Real Estate*. Routledge.
- Nguyen, Nghiep, and Al Cripps. 2001. "Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks." *Journal of Real Estate Research* 22 (3): 313–36.
- Pagourtzi, Elli, V. Assimakopoulos, Thomas Hatzichristos, and Nick French. 2003. "Real Estate Appraisal: A Review of Valuation Methods." *Journal of*

- Property Investment & Finance* 21 (4): 383–401.
doi:10.1108/14635780310483656.
- Ross, S. 2011. “A Model for Examining Foreign Direct Investment in Real Estate.” *Journal of New Business Ideas and Trends* 9 (2): 23–33.
- Santos, Cordeiro, Teles, and Serra. 2014. “Finança E Habitação Em Portugal.” *Observatório Sobre Crises E Alternativas*. Lisboa.
- Teixeira, Maria Cristina Canavarro. 2012. “Redes Neurais Artificiais: Um Estudo de Caso No Mercado Imobiliário de Castelo Branco.” In *I Ciclo de Conferências Conselho Técnico-Científico*, 23–28. IPCB.
- Tsatsaronis, Kostas, and H. Zhu. 2004. “What Drives Housing Price Dynamics : Cross-Country Evidence.” *BIS Quarterly Review*, no. March: 65–78.
- Winkler, D.T., and G. Donald Jud. 2002. “The Dynamics of Metropolitan Housing Prices.” *Journal of Real Estate Research* 23 (1/2): 29–59.
doi:10.3386/w20418.
- Zurada, J. Guan, and A. S. Levitan. 2011. “A Comparison of Regression and Artificial Intelligence Methods in a Mass Appraisal Context.” *Journal of Real Estate Research* 33 (3): 349–88. [repec:jre:issued:v:33:n:3:2011:p:349-388](#).