DETERMINING HOUSING PRICES USING THE SEMIPARAMETRIC ESTIMATION WITHIN THE HEDONIC PRICE MODEL FRAMEWORK: CASE STUDY OF ISTANBUL HOUSING MARKET EXAMPLE

Konut Fiyatlarının Hedonik Fiyat Modeli Çerçevesinde Semiparametrik Tahmin Kullanılarak Belirlenmesi: İstanbul Konut Piyasası Örneği Durum Çalışması

Tuğçe ACAR[*]

Abstract

Different properties of the houses have given heterogeneous structure to the housing markets. Therefore, the housing market can be analyzed with the hedonic price model In this direction, this study is important in terms of revealing the effects of different characteristics of houses on housing prices to follow the price changes in the housing market. The study aims to examine the effect of flat-type residential properties in İstanbul on the price within the scope of hedonic price model. For this purpose, estimation was made with semiparametric regression models for 974 flat-house, housing prices and factors affecting prices. In as much as the semiparametric estimation is superior to parametric estimation in that it doesn't force the data set to certain assumptions and takes into account the variable structure that has a nonlinear effect on the dependent variable. According to the findings, the properties that affect the housing prices (apartment) for sale are determined as the parking variable, credit eligibility variable, and elevator variable respectively. As for the nonparametric variables, it was concluded that the square meter variable had a reducing effect on house prices after 400 m2, and the age of housing had a positive effect on prices after 15 years.

Özet

Konutların sahip oldukları birbirinden farklı özellikler, konut piyasalarına heterojen yapı kazandırmıştır. Dolayısıyla konut piyasası hedonik fiyat modeli ile analiz edilebilir. Bu doğrultuda çalışma, heterojen yapıya sahip olan konut piyasasındaki fiyat değişimlerini takip edebilmek amacıyla, konutların sahip olduğu farklı karakteristik özelliklerin konut fiyatları üzerindeki etkilerinin ortaya çıkarılması bakımından önem taşımaktadır. Bu çalışmada amaçlanan ise İstanbul ili için apartman dairesi türündeki konut özelliklerinin fiyat üzerindeki etkisinin hedonik fiyat modeli kapsamında incelenmesidir. Bu amaçla 974 daire, konut fiyatları ve fiyatları etkileyen faktörler için semiparametrik regresyon modelleri ile tahmin yapılmıştır. Çünkü semiparametrik tahmin, veri setini belirli varsayımlara zorlamaması ve bağımlı değişken üzerinde doğrusal olmayan etkiye sahip değişken yapısı söz konusu olduğu durumda bu yapıyı dikkate alması bakımından parametrik tahmine göre daha üstündür. Elde edilen bulgulara göre satılık konut fiyatlarını (apartman dairesi) en çok etkileyen özellikler olarak sırayla otopark değişkeni, krediye uygunluk değişkeni ve asansör değişkeni olarak belirlenmiştir. Nonparametrik olarak ele alınan değişkenlerden ise metrekare değişkeninin 400 m^2 den sonra ev fiyatları üzerinde azaltıcı bir etki yarattığı, bina yaşının ise 15 yıldan sonra fiyatlar üzerinde pozitif etki yarattığı sonucuna ulaşılmıştır.

[*] Research Assistant, Beykent University, Faculty of Economics and Administrative Sciences, Department of Economics, tugceacar@beykent.edu.tr, ORCID: 0000-0001-9223-0089

1. Introduction

Housing is an important component of household wealth worldwide and investments in housing have always been important. On the other hand, the housing sector is unique and is significantly related to a country's economic health and wealth. Because it shows the characteristics of resilience, heterogeneity, and spatial stability. In this respect, the high demand for housing is a triggers growth in many economic sectors. The importance of housing is an important element in the psychological theory of the American psychologist Abraham Maslow, which was introduced in 1943 and subsequently developed. Maslow (1943) divides this hierarchy into 5 main categories in his theory of what he calls a hierarchy of needs. The hierarchy acknowledges that it is meaningless to meet the need in the upper category without meeting the need in the subcategory. Here, housing is considered as an element of protection, shelter, and feeling safe due to the need for security in the second category.

On the other hand, housing markets may differ according to the development levels of the countries. In economies, it is considered as an alternative to securities in some cases for investment purposes, in some cases, it is considered as a high-profit investment instrument. It is also very important as it has an effect on the general economy in general. Especially, the "Mortgage Credit Crisis" that occurred in the USA in 2005 was one of the most important reasons for the global financial crisis after 2007. With the impact of this global economic crisis, significant changes observed in housing prices in both international and national markets have demonstrated the importance of stability in the housing market. Also, it has made it mandatory to follow the housing market in terms of having a significant impact on financial stability.

The housing market has three main structural features. (Arıkan, 2008) These are:

1. Housing is distributed both spatially and sectorally. This means; the house is not only scattered for reasons of its structure (such as the number of rooms, heating system), but it also differs according to the features listed below:

- Residences in the same district may have different prices as they can belong to different residential sectors.

- Houses belonging to the same type of sector may have different prices as they can be in different districts.

- Houses belonging to different housing sectors and different districts may have similar housing prices

2. Consumers do not have complete information about the housing market.

3. Housing surveys also lead to temporary housing demand traffic in the housing sub-market. Even if the consumer does not buy the house he / she has researched, the demand for housing has been added to the housing submarket that he / she has researched due to the investigation made.

On the other hand, the application of the hedonic price model to the housing markets is based on various assumptions. The first one is towards the view that it is more accurate to handle the housing sector in a heterogeneous structure. The fact that both the structural and physical features differ from each other causes the houses to differ from each other in terms of features and thus the market is heterogeneous. With this heterogeneous structure in question, the process of estimating the price in the housing markets differs and becomes difficult compared to

other markets. (Kangallı Uyar and Yayla, 2016) This heterogeneity mentioned in the housing market can be addressed by the 'Hedonic Price Approach'. Another assumption is that the market operates under perfect competition conditions and is a large number of buyers and sellers. Therefore, individual buyers and sellers cannot significantly affect the market price.

In this this study aims to determine the effect of residential properties of the type of apartment belonging to Istanbul on the prices with the semiparametric estimation method within the framework of the hedonic price model. Unlike the parametric estimation method, the semiparametric estimation is based on the distribution of data and does not force the functional form to comply with a particular distribution and assumptions. It provides an advantage in terms of creating the most suitable model representing the data (Horowitz, 2009). Thus, the existence of a nonlinear relationship between the housing prices and each characteristic feature of the houses is taken into consideration and in this respect, superiority is provided compared to other studies in the subject area. Therefore, it is thought that determining the factors affecting the housing prices for the city of Istanbul with a semiparametric approach will benefit the housing market.

2. Hedonic Price Model

The hedonic pricing model is a model based on the consumer demand theory of Classical Economics, which shows that each characteristic of heterogeneous goods provides a different level of benefit or satisfaction to the consumer. Model-based theory suggests that people view goods as the sum of the characteristics they value. Accordingly, the first hedonic market equilibrium supply and demand model was developed by Lancester (1966) and Rosen (1974) depending on the characteristics of the good. According to this approach, while consumers are pursuing benefit maximization, manufacturers are also seeking profit maximization. The demand for the product depends on the characteristics of that product and it is possible to determine the marginal contribution of each characteristic of the product on the price of the product. In other words, the price of one house relative to another varies according to the additional unit of different properties in one house. (Chau and Chin, 2003) Thus, the decision to purchase a product will depend on the relative benefit of each of the characteristics that make up it. (Ayvaz, 2002) With this in mind, Rosen (1974) theoretically analyzed the short and long term balance in the heterogeneous goods market under the conditions of perfect competition, maximizing the benefit of the consumer and the profit of the producer.

In the model, goods (Z) are considered as the sum of n characteristics (Zi) belonging to them. Accordingly, Rosen defines the general form of the hedonic price function as follows:

$$P(Z) = Z(Z1,Z2,Z3,………,Zn)$$

Here P (Z) is the observed market price of the product. Z (Z1, Z2, Z3, …, Zn) are the vectors of objectively measurable properties. In other words, while the price of the product is considered as a dependent variable, each of its features is considered as an independent variable and the theory shows the level of benefit these features provide to the consumer.

However, the models proposed by Lancaster and Rosen have some fundamental differences. Lancaster's model assumes that the goods are composed of qualitative combinations depending on the budget constraint and they are members of the same group. Rosen's model reveals a preference ranking among the goods.

In the Lancaster model, it is possible to use the goods in the goods group together, so this approach is suitable for all consumer goods. However, this is not the case with Rosen, only for durable consumer goods. In response to Lancaster's model, Rosen expresses the existence of a nonlinear relationship between the prices and internal qualities of goods (Baldemir, Kesbiç and İnci, 2007).

3. Literature Review

The hedonic price approach, based on consumer demand theory, by Lancester (1966) and Rosen (1974), has been seen as an effective way to incorporate heterogeneity in residential markets and has become an important approach today. Therefore, many studies have been carried out in this area, which remains important at the pricing stage of the houses.

The study of Ridker and Henning (1967) was the first study in which the hedonic price approach was applied to the housing market. Using the horizontal cross-section data, Ridker and Henning emphasized that air pollution has a significant effect on household preferences in their studies, which they estimated with the linear hedonic price function they determined.

In the continuation of this study, Kain and Quigley (1970) analyzed the factors affecting the house sales price and house rent by using a semi-logarithmic linear function in their studies using horizontal cross-section data. In the analysis, (whether the building is duplex, building age, residential quality, whether there is hot water, number of bathrooms, number of rooms, parcel area, whether it is within the site, whether there is central heating, residence time, age of the residence etc.) the effects of twenty-eight different independent variables were examined. In addition to the effects created by the mentioned variables, it was determined that the schools in the vicinity had an impact on the value of housing.

Straszheim (1973, 1974), Goodman (1978), Macedo (1996), Ogwang and Wang (2002), Filho and Bin (2003), Stadelmann (2010) also, tried to explain the factors affecting the house sales price with a hedonic price approach. Since the analysis is related to the housing market in Istanbul, more studies will be included in the literature.

Ozus, Dokmeci, Kiroglu and Egdemir (2007) aimed to investigate the nature of housing price differences and for this purpose, they analyzed the spatial distribution of the housing prices of the province of Istanbul on a provincial and district basis. They found that the most important factors affecting the price at the provincial level are the residential area and the sea view, while at the district level the prices vary according to the socio-economic and property characteristics. They also stated that the results of the study are important in determining the districts that may be targeted for future capital investments.

Selim (2008) has used Turkey data in her study for determining the house price. Analysis was carried out within the framework of hedonic price theory, taking into account the 2004 Household Budget Survey Data. The size of the estimation sample determined 5741. 46 variables were used in the semi-logarithmic estimated model. According to the results obtained using the ordinary least squares method reveal that type of house, water system, number of rooms, pool, house size, locational characteristic and type of the building were the most significant variables that affect the house prices.

Alkay (2008) analyzed the properties affecting the prices of the houses using the linear regression model by using the hedonic pricing approach for the pricing of 522 houses for sale in Istanbul for 2001. First, three potential sub-housing markets were identified in Istanbul, then hedonic models were estimated for both these three sub-housing markets and the entire housing market. When an overall evaluation of the results obtained from the estimation of the model was made, the price of the house increases as the income for the sub-markets and the whole market increases. However, other factors affecting the housing price differ for each sub-market and the whole market.

Selim (2009) taking advantage of the 2004 Household Budget Survey examined housing price in Turkey with Artificial Neural Networks (ANN). He analyzed the prediction performance between the hedonic regression and artificial neural network models. Analysis of the results shows a better alternative would be to estimate the ANN in housing prices in Turkey. Thus, interpretations made according to ANN.

Çağlayan and Arikan (2009) analyzed the relationship between housing prices and properties with the Quantile Regression and Semi-Logarithmic Regression Model by using the hedonic pricing approach in the pricing of 992 houses for sale for the period October-November-December 2007 in Istanbul. According to the results obtained from the model estimations made with 12 independent variables affecting the house prices, the variables of security, heating system, garage, and cable TV, the area of the kitchen, the number of rooms increases the prices of the houses and the fact that the house is on the street decreases the price of the house. The building age variable for the buildings on the Anatolian Side positively affected the housing price.

Kaya (2012) tried to determine, the pure price changes occurring in housing prices for Turkey seeks to identify. Thus, she stated that it is important to work with the possibility that the bubble that may occur in the housing prices will be detected in advance. The review was made for Turkey firstly on a general basis and then on a city basis. According to the results obtained, for some of the provinces that are effective in the housing market, the pure value change percentages in housing prices were negative, while others were positive. Turkey ranks for the overall value of the increase that occurred in prices during the period analyzed were obtained findings that compose 6.21%. Turkey for the general population, a finding that 6:21% of the value increase occurred in prices during the period under review was obtained.

Demir and Yayar (2014) analyzed the properties affecting the prices of the houses in Turkey by using the hedonic pricing approach for the 45 variables. The data in the study, which was carried out using 30709 observations, were obtained from the 2010 household budget survey. According to the results of linear, semi-logarithmic and full-logaritmic regresion model, the house's being on basement or ground floor, construction year of the building, room floor's being tile, bathroom's being cement finish and fuel's being coal have affected the value of the house.

Kangallı Uyar and Yayla (2016) analyzed the prices of 2797 houses in Istanbul for the period October-December 2013 within the framework of the relationship between the hedonic price approach and the structural, physical, spatial, payment properties and neighborhood properties of the houses. According to the spatial Durbin Model results which take into account the spatial dependence, the basement variable which most negatively affected the house prices in the period examined in Istanbul was determined, as a rate of 22.51%. On the other hand, it

was determined that the variables that affect the housing prices positively the most were the Bosphorus view with 34.04% and the quality of life with 21.29%.

Wittowsky, Hoekveld, Welsch and Steeir (2020) determined which of the least-squares and spatial lag models made better estimates in his study, where he examined the factors that determine housing prices for Dortmund. The features of the spatial lag model are explained in detail. As a result of analyzing, they emphasized location has proved to be an important factor explaining residential property prices and the price of neighboring dwellings also plays a role. Besides, they found that the spatial lag effect is strongest for the rental apartment models.

4. Experimental Study

4.1. Data Set

In this study, 974 flat-house housing prices and factors affecting house-sales prices for the province of Istanbul were discussed within the scope of the hedonic price approach. These variables and their definitions are listed in the summary table below.

Table 1. Defining Explanatory Variables

Variables	Short Names	Definition	Type of Variable
Housing Price	LNHP (LNHP is housing price including logarithmically)	TL	Continuous
Housing Age	HousingA (HAG)	Year	Continuous
Housing Area	HousingA (HAR)	Square meters	Continuous
Housing Location	Site (ST)	If the housing is within the site 1, if not 0.	Discrete
Housing Property 1	Elevator (EL)	if the housing has an elevator 1, if not 0.	Discrete
Housing Property 2	Thermal Insulation (TI)	if the housing has a thermal insulation 1, if not 0.	Discrete
Housing Property 3	Mass Transport (TT)	If the dwelling is close to mass transportation (sea bus-subway-marmaray-bus stop-tram-minibus-subway-train station-subway-bus-minibus close to at least three of the means of transportation) 1, if not 0	Discrete
Housing Property 4	Parking (PR)	if the housing has a parking 1, if not 0.	Discrete
Housing Property 5	Steel Door (SD)	if the housing has a steel door 1, if not 0.	Discrete
Housing Property 6	Morgage Credit (MC)	if the housing is suitable for morgage kredi 1, if not 0.	Discrete

If the units differ in terms of the features to be examined, the best sampling method to be chosen to represent the population is stratified sampling*. If it is impossible or difficult to organize a random sampling of the units forming the population to be studied, the difficulty may be overcome by stratified sampling (Neyman, 1937). If subpopulations have different mean and variance values, stratified sampling creates more accurate estimators for population quantities (Ozturk and Kavlak, 2019). Thus assuming heterogeneous groups, this technique of sampling helps to increase the precision of the resultant estimates (Salinas, Sedory and Singh, 2019). In the field of statistics, stratified sampling is the acquisition of data from a population with a special shape probability sampling method. The feature that distinguishes the stratified sampling method from other probability sampling methods is that all elements in the population are composed of several groups and layers that are similar to each other according to certain characteristics. In this direction, the population elements were stratified and the proportional separation obtained by the ratio of the number of layer elements to the size of the population was used. Stratified sampling was used in the data collected with 974 sample volumes based on 39 districts for Istanbul in the residential market for sale. Random samples were determined from each district that expressed the layer and data were obtained.

The data distributed by districts are below:

Table 2. Distribution of Sample Size by District

District							
Arnavuköy	13	Beykoz	3	G.paşa	27	Şile	5
Adalar	1	B.düzü	58	Güngören	27	Silivri	10
Avcılar	30	Beyoğlu	8	Kadıköy	54	Şişli	30
Ataşehir	28	B.çekmece	20	Kağıthane	25	Sultanbeyli	10
Bağcılar	25	Çatalca	3	Kartal	25	Sultangazi	22
B.evler	40	Çekmeköy	29	K.çekmece	24	Tuzla	26
Bakırköy	20	Esenler	14	Maltepe	44	Ümraniye	36
Başakşehir	48	Esenyurt	63	Pendik	23	Üsküdar	26
B.paşa	12	E.sultan	45	Sancaktepe	26	Z.burnu	25
Beşiktaş	14	Fatih	21	Sarıyer	14	Toplam	974

* Source: www.sahibinden.com

4.2. Methodology

Ethics of research and publication were followed in this study, which did not require permission from the ethics committee and / or legal / special permission. In the hedonic price model, it is suggested that the value of a good depends on the number of qualities it contains. Thus, the price of the goods will be a function of its properties. The implicit price of the characteristics can be calculated from partial derivatives of the price function. Since these derivatives depend on the quality of these features, the selection of functional form is very important in the analysis. Therefore, econometric techniques such as nonparametric or semiparametric methods, which require very little restrictions on the functional form of the hedonic price model, provide more reliable information about implicit prices (Parmeter, 2007). Effective results are obtained when the parametric method structures used in econometric analysis have the desired properties under the assumption of linearity. But nowadays,

---

* For detailed information see: Neyman (1937), Rao and Rao (2016).

relationships between variables contain nonlinear structures. Parametric methods assume that the data are suitable for certain distribution and this distribution generally appears as a normal distribution. This creates limitations in practice. One solution that eliminates this limitation is the nonparametric or semiparametric approach. Nonparametric methods don't force data to a certain distribution. Therefore, it is more flexible than parametric methods. The only assumption in this approach is that the distribution should be continuous. We have the problem of dimension in here. As the number of variables increases, it becomes difficult to comment. Besides, all variables may not be linear with the dependent variable. While some are linear, some may have a nonlinear effect (Powel, 1994). In this study, the semiparametric estimation method will be applied, which eliminates these negativities, allows the continuous variables to be handled in a nonparametric structure, doesn't force the data to various assumptions, as in the parametric estimation process. The method provides the functional structure that will explain the relationship between the variables. For a nonlinear relationship structure, the distribution is formed on the probability density function of the data set considered for the model. The probability density function is, in the most general sense, the probability function of a continuous random variable. The first attempts to estimate probability density functions appear to have been made by Karl Pearson (1902a, 1902b) (Wegman, 1972). This function gives the probability that in a continuous set of values the variable occurs within a certain range and thanks to the density function, the closest real relationship between variables will be demonstrated in a healthier way.

Thus, the distribution of the functional shape of the model is released without being compressed to a certain mold like the assumption of linearity. Also, the fact that the method in question doesn't require a sharp assumption about the relationship between dependent and continuous independent variables, making it possible to examine both the parametric and nonparametric approaches together, makes the semiparametric approach more attractive than the parametric approach. The semiparametric estimation allows us to see the effect of each variable on the dependent variable by including the parametric and non-parametric parts in the model at the same time. While the parametric part is estimated according to the parametric rules, the nonparametric part is analyzed according to the rules of the non-parametric method. Thus, the problems that arise by handling both approaches one by one are eliminated (Robinson, 1988).

In the model discussed, the square meter and the age of the housing representing the residential area are a continuous variable and the other variables are discrete. The semiparametric approach will come to the fore as it allows us to examine the effect of continuous and discrete variables together and eliminate the negativities such as the size problem that arises in the nonparametric prediction. The semiparametric regression model in which some of the variables are parametric and the other part is nonparametric is as follows:

$$y = X\beta + m(Z) + \varepsilon \tag{1}$$

In the model, X and Z matrices are independent variable matrices consisting of parametric and nonparametric variables, respectively. β is the parameter vector, and m (.) Is the function of nonparametric variables. ε is a constant variance, normally distributed, independent error term. (Çağlayan Akay and Kangallı Uyar, 2017).

Semi-logarithmic functional form is commonly used in hedonic price modeling (Selim, 2009). This functional form is preferred because it fits the data well and because the coefficient estimates generated from the model can be interpreted as being the proportion of a good's price

that is directly attributable to the respective characteristics of this good (see, Halvorsen and Palmquist, 1980).

Thus, parametric and semiparametric model where housing area, housing age, elevator, parking and mortgage credit variables are statistically significant are as follows:

$$LNHP_i = \alpha + \beta_1 EL_i + \beta_2 PR_i + \beta_3 MC_i + \beta_4 HAG_i + \beta_5 HAR_i + \varepsilon_i \quad \varepsilon_i \sim niid(0,\sigma^2) \quad (2)$$

$$LNHP_i = \alpha + \beta_1 EL_i + \beta_2 PR_i + \beta_3 MC_i + m_1(HAG) + m_2(HAR) + \varepsilon_i \quad \varepsilon_i \sim niid(0,\sigma^2) \quad (3)$$

Table 3. Parametric and Semiparametric Estimation R Programming Results

Variables	Parametric Estimation	Semiparametric Estimation
C	11.16**	5.44934**
	(0.08087)	(0.06671)
EL	0.2377 **	0.20414 **
	(0.05497)	(0.05446)
PR	0.2614 **	0.26350**
	(0.04494)	(0.04446)
MC	0.2592**	0.24701**
	(0.06298)	(0.06190)
HAG	0.01082**	F Test Stat:
	(0.002391)	125.128 * [0.000]
		Degrees of freedom
		2.828
HAR	0.009552**	F Test Stat:
	(0.000)	9.323 * [0.000]
		Degrees of freedom
		6.383
Adjusted R -square	0.5548	0.575
Multiple R2	0.5671	
F Statistic	243.3 [0.000]*	
Deviance Explained		58%
Scale estimate		0.35119
GVC		0.35602

Note: (i) Depended variable is logarithmic housing price
(ii) *, ** and *** indicate significance at 1%, 5% and 10% statistical levels, respectively.
(iii) Numbers in parenthesis are standard errors and in square brackets are p-values

According to the prediction output above, all variables in both the parametric part and the nonparametric part are statistically significant. Sign expectations regarding all variables are in the expected direction. Only the housing age isn't meet the expectation. Among the parametric variables, while parking lot variable in question increases the housing prices by 30%, the elevator variable increases by 23% and the morgage credit variable increase by 28%.

The results of the nonparametric part of the model include the degree of smoothing of the variables and F statistical values. The nonlinear relationship between house prices and nonparametric variables are expressed in graphical representation, and the graph showing this relationship is as follows:
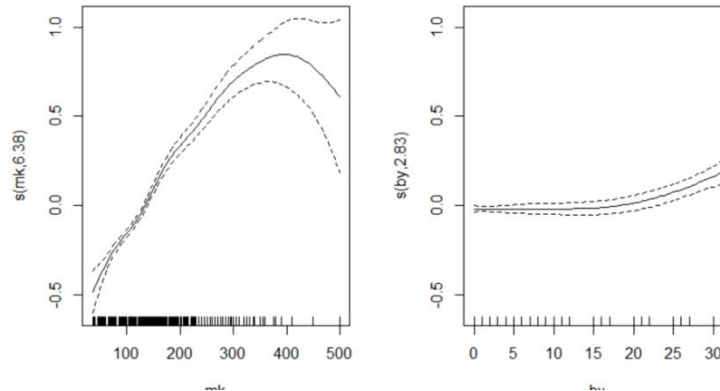
Figure 1: Semiparametric Regression Model Estimation Results, Graphs
* Dashed lines indicate 95% confidence interval

The figure shows that there is a nonlinear structure relationship between house prices and square meter and housing age variables. Graphs are important in terms of expressing how the coefficient estimates change against the change in each nonparametric variable. While the values belonging to the nonparametric variable are included in the horizontal axis, the coefficient estimates are located in the vertical axis. The dashed lines represent the 95% confidence interval. According to the graphics, it has been determined that the housing age variable has an effect on increasing housing prices after about 15 years. The reason for obtaining a positive effect may be that the positive effect of other features of the house prevents the house price from decreasing, even though the housing age increases. In other words, due to the predominance of other features, we can think that the age of the housing doesn't reflect enough to create a negative effect. When the relation between square meter and housing prices is considered nonparametric, an inverse result was obtained with the prior knowledge of "price increases as square meter increases" after a certain measurement area. According to the result obtained, a decrease in housing prices was detected after 400 square meters.

In the semiparametric model, adding the wrong variable to the linear and nonlinear components affects the reliability of the results. Therefore, variables should be tested.

The F test results show that the effect of nonparametric variables, and the LR test results that enable us to choose between parametric and semiparametric models are as follows:

Table 4. Partial F Test and Likelihood Ratio (LR) Results

HAR	**Partial F Test I**
Model 1: lnHP~EL+PR+MC+s(HAG)	F-statistic : 146.4 *
Model 2: lnHP~EL+PR+MC+s(HAG)+s(HAR)	[0.000]
HAR	**Partial F Test II**
Model 1: lnHP~EL+PR+MC+HAR+s(HAG)	F-statistic : 5.8558*
Model 2: lnHP~EL+PR+MC+s(HAG)+s(HAR)	[0.000018]
HAGE	
Model 1: lnHP~EL+PR+MC+s(HAR)	**Partial F Test I**
Model 2: lnHP~EL+PR+MC+s(HAG)+s(HAR)	F-statistic : 7.9778*
	[0.0000061]
HAGE	
Model 1: lnHP~EL+PR+MC+HAG+s(HAR)	**Partial F Test II**
Model 2: lnHP~EL+PR+MC+s(HAG)+s(HAR)	F-statistic : 5.5259*
	[0.000976]

Note: (i) *, ** and *** indicate significance at 1%, 5% and 10% statistical levels, respectively
(ii) Numbers in square brackets are probability values.
(iii) Ki-Kare test statistics $\chi^2_{7,0.05} = 14.06$

According to the partial F test I, the basic hypothesis states that the square meter and housing age variables are ineffective on housing prices. According to the result obtained, square meter and housing age have an impact on housing prices. This effect shows a nonlinear structure as explained in Figure 1.

According to the partial F test II results, the basic hypothesis that the square meter and housing age variables should be included in the model in parametric form is rejected. Therefore, it is concluded that these variables should be included in the model in a nonparametric form. This result supports the nonlinear structure of the housing age and square meters variables obtained from the graphical representation.

The basic hypothesis, which states that there is no difference between the parametric and semiparametric models according to the likelihood ratio used in determining which of the semiparametric and parametric models are suitable, is rejected. Therefore, the appropriate model is determined as the semiparametric model.

5. Conclusions

In this study, the factors affecting the housing prices in the province of Istanbul within the framework of the hedonic price approach was estimated by both parametric and semiparametric methods. Many estimation problems in econometrics involve an unknown function and an unknown finite-dimensional parameter. Parametric methods for estimating empirical models in economics and many other fields rely on strong and sharp assumptions about functional form and the distributions of variables (normal distribution etc.). In addition, the assumption that the functional form of the model is known in the parametric approach is rather restrictive and if the functional form of the model is incorrectly determined in the parametric approach, a specification error occurs. On the other hand, variables that do not have a parametric form can also be included in the model. More reliable estimates are obtained in semiparametric results

since it reduces the size problem compared to the nonparametric estimate and with the flexibility it provides for functional form, there is less risk of model building error compared to parametric estimation. According to the results all variables in both parametric and nonparametric parts are statistically significant.

The results show that the semi-parametric model should be chosen and that the square meter and housing age variables are in a non-linear relationship with the house prices. Therefore, since the model has both continuous and discrete variables, the semiparametric approach that allows examination of both nonparametric and parametric relationships was used in the analysis. Nowadays, relationship between variables contain nonlinear structures. Thus, the study differs from other studies in the literature with the method applied. Since the method applied takes into account the non-linear relationship between housing prices and features of the houses, it is superior to other studies in the same area.

Sign expectations regarding all variables were in the expected direction except for the housing age variable. For the housing age, it has been determined that the variable in question has an effect on increasing the housing prices after about 15 years. As a reason it was stated that the positive effect of other features of the house may be the reason for the positive effect to prevent the decline in the price of the house even as the age of the housing increases. Among the parametric variables, the variable that most affects the housing prices are parking lot. The variable in question increases the housing prices by 30% while the elevator variable increases by 23% and the morgage credit variable increase by 28%. On the other hand according to the result obtained, after 400 square meters, a decrease in housing prices was detected. As a reason it was stated that the positive aspects of the other features of the house may reduce the importance of the housing area. For example, instead of spending money on a house with a large area and many deficiencies, it may prefer to choose another house with a smaller area and many features. Therefore, this indicator supports the obtained result. The results of the study can be a guide for both sellers and buyers in the housing market. In addition, it is considered that determining the most prominent factors affecting house prices is very important for resource use efficiency and planning in this area.

References

Alkay, E. (2008). Housing submarkets in Istanbul. *International Real Estate Review*, *11*(1), 113-127. Retrieved from https://econpapers.repec.org

Arıkan, F. E. (2008). *Ev kiralarını etkileyen faktörlerin hedonik fiyat yöntemi ile belirlenmesi* (Unpublished doctoral dissertation). Marmara Üniversitesi Sosyal Bilimler Enstitüsü,İstanbul.

Ayvaz, Ö. (2002). *Emlak fiyatlarının hedonik model ile araştırılması: İzmir örneği* (Unpublished doctoral dissertation). Dokuz Eylül Üniversitesi Sosyal Bilimler Enstütüsü, İzmir.

Baldemir, E. K., Kesbiç, C.Y. ve İnci, M. (2007, Mayıs). *Emlak piyasasında hedonik talep parametrelerinin tahminlenmesi : Muğla örneği* [Estimating hedonic demand parameters in real estate market: The caseof Muğla]. 8. Türkiye Ekonometri ve İstatistik Kongresi'nde sunulan bildiri. Malatya. Retrieved from https://docplayer.biz.tr/7196332-Emlak-piyasasinda-hedonik-talep-parametlerinin-tahminlenmesi-mugla-ornegi.html

Chau, K. W. and Chin, T. L. (2003). A critical rewiev of literature on the hedonic price model. *International Journal for Housing and its Applications*, *27*(2), 145-165. Retrieved from https://ssrn.com

Çağlayan, E. and Arikan, E. (2011). Determinants of house prices in Istanbul: a quantile regression approach. *Qual Quant 45*, 305–317. https://doi.org/10.1007/s11135-009-9296-x

Çağlayan Akay, E. ve Kangallı Uyar, S.G. (2017). Semiparametrik tahmin yöntemi. *R Uygulamalı Ekonometri* içinde. İstanbul: Der Yayınları.

Demir, D. and Yayar, R. (2014). Hedonic estimation of housing market prices in Turkey. *Erciyes Üniversitesi İİBF Dergisi*, (43), 67-82. https://doi.org/10.18070/euiibfd.42448

Filho, C. M. and Bin, O. (2003). Estimation of hedonic price function via additive nonparametric regression. *Empirical Economics*, *30*(1),93-114. https://doi.org/10.1007/s00181-004-0224-6

Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, *5*(4), 471-484. https://doi.org/10.1016/0094-1190(78)90004-9

Halvorsen, R. and Palmquist, R. (1980, June). The interpretation of dummy variables in semilogrithmic regressions. *American Economic Review, 70*(3), 474-475. Retrieved from http://www.jstor.org

Horowitz, J. L. (2009). *Semiparametric and nonparametric methods in econometrics*. New York: Springer

Kain, J. F. and Quigley, J. M. (1970). Measuring the value of housing quality. *Journal of The American Statistical Association*, *65*(33), 532-548. Retrieved from http://www.jstor.org

Kangallı Uyar, S.G. ve Yayla, N. (2016). Konut fiyatlarının hedonik fiyat yaklaşımına göre mekansal ekonometrik modeller ile tahmini : İstanbul konut piyasası örneği [The estimation of housing prices within the frame of hedonic pricing approach by spatial econometric models: the case of Istanbul housing market]. *Social Sciences*, *11*(4), 326-342. http://dx.doi.org/10.12739/NWSA.2016.11.4.3C0153

Kaya, A. (2012). *Türkiye'de konut fiyatlarını etkileyen faktörlerin hedonik fiyat modeli ile belirlenmesi.* (Unpublished doctoral dissertation). Türkiye Cumhuriyeti Merkez Bankası İstatistik Genel Müdürlüğü, Ankara.

Lancester, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*(2),132-157. Retrieved from http://www.jstor.org

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370-396. https://doi.org/10.1037/h0054346

Macedo, R. (1996). Hedonic price models with spatial effects: an application to the housing market of Belo Horizonte, Brazil. *Revista Brasiliera de Economia*, *29*(3), 343-365. Retrieved from https://ideas.repec.org

Neyman, J. (1937). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society, 97*(4), 558-625. Retrieved from http://www.jstor.org

Ogwang, T. and Wang, B, (2002). A hedonic price function for a northhern BC community. *Social Indicators Research , 61*(3), 285- 296. Retrieved from http://www.jstor.org

Ozturk, O. and Kavlak, K. B. (2019). Statistical inference using stratified ranked set samples from finite populations. In C. N. Bouza-Herrera and A.I.F Al-Omari (Eds.), *Ranked Set Sampling* (pp. 157-168). New York: Academic Press. https://doi.org/10.1016/B978-0-12-815044-3.00012-5

Ozus, E., Dokmeci, V., Kiroglu, G. and Egdemir, G. (2007). Spatial analysis of residential prices in Istanbul. *European Planning Studies, 15*(5), 707-721. https://doi.org/10.1080/09654310701214085

Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability. *Reader in Statistics, 236*(767), 333-380. https://doi.org/10.1098/rsta.1937.0005

Parmeter, C. H. (2007). Nonparametric estimation of a hedonic price function. *Journal of Applied Econometrics, 22*(3), 695-699. Retrieved from http://www.jstor.org

Pearson, K. (1902a). On the systematic fitting of curves to observations an measurements I. *Biometrika, 1*(3), 265-303. Retrieved from http://www.jstor.org

Pearson, K. (1902b). On the systematic fitting of curves to observations an measurements II. *Biometrika, 2*(1), 1-23. Retrieved from http://www.jstor.org

Powel, J. L. (1994). Estimation of semiparametric models. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics*, (pp. 2443-2521). https://doi.org/10.1016/S1573-4412(05)80010-8

Rao, T. J. and Rao, C. R. (2016). Review of certain recent advances in randomized response techniques. In A. Chaudhuri, C. T. Christofides and C.R. Rao (Eds.), *Handbook of Statistics* (pp. 1-11). https://doi.org/10.1016/bs.host.2016.01.001

Ridker, R. K. and Henning, J. A. (1974). The determinants of residential property values with special reference to air pollution. *The Review of Economics and Statistics, 49*(2), 246-257. Retrieved from http://www.jstor.org

Robinson, P. (1988). Semiparametric econometrics: a survey. *Journal of Applied Econometrics, 3*(1), 35-51. Retrieved from http://www.jstor.org

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy, 82*(1), 34-55. Retrieved from http://www.jstor.org

Salinas, V. I., Sedory, S. A. and Singh, S. (2019). Calibrated estimator of population mean using two-stage ranked set sampling. In C. N. Bouza-Herrera and A.I.F Al-Omari (Eds.), *Ranked Set Sampling* (pp. 219-327). United Kingdom: Academic Press. https://doi.org/10.1016/B978-0-12-815044-3.00016-2

Selim, S. (2008). Derterminants of house prices in Turkey: a hedonic regression model. *Doğuş Üniversitesi Dergisi, 9*(1), 65-76. https://doi.org/10.31671/dogus.2019.223

Selim, H. (2009). Derterminants of house prices in Turkey: Hedonic regression versus artificial neural network, *Expert Systems with Applications, 36*(2), 2843-2852. https://doi.org/10.1016/j.eswa.2008.01.044

Stadelmann, D. (2010). Which factors capitalize into house prices? A bayesian averaging approach. *Journal of Housing Economics, 19*(3),180-204. https://doi.org/10.1016/j.jhe.2010.07.002

Straszheim, M. R. (1973). Estimation of the demand for urban housing services from household interview data. *The Review of Economics and Statistics, 55*(1), 1-8. Retrieved from http://www.jstor.org

Straszheim, M. R. (1974). Hedonic estimation of housing market prices: a further comment. *The Review of Economics and Statistics, 56*(3), 404-406. Retrieved from http://www.jstor.org

Wegman, E. J. (1972). Nonparametric probability density estimation: I. A summary of available methods. *Technometrics, 14*(3), 533.546. Retrieved from http://www.jstor.org

Wittowsky, D., Hoekveld, J., Welsch, J. and Steier, M. (2020). Residential housing prices: impact of housing characteristics, accessibility and neighboring apartments: A case study of Dortmund, Germany. *Urban Planning and Transport Research, 8*(1),44-70. https://doi.org/10.1080/21650020.2019.1704429