Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 36 (2009) 2843–2852
Expert Systems with Applications
www.elsevier.com/locate/eswa

# Determinants of house prices in Turkey: Hedonic regression versus artificial neural network

Hasan Selim *

*Department of Industrial Engineering, Dokuz Eylul University, Bornova, 35100 Izmir, Turkey*

**Abstract**

Determinants of house prices in Turkey are examined in this paper using the 2004 Household Budget Survey Data. In property valuation and housing market research, the locational value is usually analyzed by hedonic methods that use multiple regression techniques on large data sets and require a formality based on microeconomic theory in the analyses. Because of potential non-linearity in the hedonic functions, artificial neural network (ANN) is employed in this study as an alternative method. By comparing the prediction performance between the hedonic regression and artificial neural network models, this study demonstrates that ANN can be a better alternative for prediction of the house prices in Turkey.


*Keywords:* House price; Hedonic regression; Artificial neural networks; Turkey

## 1. Introduction

For many households, owner-occupied houses do not only offer an alternative for a place to live in, but they also represent the most important chunk of assets in these household's portfolio. Indeed, in most industrialized countries real estate is the greatest component of private households' wealth. As a consequence, the value of their house has a major impact on households' consumption and savings opportunities (Case, Clapp, Dubin, & Rodriguez, 2004). House prices are therefore of great interest to real estate developers, banks, policy makers or, in short, the general public as well as to actual and potential home owners (Schulz & Werwatz, 2004).

The housing market is defined as one where housing services are allocated by the mechanism of supply and demand. One of the characteristics of the housing market that is different from the markets of goods and services is the inelasticity of housing supply. Housing services are one of the most expensive household expenditures. Changing housing prices have been of concern to both individuals and governments in that they influence the socio-economic conditions and have a further impact on the national economic conditions. Expectations of capital gains from housing investments would affect housing prices by increasing the demand for housing which in turn would cause high volatility in housing prices. This causes increases in housing prices since the supply of housing cannot adjust in the short run. The housing market can be influenced by macro-economic variables, spatial differences, characteristics of community structure, and environmental amenities (Kim & Park, 2005).

The valuation of real estate is required to provide a quantitative measure of the benefit and liabilities accruing from the ownership of the real estate. Valuations are required, and often carried out, by a number of different players in the marketplace such as real estate agents, appraisers, assessors, mortgage lenders, brokers, property developers, investors and fund managers, lenders, market researchers and analysts and other specialists and consultants. Market value is estimated through the application of valuation methods and procedures that reflect the nature of property and the circumstances under which the given property would most likely trade in the open market

* Tel.: +90 232 388 1047; fax: +90 232 388 7869.
*E-mail address:* hasan.selim@deu.edu.tr

0957-4174/$ - see front matter 
doi:10.1016/j.eswa.2008.01.044

(Pagourtzi, Assimakopoulos, Hatzichristos, & French, 2003). Numerous methods are available to estimate the market value in the literature. Pagourtzi et al. (2003) classified these methods into two categories: traditional and advanced. It is stated in the paper that the majority of all methods will rely upon some form of comparison to assess market value, and this may be done, in its simplest form, by a direct capital comparison or may rely upon a range of observations that allow the determining of a regression model. Any such method in their paper is referred to as 'traditional'. Other models or methods that try to analyze the market by mimicking the thought processes of the players in the market in an attempt to estimate the point of exchange are referred to as 'advanced'. Herein, the traditional valuation methods are comparable method, investment/income method, profit method, development/residual method, contractor's method/cost method, multiple regression method and stepwise regression method. On the other hand, the advanced valuation methods are artificial neural networks (ANN), hedonic pricing method, spatial analysis methods, fuzzy logic and autoregressive integrated moving average.

In property valuation and housing market research, the locational value is usually analyzed by hedonic methods that use multiple regression techniques on large data sets and require a formality based on microeconomic theory in the analyses. Hedonic methodology is mainly used for the market valuation of goods for their utility-bearing characteristics. The goods under consideration embody varying amounts of attributes and are differentiated by the particular attribute composition that they possess. In most cases, the attributes themselves are not explicitly traded, so that one cannot observe the prices of these attributes directly. In such a case, hedonic pricing models are very essential in order to determine how the price of a unit of commodity varies with the set of attributes it possesses. If the prices of these attributes are known, or can be estimated, and the attribute composition of a particular differentiated good is also known, hedonic methodology will provide a framework for value estimation (Ustaoğlu, 2003).

The term *hedonic* was used to describe "the weighting of the relative importance of various components among others in constructing an index of usefulness and desirability" (Goodman, 1998: p. 292). Rosen (1974: p. 34) defined hedonic prices as "the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them". Rosen (1974) comprehensively laid down a theoretical foundation for determining the bid price, or implicit value of the attributes of a commodity for different consumers. The bid price ($\varphi$) is defined as the maximum amount of money which a consumer is willing to pay for a good under the condition that he or she retains a specific level of happiness or utility. He proposed to utilize the information from the tangent of the market price curve with which the consumers or producers share the same value of the equilibrium conditions. The methods used to identify the consumer's bid price function and the producer's offer function ($o$) were fully discussed. The offer function is defined as a function to determine the minimum value of price which a producer should accept to sell a good for a certain profit. The relationship among market price, bid price and offer functions is shown in Fig. 1 (Hidano, 2002: p. 10).

As stated above, the theory of hedonic price functions provides a framework for the analysis of differentiated goods like housing units, whose individual features do not have observable market prices. The traditional use of hedonic estimation in housing studies has been for the purpose of making inferences about non-observable values of different attributes like air quality, airport noise, commuter access (railway, subway or highway) and neighborhood amenities (Janssen, Söderberg, & Zhou, 2001).

Over the past three decades, the hedonic-based regression approach has been utilized extensively in the housing market literature to investigate the relationship between house prices and housing characteristics. The primary reasons for such an extensive application are analyzing household demand for these characteristics as well as constructing housing price indices (see, e.g. Can, 1992; Sheppard, 1999). However, this approach is subject to criticisms arising from potential problems relating to fundamental model assumptions and estimation such as the identification of supply and demand, market disequilibrium, the selection of independent variables, the choice of functional form of hedonic equation and market segmentation. These problems have been of great concern in the literature (see, Fan, Ong, & Koh, 2006; Malpezzi, 2003; Sheppard, 1999).

Most of the price studies are conducted with hedonic modeling and other methods based on multiple regression analysis. Basically, these methods are appropriate to a straightforward estimation of the relationship between price and the various characteristics. However, these techniques might become problematic if the agenda of the appraisal is widened to include aspects such as outliers, non-linearity, spatial and other kinds of dependence between observations, discontinuity, and fuzziness. There

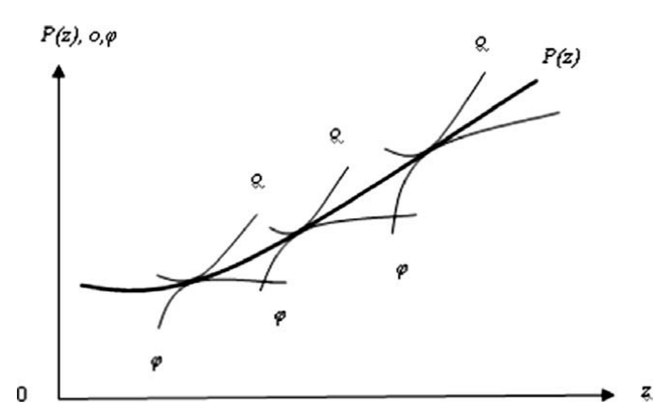


Fig. 1. Hedonic price function (Hidano, 2002: p. 11).

are, however, some plausible alternatives, one being the use of ANN, which are better suited to deal with these aspects. The neural network is, in fact, an example of a flexible regression approach. These types of methods are basically different from the standard methods. Specifically, they allow for a broader range of variation in the output than the hedonic regression model, with its spatial extensions. However, it is not clear how the coefficients in the model vary in space, and there is no straightforward functional relationship between the input and the output values (Kauko, 2003).

In this paper, determinants of house prices in Turkey are examined for the urban, rural and whole country using the 2004 Household Budget Survey Data. To our knowledge, no systematic empirical research exists in analyzing the house market in Turkey using hedonic price model. Because of potential non-linearity in the hedonic functions, ANN is employed as an alternative method in our analysis.

The remainder of this paper is organized as follows. Section 2 briefly introduces ANNs. Section 3 reviews the literature that employs hedonic regression and/or ANN models for real estate valuation. The data and functional form are introduced in Section 4. Section 5 reports the estimation results obtained by the hedonic regression and ANN models. Finally, we present some concluding remarks in Section 6.

## 2. Artificial neural networks

One important issue rising in the hedonics literature is the precise functional form to be employed and in this respect ANNs come to consideration. Fundamental to the practical application of ANN models is the concept of 'universal approximation', used, e.g. by Hornik, Stinchcombe, and White (1989). Put it very simply, this means that the networks are capable of adapting to or 'mimicking' arbitrary and unknown functional forms, with an arbitrarily specified degree of precision. Universal approximation leads us towards regarding neural networks in general as flexible non-linear statistical methods (Curry, Morgan, & Silver, 2002).

The most commonly applied network structure, the feed forward network, is employed in this study. Such a network model can be viewed as a case of flexible non-linear regression. The term 'non-parametric regression' may also be adopted. The ANN models use the same input and output parameters as in the linear models. These models have three primary components: the input data layer, the hidden layer(s) and the output measure(s) layer. Each of these layers contains nodes, and these nodes are connected to nodes at adjacent layer(s). The hidden layer(s) contain two processes: the weighted summation functions and the transformation function. Both of these functions relate the values from the input data to the output measures. The weighted summation function is typically used in a feed-forward/back propagation neural network model. Fig. 2 demonstrates a simplified neural network. Log of the property sales price is used as the output layer for the ANN model in this paper.

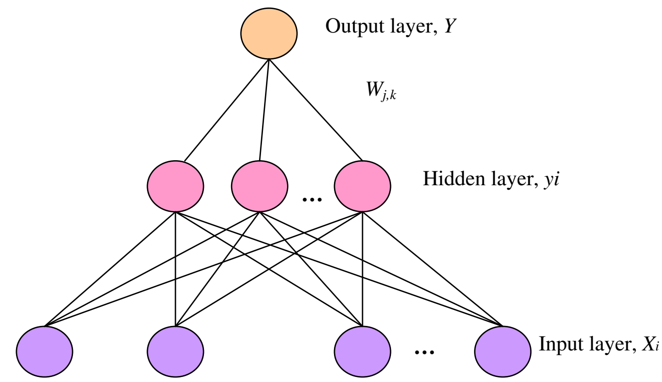


Fig. 2. A neural network model.

Among its various application areas, the ANN models have been applied to real estate valuation (see, e.g. Din, Hoesli, & Bender, 2001; Lenk, Worzala, & Silva, 1997; McGreal, Adair, Mcburney, & Patterson, 1998; Pagourtzi et al., 2003; Worzala, Lenk, & Silva, 1995).

## 3. Related literature

Hedonic price model is based on Lancaster (1966)'s consumer theory. Since this theory has been extended to the residential market by Rosen (1974), residential hedonic analysis has become widely used as an assessment tool and for property market and urban analysis. The regression of house prices on a variety of property specific and neighborhood descriptors evaluates their marginal contribution, also called implicit or hedonic prices. A comprehensive treatment of hedonic price theory is provided by Rosen (1974). A theory of hedonic prices is formulated as a problem in the economics of spatial equilibrium in which the entire set of implicit prices guides both consumer and producer locational decisions in characteristics space.

Residential housing is an important aspect of the quality of life in any community. Therefore, the appropriate valuation of specific characteristics of a residential house is in order. To achieve this objective, empirical researchers often specify hedonic price functions or hedonic models (Ogwang & Wang, 2003). Among the researches, Adair, Mcgreal, Smyth, Cooper, and Ryley (2000) focused upon factors affecting the price structure of residential property in the Belfast Urban Area, examining the relative influence of property characteristics, socio-economic factors and the impact of accessibility. The analysis highlights the importance of investigation at a sub-market level and draws conclusions regarding the complexity of relationships within an urban area. Janssen et al. (2001) compared the performance of least squares and least median of squares, a robust method, in the estimation of price/income relationships for apartment buildings. Multiplicative models with multiplicative errors are estimated by means of natural log transformations. The study confirms the importance of employing robust methods for this application. Meese and Wallace (2003) compared two methods to evaluate the effect of market fundamentals on housing price dynamics. The first method follows the traditional two-step procedures found in the literature in which one first estimates a house price index and then uses the estimated index in subsequent structural modeling. The second method applies a Kalman filter strategy that allows for the simultaneous estimation of the parameters of a dynamic hedonic price model, the price index and the parameters of a structural model for housing prices. Very similar empirical results are found for the two estimation strategies suggesting that the small efficiency gains of the simultaneous estimator may be outweighed by the relative ease of implementing the traditional methods.

Stevenson (2004) re-examined the issue of heteroscedasticity in hedonic house price models. The paper uses data for Boston, which has a high average age of dwelling. The results largely support the previous findings with the evidence of heteroscedasticity with respect to the age of dwelling. The iterative GLS correction, which is specified in terms of age, eliminates all heteroscedasticity at both aggregate and disaggregate levels. Fletcher, Mangan, and Raeburn (2004) argued that a wider range of diagnostic statistics should be used in the specification search for a good model, in particular, but not exclusively, those concerned with predictive stability. The paper illustrates this approach by examining both in-sample and out-of-sample diagnostic tests of various specifications of a hedonic house price model using data taken from the sale of over 1600 properties in the Midlands of the UK in 1999/2000. The models investigated include various specifications of the dependent and independent variables, including models that are non-linear in the parameters. The paper concludes that such statistics can often help in model selection and should be more widely employed.

Bin (2004) estimated a hedonic price function using a semi-parametric regression and compared the price prediction performance with conventional parametric models. Data from Geographic Information Systems (GIS) are incorporated to account for the locational attributes of the houses. The results show that the semi-parametric regression outperforms the parametric counterparts in both in-sample and out-of-sample price predictions, indicating that the semi-parametric model can be useful for measurement and prediction of housing sales prices. Bao and Wan (2004) illustrated how the technique of smoothing splines can be used to estimate hedonic housing price models. Their illustration takes the form of a rather limited, but very promising, application with Hong Kong data. In the forecasting comparison, the spline smoothing procedure outperforms the traditional parametric Box–Cox model in mean square error terms for out-of-sample predictions. Their results also suggest that the distortion caused by underfitting the model is smaller for spline smoothing than for the kernel and $k$-nearest-neighbor semi-parametric procedures. Kim and Park (2005) identified the spatial pattern of housing price changes and their determinants in Seoul and its neighboring new towns. The results of a cluster analysis show that the spatial pattern of housing price change rates is not correlated with housing prices. Filho and Bin (2005) modeled a hedonic price function for housing as an additive non-parametric regression. Estimation is done via a backfitting procedure in combination with a local polynomial estimator. It avoids the pitfalls of an unrestricted non-parametric estimator. Bandwidths are chosen using a novel plug-in method that minimizes the asymptotic mean average squared error of the regression. They compared their results to alternative parametric models and found evidence of the superiority of our non-parametric model. Fan et al. (2006) utilized the decision tree approach, which is an important statistical pattern recognition tool in examining the relationship between house prices and housing characteristics. Using the Singapore resale public housing market as a case study, the article demonstrates the usefulness of this technique. Kestens, Theriault, and Rosier (2006) introduced household-level data into hedonic models in order to measure the heterogeneity of implicit prices regarding household type, age, educational attainment, income, and the previous tenure status of the buyers. Two methods are used for this purpose: a first series of models uses expansion terms, whereas a second series applies geographically weighted regressions.

In recent years, the neural network modeling technique has become a serious alternative to and extension of more conventional property value modeling approaches. Several researches have been conducted in the literature. In one of these researches, Din et al. (2001) aimed to compare various real estate valuation models and the manner in which they take into account environmental variables. The reference model is taken to be a standard linear regression model including ordinal variables to measure environmental quality. They additionally found that ANN models, which are non-linear *per se*, exhibit a similar general form of the price indices, but that the detailed price behaviors of different models feature notable differences depending on the input choice of environmental variables. Kauko, Hooimeijer, and Hakfoort (2002) examined neural network modeling with an application to the housing market of Helsinki, Finland. The exercise shows how it is possible to identify various dimensions of housing sub-market formation by uncovering patterns in the data set, and also shows the classification abilities of two neural network techniques: the self-organizing map and the learning vector quantization. Curry et al. (2002) examined the potential of a neural network approach to the analysis of hedonic regressions, in which price is dependent on quality characteristics. The aim of the regressions is to measure, using objective data, the valuation consumers place on these characteristics. The results obtained provide an improvement on linear formulations, but the improvement in this case is relatively marginal. Curry et al. (2002) viewed neural network modeling as a useful mean of specification testing, and hence their results imply some support for a linear for-

Table 1
Descriptive statistics

| Variables | Full sample | | Urban | | Rural | |
|---|---|---|---|---|---|---|
| | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| Ln (house price) | 23.849 | 0.766 | 24.102 | 0.616 | 23.327 | 0.781 |
| *Locational characteristic* | | | | | | |
| Urban | 0.674 | 0.469 | | | | |
| Rural | 0.326 | 0.469 | | | | |
| *Type of house* | | | | | | |
| Detached | 0.401 | 0.490 | 0.237 | 0.425 | 0.742 | 0.438 |
| Semi-detached | 0.087 | 0.282 | 0.089 | 0.285 | 0.083 | 0.276 |
| Basement | 0.038 | 0.192 | 0.052 | 0.223 | 0.010 | 0.098 |
| Apartment | 0.442 | 0.497 | 0.583 | 0.493 | 0.152 | 0.359 |
| Shanty house | 0.021 | 0.144 | 0.027 | 0.162 | 0.009 | 0.095 |
| Duplex (base category) | 0.009 | 0.097 | 0.012 | 0.108 | 0.004 | 0.065 |
| *Age of building* | | | | | | |
| 0–5 (base category) | 0.082 | 0.274 | 0.081 | 0.273 | 0.083 | 0.276 |
| 5–10 | 0.169 | 0.375 | 0.184 | 0.388 | 0.139 | 0.346 |
| 10–15 | 0.184 | 0.387 | 0.197 | 0.398 | 0.158 | 0.364 |
| 15–20 | 0.139 | 0.346 | 0.148 | 0.355 | 0.120 | 0.325 |
| 20+ | 0.426 | 0.495 | 0.390 | 0.488 | 0.501 | 0.500 |
| *Type of building* | | | | | | |
| Ferroconcrete (base category) | 0.662 | 0.473 | 0.775 | 0.418 | 0.428 | 0.495 |
| Timber | 0.024 | 0.152 | 0.011 | 0.102 | 0.051 | 0.219 |
| Briquette | 0.086 | 0.280 | 0.075 | 0.263 | 0.107 | 0.310 |
| Stone | 0.049 | 0.216 | 0.017 | 0.129 | 0.116 | 0.320 |
| Brick | 0.131 | 0.337 | 0.099 | 0.298 | 0.198 | 0.398 |
| Mud brick | 0.049 | 0.215 | 0.024 | 0.152 | 0.100 | 0.301 |
| *Saloon floor* | | | | | | |
| Parquet (base category) | 0.167 | 0.373 | 0.229 | 0.420 | 0.040 | 0.196 |
| Board | 0.218 | 0.413 | 0.182 | 0.386 | 0.292 | 0.455 |
| Floor tile | 0.125 | 0.330 | 0.142 | 0.349 | 0.089 | 0.285 |
| Vinyl floor covering | 0.117 | 0.322 | 0.157 | 0.364 | 0.036 | 0.186 |
| Alum | 0.267 | 0.443 | 0.170 | 0.376 | 0.468 | 0.499 |
| Carpet, mosaic and marble | 0.106 | 0.308 | 0.120 | 0.326 | 0.076 | 0.265 |
| *Living room floor* | | | | | | |
| Parquet (base category) | 0.135 | 0.342 | 0.184 | 0.388 | 0.033 | 0.179 |
| Board | 0.241 | 0.428 | 0.207 | 0.405 | 0.313 | 0.464 |
| Floor tile | 0.113 | 0.316 | 0.133 | 0.339 | 0.071 | 0.257 |
| Vinyl floor covering | 0.138 | 0.345 | 0.185 | 0.388 | 0.041 | 0.199 |
| Alum | 0.269 | 0.443 | 0.173 | 0.378 | 0.467 | 0.499 |
| Carpet, mosaic and marble | 0.104 | 0.306 | 0.119 | 0.323 | 0.075 | 0.263 |
| *Bathroom floor* | | | | | | |
| Alum (base category) | 0.330 | 0.470 | 0.202 | 0.402 | 0.593 | 0.491 |
| Floor tile | 0.556 | 0.497 | 0.672 | 0.469 | 0.316 | 0.465 |
| Vinyl floor covering | 0.017 | 0.128 | 0.021 | 0.144 | 0.007 | 0.083 |
| Mosaic | 0.098 | 0.297 | 0.104 | 0.306 | 0.084 | 0.277 |
| *Heating system* | | | | | | |
| Stove (base category) | 0.786 | 0.410 | 0.703 | 0.457 | 0.957 | 0.203 |
| Central heating | 0.122 | 0.327 | 0.166 | 0.373 | 0.030 | 0.172 |
| Wall hung gas boilers | 0.092 | 0.290 | 0.131 | 0.337 | 0.013 | 0.112 |
| *Number of rooms* | | | | | | |
| 2 and under (base category) | 0.064 | 0.245 | 0.050 | 0.219 | 0.092 | 0.289 |
| 3 | 0.442 | 0.497 | 0.449 | 0.497 | 0.428 | 0.495 |
| 4 | 0.447 | 0.497 | 0.467 | 0.499 | 0.407 | 0.491 |
| 5+ | 0.047 | 0.211 | 0.034 | 0.181 | 0.073 | 0.260 |
| *Size (square meters)* | | | | | | |
| 70 and under (base category) | 0.080 | 0.271 | 0.066 | 0.248 | 0.109 | 0.312 |
| 70–110 | 0.654 | 0.476 | 0.664 | 0.473 | 0.633 | 0.482 |
| 110–150 | 0.238 | 0.426 | 0.240 | 0.427 | 0.233 | 0.423 |
| 150+ | 0.028 | 0.166 | 0.030 | 0.171 | 0.025 | 0.155 |

(*continued on next page*)

Table 1 (*continued*)

| Variables | Full sample | | Urban | | Rural | |
|---|---|---|---|---|---|---|
| | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| *Other structural characteristics* | | | | | | |
| Sauna-jacuzzi | 0.005 | 0.071 | 0.007 | 0.083 | 0.001 | 0.033 |
| Toilet | 0.891 | 0.311 | 0.966 | 0.182 | 0.738 | 0.440 |
| Garbage grinder | 0.002 | 0.049 | 0.003 | 0.056 | 0.001 | 0.033 |
| Water system | 0.961 | 0.194 | 0.998 | 0.043 | 0.883 | 0.321 |
| Hot water | 0.629 | 0.483 | 0.724 | 0.447 | 0.431 | 0.495 |
| Cable television | 0.043 | 0.204 | 0.063 | 0.244 | 0.002 | 0.046 |
| Elevator | 0.079 | 0.270 | 0.111 | 0.315 | 0.012 | 0.110 |
| Garage | 0.021 | 0.144 | 0.024 | 0.153 | 0.015 | 0.121 |
| Pool | 0.004 | 0.063 | 0.004 | 0.062 | 0.004 | 0.065 |
| Natural gas | 0.102 | 0.303 | 0.150 | 0.358 | 0.002 | 0.046 |
| Number of observation ($n$) | 5741 | | 3868 | | 1873 | |

Table 2
Hedonic model estimates

| Independent variables | Full sample | | Urban | | Rural | |
|---|---|---|---|---|---|---|
| | Coeff. | $t$ value | Coeff. | $t$ value | Coeff. | $t$ value |
| *Locational characteristic* | | | | | | |
| Urban | 0.233 | 14.140*** | | | | |
| *Type of house* | | | | | | |
| Detached | −0.494 | −8.120*** | −0.464 | −6.410*** | −0.316 | −2.780*** |
| Semi-detached | −0.478 | −7.570*** | −0.449 | −6.000*** | −0.372 | −3.150*** |
| Basement | −0.488 | −7.800*** | −0.452 | −6.250*** | −0.252 | −1.750* |
| Apartment | −0.367 | −6.320*** | −0.340 | −4.970*** | −0.066 | −0.600 |
| Shanty house | −0.398 | −5.520*** | −0.397 | −4.680*** | −0.280 | −2.120** |
| *Age of building* | | | | | | |
| 5–10 | −0.088 | −3.670*** | −0.060 | −2.160** | −0.129 | −2.710*** |
| 10–15 | −0.019 | −0.790 | 0.012 | 0.420 | −0.054 | −1.160 |
| 15–20 | −0.013 | −0.490 | 0.049 | 1.580 | −0.123 | −2.340** |
| 20+ | 0.010 | 0.430 | 0.046 | 1.620 | −0.033 | −0.760 |
| *Type of building* | | | | | | |
| Timber | −0.271 | −5.610*** | −0.108 | −1.470 | −0.340 | −5.450*** |
| Briquette | −0.087 | −3.050*** | −0.028 | −0.760 | −0.176 | −3.880*** |
| Stone | −0.228 | −5.420*** | 0.050 | 0.820 | −0.346 | −6.410*** |
| Brick | −0.030 | −1.260 | 0.023 | 0.760 | −0.104 | −2.740*** |
| Mud brick | −0.300 | −8.530*** | −0.158 | −3.180*** | −0.409 | −8.410*** |
| *Saloon floor* | | | | | | |
| Board | −0.089 | −2.260** | −0.086 | −2.040** | −0.157 | −1.480 |
| Floor Tile | −0.006 | −0.130 | −0.034 | −0.750 | −0.002 | −0.020 |
| Vinyl floor covering | −0.070 | −2.210** | −0.069 | −2.090** | −0.172 | −1.350 |
| Alum | −0.126 | −2.350** | −0.200 | −2.950*** | −0.143 | −1.200 |
| Carpet, mosaic and marble | −0.043 | −0.850 | −0.038 | −0.720 | −0.108 | −0.770 |
| *Living room floor* | | | | | | |
| Board | −0.012 | −0.290 | −0.004 | −0.090 | 0.001 | 0.010 |
| Floor tile | −0.169 | −3.750*** | −0.151 | −3.210*** | −0.080 | −0.600 |
| Vinyl floor covering | −0.118 | −3.680*** | −0.110 | −3.330*** | −0.048 | −0.380 |
| Alum | −0.132 | −2.420** | −0.129 | −1.880* | −0.100 | −0.850 |
| Carpet, mosaic and marble | −0.150 | −2.960*** | −0.161 | −2.990*** | −0.068 | −0.500 |
| *Bathroom floor* | | | | | | |
| Floor tile | 0.262 | 10.710*** | 0.216 | 7.530*** | 0.248 | 6.010*** |
| Vinyl floor covering | 0.201 | 4.570*** | 0.154 | 3.250*** | 0.191 | 1.640* |
| Mosaic | 0.070 | 2.300** | 0.001 | 0.030 | 0.117 | 2.180** |
| *Heating system* | | | | | | |
| Central heating | 0.048 | 2.690*** | 0.062 | 3.210*** | 0.012 | 0.290 |
| Wall hung gas boilers | 0.111 | 4.190*** | 0.104 | 3.910*** | 0.359 | 3.060*** |

Table 2 (*continued*)

| Independent variables | Full sample | | Urban | | Rural | |
|---|---|---|---|---|---|---|
| | Coeff. | *t* value | Coeff. | *t* value | Coeff. | *t* value |
| *Number of rooms* | | | | | | |
| 3 | 0.221 | 6.100*** | 0.275 | 6.240*** | 0.165 | 2.910*** |
| 4 | 0.310 | 8.180*** | 0.371 | 8.070*** | 0.243 | 4.000*** |
| 5+ | 0.430 | 8.690*** | 0.575 | 8.760*** | 0.302 | 4.000*** |
| *Size (square meters)* | | | | | | |
| 70–110 | 0.087 | 2.910*** | 0.034 | 0.980 | 0.142 | 2.860*** |
| 110–150 | 0.235 | 7.020*** | 0.194 | 5.050*** | 0.267 | 4.500*** |
| 150+ | 0.342 | 6.890*** | 0.312 | 5.600*** | 0.325 | 3.460*** |
| *Other structural characteristics* | | | | | | |
| Sauna-jacuzzi | 0.251 | 3.220*** | 0.291 | 3.680*** | −0.003 | −0.050 |
| Toilet | 0.339 | 12.230*** | 0.217 | 5.120*** | 0.366 | 10.420*** |
| Garbage grinder | 0.384 | 3.050*** | 0.189 | 1.510 | 0.942 | 7.380*** |
| Water system | 0.559 | 13.110*** | 0.091 | 0.520 | 0.538 | 11.740*** |
| Hot water | 0.133 | 8.850*** | 0.128 | 7.460*** | 0.128 | 4.380*** |
| Cable television | 0.370 | 13.580*** | 0.375 | 13.630*** | 0.159 | 1.040 |
| Elevator | 0.157 | 6.840*** | 0.170 | 7.170*** | 0.124 | 1.690* |
| Garage | 0.070 | 1.940** | 0.038 | 0.960 | 0.099 | 1.300 |
| Pool | 0.432 | 4.220*** | 0.594 | 4.310*** | 0.041 | 0.440 |
| Natural gas | 0.207 | 8.330*** | 0.230 | 9.200*** | 0.109 | 0.530 |
| Constant | 22.808 | 261.830*** | 23.580 | 119.460*** | 22.738 | 148.880*** |
| $R^2$ | 0.646 | | 0.551 | | 0.567 | |
| Adj. $R^2$ | 0.643 | | 0.546 | | 0.556 | |
| *F*-statistics (prob) | 225.83 | (0.000) | 104.14 | (0.000) | 53.08 | (0.000) |
| *White Test* | | | | | | |
| *F*-statistic (prob) | 10.276 | (0.000) | 6.866 | (0.000) | 3.129 | (0.000) |
| $nR^2$ (prob) | 439.945 | (0.000) | 289.315 | (0.000) | 134.033 | (0.000) |

*Note*: $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

mulation as an adequate approximation. Kauko (2003) evaluated the pros and cons of neural network models of property valuation in comparison with hedonic models, and provided some examples. Of particular interest is how the different locational, environmental, and social factors impact housing market segments and house price levels. It is argued that these objectives are conveniently handled with a method based on the self-organizing map. Some ideas for follow-up are also presented for this method. Liu, Zhang, and Wu (2006) proposed a fuzzy neural network prediction model based on hedonic price theory to estimate the appropriate price level for a new real estate. The model includes a database storing hedonic characteristics and coefficients affecting the real estate price level from recently sold projects that are representative in the local environment. The experimental result shows that the fuzzy neural network prediction model has strong function approximation ability and is suitable for real estate price prediction depending on the quality of the available data.

## 4. The data and functional form

The data set contains the 2004 Household Budget Survey Data for Turkey. The size of the estimation sample (5741) enables extensive modeling of the housing characteristics. The model contains 46 variables, which are presented in Table 1 together with the descriptive statistics. The variables include 'locational characteristic', 'type of house', 'age of the building', 'type of the building', 'saloon' and 'living rooms floors', 'bathroom floors', 'heating system', 'number of rooms', 'size' (square meters), and other structural characteristics. Because of the characteristics of the data, environmental factors cannot be considered.

While hedonic price models have been routinely used to analyze the market price of housing, selecting an appropriate functional form has been a frequent concern in the literature. The issue arises because there is little guidance from economic theory about the proper functional relationship between housing price and its attributes.

The most common functional form recommended in the hedonic literature is the semi-logarithmic form. This form is preferred because it fits the data particularly 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 & Palmquist, 1980). Thus, in this study we use the semi-logarithmic form for the models. More specifically, natural logarithm of the house price is treated as dependent variable. The model is as follows:

$$\ln P = \beta x + u.$$

here $P$, $\beta$, $x$ and $u$ denote house prices, coefficient matrix, set of independent variables and error term, respectively. Additionally, ordinary least square method is used in estimating the hedonic model. STATA 8.0 software is employed in our hedonic model analysis.

As stated previously, we employ in this study a generalized feed forward network with two hidden layers. In the ANN model, we use TanhAxon as the transfer function. The numbers of processing elements for the first and the second hidden layers are nine and four, respectively. At the solution stage, we use NEUROSOLUTIONS 4.0 software.

## 5. Results

In this section, the results obtained by the hedonic model and ANN are discussed. Table 2 provides the hedonic regression model results. As seen from the results, our specifications do not show any multicollinearity among explanatory variables, but heteroscedasticity is present as shown by White test statistics. Heteroscedasticity has long been recognized as a potential problem in hedonic house price equations. We have corrected the standard errors by using White's (1980) heteroscedasticity consistent coefficient covariance matrix. The results shown in Table 2 report that most of the variables are highly significant, and the sign of the coefficients are consistent with the expectations.

Percent effect for the hedonic model estimates is presented in Table 3. If the results are analyzed, it can be seen that house prices in urban area are higher than rural area by 26.26%. The results also denote that the prices of the other types of house are less in a range of 30–39% compared to the base category (duplex) for the urban area and full sample, while the range is 24–31% for the rural area. According to the results, the prices of houses that are between 5 and 10 years of age are less than those that are 0–5 years of age by 8%, 5.8% and 12%, respectively, for the full sample, urban and rural areas. Prices of the other types of building are less than those of the base category (ferroconcrete) in a range of 8.33–29.24%. Saloon floor and living room floor variables are insignificant in rural area. Saloon floor types of board, vinyl floor covering and alum negatively affect the house prices compared to the parquet in the full sample and urban area. On the other hand, the effect of the living room floor variables except the board on house prices is significant and negative again in the full sample and urban area. Additionally, bathroom floor variables have a positive effect on the prices in contrast to the effect of alum. The prices of the houses with central heating and wall hung gas boilers are higher than those with stove. Herein, the percent effects are 5% and 11.8%, respectively, in the full sample, while they are 6% and 11% in the urban areas. In rural area, having wall hung gas boilers increases the house prices by 43%, which is significantly higher than those in the full sample and urban area, compared to the houses with stove. As expected,

Table 3
Percent effect for hedonic model

| | Full sample | Urban | Rural |
|---|---|---|---|
| *Locational characteristic* | | | |
| Urban | 26.263 | | |
| *Type of house* | | | |
| Detached | −39.007 | −37.106 | −27.125 |
| Semi-detached | −38.011 | −36.171 | −31.077 |
| Basement | −38.604 | −36.351 | −22.283 |
| Apartment | −30.731 | −28.794 | * |
| Shanty house | −32.853 | −32.754 | −24.441 |
| *Age of building* | | | |
| 5–10 | −8.432 | −5.830 | −12.087 |
| 10–15 | * | * | * |
| 15–20 | * | * | −11.557 |
| 20+ | * | * | * |
| *Type of building* | | | |
| Timber | −23.702 | * | −28.840 |
| Briquette | −8.330 | * | −16.102 |
| Stone | −20.424 | * | −29.241 |
| Brick | * | * | −9.838 |
| Mud brick | −25.910 | −14.630 | −33.545 |
| *Saloon floor* | | | |
| Board | −8.552 | −8.269 | * |
| Floor tile | * | * | * |
| Vinyl floor covering | −6.796 | −6.672 | * |
| Alum | −11.881 | −18.104 | * |
| Carpet, mosaic and marble | * | * | * |
| *Living room floor* | | | |
| Board | * | * | * |
| Floor tile | −15.568 | −14.051 | * |
| Vinyl floor covering | −11.095 | −10.423 | * |
| Alum | −12.389 | −12.142 | * |
| Carpet, mosaic and marble | −13.953 | −14.838 | * |
| *Bathroom floor* | | | |
| Floor tile | 29.973 | 24.108 | 28.112 |
| Vinyl floor covering | 22.221 | 16.647 | 21.043 |
| Mosaic | 7.212 | * | 12.436 |
| *Heating system* | | | |
| Central heating | 4.969 | 6.387 | * |
| Wall hung gas boilers | 11.760 | 10.958 | 43.196 |
| *Number of rooms* | | | |
| 3 | 24.721 | 31.709 | 17.910 |
| 4 | 36.408 | 44.859 | 27.557 |
| 5+ | 53.723 | 77.691 | 35.286 |
| *Size (square meters)* | | | |
| 70–110 | 9.088 | * | 15.275 |
| 110–150 | 26.544 | 21.361 | 30.589 |
| 150+ | 40.764 | 36.570 | 38.370 |
| *Other structural characteristics* | | | |
| Sauna-jacuzzi | 28.479 | 33.785 | * |
| Toilet | 40.381 | 24.238 | 44.197 |
| Garbage grinder | 46.848 | * | 156.547 |
| Water system | 74.958 | * | 71.222 |
| Hot water | 14.220 | 13.700 | 13.646 |
| Cable television | 44.731 | 45.528 | * |
| Elevator | 17.024 | 18.478 | 13.194 |
| Garage | 7.292 | * | * |
| Pool | 53.961 | 81.062 | * |
| Natural gas | 22.997 | 25.886 | * |

*Coefficient is statistically insignificant.

Table 4
Predicted prices obtained by hedonic regression model and ANN

| Cases | Actual prices | Hedonic model prices | ANN prices |
|---|---|---|---|
| 1 | 24.27861 | 23.25467 | 24.06926 |
| 2 | 24.52993 | 22.85483 | 24.15476 |
| 3 | 23.43132 | 23.92889 | 23.35455 |
| 4 | 23.94214 | 24.12761 | 24.27606 |
| 5 | 25.32844 | 23.58241 | 24.45582 |
| 6 | 22.33270 | 24.74305 | 22.93157 |
| 7 | 22.59507 | 22.72321 | 23.06144 |
| 8 | 21.97603 | 22.74866 | 23.38378 |
| 9 | 21.41641 | 23.18521 | 22.82777 |
| 10 | 25.91622 | 22.38802 | 26.01527 |
| 11 | 24.12446 | 22.53630 | 24.46816 |
| 12 | 22.80271 | 22.89340 | 23.44858 |
| 13 | 23.71900 | 22.39763 | 23.22272 |
| 14 | 22.66918 | 22.72321 | 22.41153 |
| 15 | 23.28822 | 22.88452 | 23.30009 |
| 16 | 25.10529 | 23.40890 | 24.52888 |
| 17 | 25.16592 | 24.64970 | 24.64278 |
| 18 | 21.82188 | 25.05153 | 23.09273 |
| 19 | 22.92049 | 24.17086 | 22.70191 |
| 20 | 23.02585 | 23.77689 | 23.10048 |

the higher the number of rooms and house size the higher the house prices. Finally, the results indicate that most of the other structural characteristics have a significant and positive effect on the house prices. The effect changes between 7% and 156%.

The predicted prices obtained by hedonic regression model and ANN are presented in Table 4, and illustrated graphically in Fig. 3. As seen from the figure, the predictions of ANN model are more precise than those of the hedonic model. Finally, the performances of hedonic and ANN models are compared in Table 5 and in Fig. 4 in

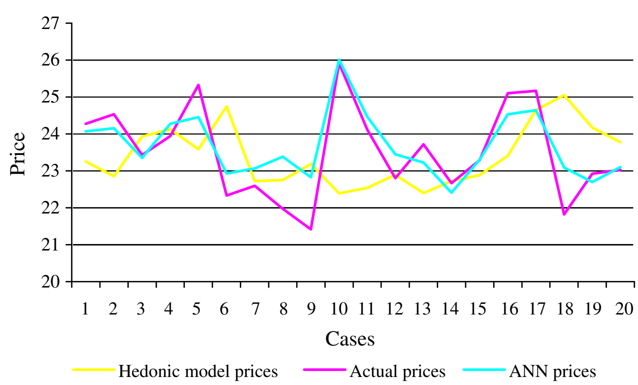


Fig. 3. Predicted prices obtained by hedonic regression model and ANN.

Table 5
Comparison of the performance of hedonic regression model and ANN

| Performance measures | Hedonic model | ANN |
|---|---|---|
| Mean squared error (MSE) | 2.4665 | 0.4373 |
| Root mean squared error (RMSE) | 1.5705 | 0.6613 |
| Mean absolute error (MAE) | 1.2319 | 0.5135 |

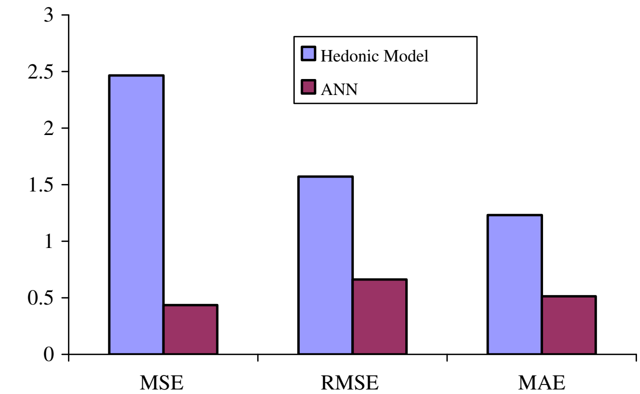


Fig. 4. Comparison of the performance of hedonic regression model and ANN.

terms of common performance measures, mean squared error (MSE), root mean squared error (RMSE) and mean absolute error (MAE). The results report that ANN is a better alternative for prediction of the house prices in Turkey in terms of all common performance measures.

## 6. Conclusions

This study examines the determinants of house prices in Turkey for the whole country, the urban and the rural areas. Two types of modeling approaches are employed in the analysis: hedonic regression model and ANN. The results of the hedonic model reveal that water system, pool, type of house, number of rooms, house size, locational characteristic and type of the building are the most significant variables that affect the house prices. Because of potential non-linearity in the hedonic functions, ANN is employed as an alternative method for the prediction. By comparing the prediction performance between the hedonic regression and ANN models, this study demonstrates that ANN can be a better alternative for prediction of the house prices in Turkey.

## References

Adair, A., Mcgreal, S., Smyth, A., Cooper, J., & Ryley, T. (2000). House prices and accessibility: The testing of relationships within the Belfast urban area. *Housing Studies, 15*(5), 699–716.

Bao, H. X. H., & Wan, A. T. K. (2004). On the use of spline smoothing in estimating hedonic housing price models: Empirical evidence using Hong Kong data. *Real Estate Economics, 32*(3), 487–507.

Bin, O. (2004). A prediction comparison of housing sales prices by parametric versus semi-parametric regressions. *Journal of Housing Economics, 13*, 68–84.

Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics, 22*, 453–474.

Case, B., Clapp, J., Dubin, R., & Rodriguez, M. (2004). Modeling spatial and temporal house price patterns: A comparison of four models. *Journal of Real Estate Finance and Economics, 29*(2), 167–191.

Curry, B., Morgan, P., & Silver, M. (2002). Neural networks and non-linear statistical methods: An application to the modelling of price-quality relationships. *Computers & Operations Research, 29*, 951–969.

Din, A., Hoesli, M., & Bender, A. (2001). Environmental variables and real estate prices. *Urban Studies, 38*(11), 1989–2000.

Fan, G., Ong, Z. S. E., & Koh, H. C. (2006). Determinants of house price: A decision tree approach. *Urban Studies, 43*(12), 2301–2315.

Fletcher, M., Mangan, J., & Raeburn, E. (2004). Comparing hedonic models for estimating and forecasting house prices. *Property Management, 22*(3), 189–200.

Filho, C. M., & Bin, O. (2005). Estimation of hedonic price functions via additive nonparametric regression. *Empirical Economics, 30*(1), 93–114.

Goodman, A. C. (1998). Andrew court and the invention of hedonic price analysis. *Journal of Urban Economics, 44*, 291–298.

Halvorsen, R., & Palmquist, R. (1980). The interpretation of dummy variables in semilogrithmic regressions. *American Economic Review, 70*(June), 474–475.

Hidano, N. (2002). The economic valuation of the environment and public policy: A hedonic approach. In W. E. Oates & H. Folmer (Eds.), *New Horizons in Environmental Economics*. Aldershot: Edward Elgar Publishing.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multi-layer feedforward networks are universal approximators. *Neural Networks, 2*(5), 359–366.

Janssen, C., Söderberg, B., & Zhou, J. (2001). Robust estimation of hedonic models of price and income for investment property. *Journal of Property Investment & Finance, 19*(4), 342–360.

Kauko, T. (2003). On current neural network applications involving spatial modelling of property prices. *Journal of Housing and the Built Environment, 18*(2), 159–181.

Kauko, T., Hooimeijer, P., & Hakfoort, J. (2002). Capturing housing market segmentation: An alternative approach based on neural network modeling. *Housing Studies, 17*(6), 875–894.

Kim, K., & Park, J. (2005). Segmentation of the housing market and its determinants: Seoul and its neighbouring new towns in Korea. *Australian Geographer, 36*(2), 221–232.

Kestens, Y., Theriault, M., & Rosier, F. D. (2006). Heterogeneity in hedonic modelling of house prices: Looking at buyers' household profiles. *Journal of Geographical Systems, 8*(1), 61–96.

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*, 132–157.

Lenk, M. M., Worzala, E. M., & Silva, A. (1997). High-tech valuation: Should artificial neural networks bypass the human valuer. *Journal of Property Valuation and Investment, 15*, 8–26.

Liu, J., Zhang, G. X. L., & Wu, W. P. (2006). Application of fuzzy neural network for real estate prediction. *LNCS, 3973*, 1187–1191.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In T. O'Sullivan & K. Gibb (Eds.), *Housing economics and public policy* (pp. 67–89). Malden, MA: Blackwell Science.

McGreal, S., Adair, A., Mcburney, D., & Patterson, D. (1998). Neural networks: The prediction of residential values. *Journal of Property Valuation and Investment, 16*, 57–70.

Meese, R., & Wallace, N. (2003). House price dynamics and market fundamentals: The Parisian housing market. *Urban Studies, 40*(5–6), 1027–1045.

Ogwang, T., & Wang, B. (2003). The hedonic price function for a Northern BC Community. *Social Indicators Research, 61*, 285–296.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: A review of valuation methods. *Journal of Property Investment & Finance, 21*(4), 383–401.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*, 34–55.

Schulz, R., & Werwatz, A. (2004). A state space model for Berlin House prices: Estimation and economic interpretation. *Journal of Real Estate Finance and Economics, 28*(1), 37–57.

Sheppard, S. (1999). Hedonic analysis of housing markets. In P. C. Cheshire & E. S. Mills (Eds.), *Handbook of regional and urban economics*. Amsterdam: Elsevier.

Stevenson, S. (2004). New empirical evidence on heteroscedasticity in hedonic housing models. *Journal of Housing Economics, 13*, 136–153.

Ustaoğlu, E. (2003). Hedonic price analysis of office rents: A case study of the office market in Ankara. Master of Science Thesis, Middle East Technical University, Turkey.

White, H. (1980). Heteroskedasticity-consistent covariance matrix and a direct test for heteroscedasticity. *Econometrica, 48*, 817–838.

Worzala, E., Lenk, M., & Silva, A. (1995). An exploration of neural networks and its application to real estate valuation. *Journal of Real Estate Research, 10*, 185–201.