*Asian Journal of Economics, Business and Accounting*

***21(1): 130-148, 2021; Article no.AJEBA.65400***
***ISSN: 2456-639X***

# Combining Housing Price Forecasts Generated Separately by Hedonic and Artificial Neural Network Models

**Salvatore Joseph Terregrossa[1*] and Mohammed Hussein Ibadi[1]**

[1]*Department of Business Administration, Istanbul Aydin University, Istanbul 34295, Turkey.*

***Authors' contributions***

*This work was carried out in collaboration between both authors. Author SJT designed the overall study (especially regarding the combining models analyzed and the testing and comparison of all forecasting models), managed all analyses of the study (including all of the in sample and out of sample analyses), helped to manage the literature searches, wrote the protocol and constructed the manuscript (in large part). Author MHI also helped with the design of the study (especially regarding the hedonic and artificial neural network models analyses), helped manage the literature searches, collected the data, and performed the statistical analysis. Both authors read and approved the final manuscript.*

***Article Information***

DOI: 10.9734/AJEBA/2021/v21i130345
*Editor(s):*
(1) Dr. Fang Xiang, University of International and Business Economics, China.
*Reviewers:*
(1) Damian Boniface Sambuo, Moshi Co-operative University (MoCU), Tanzania.
(2) Ramjeet Singh Yadav, Ashoka Institute of Technology and Management, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/65400

***Original Research Article***

***Received 01 December 2020***
***Accepted 06 February 2021***
***Published 13 February 2021***

## ABSTRACT

**Aims:** A) To enhance accuracy in forecasting *housing unit* prices by forming *combinations of component forecasts* generated separately by *hedonic* and *artificial neural network* models; B) To help ascertain whether a *constrained* or *unconstrained* linear combining model achieves superior forecasting performance.
**Place and Duration of the Study:** Department of Business Administration, Istanbul Aydin University, Istanbul 34295, Turkey**;** from 2019 to 2020.
**Study Design:** A *cross sectional* data set of *housing unit* prices and corresponding *housing unit* attributes and characteristics is formed and then randomly divided into two segments: *in sample* (80%) and *out of sample* (20%). Three different methods (hedonic, artificial neural network and combining) are then employed to process the same *in sample* data set, and generate *out of sample* forecasts. The three forecasting methods are then tested and compared.

**Corresponding author: E-mail: salvatore@aydin.edu.tr;*

**Methodology:** *Out of sample* combination forecasts are formed with component forecast weights generated by *in sample* weighted least squares (WLS) regression of realized price against *in sample* component forecasts. Four types of regressions are run: unconstrained, with and without a constant; constrained, with and without a constant. Then the *mean absolute forecast error* of each forecasting method is calculated and the mean difference in *absolute forecast error* between all pairs of models are compared and tested with a nonparametric *Wilcoxon sign rank test*.
**Results:** The combining model formed with component forecast weights generated by weighted least squares (WLS) regression with the constant term suppressed and the sum-of-the-coefficients *constrained to equal one*, generally performs the best, in comparison with all other forecasting models (component and combination) examined in the study.
**Conclusion:** The findings represent further evidence regarding the benefits of applying constraints on the linear combining forecast model; and demonstrate that a constrained linear combining model can be a successful technique for enhancing the forecast accuracy of *housing unit* prices.

*Keywords: Housing price forecasts; hedonic model; artificial neural network model; constrained, linear combining model.*

## 1. INTRODUCTION

Enhanced accuracy in forecasting *housing unit* prices is achieved by forming linear combinations of individual forecasts generated separately by *hedonic* and *artificial neural network* (ANN) models. Accuracy in estimating *housing unit* values is important for at least two groups: Prospective home-owners seeking a dwelling; and investors (in a particular housing market) seeking to add *real estate* assets to their portfolios. Members of each group certainly have a vested interest, first in determining the intrinsic value of a given housing unit, based on some established, generally accepted, objective set of characteristics and attributes. And then second, in making a comparison of the estimated intrinsic value with market price in order to help identify under-, over- and correctly-valued units.

This central aim of this empirical analysis is to develop a method of processing information, contained in an objective set of housing characteristics and attributes, that enhances accuracy of estimates of *housing unit* intrinsic value. Our empirical analysis begins with estimates of *housing unit* values generated separately by two different models: *hedonic* and *artificial neural network* (ANN). Then, combination forecasts are constructed, by forming weighted averages of the individual component forecasts generated separately by the *hedonic* and ANN models. This combining method then, in effect, is a third way to process the same information contained in the objective set of *housing unit* characteristics and attributes. A previous combining model study [1] demonstrates that combining forecasts generated by two different models may be effective if each model contributes *independent information* with regard to movement of the forecast variable. As has been explained previously [1], two individual forecasting methods may provide independent information (with regard to movement of a target variable) if each method processes different data; or if each method models the same data, but processes that data differently. In the present study, it is the latter case: The *hedonic* and ANN methods both model the same data (in the form of a particular set of characteristics and attributes of *housing units*), but process that data differently. A test of independent information may be achieved with a regression of *in sample* realized values against *in sample* component-model (*hedonic* and ANN) forecasts of the target variable (housing price). If the regression coefficients are all nonzero and statistically significant (which is the case in the present analysis), then this would imply that each of the individual component forecasts contain independent information, and thus forming an *out of sample* combination of the individual forecasts may lead to improved forecast accuracy. A combining model may be formed as a weighted average of the component forecasts with the *in sample* estimated regression coefficients serving as weights for the *out of sample* combination forecasts.

A limited number of other studies have also constructed combination forecasts of housing prices and real estate values, with an employment of different types of component model forecasts (structural and time series), and with different approaches to forming combinations. (See for example, [2-6]; these studies are summarized in section 1.1.3 below). While our study has some similarities, there are also some differences with each of these previous studies (as indicated in section 1.1.3), including one fundamental aspect that is germane to our analysis: We investigate, and provide additional insight, with regard to the unresolved debate [1] in the literature regarding whether *restricted- or unrestricted linear combination forecasting* achieves superior forecast results, for a given forecast variable. Briefly, the issue (revisited) is this: Unconstrained (unrestricted) regressions will result in unbiased estimators and minimum sum-of-squared errors for the data employed to fit the regression (of realized values against component model *in sample* forecasts) [7]. However, the objective is not to minimize the squared errors within the *in-sample* fitting data, but to enhance the accuracy of the *out-of-sample* forecasts [8]. Thus, if the process of constraining the linear combination leads to somewhat biased estimators, it may be worthwhile to trade off some incurred bias for more efficient estimators to enhance the accuracy of the *out-of-sample* forecasts. An estimator with lower dispersion about the mean (more efficient) and some bias will more closely approximate the true parameter than will an unbiased estimator with a larger dispersion about the mean [9].

The empirical findings of the current analysis are supportive of the method of the restricted linear combining model.

The organization of the paper is as follows: Following the section 1 above, section 1.1 provides a description of each of the three forecasting methods (section1.1.1: *hedonic* model; 1.1.2: *ANN* model; 1.1.3: *combining* model), inclusive of an essential, brief review of related studies. Section 2 presents the methodology, starting with an explanation of a) the *data* in section 2.1, and b) the *sample* in section 2.2; followed by an explanation of the estimation of each of the three *forecasting methods* in section 2.3 (section 2.3.1: *hedonic* method; 2.3.2: *ANN* method; 2.3.3: *combining* method).Section 2.3.4 explains the method by which the performances of the respective forecasting methods are measured, tested and compared. Section 3 presents and discusses the empirical results (section 3.1: *hedonic* regression analysis; 3.2: *ANN* model estimation analysis; 3.3: *combining* model estimation analysis).The different forecasting methods are compared in section 3.4.The paper closes with summary remarks and conclusions in section 4.

## 1.1 Background

### 1.1.1 Hedonic model

The *hedonic* model is a method to determine the value of an asset (such as a *housing-unit*), based on an underlying set of distinguishing characteristics and attributes. The *hedonic* theory supposes that the overall value of an asset can be considered as a collection of the values of specific, underlying attributes or characteristics of that asset [12]. Correspondingly, asset values within a group usually reveal the differences in quality, depending on how the attributes of each asset in the group are customized according to the customers' desires [13]. The advantage of *hedonic* models is that they have the ability to control for an asset's characteristics, and thus are able to distinguish the impact of marginal change in one of these attributes on the asset's intrinsic value [14]. These attribute parameters detect a housing-price's proportional change in relation to the proportional changes in its characteristics. Therefore, the *hedonic* price can be interpreted as an indication of the additional cost of buying a *housing unit* with a slightly better characteristic, *ceteris paribus* [15]. The *hedonic* model has been recognized and used in many *price prediction* settings, including housing values. However, many issues can affect the *hedonic* model's performance such as heteroscedasticity, multicollinearity, interactions of the independent variables, outlier data points and nonlinearity [16]. The *artificial neural network* (ANN) has been advanced an alternative model that avoids many of these issues (see for example [17] and [18].

### 1.1.2 Artificial neural network (ANN) model

The concept of *artificial neural network* (ANN) models stems from the universal approximation concept (see for example [19]). This concept holds that *artificial neural networks* have the ability to adapt to or mimic unknown (and perhaps arbitrary) *functional forms* to uncover relations among an asset's set of attributes and characteristics; and then use those discovered relations to forecast results [20]. Data is fed into an *artificial neural network* model with the aim of finding a function that can accurately map a particular set of inputs (such as housing characteristics and attributes) to an output (such as predicted housing value). The *universal approximation concept* holds that regardless of

[1] *With regard to this debate in the literature, see in particular the empirical studies of [9-11].*

the particular form a function (that relates inputs to outputs) may take, there exists some *neural network* that can accurately approximate that function.The concept of *universal approximation* has led to the use of neural networks in general as non-linear statistical methods that are a flexible (i.e., non/semi-parametric, model-free) regression technique not requiring predication on a prior, specific theory on which to advance [21, 22]. The structural system of an ANN model is complex, consisting of a group of primary units, called neurons, which are joined in netting structures consisted of interconnecting layers. The degree of complexity of a *neural network's* structure depends on the total number of neurons and the existing connections [23]. An *artificial neural network* (ANN) model structure has three main layers: the input data layer (characteristics and attributes of the entity under analysis), the hidden functional layer or layers (referred as the “black box”), and the output layer (such as *housing unit* price forecasts, as in the present study) [14]. Further, an ANN is considered as an “interconnected network” consisting of artificial neurons that have the ability to adjust the units' *connection weights and strength* according to the data externally provided. Within the *neural network matrix*, every neuron has *connecting units* to some of its neighbors. The sum of these *weighted input connections* will be transformed by a transfer function into an output (such as forecasts of housing prices, in the present analysis) [24].

### 1.1.3 Combination forecasting model

The concept of *combination forecasting* has been demonstrated to be an effective method for improving forecast accuracy, as indicated by many published studies, including those referenced in the present analysis. As is widely known, the central idea of the *combining model* is to form a weighted average of forecasts generated by two or more different models, with the aim of creating a forecast that is more accurate than any of the individual component forecasts. The initial (and conventional) approach is to combine one or more *structural model* forecasts with one or more *time-series model* forecasts of a given forecast variable. The idea is to offer a *structural explanation* of the variance of the forecast variable, in conjunction with a *time-series explanation* of that part of the variance that cannot be explained by the structural model [or models]. (See for example [25].) In effect, the conventional technique combines two or more models (*structural* and *time-series*, respectively) that contribute different or independent information as to the sources of movement of the forecast variable.

However, of particular interest to the present analysis is a combining model study [1] which pivots from the *conventional approach* by first hypothesizing, and then demonstrating that combining forecasts generated by two (or more) individual models may be effective if each model contributes *independent information* with regard to movement of the forecast variable, regardless of the types of models employed. In other words, a successful combining model does not need to exclusively combine a *structural model* forecast with a *time-series model* forecast. This alternative approach has been successfully applied by others, subsequently (including [2-3] and [5] as indicated below). The present analysis also follows this alternative tack, in forming combinations of forecasts of housing prices, generated separately by *hedonic* and ANN structural models. As mentioned above, a test of *independent information* may be achieved with a regression of *in sample* realized values against *in sample* component-model forecasts of the target variable. If the regression coefficients are all nonzero and statistically significant, then this would imply that each of the individual component forecasts contain independent information, and thus forming an *out of sample* combination of the individual forecasts may lead to improved forecast accuracy. The estimated *in sample* regression coefficients, if found to be nonzero and separately identified, give indication as to the appropriate proportional values of the weights to assign to each component forecast in the *out of sample* combination model (see for example [1]).

Lately, the combining model technique has been applied over a wide range of areas of interest, with various types of forecast variables. For example, [26] successfully applied the combination forecasting method in improving hydrological operational predications (i.e. flood forecasting). Also, the combining method has been applied in the prediction of energy consumption [27]; in solar radiation forecasting [28]; in tourism demand forecasting [29]; in electrical load forecasting [30] and in forecasting wind speed [31-33].

To date, (as mentioned above) the combination forecasting method has also been applied in a limited number of empirical studies that utilize real estate- and housing-pricing models. These studies are summarized as follows:

Beginning with [2], which forms combinations of forecasts of *real estate* value generated by two types of structural models: a repeat sales model and a high-speed net connect (HNC) model (from HNC software Inc.). Component structural model forecasts are generated with employment of an estimation data set (*in sample*), and forecast errors are measured. A linear combining model is formed with a weighted average of the *out of sample* component forecasts; with the forecast weights based on the inverse of the forecast errors of the *in sample* component model forecasts. This method was found to be successful in enhancing forecast accuracy of *real estate* value.

Another previous study [3] forms combinations of *real estate* value forecasts generated from structural models: the *repeat sales* model, the *tax assessment* method, the *hedonic* model and the *neural network* model (ANN). Using an approach similar to [2], each component model *in sample* forecast is assessed in terms of forecast error; and *out of sample* forecast weights are based on the inverse of the respective component model forecast errors. Different combinations are formed of forecasts by at least two models, those with the lowest component *in sample* forecast error. This method also proved successful in improving the accuracy of property values prediction (*out of sample*), over that of each of the component models.

In estimating housing prices, a study from New Zealand [4] employed the conventional (traditional) approach (as explained above) in generating a combining model using forecasts from a wide range of both structural and time-series models. An *out of sample* combining model is constructed three ways: a) With an equally-weighted average of component forecasts; b) With an *unrestricted* ordinary least squares (OLS) regression of *in sample* realized prices against component model *in sample* forecasts; with the estimated regression coefficients serving as weights for the *out of sample* combination forecasts; and c) With an approach similar to that of [2] and [3], with forecast weights calculated as the inverse of component model *in sample* forecast errors. While all three types of combining models were successful in improving forecast accuracy over the component model forecasts, the combining model formed with *unrestricted* OLS estimated coefficients as forecast weights proved superior.

A study from US [5] also experimented with the conventional approach by combining forecasts from a range of structural and time-series models. However, this study found that equally-weighted combinations formed using only forecasts from different, component structural-models (with no time-series forecasts included) outperformed all component model forecasts (structural and time-series); and also proved superior to a combination of structural and time-series model forecasts.

A study forecasting real estate market returns across nine nations spanning UK, EU and Asia [6] also experimented with combinations of time-series and structural model (including ANN) forecasts. While none of the structural, non-linear models outperformed the study's benchmark linear time-series model (martingale), the combining models formed with an equally-weighted average of the non-linear structural model forecasts and the linear time-series model forecasts (martingale) were superior to all component model forecasts (linear and non-linear).

In summary, all six of these previous studies forming *combinations* of *housing price* or *real estate value forecasts* (generated by time-series and/or structural models) have proved successful in increasing forecast accuracy over the respective component forecast models. In three of these studies [2-4] forecast weights for *out of sample* combinations were determined on the basis of the inverse of the forecast errors of the *in sample* component model forecasts. In three of these studies [4-6] combinations were formed with a simple, equally weighted average. One of these studies [4] conducted a controlled experiment by forming combinations with three different *weighting schemes*, and found that combinations formed with weights generated by *unrestricted* OLS regression were superior to combinations formed with weights determined on the basis of the inverse of the component model forecast errors, and also superior to combinations formed with a simple, equally weighted average.

Our study takes the analysis in [4] a step further, by demonstrating that a *constrained* linear regression combining model is superior to an *unconstrained* (i.e. *unrestricted*) linear regression combining model in improving forecast accuracy over component model forecasts of housing-price.

# 2. METHODOLOGY

## 2.1 Data

The area of study is the *real estate housing market* in the European area of Istanbul, Turkey. A sample of 100 housing units in Istanbul were randomly selected and retrieved through various real estate websites. [2] The sample was distributed between the residential areas and the city centre and other parts of the city, as well. The data covers the period from May 2019 to July 2019. Due to the fact that majority of the housing units in Istanbul are in the form of *multi dwelling* residential buildings, the study only included *housing units* in the form of *apartment units* in the data collection. For each apartment, the following information was taken: the sale price (*realized price*) in Turkish Lira (each 1 USD = 5.71 TL, as of the 15$^{th}$ of July, 2019); the *geographical location*; the *land size* in square meters; the *number* of *bedrooms* and *bathrooms*; the *property's age* in years; and the *apartment floor* within the building.

## 2.2 The Study Sample

According to standard analytical practice (see for example [16]), the study sample was divided randomly into two sets: The "estimation set" (*in sample*), and the "forecasting set" (*out of sample*), as known in *regression analysis* literature. (Or the "training set" [*in sample*] and the "production set" [*out of sample*], as known in *neural network* literature.) The *in sample* data set contains 80% of the test data and the *out of sample* data set contains the remaining 20% of data.

The *in sample* data set is used: a) to test the effects of a *housing-unit's* attributes and characteristics on its price; b) to generate component model (*Hedonic* and ANN) *in sample* housing-unit price-forecasts, with which to generate *forecast weights* (with WLS regressions) for the *out of sample* combination forecasts; and c) to help generate component model (*Hedonic* and ANN) *out of sample* housing unit price forecasts, as well, for use in the construction of *out of sample* weighted average combination forecasts.

## 2.3 Data Analysis

This study has employed (and evaluated) three different methods to process information contained in an objective set of housing characteristics and attributes, to generate estimates of intrinsic value of *housing units* (that is to say, to generate forecasts of *housing unit* prices):The *hedonic* and *artificial neural network* (ANN) models, and a *linear combining* model.

As mentioned, the objective set of housing characteristics and attributes includes: Geographical location; land size; age of the house; number of bedrooms and bathrooms; and the building floor of the unit.

### 2.3.1 The hedonic method

In the regression analysis in the estimation of the *hedonic* model, the *semi logarithmic* form has been the most commonly used functional form of the *hedonic* model (see for example [20,16] and [40]). This form is preferred, since it typically fits the data very well. The resultant estimated regression coefficients can be interpreted as being proportional to the property's price that is directly correlated to its characteristics and attributes. The present analysis follows this approach in using the *natural logarithm* of the *housing unit* price as the dependant variable.

The *hedonic model* implementation is carried out in three stages: First, with a linear regression analysis on a random 80% of the housing units (*in sample*): The housing units' (*in sample*) log prices (the dependant variable) are regressed against their characteristics and attributes (the independent variables), as described below in Eq. 2.

To begin, the implicit model for the *hedonic* price function (*f*) in the present analysis is very similar to the model employed by [20], and is formulated as follows:

*Price* = *f* (L, S, BD, BA, A, FL) *Equation (1)*

With the *housing unit* characteristics and attributes identified as follows:

(*L*) = the location
(*S*) = the land size in square meters (m$^2$)
(*BD*) = number of bedrooms

[2] *The following real estate websites were used to collect the study dataset:*

1. *[34]*
2. *[35]*
3. *[36]*
4. *[37]*
5. *[38]*
6. *[39]*

($BA$) = number of bathrooms
($A$) = house age in years
($FL$) = apartment floor within the building.

The *hedonic* model in *regression form* is stated as follows:

$$P_{it}(in\ sample) = \beta_0 + \beta_1 L_{it} + \beta_2 S_{it} + \beta_3 BD_{it} + \beta_4 BA_{it} + \beta_5 A_{it} + \beta_6 FL_{it} + \varepsilon_{it}$$

*Equation (2)*

Where $P_{it}(in\ sample)$ is the observed price of housing unit *i* (*in sample*); L, S, BD, BA, A, FL are the housing unit characteristics and attributes [3]; parameters $\beta_0 - \beta_6$ are fixed; $\varepsilon_{it}$ = error term; E [$\varepsilon_{it}$] = 0.

(See Table 1 for the hedonic model Eq. 2 regression results; and Appendix 1 for additional regression results of Eq. 2.)

Then, secondly, these estimated coefficient estimates (*in sample*) are used to generate *housing unit* price forecasts for the *in sample* housing units (80%), as follows;

$$\hat{P}_{it}(in\ sample) = \hat{\beta}_0 + \hat{\beta}_1 L_{it} + \hat{\beta}_2 S_{it} + \hat{\beta}_3 BD_{it} + \hat{\beta}_4 BA_{it} + \hat{\beta}_5 A_{it} + \hat{\beta}_6 FL_{it}$$

*Equation (3)*

Where,

$\hat{P}_{it}$ = the predicted price of *housing unit i* (*in sample*);
$\hat{\beta}_0$ = the *in sample* regression (Eq. 2) estimated *constant coefficient*;
$\hat{\beta}_1 - \hat{\beta}_6$ = the *in sample* regression (Eq. 2) estimated variable *coefficients*.

Thirdly, *housing unit* price forecasts for the remaining 20% housing units (*out of sample*) are generated using the *in sample* regression (eq. 2) estimated coefficients, as follows:

$$\hat{P}_{it}(out\ of\ sample) = \hat{\beta}_0 + \hat{\beta}_1 L_{it} + \hat{\beta}_2 S_{it} + \hat{\beta}_3 BD_{it} + \hat{\beta}_4 BA_{it} + \hat{\beta}_5 A_{it} + \hat{\beta}_6 FL_{it}$$

*Equation (4)*

Where,

$\hat{P}_{it}$ = the predicted price of *housing unit i* (*out of sample*);
$\hat{\beta}_0$ = the *in sample* regression (Eq. 2) estimated *constant coefficient*;
$\hat{\beta}_1 - \hat{\beta}_6$ = the *in sample* regression (Eq. 2) estimated variable *coefficients*.

In the present analysis, the issue of heteroscedasticity is anticipated, based on the evidence from previous studies regarding *hedonic housing price models*. For example, a property's age has been found to be a primary cause of heteroscedasticity [41-43]. A property's characteristic of *external area* has also been found to cause heteroscedasticity [44]. We test for the presence of heteroscedasticity and find that the spread of data along the regression line is heteroscedastic. As a result, we employ the *Weighted Least Squares* (WLS) technique in the above regression analysis.

### 2.3.2 The ANN Method

For the *artificial neural network* (ANN) model, *relative contribution factors* (for each of the *housing unit* explanatory variables) to the *housing unit* price were identified using the *multilayer perception* neural network analysis with one hidden layer on a random 80% of the studied sample (*in sample*). The resultant information from the *in sample* analysis is used to first generate *in sample* (80%) housing forecasts. Then secondly the resultant information from the *in sample* analysis is used to generate *out of sample* (20%) housing forecasts.

The *neural network* application method is a similar process to the *hedonic* price model, with the logarithm *housing unit* price as the dependant variable, and the *housing unit* set of attributes and characteristics (location; size; house age; number of bedrooms and bathrooms; building floor) the explanatory variables. However, the *neural network model* differs from the *hedonic model* in that, for a specific input (set of attributes and characteristics), an output (*housing unit* price forecasts) [both *in sample* forecasts and *out of sample* forecasts] is directly

[3] *The housing unit characteristics and attributes were represented as numerical values in the statistical analysis (using the Statistical Package for the Social Sciences [SPSS] software [IBM, Version 24] for Microsoft Windows). For example, numerical codes were given (such as 1, 2, 3…etc.) for the properties' locations, where each number represented a particular neighbourhood in Istanbul. And this coding system was applied for the rest of the housing unit attributes and characteristics, as well. All the information concerning the details of the codes for the housing unit characteristics is available upon request.*

generated from the model. (See Appendix 2 for additional description of the estimation of the ANN model, in the present study.)

### 2.3.3 The combination forecast method

In the present study, the *combining model* technique is implemented as follows: First, a test of *independent information* in the *hedonic* and *ANN* models is done with a WLS regression of actual (realized) housing prices (*in sample* values [80%]) against the *in sample* housing price forecasts generated separately by the *hedonic* and the *ANN* models (employing *in sample* data), in the following fashion:

$$P_{it}(in\ sample) = \alpha + \beta(\hat{P}_{it}[hedonic]) + \gamma(\hat{P}_{it}[ANN]) + \varepsilon_{it} \qquad \textit{Equation (5)}$$

Where,

$P_{it}(in\ sample)$ = the actual (realized) housing unit price from the *in sample* data set; $\hat{P}_{it}[hedonic]$ = the *in sample* forecasts generated by the *hedonic* model using *in sample* data; $\hat{P}_{it}[ANN]$ = the *in sample* forecasts generated by the ANN model using *in sample* data; parameters $\alpha, \beta$ and $\gamma$ are fixed; $\varepsilon_{it}$= error term; E [$\varepsilon_{it}$] = 0.

If the estimated regression coefficients $\widehat{\beta}$ and $\widehat{\gamma}$ are both nonzero and separately identified, then this would indicate that both models contain independent information and can be useful in generating combination forecasts. Indeed, this is the case in the present analysis (see Table 3). Subsequently, the *in sample* estimated regression coefficients are used as weights for the *out of sample* component *hedonic* and *neural network* forecasts to generate *out of sample* combination forecasts.

In a similar vein as the studies of [9-11], Eq. 5 is estimated in the present study by using *weighted least squares* (WLS) (to overcome the issue of heteroscedasticity), and applying in turn four variations with regard to regression restrictions, as follows:

1. WLS with a constant term and unrestricted coefficients;
2. WLS with a suppressed constant term and unrestricted coefficients;
3. WLS with a constant term and the *sum-of-the-coefficients* constrained to equal one;
4. WLS with a suppressed constant term and the *sum-of-the-coefficients* constrained to equal one.

(See Table 3 for the four sets of Eq.5 regression results; and Appendix 3 for further details.)

In the present analysis, each set of estimated regression coefficients $\widehat{\beta}$ and $\hat{\gamma}$ is then alternately employed as forecast-weights to form *out-of-sample* combination forecasts of housing-price for each Istanbul *housing unit* in the *out of sample* data production set.The *combining model* is constructed by forming a weighted average of both hedonic and ANN model forecasts (*out of sample*) as follows:

$$F_c = w_1\,(hedonic) + w_2\,(ANN) \qquad \textit{Equation (6)}$$

Where,

$F_c$ stands for *combination forecast*;

*Hedonic* refers to the *hedonic* model *out of sample* forecast;

*ANN* refers to the *out of sample* forecast from *artificial neural network* model;

$w_1$, $w_2$ are the proportional weights, which are the estimated regression coefficients $\widehat{\beta}$ and $\hat{\gamma}$ from the *in sample* independent information test WLS regressions (unrestricted and restricted; with and without a constant, respectively) of Eq. 5.

The four combining models' forecasts are generated and compared to each other (and to the component model forecasts) to determine the superior method.

### 2.3.4 Comparing forecasting model performance

To determine superior performance, *mean absolute forecasting errors* (MABE) of each model (component- and combining-) are calculated and compared. MABE is calculated by first measuring the *absolute value* of the difference between the actual housing unit price and the predicted housing price. The average of these absolute values is the MABE:

$$MABE = \frac{1}{n}\sum\left|P - P_y\right| \qquad \textit{Equation (7)}$$

Where $P$ is the actual (*realized*) *housing unit* price, $P_y$ is the predicted price and *n* is the data number.

The *Wilcoxon signed rank test* is employed as nonparametric test of the mean difference between the *absolute forecast errors* of two different models, for all pairs of the estimated models (*hedonic, ANN, and combining* models) and across all *housing units* in the *out of sample* data set.

## 3. RESULTS AND DISCUSSION

### 3.1 Hedonic Regression Analysis

The *hedonic pricing model* results presented in Table 1 and generated with *weighted least squares* (WLS) analysis on the estimation set (*in sample*), indicate that the number of bedrooms, the number of bathrooms and the unit floor, respectively are significantly, positively related to *housing unit* price. The *housing unit* location within the city is also positively related, but not significantly. The *housing unit's* land size is significantly, negatively related to *housing unit* price. The age of the *unit* is also negatively related, but not significantly. (Appendix 1 for additional *hedonic* model estimation results.)

### 3.2 Artificial Neural Network Analysis

The *neural network analysis*, by the *multilayer perception with one hidden layer*, was done on the *in sample* (80%) set. The *log prices* of the *housing units* were used as the dependant variable, in order to eliminate data skewness and outliers. The *relative contribution factors* of the *artificial neural network* analysis (indicating the relative importance of inputs) are shown in Table 2 and Fig. 1. (Appendix 2 for additional ANN model estimation results.)

### 3.3 Combining Model Analysis

As outlined above, *independent information* tests are carried out with WLS regression analysis employing *in sample* information (Eq.5). Each of the four, respective WLS *regression model* specifications (unrestricted, with and without a constant; and constrained, with and without a constant) generate significantly positive coefficients with regard to both the *Hedonic* and *ANN* component model forecasts (*in sample*), indicating that both component model forecasts contain independent information; and thus are useful in forming *out of sample* combination forecasts. (Table 3; and Appendix 3 for additional Eq. 5 estimation results.) The Eq. 5 estimated *in sample* regression coefficients $\widehat{\beta}$ and $\hat{\gamma}$ then serve as *forecast weights* for the *out of sample* forecasts.

### 3.4 Model Comparisons

#### 3.4.1 Hedonic vs ANN forecasts

Tables 4 and 5 indicate that that the *neural network (ANN)* model forecasts (*out of sample*) are superior to those of the *hedonic* model (*out of sample*), in terms of lower *mean absolute error* (MABE) (Table 4); with the *Wilcoxon signed rank test* (of the average difference between the absolute errors of the hedonic and ANN models) significantly positive. (Table 5; Appendix 4.)

#### 3.4.2 Combining model forecasts vs component model (hedonic and ANN) forecasts

In comparison with each of the *component forecast* models (*Hedonic* and *ANN*), and each of the other combining models, the combining model with weights generated by constrained WLS with a suppressed constant (Model 4) achieves the lowest MABE; followed by the combining model with weights generated by constrained WLS with a constant (Model 3). (Table 4.)

Further, the *Wilcoxon signed rank test* of the average difference between the absolute forecast errors (ABE) of the ANN model and Model 3, and between the ANN model and Model 4, are positive and statistically significant. Which indicates that both of the *constrained combining models* (with the *sum of the coefficients constrained to equal one*; with and without a constant, respectively) significantly outperform the ANN model in forecasting *housing unit* price. (Table 5; Appendix 4.)

Additionally, the *Wilcoxon signed rank test* of the average difference between the absolute forecast errors (ABE) of the two constrained combining models (Model 3 and Model 4) is positive and statistically significant. Which indicates that the combining model with weights generated by constrained WLS with the constant suppressed (Model 4) outperforms the combining model with weights generated by constrained WLS with a constant (Model 3). (Table 5; Appendix 4.)

Our study also finds that the unrestricted *combining models*, with and without a constant, respectively (Model 1 and Model 2) failed to enhance the forecast accuracy over the ANN model (Tables 4 and 5).

These empirical findings are in stark contrast with the findings of both the [10] and [11] studies, which found that *unrestricted regression models* performed best in leading to a *combining model* with superior *forecast accuracy* (over each of the individual, component model forecasts in those respective studies). Instead, the empirical findings of the present study are supportive of the results of [9], in that a *combining model* formed with weights generated by WLS with the constant suppressed, and the *sum of the coefficients* constrained to equal one, generally performed best, in leading to enhanced *forecast accuracy* over both of the individual, component model (*ANN* and *hedonic*) forecasts of housing prices in Istanbul.

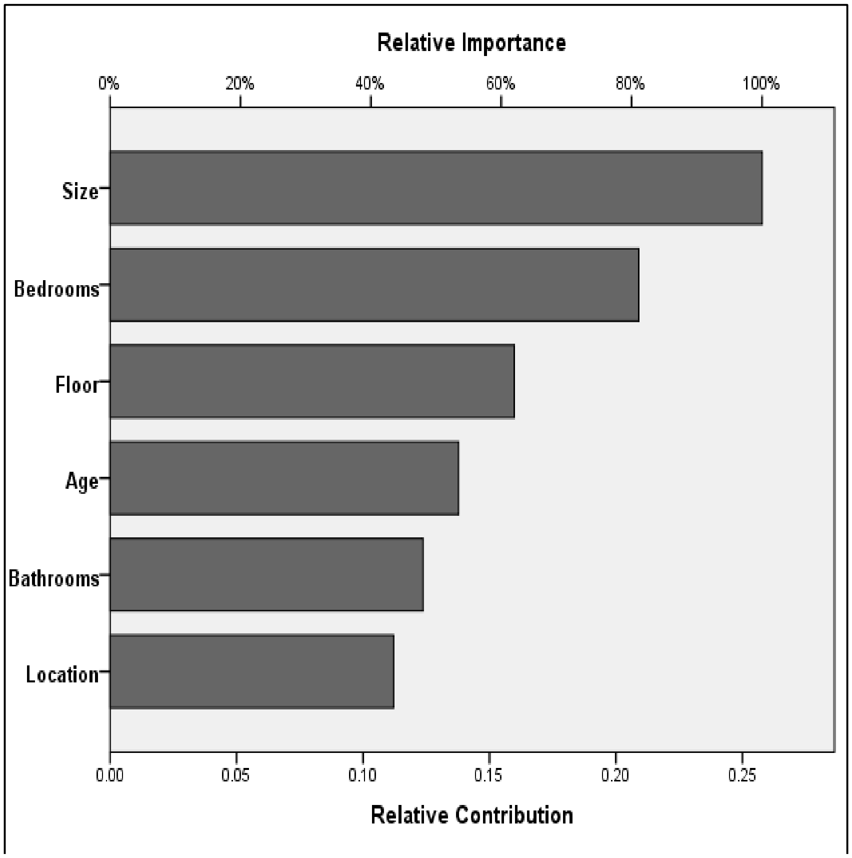



**Fig. 1. Relative contribution and importance of housing unit characteristics on price by ANN (in sample)**

**Table 1. The hedonic model *in sample* regression analysis (Weighted Least Squares) Eq.2:**

$$P_{it}(in\ sample) = \beta_0 + \beta_1 L_{it} + \beta_2 S_{it} + \beta_3 BD_{it} + \beta_4 BA_{it} + \beta_5 A_{it} + \beta_6 FL_{it} + \varepsilon_{it}$$

| Variables | Coefficient | t-value | p-value |
|---|---|---|---|
| Constant ($\beta_0$) | 5.207 | 16.684 | .000** |
| Location (L) | 0.027 | 1.167 | .247 |
| Size (S) | -0.007 | -6.953 | .000** |
| Bedrooms (BD) | 0.637 | 5.530 | .000** |
| Bathrooms (BA) | 0.421 | 6.620 | .000** |
| Age (A) | -0.034 | -1.605 | .113 |
| Floor (F) | 0.046 | 3.237 | .002* |
| n = 80 | | | |

**Estimated WLS regression coefficient significant at the 0.05 level;*
*** Estimated WLS regression coefficient significant at the 0.01 level*

**Table 2. Neural network (ANN) relative contribution factors (in sample)**

| Factors | Relative contribution | Importance |
|---|---|---|
| Location | 0.112 | 43.5% |
| Land size | 0.258 | 100% |
| Bedrooms | 0.209 | 81.1% |
| Bathrooms | 0.124 | 48% |
| Age | 0.138 | 53.4% |
| Floor | 0.160 | 62% |
| n =80 | | |

**Table 3. *In sample* WLS regression results (Eq.5):**

$$P_{it}(in\ sample) = \alpha + \beta\left(\hat{P}_{it}[hedonic]\right) + \gamma\left(\hat{P}_{it}[ANN]\right) + \varepsilon_{it}$$

| | $\boldsymbol{\alpha}$ | $\boldsymbol{\beta}$ | $\boldsymbol{\gamma}$ |
|---|---|---|---|
| ***Unrestricted WLS*** | | | |
| Estimated coefficients | -1.267 | 0.415 | 0.780 |
| standard error | 0.273 | 0.070 | 0.086 |
| t-statistic | -4.636** | 5.953** | 9.095** |
| ***Unrestricted WLS with the constant suppressed*** | | | |
| Estimated coefficients | NC | 0.459 | 1.067 |
| standard error | | 0.07 | 0.001 |
| t-statistic | | 5.910** | 1309.55** |
| ***Restricted WLS (sum-of-the-coefficients constrained to equal one)*** | | | |
| Estimated coefficients | -0.01 | 0.028 | 0.972 |
| standard error | 0.595 | 0.080 | 0.079 |
| t-statistic | -9.33* | 5.019* | 3718.515* |
| ***Restricted WLS with the constant suppressed*** | | | |
| Estimated coefficients | NC | 0.006 | 0.994 |
| standard error | | 0.05 | 0.77 |
| t-statistic | | 4.2* | 12.011** |

** Significant at 0.05 level; ** Significant at 0.000 level*

In comparison with the *hedonic* model, the *ANN* model clearly does an overall better job of processing the same data set (of housing unit characteristics and attributes), and relating it to price. Still, the *hedonic* model contains a sufficient amount of independent information such that a constrained, linear combination is superior to the ANN model in predicting housing value.

Perhaps the most insightful aspect of this empirical finding is that our study, in comparison with [9], forecasts a different target variable (*housing price* vs. *earnings growth*), and combines forecasts generated by a completely different set of component forecast models (*hedonic* and *ANN models* vs. *CAPM* and *International Brokers Estimate System Inc. [IBES] structural model*), and of course utilizes a completely different data set, as well as a different type (*cross sectional* vs. *time series*). This common finding, then, suggests a certain level of robustness to the *constrained, linear combining model* method of forecasting. The implication, as previously stated in [9] and further established in the present analysis, is that if the process of restricting the *linear combining model* leads to somewhat biased predictors, trading off incurred bias for more efficient predictors may be worthwhile to enhance the *out of sample* forecasts' accuracy. The findings of the present analysis, then, represent further (and the latest) evidence regarding the benefits of applying constraints on the *linear combining forecast model*. In the process, our study demonstrates that a *constrained linear combining model* can be a successful technique for enhancing the forecast accuracy of *housing unit* prices.

Future work may involve a larger, multiyear sample; along with inclusion of different *housing unit* types, such as *houses* and *studio dwellings*, in addition to *apartment dwellings*. Also, additional and/or different *component forecasting models* may be included in the *constrained linear*

**Table 4. Mean absolute forecast error (MABE)**

| Forecast model | MABE |
|---|---|
| Model A | 0.740 |
| Model B | 0.158 |
| Model 1 | 0.197 |
| Model 2 | 0.271 |
| Model 3 | 0.026 |
| Model 4 | 0.019 |

*Notes:*
*Model A: Hedonic forecasting model.*
*Model B: Neural network (ANN) forecasting model.*
*Model 1: Combining model with weights generated by unrestricted regression with a constant.*
*Model 2: Combining model with weights generated by unrestricted regression with a suppressed constant.*
*Model 3: Combining model with weights generated by constrained WLS with a constant.*
*Model 4: Combining model with weights generated by constrained WLS with a suppressed constant*

**Table 5. Mean difference in Absolute Forecast Error (ABE)**

| Forecast models | ABE |
|---|---|
| Model A – Model B | 0.582* |
| Model 1 – Model 3 | 0.171** |
| Model 2 – Model 4 | 0.251** |
| Model A – Model 1 | 0.543* |
| Model A – Model 2 | 0.470* |
| Model A – Model 3 | 0.714** |
| Model A – Model 4 | 0.720** |
| Model B – Model 1 | -0.039** |
| Model B – Model 2 | -0.113** |
| Model B – Model 3 | 0.132* |
| Model B – Model 4 | 0.138** |

*Notes:*
*Model A: Hedonic forecasting model*
*Model B: Neural network (ANN) forecasting model*
*Model 1: Combining model with weights generated by unrestricted regression with a constant.*
*Model 2: Combining model with weights generated by unrestricted regression with a suppressed constant.*
*Model 3: Combining model with weights generated by constrained WLS with a constant.*
*Model 4: Combining model with weights generated by constrained WLS with a suppressed constant.*
** Wilcoxon signed rank test significant at 0.05 level.*
*** Wilcoxon signed rank test significant at 0.000 level*

*-combination*, other than, or in conjunction with, the *artificial neural network* and *hedonic* models.

## 4. CONCLUSIONS

This study generates forecasts of *housing unit* prices, with employment of data from a *cross sectional*, *random sample* of apartment dwellings in Istanbul. Consistent with the findings of previous studies, our results indicate that the *neural network* (ANN) model has greater forecast accuracy than the *hedonic* model. However, our study constructs a forecasting method that we demonstrate to have superior forecast accuracy over the ANN model, by forming *weighted average combinations* of *out of sample* housing-unit price forecasts generated separately by the *hedonic* and *ANN* models. Our empirical analysis finds that the *combining model* formed with weights generated by *in-sample* weighted least squares (WLS) regression with the constant term suppressed and the *sum-of-the-coefficients* constrained to equal one, generally outperformed all other *forecasting models* tested in our study.

## ACKNOWLEDGEMENT

The authors are grateful and wish to thank the two anonymous reviewers of the manuscript for their astute observations, comments and suggestions; which led to a significant improvement.

## COMPETING INTERESTS

Authors have declared that no competing interests exist.

## REFERENCES


1. Terregrossa SJ. Combining analysts' forecasts with causal model forecasts of earnings growth. Applied Financial Economics. 1999; 9:143-153. DOI: 10.1080/096031099332401.
2. Bradley M, Gordon J, Mc Manus D. Method for combining house price forecasts. United States Patent. 2003; 660 9109.
3. Fleming M, Kuo CL. Method and apparatus for predicting and reporting a real estate value based on a weighted average of predicted values. United States Patent. 2007; 7305328.
4. Drought S, Mc Donald C. Forecasting house price inflation: A model combination approach. Reserve Bank of New Zealand. Discussion Paper Series 32; 2011.
5. Gupta R, Kabundi A, Miller S. Forecasting the US real house price index: Structural and non-structural models with and without fundamentals. Economic Modelling. 2011; 28(4):2013-2021.
6. Cabrera J, Wang T, Yang J. Linear and nonlinear predictability of international securitized real estate returns: A reality check. Journal of Real Estate Research. 2011; 33(4):565-594. DOI: 10.1080/10835547.2011.12091317
7. Granger CWJ, Ramanathan R. Improved methods of combining forecasts. Journal of Forecasting. 1984; 3: 197-204. http://dx.doi.org/10.1002/for.3980030207
8. Clemen RT. Linear-constraints and the efficiency of combined forecasts. Journal of Forecasting. 1986; 5:31-8. https://doi.org/10.1002/for.3980050104
9. Terregrossa SJ. On the efficacy of constraints on the linear combination forecast model. Applied Economics Letters. 2005; 12:19-28. DOI: 10.1080/1350485042000307062.
10. Guerard JB. Linear constraints, robust weighing and e□cient composite modeling. Journal of Forecasting. 1987; 6:193-9. DOI: doi.org/10.1002/for.3980060305
11. Lobo G. Alternative methods of combining security analysts' and statistical forecasts of annual corporate earnings. International Journal of Forecasting. 1991; 7:57-63. https://doi.org/10.1016/0169-2070(91)90033-R
12. Griliches Z. ed. Price indexes and quality changes: Studies in new methods of measurement. Cambridge, Mass: Harvard U. Press; 1971.
13. Rosen S. Hedonic prices and implicit markets. Product differentiation in Pure Competition, J. Political Econ. 1974; 82(1): 34-55. https://doi.org/10.1086/260169
14. Calhoun, C A. Property valuation methods and data in the United States. Housing. Finance International. 2001; 16:12-23.
15. Boardman AE, Greenberg DH, Vining AR, Weiner DL. Cost-benefit analysis: Concepts and practice. 2nd edition. Upper Saddle River, NJ: [Great Britain]: Prentice Hall. 2002; 349-352.
16. Limsombunchai V, Gan C, Lee M. House price prediction: Hedonic price model vs. artificial neural network. American Journal of Applied Sciences. 2004; 1(3):193-201. DOI: 10.3844/ajassp.2004.193.201
17. Lenk MM, Worzala EM, Silva A. High-tech Valuation: Should artificial neural networks bypass the human valuer? J Property Valuation & Investment. 1997; 15:8-26. DOI: 1108/14635789710163775
18. Owen C, Howard J. Estimation Realisation Price (ERP) by Neural Networks: Forecasting commercial property values. J Property Valuation & Investment. 1998; 16: 71-86.
19. Hornik K, Stinchcombe M, White H. Multi-layer feedforward networks are universal approximators. Neural Networks. 1989; 2(5):359-366. https://doi.org/10.1016/0893-6080(89)90020-8
20. Selim H. Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. Expert Systems with Applications. 2009; 36:2843-2852. DOI: 10.1016/j.eswa.2008.01.044
21. Kauko T. On current neural network applications involving spatial modelling of property prices. Journal of Housing and the Built Environment. 2003; 18(2):159-181. https://doi.org/10.1023/A:1023977111302
22. Curry B, Morgan P, Silver M. Neural networks and nonlinear statistical methods: An application to the modelling of price quality relationships. Computers & Operations Research. 2002; 29:951-969. DOI: 10.1016/S0305-0548(00)00096-4
23. Morano P, Tajani F. Bare ownership evaluation. Hedonic price model vs.

artificial neural network. Int. J. Business Intelligence and Data Mining. 2013; 8-4. DOI: 10.1504/IJBIDM.2013.059263

24. Stanley M, Alastair A, Dylan M, Patterson D. Neural networks: The prediction of residential values. J Property Valuation & Investment. 1998; 16: 57-70.
25. Bischoff C W. The combination of macroeconomic forecasts. Journal of Forecasting. 1989; 8: 293-314. https://doi.org/10.1002/for.3980080312
26. Wu J, Zhou J, Chen L, Ye L. Coupling forecast methods of multiple rainfall–runoff models for improving the precision of hydrological forecasting. Water Resources Management. 2015; 29: 5091-5108. DOI: 10.1007/s11269-015-1106-8
27. Liu X, Moreno B and García S. A grey neural network and input-output combined forecasting model. Primary energy consumption forecasts in Spanish economic sectors. Energy. 2016; 115(1):1042-1054. DOI: 10.1016/j.energy.2016.09.017
28. Heng J, Wang J, Xiao L and Lu H. Research and application of a combined model based on frequent pattern growth algorithm and multi-objective optimization for solar radiation forecasting. Applied Energy. 2017; 208:845-866. DOI: 10.1016/j.apenergy.2017.09.063
29. Jun W, Yuyan L, Lingyu T, Peng G. Modelling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting. Chaos, Solitons & Fractals. Elsevier. 2018; 108:136-147. DOI: 10.1016/j.chaos.2018.01.028
30. Wang R, Wang J, Xu Y. A novel combined model based on hybrid optimization algorithm for electrical load forecasting. Applied Soft Computing. 2019; 82:105 548.
31. Liu Y, Zhang S, Chen X, Wang J. Artificial combined model based on hybrid nonlinear neural network models and statistics linear models-research and application for wind speed forecasting. Sustainability MDPI, Open Access Journal. 2018; 10(12):1-30. https://doi.org/10.3390/su10124601
32. Niu X, Wang J. A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. Applied Energy. Elsevier. 2019; 241:519–39. DOI: 10.1016/j.apenergy.2019.03.097
33. Liu Z, Jiang P, Zhang L, Niu X. A combined forecasting model for time series: Application to short-term wind speed forecasting. Applied Energy Elsevier. 2020; 259:1-25. DOI: 10.1016/j.apenergy.2019.114137
34. Istanbul real estate. Available: www.istanbulrealestate.com. Accessed: 15th, 16th and 17th May 2019.
35. Istanbul Property World. Available:www.istanbulpropertyworld.com. Accessed: 19th to 22nd May 2019].
36. Property Turkey. Available:www.propertyturkey.com. Accessed: 29th May to 10th June 2019.
37. Turkey Expert. Available:www.turkeyexpert.com. Accessed: 20th, 21st&30th June 2019.
38. Istanbul Homes. Available:www.istanbulhomes.com. Accessed from 3rd to 6th June 2019.
39. Turkey Homes. Available: www.turkeyhomes.com. Accessed: 1st to 9th July 2019.
40. Halvorsen R, Palmquist R. The interpretation of dummy variables in semi-logarithmic equations. American Economic Review. 1980; 70(3):474-475.
41. Goodman A, Thibodeau T. Dwelling age heteroscedasticity in hedonic house price equations. J Housing Res. 1995; 6:25-42.
42. Goodman A, Thibodeau T. Dwelling age heteroscedasticity in hedonic house price equations: An extension. J Housing Res. 1997; 8:299-317.
43. Stevenson S. New empirical evidence on heteroscedasticity in hedonic housing models. Journal of Housing Economics. 2004; 13:136-153. DOI: 10.1016/j.jhe.2004.04.004
44. Fletcher M, Gallimore P, Mangan J. Heteroscedasticity in hedonic house price models. J Property Res. 2000a; 17:93-108. DOI: 10.1080/095999100367930

## Appendix 1. Hedonic regression (*in sample*): *Equation (2)*

$$P_{it}(in\ sample) = \beta_0 + \beta_1 L_{it} + \beta_2 S_{it} + \beta_3 BD_{it} + \beta_4 BA_{it} + \beta_5 A_{it} + \beta_6 FL_{it} + \varepsilon_{it}$$

**Model Summary**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .946[a] | .895 | .886 | 1.17508526 |

*a. Predictors: (Constant), Floor, Location, Age, Size, Bathrooms, Bedrooms*

**ANOVA[a,b]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 857.882 | 6 | 142.980 | 103.547 | .000[c] |
| | Residual | 100.800 | 73 | 1.381 | | |
| | Total | 958.683 | 79 | | | |

*a. Dependent Variable: log price*
*b. Weighted Least Squares Regression - Weighted by weight*
*c. Predictors: (Constant), Floor, Location, Age, Size, Bathrooms, Bedrooms*

**Coefficients[a,b]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 5.207 | .312 | | 16.684 | .000 |
| | Location | .027 | .023 | .049 | 1.167 | .247 |
| | Size | -.007 | .001 | -.828 | -6.953 | .000 |
| | Bedrooms | .637 | .115 | .777 | 5.530 | .000 |
| | Bathrooms | .421 | .064 | .603 | 6.620 | .000 |
| | Age | -.034 | .021 | -.076 | -1.605 | .113 |
| | Floor | .046 | .014 | .163 | 3.237 | .002 |

*a. Dependent Variable: log price*
*b. Weighted Least Squares Regression - Weighted by weight*

## Appendix 2. ANN estimation

### A. ANN (*in sample*)

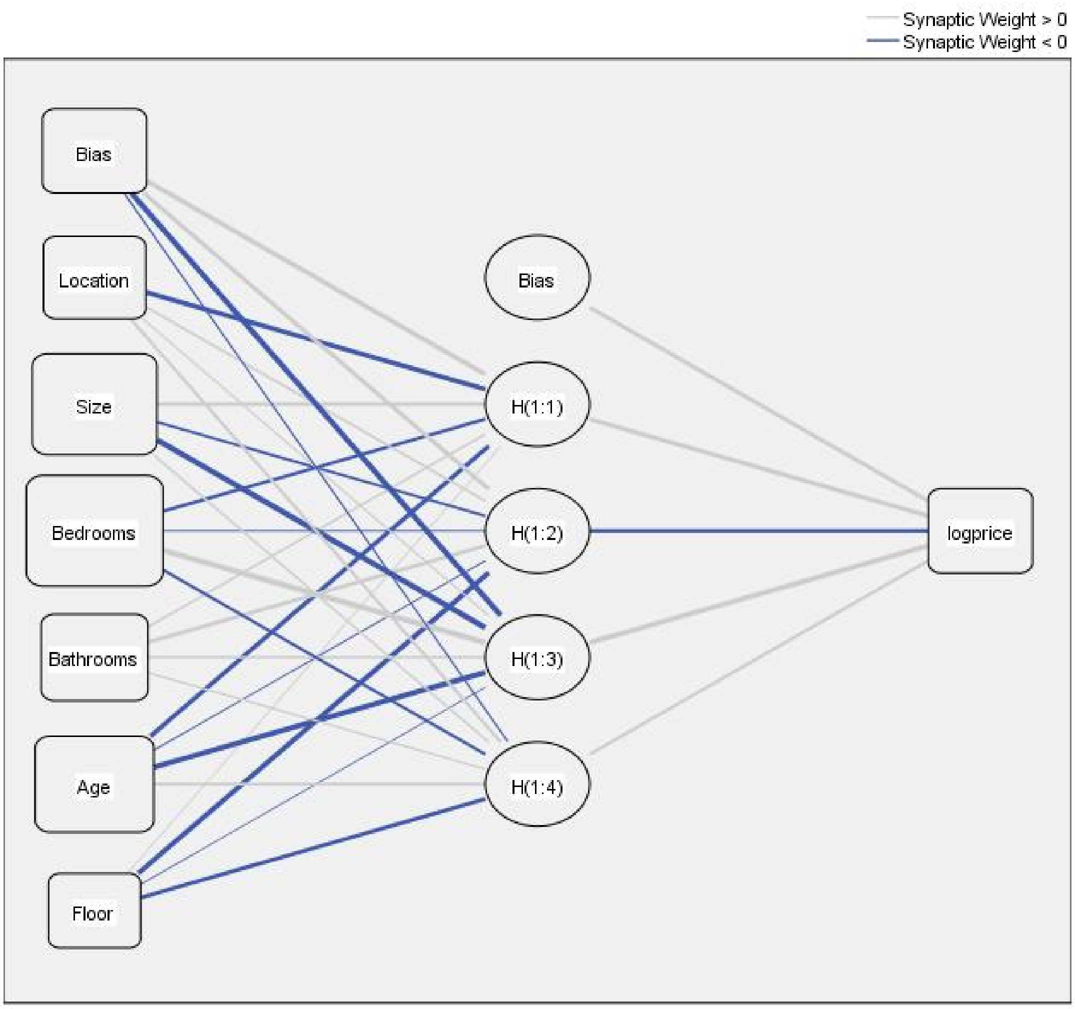
**Network Information**

| | | | |
|---|---|---|---|
| Input Layer | Factors | 1 | Location |
| | | 2 | Size |
| | | 3 | Bedrooms |
| | | 4 | Bathrooms |
| | | 5 | Age |
| | | 6 | Floor |
| | Number of Units[a] | | 104 |
| Hidden Layer(s) | Number of Hidden Layers | | 1 |
| | Number of Units in Hidden Layer 1[a] | | 8 |
| | Activation Function | | Hyperbolic tangent |
| Output Layer | Dependent Variables | 1 | Log price |
| | Number of Units | | 1 |
| | Rescaling Method for Scale Dependents | | Standardized |
| | Activation Function | | Identity |
| | Error Function | | Sum of Squares |

*a. Excluding the bias unit*

**Independent Variable Importance**

| | Importance | Normalized Importance |
|---|---|---|
| Location | .112 | 43.5% |
| Size | .258 | 100.0% |
| Bedrooms | .209 | 81.1% |
| Bathrooms | .124 | 48.0% |
| Age | .138 | 53.4% |
| Floor | .160 | 62.0% |




Hidden layer activation function: Hyperbolic tangent

Output layer activation function: Identity

**B. ANN *(out of sample*)**

| **Network Information** | | | |
|---|---|---|---|
| Input Layer | Factors | 1 | location |
| | | 2 | size |
| | | 3 | bedroom |
| | | 4 | bathroom |
| | | 5 | age |
| | | 6 | floor |
| | Number of Units[a] | | 52 |
| Hidden Layer(s) | Number of Hidden Layers | | 1 |
| | Number of Units in Hidden Layer 1[a] | | 8 |
| | Activation Function | | Hyperbolic tangent |
| Output Layer | Dependent Variables | 1 | Log price |
| | Number of Units | | 1 |
| | Rescaling Method for Scale Dependents | | Standardized |
| | Activation Function | | Identity |
| | Error Function | | Sum of Squares |

*a. Excluding the bias unit*

| | **Independent Variable Importance** | |
|---|---|---|
| | **Importance** | **Normalized Importance** |
| location | .198 | 72.2% |
| size | .274 | 100.0% |
| bedroom | .092 | 33.8% |
| bathroom | .014 | 5.3% |
| age | .223 | 81.6% |
| floor | .199 | 72.7% |

**Appendix 3. Combination weights estimation (WLS regression): *Equation (5):***

$$P_{it}(in\ sample) = \alpha + \ \beta\big(\hat{P}_{it}[hedonic]\big) + \ \gamma\big(\hat{P}_{it}[ANN]\big) + \varepsilon_{it}$$

**A. Unrestricted regression with constant term**

| **Coefficients[a,b]** | | | | | | |
|---|---|---|---|---|---|---|
| **Model** | | **Unstandardized Coefficients** | | **Standardized Coefficients** | **t** | **Sig.** |
| | | **B** | **Std. Error** | **Beta** | | |
| 1 | (Constant) | -1.267- | .273 | | -4.636- | .000 |
| | Hedonic | .4-+][=15 | .070 | .393 | 5.953 | .000 |
| | ANN | .780 | .086 | .600 | 9.095 | .000 |

*a. Dependent Variable: log price*
*b. Weighted Least Squares Regression - Weighted by weight*

**B. Unrestricted regression with constant term supressed**

| **Coefficients[a,b,c]** | | | | | | |
|---|---|---|---|---|---|---|
| **Model** | | **Unstandardized Coefficients** | | **Standardized Coefficients** | **t** | **Sig.** |
| | | **B** | **Std. Error** | **Beta** | | |
| 1 | ANN | 1.067 | .001 | 1.000 | 1309.547 | .000 |

*a. Dependent Variable: log price*
*b. Linear Regression through the Origin*
*c. Weighted Least Squares Regression - Weighted by weight*

**Excluded Variables[a,b,c]**

| Model | | Beta In | t | Sig. | Partial Correlation | Collinearity Statistics |
|---|---|---|---|---|---|---|
| | | | | | | Tolerance |
| 1 | Hedonic | .459[d] | 5.910 | .000 | .556 | 6.760E-5 |

*a. Dependent Variable: log price*
*b. Linear Regression through the Origin*
*c. Weighted Least Squares Regression - Weighted by weight*
*d. Predictors in the Model: ANN*

**C. Restricted regression (sum-of-the-coefficients constrained to = 1) with constant term**

**Parameter Estimates**

| Parameter | Estimate | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| a | .028 | .108 | -.187- | .244 |
| b | .972 | .081 | .810 | 1.134 |
| c | -.001- | .595 | -1.187- | 1.185 |

**D. Restricted regression (sum-of-the-coefficients constrained to = 1) with constant term suppressed**

**Parameter Estimates**

| Parameter | Estimate | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| a | .006 | .050 | -.132- | .188 |
| b | .994 | .077 | .813 | 1.131 |

**Appendix 4. Wilcoxon signed rank test**

**Test Statistics[a]**

| | MAE_ANN - MAE_Hedonic |
|---|---|
| Z | -3.173-[b] |
| Asymp. Sig. (2-tailed) | .002 |

*a. Wilcoxon Signed Rank Test*
*b. Based on positive ranks*

**Test Statistics[a]**

| | MAE3 - MAE1 | MAE4 - MAE2 |
|---|---|---|
| Z | -3.920-[b] | -3.920-[c] |
| Asymp. Sig. (2-tailed) | .000 | .000 |

*a. Wilcoxon Signed Rank Test*
*b. Based on positive ranks*
*c. Based on negative ranks*

**Test Statistics[a]**

| | MAE1 - MAE_ANN | MAE2 - MAE_ANN | MAE3 - MAE_ANN | MAE4 - MAE_ANN |
|---|---|---|---|---|
| Z | -3.920-[b] | -3.920-[b] | -3.323-[c] | -3.733-[c] |
| Asymp. Sig. (2-tailed) | .000 | .000 | .001 | .000 |

*a. Wilcoxon Signed Ranks Test*
*b. Based on negative ranks*
*c. Based on positive ranks*

**Test Statistics[a]**

| | MAE1 - MAE_Hedonic | MAE2 - MAE_Hedonic | MAE3 - MAE_Hedonic | MAE4 - MAE_Hedonic |
|---|---|---|---|---|
| Z | -3.360 [c] - | -3.920- [b] | -3.845- | -3.845-[c] |
| Asymp. Sig. (2-tailed) | .001 | .000 | .000 | .000 |

*a. Wilcoxon Signed Ranks Test*
*b. Based on negative ranks*
*c. Based on positive ranks*



*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*http://www.sdiarticle4.com/review-history/65400*