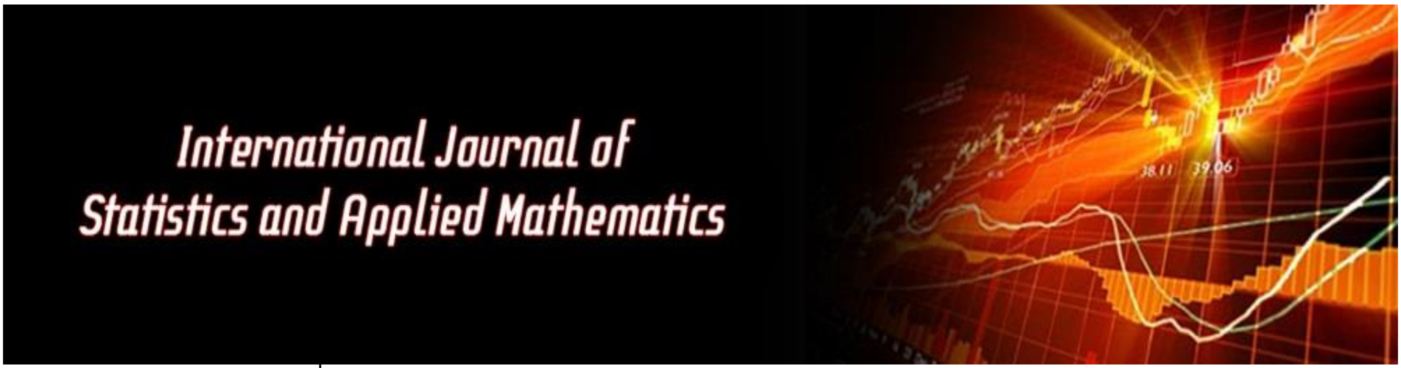

ISSN: 2456-1452
Maths 2022; 7(4): 101-104

www.mathsjournal.com
Received: 01-04-2022
Accepted: 04-05-2022

Nguyen Hong Nhat
National Economic University, Hanoi, Vietnam

Phung Duy Quang
Foreign Trade University, Hanoi, Vietnam

Corresponding Author:
Phung Duy Quang
Foreign Trade University, Hanoi, Vietnam

Approach to hedonic model and analyzing factors affecting real estate prices in Hanoi

Nguyen Hong Nhat and Phung Duy Quang

DOI: https://doi.org/10.22271/maths.2022.v7.i4b.849

Abstract
In real estate valuation techniques and housing market research, real estate prices are often analyzed using The Hedonic model based on factors that affect real estate prices. The Hedonic model shows the factors that affect real estate prices. The determinants of real estate prices in Hanoi in this article use 240 observations. The variables that have the greatest influence on real estate prices are the location of the property, the area of the land, the area of the house, the distance from the property to the city center, and the distance from the property to the street frontage.

Keywords: Real estate price, Hedonic model, land area, house area, distance

1. Introduction

For many households, housing not only provides them shelter, but it is also a major asset. Real estate is one of the most important assets in the investment and buying and selling portfolio of households. Indeed, in industrialized countries, real estate is the asset that makes up the largest share of affluent households. As a result, the value of their apartments has a great impact on consumption, household savings and the supply-demand mechanism. Thus, home prices are of great concern to property developers, banks, policy makers or the general public as well as existing and potential owners (Ustaoilu, E. 2003) [10].

The real estate market can be influenced by macroeconomic variables, spatial differences, community structural features and properties of real estate. Real estate valuation is necessary to provide a quantitative measure of the benefits and liabilities accrued from ownership of real estate. Valuations are performed by real estate agents, appraisal and inspection specialists, mortgage lenders, brokers, property developers, investors and fund managers, lending companies, market researchers, analysts, consultants and other professionals.

2. The Hedonic model and experimental studies

The Hedonic model is one of the oldest techniques of economic valuation developed by Lancaster (1966), Ridker (1967) [7], Grilihces (1971), Rosen (1974) and others. Originally used to study the relationship between air pollution reductions and property values, the Hedonic model became an important study in the second half of the 1970s and throughout the 1980s. During this period, the Hedonic pricing method has become a very important tool for academic research, used in theoretical and experimental studies to determine the monetary value of goods in relation to the environmental and physical characteristics of the property.

The first basic Hedonic model was presented by Ridker (1967) [7]:

$$P_i = f(S_{1i}, \ldots, S_{ki}; N_{1i}, \ldots, N_{mi}, Z_{1i}, Z_{2i}, \ldots, Z_{ni})$$

In which: P_i is the house price; S is the structural characteristics of the house (1 2 … k) such as the area of the house, the number of rooms, the type of construction and some other factors.

N is neighborhood characteristics (1….m), such as distance to workplace, quality of school, local crime rate and other factors.

Z = typical environment (1…n), such as air quality, water source, noise and other factors.

$$P_i = \alpha_o + \alpha_1 S_{1i} + \ldots + \alpha_{ki} S_{ki} + \beta_1 N_{1i} + \ldots + \beta_m N_{mi} + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + \ldots + \gamma_n Z_{ni}$$

Some recent empirical studies apply the Hedonic model to determine real estate prices. Sérgio A. B. *et al.* (2002) [8], in a study applying the Hedonic model to evaluate the adverse environmental effects of "odor" emitted from a wastewater treatment plant in Brasilia (a city of Brazil). He identified 20 variables and built 4 models to analyze the influence of the atmospheric environment on the prices of apartments. The model results show that "air quality has an effect on apartment value", meaning that the closer the apartment is located to the wastewater treatment plant, the lower the value.

In Selim S.'s study (2008) [9] about the intrinsic elements of the house made in Turkey, he built a house model as follows:

LogP = bx + u

The results of the Hedonic regression model show that the area of the house, the number of rooms, the type of house, the water system, the swimming pool, the location characteristics and the type of the building are the most important variables affecting the house price. This model is continued and extended by Selim. H (2009) [2]. With two types of approaches, the models used in the analysis are: The Hedonic and ANN regression models. The results of the regression model show that water system, swimming pool, house type, number of rooms, house size, location characteristics and building type are the most important variables affecting house price. It can be seen that house prices in urban areas are 26.26% higher than that in rural areas. By comparing the predictive performance between Hedonic regressions and ANN models, this study demonstrates that ANNs can be a better alternative to Turkish house price predictions.

Research results of Gabriel K. B (2011) [1] also show that external factors affecting real estate also affect the price of that real estate. He uses the Hedonic model considering factors such as: distance from property to church, distance from property to workplace, security, parking spaces.

The model is as follows:

$$P_i = \beta_o + \beta_1 X_j + \beta_d X_d + u$$

The results of the regression model show that external factors affecting the church have a negative effect on real estate prices. The further away the property is from the church, the higher the price.

In addition, there are a number of studies on real estate prices in Vietnam, which are also based on Hedonic regression model. According to the recently developed Status – Quality Trade-off theory, Authors Hoang Huu Phe and Patrick Wakely (2000) [3], have given some suggestions to orient the construction of a scientific basis and method of real estate valuation in line with market economic institutions. The study shows that real estate prices depend on quality and status factors: Y = f(CL, VT). Regression results show that the variable campus area has the strongest influence on real estate prices, followed by the variables of distance to the center of the city, number of building floors, location of houses being in the frontage or in alleys. In addition, Kim (2007) [4] also considers the legal impact on house prices in Ho Chi Minh City. In Ho Chi Minh City and Hanoi, the author builds a model with legal representative variables such as legal documents, house owner certificates, certificate of land use rights, and related legal rights.

3. Description of variables and regression models

In this article, 240 apartment samples are studied within Hanoi on 6 routes: Nguyen Chi Thanh, Tran Duy Hung, Nguyen Trai, Tay Son, Huynh Thuc Khang, Truong Chinh (data collected by the authors). The method of using regression according to Hedonic model and Status - Quality Trade-off Theory is extended to fit the variables and data of this article.

Sampling process was randomly collected on 6 routes. The above routes are all part of the main route of the city, which has been formed for a long time and is the place where busy business and trade activities take place. The sample in the article has all the documents: land use rights, blueprints, and registration fee records.

Due to the volatility of the real estate market, the relatively short survey period from April 1, 2021 to December 31, 2021 helped minimize macro fluctuations in the market, such as fluctuations in capital investments, interest rates, and exchange rates.

The regression model is thus as follows: $\text{Log}\,P = \beta x + u$, in which P is the house price, β is the regression coefficient, x is the explanatory variables and u is the random error. Regression model with 10 explanatory variables, including 5 position variables of real estate: the location of the property, the distance from the property to the center (Hanoi), the distance of the property to the street frontage, the width of the facade and the intended use of the property (for business or for residence). At the same time, 5 additional variables are considered for real estate quality: land area, house area, number of floors, terrace, shape. The model we study is:

$$\text{LogPRICE} = \beta_0 + \beta_1\text{AREA} + \beta_2\text{H} - \text{AREA} + \beta_3\text{DIS1} + \beta_4\text{DIS2} + \beta_5\text{WITHD} + \beta_6\text{SHAPE} + \beta_7\text{NOF} + \beta_8\text{TERRACE} + \beta_9\text{LOCATION} + \beta_{10}\text{USE} + \varepsilon$$

Table 1: Description of variables

Variable	Description	Expectation of Sign
LogPRICE	The logarithm of real estate prices	
AREA	Land area (m^2)	+
HAREA	House area (m^2)	+
DIS1:	Distance from the property to the center (km)	-
DIS2:	Distance from the property to street frontage (km)	-
WIDTH	Front pavement width (m)	+
SHAPE	Real estate shape: dummy variable (gets the value 1 if the back is wider or equal to the front of the land; value 0 if the back is narrower than the front)	+
NOF	Number of floors: dummy variable (gets the value 1 if the number of floors <= 4; value 0 if > 4)	-
TERRACE	Terrace: dummy variable (gets value 1 if there is a terrace; value 0 if there is no terrace)	+
LOCATION	Location: dummy variable (gets value 1 if property is in frontage; value 0 if property is in alleys)	+
USE	Usage: dummy variable (gets value 1 if the property is for business, office, or residential use; gets value 0 for residential use only)	+

Table 2: Descriptive statistics of variables

Variable	Average value	Maximum value	Minimum value	Standard deviation
LogPRICE	8.20	12.65	7.24	0.84
AREA	56.50	260.60	52.10	24.3
HAREA	2.780	4.30	2.12	1.23
DIS1:	24.60	30.30	22.40	4.62
DIS2:	4.60	6.60	3.20	1.26
WIDTH	6.30	8.65	5.40	1.68
SHAPE	0.42	1.00	0.00	0.12
NOF	4.20	1.00	0.00	0.36
TERRACE	0.65	1.00	0.00	0.21
LOCATION	0.43	1.00	0.00	0.36
USE	0.54	1.00	0.00	0.42

Table 3: Results of 6 regression models

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
AREA	0.0142*** (9.426)	0.0143*** (9.426)	0.0145*** (9.864)	0.0141*** (10.124)	0,.0148*** (10.012)	0,.0146*** (9.942)
HAREA	0.0008 (1.432)	0.0008 (1.420)	0.0008 (1.274)	0.0009* (2.524)	0,.0012* (2.788)	0.0012* (2.982)
DIS1:	-0.045* (-1.984)	-0.065* (-1.928)	-0.0665* (-1.948)	-0.0648* (-2.114)	-0.0684* (-2.682)	-0.0840* (-2.683)
DIS2:	-0.0016* (-1.914)	-0.0016* (-1.976)	-0.0016* (-1.978)	-0.0016* (-1.928)	-0.0017* (-1.926)	-0.0015* (-1.968)
LOCATION	0.9284*** (7.192)	0.8294*** (7.899)	0.8368*** (9.827)	0.8426*** (9.248)	0.8424*** (9.842)	0.8468*** (9.468)
WIDTH	-0.0085* (-1.826)	-0.0088* (-1.824)	-0.0086* (-1.946)	-0.0084* (-1.964)	-0.0094* (-1.848)	
SHAPE	0.1283 (0.984)	0.1078 (0.872)	0.1268 (0.924)	0.1046 (0.924)		
NOF	-0.1193 (-0.742)	-0.1428 (-0.846)	-0.1464 (-0.846)			
TERRACE	-0.0426 (-0.492)	-0.0428 (-0.548)				
USE	0.0602 (0.620)					
R^2	0.7824	0.7918	0.7824	0.7784	0.7689	0.7846
R^2 correction	0.7739	0.7846	0.7758	0.7727	0.7640	0.7809
Statistical value F	91.8873***	109.8138***	119.1681***	136.4073***	155.7097***	213.9984***

Robust standard errors in parentheses
*** $p<0.001$, ** $p<0.05$, * $p<0.1$
Source: Estimated results from the model

4. Result

Consider model 1

$$\text{LogPRICE} = 12.45 + 0.0142\text{AREA} + 0.0008\text{HAREA} - 0.045\text{DIS1} - 0.0016\text{DIS2}$$

$$+0.9284\text{LOCATION} - 0.0085\text{WIDTH} + 0.1283\text{SHAP} - 0.1193\text{NOF-} \ 0.0426\text{TERRACE} + 0.0602\text{USE} + \varepsilon \quad (1)$$

Model 1 has multicollinearity between variables and autocorrelation between USE and LOCATION, because the correlation coefficient between these two variables is 0.78. At the same time, the USE variable was not statistically significant, so the USE variable was excluded from model 1.

Consider model 2

$$\text{LogPRICE} = 12.45 + 0.0143\text{AREA} + 0.0008\text{HAREA} - 0.0605\text{DIS1} - 0.0016\text{DIS2}$$

$$+0.8294\text{LOCATION} - 0.0088\text{WIDTH} + 0.1078\text{SHAP} - 0.1428\text{NOF-}\ 0.0428\text{TERRACE} + \varepsilon \quad (2)$$

This model has eliminated autocorrelation but the test still shows signs of collinearity because R_i^2 is high. Hence model 2's variables will be omitted. The TERRACE variable has low statistical significance and the sign is opposite to our expectation, so we remove the variable TERRACE from model 2.
Similarly, when considering model 3, 4, we also remove variables NOF, SHAPE from the model.

Consider model 3

$$\text{LogPRICE} = 12.45 + 0.0148\text{AREA} + 0.0012\text{HAREA} - 0.0684\text{DIS1} - 0.0017\text{DIS2}$$

$$+0.8424\text{LOCATION} - 0.0094\text{WIDTH} + \varepsilon \quad (3)$$

This model has eliminated autocorrelation but the test still shows collinearity because R_i^2 is high. Hence model 5's variables will be omitted. The WIDTH variable has low statistical significance and the opposite sign is desired, so we remove the TERRACE variable from model 5.

Consider model 4

$$\text{LogPRICE} = 9.45 + 0.0146\text{AREA} + 0.0012\text{HAREA} - 0.0840\text{DIS1} - 0.0015\text{DIS2}$$

$$+0.8468\text{LOCATION} + \varepsilon \quad (4)$$

The variables of model 6 are all statistically significant, the signs of the coefficients are consistent with the expectations. The analysis results show that the price of real estate in the frontage is usually 84.68% higher than in alleys. For real estate in the alley, 1m further from the frontage causes the average price to decrease by 0.12%. Similarly, if real estate is 1km away from the center, the price will decrease by 8.4% on average. If the land area increases by $1m^2$, the price will increase by 1.46%.
We illustrate the following two cases:
*Case 1: Real estate in the frontage LOCATION =1; DIS1= 0; AREA =$60m^2$; HAREA = $186.2m^2$; DIS2=3.2 receives lnPRICE = 11.39144. Then PRICE = 88.56 (equivalent to 86 billion 560 million VND). With a frontage house 3.2 km from the center, with an area of 60m2, the house area is $186.2m^2$ (4 floors), the price is 86 billion 560 million VND.
*Case 2: real estate in alleys. LOCATION =0; DIS1= 0; AREA =$60m^2$; HAREA = $186.2m^2$; DIS2=3.2 receives lnPRICE = 10.54464. Then PRICE = 37.973 (equivalent to 37 billion 973 million). With houses in alleys 3.2 km from the center, with an area of 60m2, the house area is $186.2m^2$ (4 floors), the price has decreased significantly, at only 37 billion 973 million VND.

5. Conclusion

This article applies Hedonic model and the Status - Quality Trade-off Theory to determine real estate prices in Hanoi area. The results show that location variables, distance from property to street frontage, distance from property to center, land area and house area are important variables affecting property prices.

6. References

1. Gabriel Kayode Babawale. The Impact of Neighbourhood Churches on House Prices, Journal of Sustainable Development. 3/2011;4(1):246-253.
2. Hasan Selim. Determinants of house prices in Turkey: Hedonic regression versus artificial neural network, Expert Systems with Applications. 2009;36:2843-2852.
3. Hoàng Hữu Phê, Patrick Wakely. Vị thế, chất lượng và sự lựa chọn khác: Tiến tới một lý thuyết mới về vị trí dân cư đô thị, tạp chí Đô thị học, xuất bản tại Vương quốc Anh. 1/2000;37:1.
4. Kim. North Versus South: The Impact of Social Norms in the Market Pricing of Private Property Rights in Vietnam, World Development. 2007;35(12):2079-2095.
5. Nguyễn Mạnh Hùng, Trần Văn Trọng, Lý Hưng Thành, Trần Thanh Hùng, Hoàng Hữu Phê, Phương pháp định giá BĐS ứng dụng thuyết vị thế - chất lượng, trang 1-8.
6. Nguyễn Trọng Hoài, Phùng Thanh Bình, Nguyễn Khánh Duy. Dự báo và phân tích dữ liệu trong kinh tế và tài chính, NXB Thống kê, 2009.
7. Ridker Ronald G, John Henning A. The Determinants of Residential Property Values with Special Reference to Air Pollution, The Review of Economics and Statistics. 1967;49(2):246-57.
8. Sérgio AB, *et al*. Economics of Air Pollution: Hedonic Price Model and Smell Consequences of Sewage Treatment Plants in Urban Areas, Economics of Air Pollution. 8/2002;234:1-25.
9. Sibel Selim. Determinants of house prices in Turkey: Hedonic regression model, Doðuþ niversitesi Dergisi, 2008, 65-76.
10. Ustaoðlu E. Hedonic Price Analysis of Office Rents: A Case Study of the Office Market in Ankara, Middle East Technical University, Unpublished MSc, Thesis, 2003.