International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 12, Issue 12, 2020

Comparative Study on the Efficiency of Logarithmic and Linear Models in Determining the Variables Affecting the Demand for Housing

Sameer Abdul Jabbar Abdul Majeed[a], [a]University of Iraqia, College of Media, Department of Journalism, Email: [a]Iraq1963hony@gmail.com

The quantitative analysis of the functions of demand in the housing market with standard methods is a very important topics; those models are an explanatory tool that helps predict and set up housing policies and strategies aimed at eliminating the problem of housing shortage or contribute to mitigation. The data and information in the quantitative studies fundamental mathematical structure preparation precedes form, give us that information and indicators thus giving a clear picture of the nature of the variables affecting the housing sector and the extent of interdependence and our expectations. Backed by the views known in advance about the shape and nature of those variables thus requires that these totalitarian models have been possible to ensure the success of the objectives of the housing policy and succeed in forecasting the future size of the phenomenon and its implications for the future.

Key words: *Logarithm - linear model - housing - shape determination.*

Introduction

The characterisation of the demand function for housing faces several difficulties, including the inability to measure some of the affecting variables in the housing price because it is non-quantified data and there is an absence of units to quantify it***(Zaemah, 2018).*** Therefore, we see researchers who are unable to provide alternative solutions, and this is reflected in the problem to be studied. So the results are distorted and do not give a clear image about the problem. ***(Mahboob, 1998)*** The problem in this study lies in the scarcity of studies which use the quantitative analysis method in examining the problems that interfere with the analysis of housing demand and the variables that affect it***(Van, 2010).*** Studies which use the classical descriptive analysis method affect its shortcomings because of its inability to address the problem from all its economic, demographic and financial aspects, this reflects negatively on

the researcher's ability to analyse and accurately cover the dimensions of the problem in all its aspects and put solutions to it ***(Kazem, 1988).***

Literature Review

Characterization of the Demand Function on Housing

1- Per Capita Income (X1) most of the housing demand can be determined through the entry of individuals, and it is expected that the income elasticity of demand that reflects the degree of demand response to changes in income is positive. According to the economic theory, the demand for housing takes a specific direction and then stabilises higher income as a result of approaching the saturation limit as a slow-moving commodity ***(Baqir, 1986).***

2- The Families' Number of Baghdad City (X2), any change in the families' number generates a change in the demand for housing units in the same direction as that change and may be greater than the increase in the families number. Published statistics indicate that the number of families of Baghdad city is increasing, so it is expected that this will be positive ***(Jatterjee, 1990).***

3- Housing Balance (X3) The increase in the housing balance leads to significant positive structural changes in the housing demand, as it increases the desire of individuals who share other families in their housing to have independence with housing ***(Mahboob, 1989).***

4- The Ratio Of Granted Loans / Construction Cost Per Square Meter (X4) The size of the loan will not be an effective tool in indicating the effect unless it is compared to the construction cost per square meter by the ratio of the loan size, ***(Zaemah, 2018). W***henever the ratio was higher, the demand for housing increased and there will be a decline in this demand, indicating insufficient real estate bank loans to build housing ***(Kazem, 1999).***

5- The Values of Construction Materials (X5), the relative change represents in the consumption shares of construction materials that the private and public sectors compete with through the ratio of the total value of construction materials used in the public sector to the total value of construction materials ***(Wang, 2011).***

6- The Cost per Meter (X6). Skilled construction workers (callipers) change the prices of their services with change in the general level of prices, so the housing demand fluctuates between recovery and stagnation whenever the cost of building a meter changes ***(Zietz, 2008).***

7-Population of Baghdad city (X7) population increases determine most of the increase in the housing demand, whether in Baghdad or elsewhere. Population lead to stimulating demand and becomes an urgent need as the population ***(Zietz, 2008).***

8- The Average Size of the Family (X8) the size of the family varies between males and females to search for new housing whenever children grow older, as males and females desire individual rooms ***(Oktay, 2014).***

9- Number of Marriages (X9) As for marriages in the Baghdad city, it is expected to have a significant effect on increasing the housing demand, with assuming the couple remains in Baghdad and not move, as well as their desire to live independently of the family microkernel ***(Beltratti, 2010)***.

10- The Response Variable (Yd) The number of building permits granted to individuals will be considered as a response variable that can be inferred from the demand as its statistical data are available, see Appendix (1).

$$Y_d = f_1(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, U)....1$$

$$Y_d = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 - b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + U.....2$$

Where: (b 0) is the intercept parameter. (U) Stochastic disturbance limit ***(Elbourne, 2008).***

The Function Shape Determination

The multiplicity of demand functions types and its variations in terms of characteristics require attention to the issue of selecting and determining the demand function that is most appropriate and consistent with consumer behaviour towards the commodity, ***(Chengjie, 2015).*** If the demand is expected to take a single direction that it does not change with an increase in income. A linear or double logarithmic function or a semi-logarithmic function can be chosen to represent the demand function, if the demand takes a certain direction and then stabilizes with higher income as a result of approaching the saturation limit, the logarithmic or natural logarithmic function can be chosen, ***(Zaemah, 2018).*** But if the demand rises first and then reaches a higher limit and then decreases with higher income then the inverse or relative logarithm function can be chosen as the logarithm or function squared ***(Tsai, 2010).*** Of course, consumer behaviour may not be known, which calls the use of the prevalence and expansion of the number of chosen functions (3). As for the statistical standards, it is standers closely related with the measures of the model, the extent of the hypothesis matches which the model was built, corresponds the model with the results and the number of explanatory variables ***(Sommer, 2013)*** . The independent variable from the dependent variable where the amount of

interpretation can be recognised from the corrected determination coefficient (R2), so the demand function will be estimated in three types ***(Adams, 2008)***:

Linear function: $Y_d = a + b_1 X_1 + b_2 X_2 + ... + b_n X_K + U .....3$

Logarithmic function: $LnY_d = a + b_1 LnX_1 + ...b_n LnX_K + U .....4$

Semi-logarithmic function: $LnY_d = a + bX_1 + U .....5$

Analysis method Determination

*** Forward stepwise**: the relationship between the dependent variable and independent variable is one of a number of specific independent variables can be studied as a first step. Then the study can be expanded to include another variable from the independent variables in addition to the founded variables in the previous model ***(Tsai, 2010)*** and include or delete the variables depends on passing the statistical tests. The model can be expanded to include another variable as a next step, and thus variables are accepted or approved based on the results of statistical tests, this analysis method is called the Forward stepwise method ***(Sinai, 2013)***.

Backward stepwise**: In this method, it is possible to study the relationship between the dependent variable and the specific variables at once. Depending on the results of the statistical test and the partial (F) value of all variables, any of them is accepted or excluded, and is known as gradual regression, or the Back, Backward stepwise method ***(Sommer, 2013).

Methods and Results

In light of what has been described, and a formulation of the expected relationship between the explanatory variables of the housing demand function that includes economic, population and financial variables, the response variable and the election procedure method and the backward step wise linear to estimate the values of this relationship measures. As a first step, all the explanatory variables will be entered once then we start by deleting the independent variables that have no significant effect one after the other until we reach the final formula which contains the significant effect variables and by extracting the partial (F) value for each of the variables according to the following formula:

$F == (\frac{bx_n}{Sbx_n})^2$.....6 Where 1, 2, 3 = n

Then we choose the lowest value of the partial F values and test its significance at the significance level of (α) , if the calculated (F) is less than the tabular we delete the variable that gave the lowest F partial and then we move to the second step which is estimating the

functions with the remaining variables(6). Until obtaining variables in which the calculated partial F is greater than tabular, then we "finally stop testing", and three types of functions, linear function, logarithmic function and semi-logarithmic function have been estimated, as follows:

1- The results of estimating the linear function with all variables:
Y_d=-2330.3+0.222X_1+0.145X_3-3891.6X_4-0.00009X_5-1.46X_6+7.887X_7+433.49X_8-0.386X_9……6
2- The results of estimating the logarithmic function with all variables:
Y_d=-5.07-0.17X_1-9.45X_2+0.56X_3-0.097X_4+0.056X_5-0.142X_6+10.25X_7-7.72X_8-0.549X_9……7
3- The results of estimating the semi-logarithmic function of all variables:
Y_d=7.106+7.544X_1+1.880X_3-0.290X_4-3.486X_5-0.00053X_6+1.935X_7+0.29X_8-2.76X_9……..…8

Table 1: Calculating Partial (F) of the Linear Model

Variables	Partial F	Variables	Partial (F) lower
average	0.01	Baghdad residents	0.0016
Families number	Neglected	average	0.01
balance	1.69	building cost	0.01
loan	5.76	construction materials value	0.62
building cost	0.01	Balance	1.69
construction materials value	0.62	cases	4.41
Baghdad residents	0.0016	The volume rate	5.29
The volume rate	5.29	Loan/ cost	5.76
Cases number	4.41		

Table 2: calculating partial (f) of the logarithmic model

Variables	**Partial F**	**Variables**	**Partial (F) lower**
Per capita income	0.1764	Building cost	0.1764
Families number	2.56	Per capita income	0.1764
balance	4.2025	construction materials value	0.1849
Loan / cost	0.9801	Cases	0.9409
construction materials value	0.1849	Loan/ cost	0.9801
Building cost	0.1764	The volume rate	1.44
Baghdad residents	3.4225	Families number	2.56
The volume rate	1.44	Baghdad residents	3.4225
Cases number	0.9409	Balance	4.2025

Table 3: Calculating Partial (F) of the Semi-Logarithmic Model

Variables	Partial F	Variables	Partial (F) lower
Cases number	3.6	Baghdad residents	0.13
Loan/cost	4.8	construction materials value	0.18
Per capita income	0.1849	Cases	0.30
construction materials value	0.13	The volume rate	1.6
building cost	0.30	Loan/cost	3.6
Baghdad residents	1.6	balance	3.6
The volume rate	3.6	Per capita income	4.8
Families number	Neglected	building cost	4.8
balance	4.8		

Tables sourced from the researcher's work, based on the results of estimating the linear, logarithmic, and semi-logarithmic models, and then calculating the partial (F) value of the variables.

The researcher obtained the result of estimation and continuous deletion of the variables that record the lowest value of the partial (F) for the three models on the final selection of the variables for the three models, which are as in tables (4) (5) (6) below.

Results of the Last Estimate of the Linear Model

$$Y_d = 3130.4 - 2686.2X_3 - 0.0016X_4 + 5534.01X_7 - 0.57X_8$$

(SE) (8497.1) (1162.9) (0) (1384.8) (090)

$R^2 = 0.89$ $\bar{R}^2 = 0.8484$ F = 20.54 D-W = 2.25

Results of the Last Estimate of the Logarithmic Model

$$Y_d = -10 + 12.7X_2 + 0.90X_3 + 13.25X_7 - 11.8X_8$$

(SE) (7.71) (3.28) (0.08) (3.41) (3.38)

$R^2 = 0.93$ $\bar{R}2 = 0.90$ F = 35.64 D-W = 2.50

Results of the Last Estimate for the Semi- Logarithmic Model

$$Y_d = 8.82 + 2.736X_3 - 0.35X_4 - 0.00015X_6 - 1.281X_9$$

(SE) (0.66) (0) (0.13) (0) (0)

$R^2 = 0.90$ $\bar{R}^2 = 0.85$ F = 22.39 D-W = 1.12

Table 4: Calculating Partial (F) of the Final Linear Model

Variables	Partial F	Variables	Partial (F) lower
Loan/cost	5.33	Loan/cost	5.3361
value	14.44	Materials value	14.44
rate	15.92	The volume rate	15.9201
number	40.83	Marriages cases	40.8321

Table 5: Calculating partial (F) of the final logarithmic model

Variables	Partial F	Variables	Partial (F) lower
number	14.89	The volume rate	12.3201
balance	114.27	Families number	14.8996
Population	15.05	Baghdad residents	15.0544
rate	12.32	balance	114.2761

Table 6: Calculating partial (F) of the final semi- logarithm model

Variables	Partial F	Variables	Partial (F) lower
balance	11.66	balance	14.66
Loan/cost	7.39	Loan/cost	7.39
Building cost	6.96	Building cost	6.96
cases	0.94	cases	0.94

The tables source from the researcher's work, based on the results of the latest assessment of the linear models,

Logarithmic, semi-logarithmic, and then calculate the (F) partial value

The researcher stops at this stage of the estimation courses, where the following observations can be confirmed based on the results of tables (4) (5) (6)

1- Partial (F) values calculated at the significance level (0.05), and the degree of freedom (3,14) and for all variables of the linear and logarithmic models are more significant than the tabular (F) values, which reached at the same level of significance and freedom degree (3,41)
2- Partial (F) results of the semi-logarithmic model variables calculated (F) model at a significant level (0.05) and a freedom degree (3,14) which is more significant than tabular except for one variable, this means, another must be conducted to estimate the measures, this is from the researcher's view not possible because losing most variables as a result of repeated deletion of this model. Therefore, the researcher neglects this model and focusing on the other two models

3- The researcher judges the uselessness of conducting another assessment course, especially since the variables of the linear and logarithmic models have fulfilled the partial (F) test, so they stop there

Discussion

By looking at the linear and logarithmic models, we notice that the estimation results came in contrast to the economic logic, such as the housing balance and the loan with the cost and the average of family size, but from a statistical side,the linear model explained (89%) of the changes taking place in the housing demand and from the statistic (F),thus rejecting the null hypothesis because the variables are not significant and accepting the alternative hypothesis.

The level of significance (0.05) and freedom of degree (3,14), its statistic (F) is equal to (20.54), note that (F) tabular is at the same level of significance and freedom degree (3.34), but the logarithmic model has included four variables, all of which are population variables: the families number, population of Baghdad city, the family size average and one economic variable is the housing balance, and the estimate results showed that its agreed with what the economic logic dictates regarding to the variables of the families number, the housing balance, and the Baghdad residents number. As for the variable of the family size average, the estimating results showed that the estimate contrary of the logic, but from a statistical side, the model has interpreted (93%) of the occurred changes in demand and from its F statistic. Therefore, we reject the null hypothesis of the insignificance of the variables and accept the alternative hypothesis with the significance of the variables under the level of significance (0.05), and the freedom degree (14.3) equals (35.64), note that (F) tabular is at the same level of significance, and the freedom degree (3,34). The differentiation will be through three criteria are, the first criterion is the constant term, and the second is the matching with logic extent, while the third criterion is the illustrative power of the linear and logarithmic demand function with a comparison between the real (Y) and (Ý) estimated in the graph

1-The Constant Term and the Mean Square: The stepwise line equation consists of three terms , the constant term, the slope of the regression line, and the coincidental disturbances term, and that the linear and logarithmic models have used a partial set of variables described in a previous stage, which are variables that explain the demand structure of the housing and part of these variables have been excluded in the estimation stage because of decrease the partial (F) value. This means that deleting some variables from the studied function or variables that were excluded during the estimation has an effect on the response variable. They also affect the constant term thus resulting in the models' configuration errors, as the learner does not have the ability to recognise everything, however successful in describing the form of the relationship, there is no inevitable to be there an effect not been taken into consideration. Also, the learner desires to have a regression line with good alignment that passes all observations.

The regression analysis depends on the arithmetic mean, and this average is strongly influenced by the extreme or anomalous values of some observations, such as political, social and economic shocks. These shocks result in extreme values that greatly affect the regression line path to be estimated. If the extreme values are excluded, we will get a value for the positive constant term and vice versa, the higher side of the regression line will rise up and down to this line cutting the horizontal axis causing a negative value for the term. Returning to our linear and logarithmic models, the estimate of the linear model showed that the constant term is of a positive high value (3130.40), referring that the partial set of excluded explanatory variables has exerted a strong influence on housing demand. There are explanatory variables that have an effect on that demand, but are not taken into consideration, and the final variables does not adequately explain the relationship between them and the housing demand. As for the logarithmic model, the constant term has been distinguished by its small size. As its value reached (10), but it is a negative sign, and indicates to the effect of excluded variables is minimal, and the four variables which represented the structure of housing demand are the best selection of demographic variables that explain the relationship between them and housing demand. Therefore, a logarithmic model is more efficient than a linear model in explaining the relationship as much as the analysis relates to the constant term.

2- The Match Extent of Models Measures with the Economic Logic: It should be noted those who use the forward stepwise method "that moving away from the use of economic theory creates a state of bias from mischaracterising the capabilities of the least-squares of the measures, so it is never advisable to move away from the economic theory of correct logic supported by previous convictions. As the statistical results may be the result of false or coincidental connections, and thus are not completely dependent on the issuance of decisions. Therefore, the decision based on statistical results does not outperform the decision based on prior convictions based on a causal explanatory theory or on logical relationships supported by previous convictions (9).

3-The Illustrative Power of two Models: The illustrative power of the model can be identified. The best models that represent the phenomenon under study the best representation, by matching the estimated regression line with the regression line with real values. Then this comparison can be a beneficial way to identify the explanatory power of the model side by side with economic and statistical tests. It is noted that the linear model has achieved a large value for the Mean Square value of (8497.1). In general, the mean value is important in determining the efficiency of the model as it represents the difference between the real value and the estimated value of the error term (U), in other words ($e=Y\_ \hat{Y}$) and whenever the (e) value was smaller when the estimated value it is an indicator of the model efficiency, and the logarithmic model has achieved a small value for the Mean Square of (0.66).

Conclusions

1- The great importance of the quantitative analysis method in the analysis of economic functions in general and the function of housing demand in particular, compared to using the classic analysis method.
2- The performance which characterised the logarithmic function in describing the causal relationship between the explained variables and the housing demand, was better than the ability of the linear and semi-logarithmic models to describe this relationship, and the ability embodied by the excellent matching of the results obtained with economic logic, statistical standards, and illustrations power.
3- The great matching between the logarithmic function behaviour in interpreting the demand function for housing and individuals' behaviour. As the nature of these functions and their curves continue to rise even when satisfaction levels are met, this may sometimes apply to the individuals' behaviour, where some families when increasing their incomes, may resort to replacing their housing and moving to areas with high social classes, luxury housing and better quality.
4- The demographic variables were better able to explain the phenomenon, than the ability shown by economic variables to clarify and explain the causal relationship between them and the housing demand. This housing demand is consistent with reality, not every change in incomes pushes individuals towards changing their housing, but every increase in the families numbers and population numbers must necessarily be accompanied by changes in the housing demand.
5- The logarithmic model that excluded economic variables but included the most important population variables, could fit as a predictive model. Adopting housing policies that dealt with the phenomenon, and thus gives a clear image of the shape and effect of population variables on the housing demand with the possibility of modelling studies includes only economic variables.

REFERENCES

Adams, F. G. and Y. Shachmurove, 2008, Modeling and forecasting energy consumption in China: Implications for Chinese Energy Demand and Imports in 2020, Energy Economics, 30, 1263-78.

Baqir, Muhammad Hussein, 1986, Applied Economic Measurement, National Planning Institute, Ministry of Planning, Baghdad.

Beltratti, A. and C. Morana. 2010, International House Prices and Macroeconomic Fluctuations. Journal of Banking and Finance, , 34, 533–45

Chengjie Hea , Zhen Wanga , Huaicheng Guo, Hu Shenga , Rui Zhoub , Yonghui Yanga, 2015, Driving Forces Analysis for Residential Housing Price in Beijing, Procedia Environmental Sciences 2 (2015) 925–936

Elbourne, A. 2008, The U.K. Housing Market and the Monetary Policy Transmission Mechanism: An SVAR Approach. Journal of Housing Economics, 17, 65–87.

Jatterjee, Sambrett, and Bertram Price, 1990, regression analysis with examples, translated by Muhammad Munajid, Higher Education Press.

Kazem, Amouri Hadi, Essam Khudair Mahmoud, 1999,the nature of statistical data and building for standard models, Wael Publishing House, first edition, Amman.

Kazem, Amuri Hadi, and Muhammad Munajid, 1988, Introduction to Linear Regression Analysis, Dar Al Kutub for Printing and Publishing, Baghdad.

Mahboob, Adel Abdul-Ghani, 1989, on the use of the gradual regression method with an applied study on the demand for housing in Iraq, published study, Journal of the College of Administration and Economics, Al-Mustansiriya University, Issue 8, seventh year, January 1989, Baghdad.

Mahboob, Adel Abdul-Ghani, 1998, The Origins of Econometrics Theoretical and Application, Al-Modal Technical Printing Co., Ltd., first edition, Baghdad.

Oktay, Erkan & Karaaslan, Abdulkerim & Alkan, Ömer & Çelik, Ali Kemal. (2014). Determinants of housing demand in the Erzurum province, Turkey. International Journal of Housing Markets and Analysis. 7. 586-602. 10.1108/IJHMA-11-2013-0056.

Sinai, Todd and Nicholas Souleles 2013. “Can Owning a Home Hedge the Risk of Moving?” American Economic Journal-Economic Policy 5(2), pp. 282-312.

Sommer, Kamila, Paul Sullivan and Randal Verbrugge 2013. "The equilibrium effect of fundamentals on house prices and rents." Journal of Monetary Economics 60, pp. 854-870.

Tsai, I.C., M.C. Chen, and T. Ma. 2010, Modelling House Price Volatility in the United Kingdom by Switching ARCH Models. Applied Economics, 42, 1145–53

Van Nieuwerburgh, Stijn, and Pierre-Olivier Weill (2010): "Why Has House Price Dispersion Gone Up?" Review of Economic Studies, 77(4), 1567—1606.

Wang, Y, (2011). Urban housing reform and finance in China: a case study of Beijing. Urban Affairs Review, 36, 620 – 645.

Zaemah Zainuddin, 2018, An Empirical Analysis of Malaysian Housing Market: Switching and Non-Switching Models, doctoral thesis, Philospohy in Financ, Lincoln University

Zietz, Joachim & Zietz, Emily & Sirmans, G.. (2008). Determinants of House Prices: A Quantile Regression Approach. The Journal of Real Estate Finance and Economics. 37. 317-333. 10.1007/s11146-007-9053-7.