THE CRISIS EFFECT ON THE AVERAGE HOUSING VALUE VARIATION IN THE CITY OF LISBON

FILIPE ANDRÉ AMARAL MARTINS, INSTITUTO SUPERIOR TÉCNICO, LISBON

NOVEMBER 2019

EXTENDED ABSTRACT

1 INTRODUCTION

A key aspect for all parts involved in the real estate sector is a forecast of the value of property. The numerous studies conducted on this topic show that there is a multitude of variables that affect the value of a given property and they are dynamic over time. Models developed to estimate real estate value have revealed the importance of economic and demographic factors in their deduction. The 2008 global economic "crisis" has highlighted the influence of the general state of the economy on real estate value and once again revealed the global dimension of the economy and markets today. In Portugal, like many other countries around the world, the real estate sector saw a substantial change in the average price per square meter of real estate. Lisbon is the Portuguese city whose average housing price per square meter is the highest in the country and has remained so, despite the large variations suffered over the last decade.

Several authors have noted how the situation in the real estate sector has substantial effects on overall economic activity (e.g., Muellbauer & Murphy 2008) or how the dynamics of the housing market have been following the macroeconomic situation of a country (Beltratti & Morana 2010). These and other aspects have motivated several studies on the relationship between housing prices and economic and / or demographic variables in various countries and cities over the last decades (e.g., Clapp & Giaccotto 1994). More recently, the 2008 economic "crisis" has motivated some of these studies by how it has affected the real estate sector internationally. In this context, in addition to investigating the relationship between different factors related to the general economy and the price of housing, various approaches and tools have been adopted. Some authors approached the real estate market as cyclical and studied the change in the value of real estate based on this assumption (e.g., Edelstein & Tsang 2007; Nneji et. al. 2013). The existence of bubbles in the real estate market was another approach explored, as the discrepancy between real and forecast values leads to a search for other price determinants in the real estate market (Hui & Yue 2006). However, the most direct and most widely used approach is to study how one factor or set of economic factors influences the real estate market in general and the value of real estate in particular (e.g., Kakes & Van Den End 2004; Shen et. al. 2016; Poon 2017). Direct approaches seek to identify the determinants that most influence market fluctuations (e.g., Glindro et. al. 2011; Leszczyński & Olszewski 2017; Tupenaite et. al. 2017; Wiśniewski 2017; Asal 2018) or that directly influence the conditions for buying a dwelling from the consumer's point of view (e.g., Dua 2008). Some studies also explore where economic factors influence spillover effects between real estate markets in various countries (e.g., Nanda & Yeh 2016) or between regional markets within the same country (e.g., Weng & Gong 2017). The tools used in direct approaches are fundamentally statistical. These range from traditional statistical tools, such as linear

regression, to more recent alternatives, from the field of artificial intelligence, and various tools commonly used in econometrics. The tools used and the explanatory variables considered are intrinsically interconnected, since the tool's ability to consider nonlinearity and / or interaction between the variables will influence the statistical significance and the relative importance of the variables. The main observations and conclusions of the literature review are that there is a relationship between economic factors and the housing market (e.g., Clapp & Giaccotto 1994; Tupenaite et. al. 2017) and there are a large number of economic determinants that influence housing prices, but some stand out and may vary from market to market (e.g., Beltratti & Morana 2010; Glindro et. al. 2011).

This paper deals with the influence of the state of the economy on the variation of the value of residential properties over time. In particular, the impact of the 2008 global "crisis" is analyzed using as a case study the city of Lisbon from 2008 to 2018. The objective is to identify and quantify the importance of economic variables in the average value of residential properties in the city of Lisbon and to provide statistical models that allow their estimation or prediction.

2 METHODS

2.1 MODEL

The model chosen to elaborate the statistical analysis of this paper is the Multiple Linear Regression. It is a simple model based on time series regression. Regression analysis studies the link between dependent and independent variables through a mathematical model. Simple linear regression involves one dependent and one independent variable while multiple involves one dependent and several independent. Thus it is assumed for multiple linear regression that between a dependent variable named Y and a quantity k of independent variables named X_j (where j = 1,2,…, k) there is a linear relationship. A multiple linear regression model is represented with a mathematical equation such as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \qquad [1]$$

where the relationship between the dependent variable Y and the k independent variables X_j is described. The parameters β_j (where j = 0,1,2,…, k) and the symbol ε are respectively the partial regression coefficients and the random error. The values of the parameters β_j (with j = 1,2,… k) are an illustration of the variation expected in the solution Y for each of the units of variation in X_j.

First step is to create a correlation matrix, to observe the possibility of multicollinearity existence, relations between variables and to check if there is correlation with the dependent variable. Secondly, it is important to realize that normally the units of the independent variables are distinct and so the interpretation of the estimated parameters cannot occur in line with the degree of contribution of each independent variable to explain the variations of the dependent variable. To do this, the values of the variables can be normalized. Regression coefficients no longer express the rate of change in the initial units and, through normalization, express the rate of change in standard deviation units. Thus the coefficient values can already be compared

directly, because all variables have the same unit of measure, indicating which independent variables have the greatest contribution in the model. Normalization also minimizes potencial multicollinearity problems.

In order to adjust and improve the model, the Stepwise-backward method is applied to detect and eliminate non-significant independent variables. The p-value statistic evaluates the significance level of each independent variable, ie, the correlation level between both dependent and independent variables. If its value is greater than 5%, the variable is not significant, but if it is lower, then the variable is significant with a confidence level greater than 95%. The method consists of introducing all variables into the model and then removing one by one, successively, variables with the highest level of significance. It ends when there are only variables in the model with a significance level lower than the considered.

The coefficient of determination R^2, which ranges from zero to one, represents the proportion of the dependent variable that is explained by the independent variables included in the model. Values closer to one indicate a very suitable model and values near zero indicate a poor model. It is possible to adjust the coefficient to take into account the number of variables and obtain a more efficient result. It is a great model validation method.

2.2 DATA

In a statistical model, the dependent variable is the variable that will depend on the development of the independent variable(s), ie, it is the variable studied to understand how its variation is affected / influenced by the behavior of other factors. For the construction of the model described in the previous point, the normalized value of the real average price per square meter of housing in the city of Lisbon is used as the dependent variable (*plx*), excluding data from Nations Park (in Portuguese) parish, and the same value in 7 city zones (*pz1* to *pz7*) on a quarterly basis. These data were obtained through the Lisbon City Council at a confidential level. The values were calculated through a database consisting of all real estate transactions carried out in the city (in 23 of 24 parishes) between 2008 and 2018. The normalization of the values was calculated based on the second quarter of 2008. Figure 1 ilustrates the city map with all 7 of the different zones.

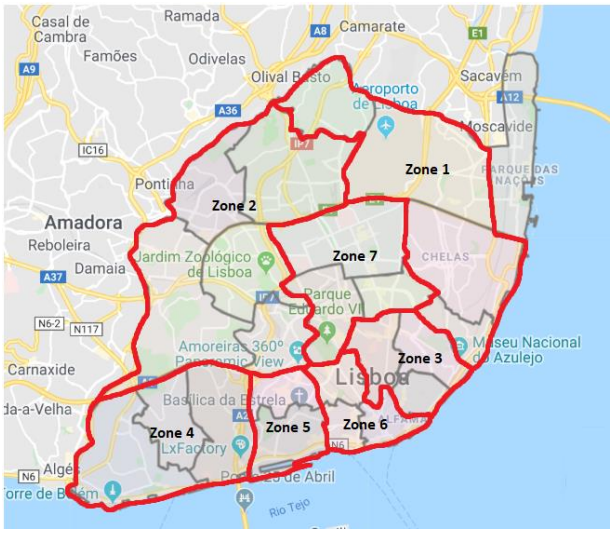

Figure 1 – The 7 different zones of the city of Lisbon

The first zone is represented by peripheral parishes, the second by expanding parishes, the third by historic parishes, the fourth by river side parishes, the fifth by two specific parishes, the sixth is the historic city centre and seventh is called the New Avenues.

Independent, or explanatory, variables are studied to understand how their variation interferes with fluctuations in the dependent variable. For this paper, 15 variables were selected to analyze and they were: unemployment rate (tdemp), gross domestic product (pib), net foreign debt (dext), gross fixed capital formation (fbcf), average monthly income (rendm), exports (expt), imports (impt), mortage rate (tjchab), tourism (htur), private consumption (cpriv), loans granted (econc), loan rates (tjebank), population (pop), licensed dwellings (lconst) and inflaction (ihpc). Data were collected on a quarterly basis from reputable institutions, namely the National Institute of Statistics and Bank of Portugal for the period 2008 to 2018. The selection of each variable was based on economic and demographic indicators hit by the crisis and consequent recessive policies. It is possible to have been included irrelevant variables in the model and omited relevant variables, where some of the reasons, for example, are limited by access to data or lack of knowledge of potential variables or variables with reduced influence. The variables were chosen based on the authors described in chapter 1 and some adapted to the Portuguese case.

Usually samples with a range between 30 and 500 observations are sufficient and appropriate for most studies (Roscoe 1975), but one author argues that for Multiple Linear Regression models with more than 50 observations, better results are obtained (Green 1991). Therefore, the choice and use of this model in the elaboration of this paper remains valid. There is a possibility of including irrelevant variables in the model and omitting relevant variables, where some of the reasons, for example, are limited by access to data or lack of knowledge of potential variables or variables with reduced influence.

3 RESULTS AND DISCUSSION

In this chapter, the multiple linear regression model is built and analyzed. The computer program used is the SAS Studio university edition. The database was introduced in the program and the first step was to verify the correlation between the various independent variables and also the dependent variable that represents the housing price index in the total city of Lisbon. There is a great correlation between some variables and when focusing on the variable plx there is a high correlation with three variables at a significant level of significance. The second step was to introduce all independent variables and the dependent variable plx into a multiple linear regression model in the simplest way and to verify the statistical confidence level of the model. The result was a high significance level for all variables thus making the model unreliable due to the high number of variables and / or the multicollinearity of the data. This proved necessary to use a method such as stepwise-backward to adjust the models. The third step was the normalization of the variables, as different units of measurement generate multicollinearity problems in the data. It consists of converting the values of each variable to a common measurement of all, with mean equal to zero and standard deviation equal to 1. After this pre-analysis steps, the adjustment and analysis of the models for each of the dependent variables began.

For the average price of housing in the total of the city of Lisbon and after applying the methods described, the final results are demonstrated in Table 1 wich shows the estimation of parameters for the resulting model after adjustment. The number of independent variable was high so the first model lacked explanatory value and significance of the included variables. The stepwise method in its backward function was used to eliminate variables with the highest significance level from the model. The process consists of running the model, identifying the variable with the highest p-value and above the significance value 0.05, eliminating this variable from the model and running the model again without the identified variable. These steps are repeated until only variables with p-value less than 0.05 remain in the model.

Table 1 – Ordinary Least Squares method for the total of the city

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	3	33.64018	11.21339	52.31	<.0001
Error	39	8.35982	0.21435		
Corrected Total	42	42.00000			
Parameter Estimates					
Variable	**Parameter Estimate**	**Standard Error**	**Type II SS**	**F Value**	**Pr > F**
Intercept	0.00537	0.07072	0.00124	0.01	0.9398
pib	0.66468	0.12524	6.03776	28.17	<.0001
tdemp	-0.37886	0.11428	2.35595	10.99	0.0020
cpriv	-0.18529	0.09042	0.90004	4.20	0.0472
Durbin-Watson D	1.716	**Root MSE**	0.46298		
Pr < DW	0.0759	**Dependent Mean**	9.50144E-16		
Pr > DW	0.9241	**Coeff Var**	4.872777E16		
Number of Observations	43	**R-Square**	0.8010		
1st Order Autocorrelation	0.141	**Adj R-Sq**	0.7856		

Figures indicate that gross domestic product contributed positively to housing prices in Lisbon and, both unemployment rate and private consumption, contributed negatively. The signals from the first 2 are as expected based on the literature but the third presented an unexpected signal. Gross domestic product growth can benefit housing price values so that their growth may be associated with a recessionary economy exit and a more favorable economic growth outlook. Lisbon is located in the region that most contributes to national gross domestic product. Unemployment rate growth could be associated with a poor outlook for the future of households, reducing housing demand and increasing the supply of those who need to retrieve their real estate investment. The private consumption case may be associated with household investment being more focused on short-term assets rather than on long-term investments such as housing, possibly related to more controlled future prospects as well. The model obtained an adjusted value of R^2 equal to 0.79 which is a good result, considering that the variables are nationally, since it indicates that 79% of the variation of the dependent variable values are explained by these 3 variables. The Durbin-Watson coefficient indicates the insignificance of the

temporal correlation for a time lag equal to 1 (Durbin & Watson 1951). Due to limitations of the sample size of observations (being less than 50), a time lag of more than 1 would result in a reduction of observations used in the model, thus leading to a decision to adjust only one first order because the loss of observations in a small sample may not compensate for possible "gains" in a possible analysis of more than 1.

The same models were built to analyse the distinct 7 zones of the city as well. The results obtained were as expected for the signals of each variable in each final model, as observed in the literature. The only variable to obtain a variation in the signal was inflation, which the literature reveals to be inconsistent but expected, whereas a distinction between expected and unexpected inflation reveals distinct signals (Clapp & Giaccotto 1994). Table 2 summarizes the final results obtained through the models constructed for the city of Lisbon and its 7 distinct zones, including the signs and coefficients of each independent variable.

Table 2 – Summary of results obtained for the total and each zone of the city of Lisbon

		Dependent Variables							
		plx	**pz1**	**pz2**	**pz3**	**pz4**	**pz5**	**pz6**	**pz7**
Independent Variables	**pib**	0.66468	-2.65463				1.19707		0.66444
	dext		2.34900				-0.48914		
	fbcf						-0.24478		
	tdemp	-0.37886		-1.53772	-0.55345	-0.67023		-0.63358	
	rendm		1.70085		1.07527	0.32920			
	expt								
	impt		1.19489	-0.84197					
	tjchab						0.49037		
	htur			-0.37379				-0.31017	
	cpriv	-0.18529							
	econc								
	tjebank		-0.72309						
	pop		-0.87900	0.65749					
	lconst		-1.06609						
	ihpc		-1.38751	1.38833	-0.85425			0.75148	

For the whole of the city of Lisbon and its zones, the main determinants for the change in the average housing price are the gross domestic product, the unemployment rate, the average monthly income and inflation. For the divided city, housing prices benefit from gross domestic product growth and are moderated by the change in the unemployment rate and the increase in expected inflation. These were interesting results because Lisbon has a real estate market with its own dynamics in relation to the country and it was possible to obtain models with a satisfactory explanatory grade. The different areas of the city reveal different dynamics and some similarities with the total of the city.

The first zone had a 40% explanatory level with 8 variables. The reasons for obtaining this low confidence level for this dependent variable may be related to the fact that the independent data originated at national level and the housing values in this city zone are not correlated or there are other factors / variables that better explain this variation. and were not included in the model, such as social or location factors. The model indicates that

the housing prices for this area were negatively affected by gross domestic product growth, bank loan interest rates, population, building permits and inflation, while benefiting from the country's external debt growth, imports and average household income. The results had little relevance, leading to an inconclusive analysis result, because the explanatory level of the model is low and thus indicating that there will be / existed factors with much higher explanatory level that were not included in this analysis. Since the model is considered invalid, there is no analysis of possible causes to obtain these results / determinants.

The resulting model for zone 2 obtained, as the model for zone 1, a low explanatory value of 47% with 5 variables. The same possibilities as the previous point model apply here to justify the low relevance of the variables. The appreciation of the average housing price for this area benefited from population growth and inflation, while growth in unemployment, tourism and imports negatively affected it. All signals obtained were as expected, considering inflation as unexpected. Again, no analysis is made of the possible causes of these results because the model is considered invalid due to the poor explanatory performance and thus the reduced relevance of the resulting determinants.

The model of zone 3 revelead a 70% explanatory level. The result indicates that an increase in average household income has brought benefits to the housing market in this area and both a rise in unemployment and inflation have had negative effects. Inflation is expected to bring benefits to the market as it increases the value of property but in this analysis was the opposite. This may have reduced demand resulting in a devaluation of the market as inflation coincides with expected inflation. Average income may indicate that household investment possibilities are favorable for long-term assets. The unemployment rate variable has already been mentioned in the analysis of the city of Lisbon in its total for the same trend. The signs of each variable were as expected, as found in the literature, considering inflation as expected.

Zone 4 obtained a satisfactory explanatory level, around 60%. It indicates that an increase in the unemployment rate hurt the area's housing market and an increase in average household income benefited the same. The signs coincide with the literature analyzed.

The model for the fifth zone revealed a high explanatory level, close to the level for the total of the city. The result concludes that the value of housing prices in this zone was positively affected by gross domestic product growth and housing credit interest rates and was negatively affected by growth in national external debt and gross fixed capital formation. Gross domestic product has already been mentioned in the analysis of the city as a whole with the same trend. An increase in the interest rate on home loans increases the final value invested by the buyer, thus increasing the value of the property. Gross fixed capital formation can negatively affect average housing prices considering, for example, that this formation can be achieved by obtaining other non-housing assets and may reduce demand. Growth in foreign debt may affect the country's future prospects resulting in households distrusting large investments such as home buying. The sign of the gross domestic product variable was once again expected and for the remaining variables the results vary or are not significant or not in the analyzed literature.

The sixth zone revealed a close to inconclusive explanatory level, about 56%. Housing prices in this area are hampered by rising unemployment and tourism, while benefiting from rising inflation. The unemployment rate and inflation have already been mentioned in previous points. The increase in tourism can bring disturbances to those who are looking for a home and have a preference for more peaceful places, which will be the majority of the population. The signs of the variables are similar to those observed in the literature.

And finaly, the model for the seventh zone obtained an unsatisfactory value of R^2, about 44%, but considering that only one variable can explain 44% of the variation of the dependent variable, reveals that this variable is a good predictor for this zone. Therefore indicating that gross domestic product growth benefits the appreciation of the average housing price in this area of the city.

The methodology applied was adequate but revealed some limitations, particularly when using small samples. The choice of variables and data regionality may be important factors in obtaining a higher explanatory grade. There was no evidence, in all models, of autocorrelation for a time lag equal to 1.

4 CONCLUSIONS

The initial analysis revealed a high correlation between the various variables studied, such as inflation with exports or loans granted with licensed dwellings or gross fixed capital formation with private consumption, among others. It also showed some strong correlations with the variation in the average housing price in Lisbon. One possibility for this evidence is that in the last decade there has been a similar trend arising from the global crisis among the various indicators of various different economic sectors. The price of housing in Lisbon may be due to its lower correlation, in terms of quantity of explanatory factors, to the fact that the statistics analyzed are at national level when compared with a dynamic market of its own. Regarding the built models, it was shown that the normalization of all variables brought improvements to the model and the choice of a method such as Stepwise - backward revealed an increase in the explanatory level of the model, as it limited the permanence of only variables with significance power lower than 5%. The choice of independent variables proved to be important, given the explanatory level of the models. The analysis itself indicated that the housing price in Lisbon, during this decade affected by the global crisis, was positively influenced by the development of the national gross domestic product and negatively by the variation of the unemployment rate and private consumption of the country. It was also observed that each area of the city had a different development regarding the variation of the average housing price, revealing some internal dynamism and the influence of other indicators besides the three found for Lisbon itself, such as inflation, the average monthly income of the households, tourism, mortage rate, gross fixed capital formation and national foreign debt. The trend of the results was as expected with the exception of inflation which obtained the two trends for different zones. It is a variable that can be divided into two parts, as found in the literature, as a different behavior is observed between expected inflation and unforeseen inflation, where respectively can hurt the price of housing in the way real estate investment is reduced due to future expectations. and can benefit from the view that an increase in prices will increase the

value of the property. Internal dynamism can be justified by various factors at local level, such as different sociodemographic and economic conditions at both national and local level.

The elaboration of this paper helped to understand and identify the factors that contributed to the variation of the average housing price in Lisbon, in the period coincident with the 2008 world crisis, and the resulting dynamics. It may have an influence on future political decision-making, investors and real estate agents decisions in the event of large-scale economic events.

REFERENCES

Dua, Pami. 2008. "Analysis of Consumers' Perceptions of Buying Conditions for Houses." *The Journal of Real Estate Finance and Economics* 37: 335-350.

Muellbauer, John, and Anthony Murphy. 2008. "Housing markets and the economy: the assessment." *Oxford Review of Economic Policy* 24 (1): 1-33.

Weng, Yingliang, and Pu Gong. 2017. "On price co-movement and volatility spillover effects in China's housing markets." *International Journal of Strategic Property Management* 21 (3): 240-255.

Poon, Joanna. 2017. "Foreign direct investment in the UK real estate market" *Pacific Rim Property Research Journal* 23 (3): 249-266.

Leszczyński, Robert, and Krzysztof Olszewski. 2017. "An analysis of the primary and secondary housing market in Poland: evidence from the 17 largest cities." *Baltic Journal of Economics* 17 (2): 136-151.

Tupenaite, Laura, Loreta Kanapeckiene, and Jurga Naimaviciene. 2017. "Determinants of Housing Market Fluctuations: Case Study of Lithuania." *Modern Building Materials, Structures and Techniques* 172: 1169-1175.

Wiśniewski, Radosław. 2017. "Modeling of residential property prices index using committees of artificial neural networks for PIGS , THE EUROPEAN-G82 , and Poland." *Argumenta Oeconomica* 38 (1): 145-194.

Nanda, Anupam, and Jia-Huey Yeh. 2016. "International Transmission Mechanisms and Contagion in Housing Markets." *The World Economy* 39 (7): 1005-1024.

Shen, Chung-Hua, Lin Bu, Kun-Li Lin, and Meng-Wen Wu. 2016. "Twin Booms: The Lead-Lag Relation Between Credit and Housing Booms." *Emerging Markets Finance and Trade* 52 (2): 522-537.

Glindro, Eloisa T., Tientip Subhanij, Jessica Szeto, and Haibin Zhu. 2011. "Determinants of House Prices in Nine Asia-Pacific Economies." *International Journal of Central Banking* 7 (3): 163-204.

Clapp, John M., and Carmelo Giaccotto. 1994. "The influence of economic variables on local house price dynamics." *Journal of Urban Economics* 36 (2): 161-183.

Beltratti, Andrea, and Claudio Morana. 2010. “International house prices and macroeconomic fluctuations.” *Journal of Banking and Finance* 34 (3): 533-545.

Edelstein, Robert H., and Desmond Tsang. 2007. “Dynamic residential housing cycles analysis.” *Journal of Real Estate Finance and Economics* 35 (3): 295-313.

Hui, Eddie C. M., and Shen Yue. 2006. “Housing price bubbles in Hong Kong, Beijing and Shanghai: A comparative study.” *Journal of Real Estate Finance and Economics* 33 (4): 299-327.

Kakes, Jan, and Jan Willem Van Den End. 2004. “Do stock prices affect house prices? Evidence for the Netherlands.” *Applied Economics Letters* 11 (12): 741-744.

Asal, Maher. 2018. “Long-run drivers and short-term dynamics of Swedish real house prices.” *International Journal of Housing Markets and Analysis* 11 (1): 45-72.

Nneji, Ogonna, Chris Brooks, and Charles W.R. Ward. 2013. “House price dynamics and their reaction to macroeconomic changes.” *Economic Modelling* 32 (1): 172-178.

Durbin, J., and G. S. Watson. 1951. “Testing for serial correlation in least squares regression II.” *Biometrika* 38 (1/2): 159-177.

Roscoe, J. T. (1975). Fundamental Research Statistics for the Behavioral Sciences (2nd ed.). New York: Holt, Rinehart and Winston.

Christoffersen, Peter F. 2012. “A Primer on Financial Time Series Analysis.” *Elements of Financial Risk Management* 39-64.

Green, Samuel B. 1991. “How Many Subjects Does It Take To Do A Regression Analysis.” *Multivariate Behavioral Research* 26 (3): 499-510.