

*Precios implícitos en  
valoración inmobiliaria  
urbana*

# Implicit Prices in Urban Real Estate Valuation

## Autores

NÚÑEZ TABALES, J. University of Córdoba, Faculty of Economics  
Puerta Nueva s/n, 14071, Córdoba, Spain  
es2nutaj@uco.es

REY CARMONA, F. University of Córdoba, Faculty of Economics  
Puerta Nueva s/n, 14071, Córdoba, Spain

CARIDAD Y OCERIN, J. M<sup>a</sup>. University of Córdoba, Faculty of Economics  
Puerta Nueva s/n, 14071, Córdoba, Spain

**Fecha de recepción** 19/6/2013

**Fecha de aceptación** 1/8/2013

Código interno RDLC 0146

## Abstract

Econometric hedonic models encounter several theoretical and practical difficulties when applied to the real estate market, such as downward biases in the estimation of hedonic prices, subjective decisions in the measurement process of categorical attributes, frontier problems related to an imperfect information framework and unequational specification. Many of these are linked to the parametric approach.

Artificial Neural Networks (ANN) provide an attractive alternative: better

dwelling prices estimates, avoidance of bias at different market segments, direct use of categorical data and full use of the information available. The price to be paid is the difficulties in the economic interpretation of network parameters. Nowadays, if the final objective to produce better estimates of the transaction prices, this methodology show lower errors, provided of a broad representative database of sales are recorded. A case study is presented for a medium size city in the South of Spain.

**Keywords:** Urban Economics; Hedonic Models; Neural Networks.

## Resumen

*La aplicación en econometría de modelos hedónicos al mercado inmobiliario tiene diversas dificultades teóricas y prácticas, tales como subestimación de los precios hedónicos, decisiones subjetivas en el proceso de medición de características, problemas relacionados con un marco de información imperfecta o con la especificación de la ecuación. Muchos de estos están relacionados con el enfoque paramétrico. Las Redes Neuronales Artificiales (RNA) proporcionan una alternativa atractiva: mejores estimaciones del precio del bien inmueble, evita el sesgo en diferentes*

*segmentos del mercado, permite utilización de datos de características sin la construcción de índices y el uso completo de la información disponible. El precio a pagar es la dificultad en la interpretación económica de los parámetros estimados. Si el objetivo final es obtener una mejor estimación de los precios de transacción, bajo esta última metodología se registran menores errores, siempre que se disponga de una amplia base de datos representativos. Se presenta un estudio de caso para una ciudad de tamaño medio en el sur de España.*

**Palabras clave:** Economía urbana; modelos hedónicos; redes neuronales.

## 1. Introduction

During last decade, Spain presented a seller's real estate market. There is a strong tradition of home ownership in the country, coupled to fiscal stimuli and low interest rates associated to the euro zone. Prices have continuously gone up, and a growing part of the family's income was devoted to attend mortgages and housing related costs. This trend has come abruptly to an end in 2008, causing turmoil in the country economy. In this context, an objective way to assess the prices of properties coming to the market is needed. Its demand comes from different actors: buyers and sellers of family flats and second homes, agents in the real estate sector, investors, value agency and financial institutions, and fiscal authorities at national and municipal levels. Hedonic and Neural Networks (NN) models could fill the gap, although they must be updated and maintained, with the additional benefit of their usefulness in detecting and measuring price changes.

Real estate prices were studied since the fifties, and two decades later, Rosen (1974) proposed the use of regression models, called 'hedonic models', and applied them in urban areas. In Spain there are macroeconomic and sectorial statistics that try to explain investments in building and home buying.

In this work two methodologies are in use: Hedonic Methodology and NN for determining empirically the valuation of housing in a city of average size as Cordova (Spain), using a random sample of 2888 transactions of flats and apartments in a year period. A comparison is realized between the power of prediction of both technologies later to effect a calculation and graphical comparative analysis of the implicit prices corresponding to the determinant attributes of the value of market of a real estate.

## 2. Hedonic Methodology

The first papers about hedonic models were introduced by economists in order to study the critical evaluation of the price indices, related to their possible bias in estimating the true evolution of the real value levels. In the 60s. The hedonic methodology, first used in Agricultural Economics (Wallace, 1926 and Waugh, 1929) and later in the car market (Court, 1939), was toughly developed in the 60s (Sirmans, 2005).

The estimation of housing prices has been one of the most analysed topics in urban economics (Freeman, 1979 and Richardson, 1978). Some authors, as the New Urban Economic (N.U.E.) followers (Solow, 1972), have developed their work through monocentric or distance

models. Others (Tiebout, 1956 and Rosen, 1974) have tried less restrictive models, with several attributes as real estate price determinants. Since the 1970's, under the particular influence of the Rosen methodology, these works have evolved into more complex models. It is possible to include a greater number of variables under a less restrictive framework (Azqueta, 1994). This improvement has been positively influenced by the hedonic price methodology (Lancaster, 1966; Muth, 1969; Griliches, 1971; and others). In the case of the Spanish real estate market, several works have been published since Caridad and Brañas (1996), Bilbao-Terol (2000), Bover and Velilla (2001), Aguiló-Segura (2002), Bengoechea Morancho (2003) and others.

In housing price estimation it is very common to take into account some external factors (distance to the central district, schools, parks, etc.) as important determinants of the price. Even more, is not unusual, to consider this kind of factors as the main component of the price, where internal flat characteristics are left out.

The hedonic methodology aims at estimating the value of a complex property related to its characteristics. It is assumed that the consumer has to purchase one good that is composed by a basket of items that generate its final price, but, that cannot be separated into independent parts, so it has to be bought as a whole. Some of these attributes will tend to increase the final price, and thus will be called positive attributes, while some others should be a burden, devaluating the good. The hedonic prices are the implicit values per unit of each attribute; these values are evaluated by the consumers when a property is sold in a market that is supposed to be in equilibrium.

Rosen (1974) developed the general theory for offer and demand functions of multi-attribute complex goods. The value of a property is perceived in different levels by several potential consumers, so the hedonic equation is an envelope of different functions. The same can be established for the offer functions of several sellers.

Frontier-function models are also related to the real estate market, where there are permanently different supply and demand functions. In this context, the equilibrium is determined by only part of the agents operating in an efficient-non efficient environment (Greene, 1993).

The utility function of a house or dwelling depends of its internal attributes,  $x_1, x_2, \dots, x_k$ , and also of some external characteristics,  $z_1, z_2, \dots, z_m$ . In the marketplace each consumer has a different utility function in which the income,  $I$ , the wealth,  $W$ , and the individual tastes,

$T$ , influence its shape. Thus the utility function for an individual person is  $U = U(I, W, T, \mathbf{x}, \mathbf{z})$ ; this will identify its offer and demand functions for a property, and the corresponding envelopes for all the consumer is the base for the market offer and demand functions.

As each utility function is concave, the bid function for a dwelling's basket of attributes will increase with the positive ones, but its marginal variation will be less than proportional, thus introducing a downward bias in the interpretation of hedonic prices. The partial derivative of the demand function related to each attribute measures its relative importance of the consumer evaluation.

The offer function will be dependent of the set  $(\mathbf{x}, \mathbf{z})$  of internal and external attributes, but also of the market conditions, like the number of properties been sold at one particular moment. This offer function is convex, so an increase of any of the positive attributes will enhance the offer price. The relative importance of the  $(\mathbf{x}, \mathbf{z})$  attributes, from the offer point of view, is measured by the partial derivatives of the offer function.

In the marketplace, the offer and demand functions define the transaction price of a property, although, the relative values of the attribute, in both the demand and offer sides, should not be identical. This fact can be a global objection to the hedonic methodology when it is not based on multi-equation models, as there is an identification problem of the hedonic prices. Nevertheless, one can assume that the main interests of the seller and buyer, when an agreement is established, are fulfilled by transaction price, so it is useful to estimate this transaction price as accurate as possible, related to the  $(\mathbf{x}, \mathbf{z})$  attributes of each dwelling, so both markets agents, seller and buyer, will be informed how far or near are their individual demand and offer functions from the market reality, that is, from the envelope functions of all the markets agents. The knowledge of this price function,  $P = P(\mathbf{x}, \mathbf{z}) + e$ , is useful for buyers and sellers, and its derivatives will be assumed related to the hedonic prices of each attribute.

### 3. Attributes and their prior treatment

To proceed with the estimation of  $P(\mathbf{x}, \mathbf{z})$ , it is necessary to decide the attributes  $\mathbf{x}$  and  $\mathbf{z}$  to be used. Some of these are numerical values that can be measured as such, while some others are categorical data, which cause some difficulties in dealing with them in a model.

It is quite a forward procedure to list the main internal attributes,  $\mathbf{x}$ . For example, surface related variables are the size, the number of rooms, bathrooms and embedded wardrobes, and surface of different parts. Some internal attributes measured in a categorical scale, can be transformed into numerical indices by some aggregation and valuation of scale procedures; a facilities index is formed with the data about the presence or absence of general state of the dwelling's facilities with reference to the construction date; it is obtained adding the variables related to water reform, electricity and shutting, weighted by the housing's age; a conservation index related to the general state of the dwelling at first sight, is the sum of the variables related to the aspect of the kitchen, the bathroom and the pavement; the improvement index shows the complements and add-ons of the property, like if there is lumber, washing place, pre-installation of conditioned air and pantry, and is composed by the sum of the binary variables associated with this equipment; the comfort index, where the very wished characteristics of a potential buyer are included, like exterior housing, air-conditioned and direct access to the garage, is elaborated, like previous indexes, by summing the correspondent binary attributes. Another index includes the characteristics that seems to be demanded for new-build apartments: a good and ostentatious entrance, swimming pool, green zones and satellite television; a luxury index is related to the equipment available in each property, such Jacuzzi, sauna, and similar. The amount paid as community charges has also to be taken into account.

Related to the external attributes,  $\mathbf{z}$ , an income index reflects the existence of different central business districts, i.e. distant zones that have high prices; the location index measures the position of the house inside the zone, is also important for the prospective customers; a parking index is a measure of the capacity of a zone to absorb cars without need of a private indoor parking place, and it could be evaluated using spatial sampling; a general quality index is related to the kind of neighbourhood level; the view index is constructed to evaluate the position of the property in the city, and its views of the surroundings areas, like a mountain, or the sea.

Considering the amount of variables to be introduced in a model, it is no surprise to forecast the usual problems of multicollinearity, increased in this case by the presence of many binary exogenous variables. The index building process can be a way to reduce this inconvenience, but at a price: losing information about individual attributes. The principal component method to reduce the dimension can also be used, but, as well, with the corresponding loss of information. As it can

be expected, some interactions do occur, increasing these classical econometric problems (Saura, 1995).

As a final solution, the hedonic methodology produces some alternative models with just a small proportion of the original variables that do influence the transaction prices. As the exogenous variables are almost always related, the economic interpretation of the hedonic prices is subjective, and, different models, with common variables, will produce different hedonic prices, in the same market.

#### 4. Alternative methods for hedonic modelling: the neural networks (NN) approach

It has been show that the hedonic methodology has some inherent fallout. Some related to the identification of offer and demand functions, that has been rounded, proposing a *price function*,  $P(\mathbf{x}, \mathbf{z})$ , that should provide a *market hedonic prices*, avoiding to identify these with the price perceptions of buyers and sellers. Others problems are purely mathematical: many related variables in a model produce ill-conditioned normal equations, and, thus, non-stability of model coefficients. The later can be partially solved with a price: the loss of some information, not using part of the exogenous variables that conforms the price, or, reducing the dimensionality by an aggregation process that generates several summary indices, or by a statistical procedure, like principal component analysis.

There is an alternative in this model building process: that is to use another statistical estimation procedure based on NN methodology (Martín del Brío y Sanz de Molina, 1997). Of course there is a different kind of price to be paid: the abandon of the hedonic price estimation, which, in any case, has many caveats, and thus is of relative usefulness. The NN approach will produce better results in estimating the price function,  $P(\mathbf{x}, \mathbf{z})$ , reducing the residuals in the price model, and thus, been finally more practical for buyers and sellers, that want to compare their desires with the reality of the market.

NN is a model build by tuning a set of parameters, the weights, using several simple functions or units, linked together by the weights, to predict some output variables from the inputs variables of the system. In the Multilayer Perceptron, MLP, the input units take the input values into the network, and the output units report the final answers. In between there are the hidden units that process the data and link the input

with the output units, extracting useful information from the original data, and use it to predict the output data. If there is no hidden layer, the relation between inputs and outputs is linear, while the presence of one or several hidden layer, results in non-linear relations to predict the outputs from the input (Rosemblatt, 1958). In our case, there is only one output, the price,  $P$ , of a property, and several inputs,  $(\mathbf{x}, \mathbf{z})$ , its internal and external attributes. Every unit receive information from the previous layer; if  $r$  units are linked to a hidden unit, this one process the data received,  $a_1, a_2, \dots, a_r$ , and produces an output,  $o = f(\sum w_r a_r)$ , defined by an activation function, using the weights associated with the links between these units. The same procedure applies to obtain the final price,  $P$ , from the hidden layer units. Several types of activation functions are commonly used, as the linear, the logistic, or the hyperbolic tangent, depending of the type of data.

The MLP network is trained using a set of input data, and its corresponding outputs, by the method of back propagation of error. The weights are changed so that the error of the same input-output data will be smaller. There are several problems to be solved in any NN use; some are related to the data preparation process, like the scaling and encoding, or the outlier detection, or the possibility of dimensionality reduction, while others topics involve decisions about the specification.

The hidden layer size has to be sufficient for accurate reproduction of the output, but should be kept small to allow generalisations, that is, it should be able to forecast a dwelling price when confronted with its attributes  $(\mathbf{x}, \mathbf{z})$ . As it is possible to reduce the dimensionality of the internal and external attributes, the hidden layer will have considerably fewer units that the number of input attributes. So in this case there is no need to use the Kolmogorov theorem to put an upper limit in the number of hidden units. In the real estate market it is not straightforward to define a set of patters, to define a training strategy, but it is possible to have quite large sample size of data, so there are no practical problems in the estimation process, and the number of hidden units can be quite low.

In Spain, several contributions have appeared in the last decade: Ceular and Caridad (2001), Mohamed (2002), Fuentes (2004), García Rubio (2004), Gallego (2005), Lara (2005) and others in different regions, as in Liu, Zhang and Wu (2006), Khalafallah (2008), Selim (2009) and Hamzaoui and Hernández (2011).

The hedonic methodology aims at estimating the hedonic prices of attributes in a complex good. If the relations between the price,  $P$ , and one of the exogenous variables from the attribute set  $(\mathbf{x}, \mathbf{z})$ , is

non-linear, it is more difficult to provide an economic interpretation of the estimated parameters. So, one has to focus on a small range of values of a precise attribute, and analyse the price variation originated in this range. In a NN, the same problem arouses, and it is possible to spot non-linearities and linear relationships. For example, the former can be checked calculating the correlation coefficients between each input variable and the sensitivity of each output unit. Non zero correlations are caused by non-linear relations. There are several methods to decide if an input variable has a linear effect on the output price. At each small range of prices, it is possible to calculate the derivatives of the network function. If the variance of these derivatives with respect to each input unit is low, the effect can be accepted to be linear.

The relevance of each variable in an hedonic model is checked by classical econometric methods, such as T tests, information criteria, and other statistical procedures. In a NN it is also possible to check if an input variable is relevant to the estimation of the final price of a property. It is possible to estimate, by back propagation, the gradient function over a range of input values and find how a variable influence the final price. The rate of change in the hedonic prices, or their elasticity, could be estimated using the gradient of each input. These functions could also be used as a marketing tool to elaborate strategies for increasing the value of a property.

### 5. A case study: Hedonic Modelling versus MLP Neural Network in a medium size city

To illustrate the possibilities of NN in property price estimation, a case study has been developed for a medium size city in the South of Spain, Córdoba, using a random sample of 2888 transactions of flats and apartments in the real estate market, in a year period. The data were obtained by the real estate brokerage companies operating in the city, as most of the sales are done using these firms.

**Table 1.** Internal data of the property and external variables

Internal data of the property		External variables	
<b>Basic</b>	Area	<b>General</b>	Building year
	Bedrooms		Lift(*)
	Bath		Laundry(*)
	Complimentary baths		
	Terrace (*)		
	Communications(*)		
	Wardrobes(*)		
	Garage(*)		
	Storage room(*)		
	Climate control		
<b>General Quality</b>	Floors	<b>Extras</b>	Pool (*)
	Windows type		Tennis Court(*)
	Interior wooden		Garden(*)
	Furniture(*)		
	Kitchen furnit (*)		
<b>Reforms</b>	Reformed(*)		
<b>Orientation</b>	Orientation(*)		
<b>Economics</b>	Community expenses	<b>Location</b>	Neighborhood
	Sales price		

(\*) Theses variables are binary

The sample data cover well the urban area, and describes well the local market. It has to be stated that the housing structure in the city is similar to others Spanish towns: most of the residences are apartments, and there is only a very small proportion of individual houses. In the sample are included second sales, avoiding the analysis of new build properties.

The variables included are presented in Table 1. These represent internal characteristics and external and location data. Some of them are of binary type, as the availability of climatisation, the quality of the pavement and of the carpentry. The urban area is classified in 33 zones, according of the income level.

To build a hedonic model, it is necessary to define several indices that can include categorical information available for each dwelling (Table 2). Of course, in theory, it is possible to associate binary variables to qualitative data, and include these binary variables



directly in a model. But the statistical problems like multicollinearity make impossible to proceed in this way. To tackle these difficulties, several indices have been built, as a way to obtain quantitative information from the original attributes. In the process, it has to be kept in mind that the interpretation of these new variables has to be straightforward, and that it has to be avoided to use complicated weights that introduce subjective beliefs in the indices. The indices have to be compatible with the consumer's perceptions of the characteristics of each property.

The indices defined have been scaled to fit in the interval [0; 1]. Values near the upper end refer to optimum values in relation to higher prices, and low values of an index reflect lack of importance of this variable in the final price. These can be treated as numerical data in the modelling processes, as in Saura (1995) and Jaén and Molina (1995).

**Table 2.** Indexes associated to each property

Index	Variables included
Quality Index	<ul style="list-style-type: none"> <li>Age, accesses, balconies and terrace</li> </ul>
External Building index	<ul style="list-style-type: none"> <li>Pavement, carpentry, kitchen furniture and reforms</li> </ul>
Internal building index	<ul style="list-style-type: none"> <li>Pool, tennis court and garden</li> </ul>
Orientation index	<ul style="list-style-type: none"> <li>Orientation of building and balconies</li> </ul>
Annex index	<ul style="list-style-type: none"> <li>Garage and store room</li> </ul>
Location index	<ul style="list-style-type: none"> <li>Location in the city and residents income</li> </ul>

The conservation index reflects the general state of the property, at first sight. It is constructed from several categorical data, like the maintenance condition of the kitchen, baths, floor, plumbing, and electric wiring, if these have been updated. The complementary size index includes the information about some additional facilities, like a washing room, embedded wardrobes, and food storage place. The situation index reflects the position of the apartment or flat in the building, the availability of a lift, and the external views. The building index includes information about the quality of the building, its age, the entrance, and additional facilities like swimming pool or satellite and cable

television. The location index includes data about the situation of the property in the city, the surroundings and the income level of the zone.

Price of transaction is the endogenous variable in both hedonic and NN models. The decision about the functional form is treated in the literature. Here is used a quadratic model, preferring it over the semi-logarithmic usual transformation. The explanatory variables included show a high level of collinearity, introducing practical difficulties in the estimation process. Using simplicity criteria, parsimonious models are preferred over more complicated specifications. Some interactions between exogenous variables are considered, using quadratic forms as shown in formulae 1.

$$Price = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \beta_{11} x_1^2 + \dots + \beta_{kk} x_k^2 + \beta_{12} x_1 x_2 + \dots + \varepsilon \quad [1]$$

although the final model used is linear but for an interaction term. The variables finally selected to be included in the model are the following: size of the flat (in square metres, including common areas), age of the building, index of location (based on its position, characteristics of the area and income of the zone), index of additional facilities (as garages or store room), common expenses (in euros), and the interaction between qualities variables regarding the pavement condition and the external carpentry, as show in formulae 2.

$$Price = 193.679 + 1109.951 Surface - 1067.449 Age + 64297.29 LocationIndex + 18458.66 AnnexIndex + 1296.708 ComonExpenses + 5117.270 Pavement \times Carpentry + e \quad [2]$$

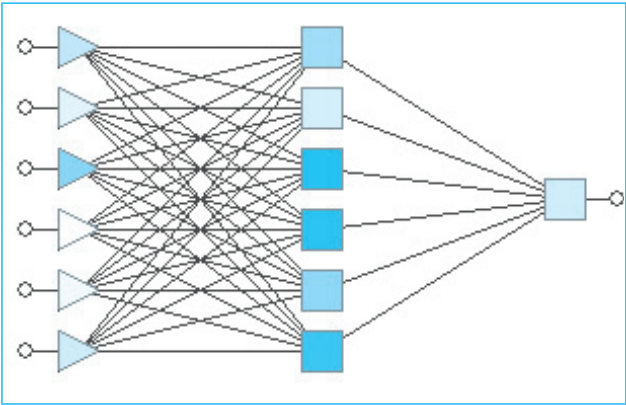
**Table 3.** Hedonic model

Endogenous variable: Sales price				
Variable	Coeficients	Standard error	t-Student	Prob.
Intercept	193.6790	10237.53	0.018919	0.9849
Surface	1109.951	83.62050	13.27368	0.0000
Age	-1067.449	173.5725	-6.149872	0.0000
Location	64297.29	5494.403	11.70232	0.0000
Annexes	18458.66	4572.246	4.037111	0.0001
Expenses	1296.708	105.9213	12.24218	0.0000
Pav xCarp	5117.270	504.0529	10.15225	0.0000

The usual diagnostics tests have been successfully applied. The stability of the model is supported by the Chow test, and the condition index is  $k = 16.6$ , supporting that the multicollinearity is not worrisome. The goodness of fit is similar to other hedonic modelling,  $R^2 = 0.7738$  and the U Theil index is 0.09 (Table 3).

NN provide a non linear alternative to estimate price models. The usual MLP is applied, as in several previous works, ad in Haykin (1999), Freeman and Skapura (1993) and Garcia Rubio (2004). To facilitate comparisons with the proposed hedonic model, the exogenous variables selected have been used as inputs to the different NN employed. Different topologies have been tried, with different number of hidden layers and neurons, as several activation functions and learning algorithms. Finally a 6:6-6-1:1 structured is selected (Figure 1), and the Trajan Software for NN is employed, with a lineal activation function, in the input, while in the other layers a sigmoid function is applied. The training set is composed of 80% of the sample, and the remaining data is used as test to evaluate the forecasting power. The back propagation method employed was limited to 5000 iterations, with a learning ratio of 0.1.

Figure 1. ANN 6:6-6-1:1Topology



The NN increase the determination coefficient value to  $R^2 = 0.8605$ , higher than in the hedonic model, although one has to take into account the larger number of parameters in the network. The root mean square error is lower in the network (39540.36) than in the hedonic model (41645.43), and similar results are observed with other measures, at can be seen in the Table 4.

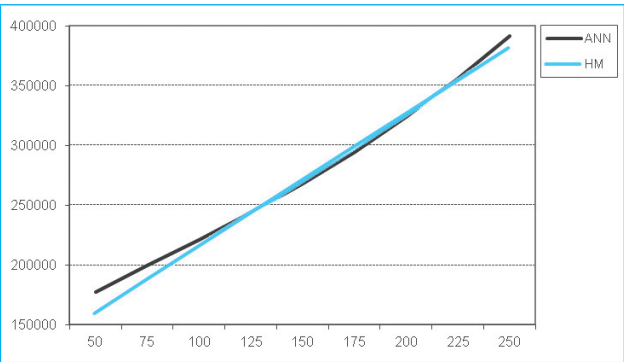
Table 4. Comparison of hedonic a ANN modelling

	MH	NN
Determination coefficient ( $R^2$ )	77,38%	86,05%
Root mean square error ( $RMSE$ )	41.645,43	39.540,36
Residual standard error	41.911,91	39.102,13
Mean absolute error	30.579,18	28.551,34
Mean relative error	14,45%	13,69%

The implicit prices are constant in the hedonic model, while in the NN, these marginal effects are non linear functions of the explanatory variables. To obtain the marginal or implicit prices, it is necessary to evaluate the derivatives  $\partial Price / \partial x_j$  for each explanatory variable, it is a cumbersome process that, nevertheless, allows to study the evolution of the hedonic prices attributed to each exogenous variable. For example, to estimate the influence of the area upon the marginal price of an additional square metre.

The hedonic model shows a linear influence of area upon de value attribute (Figure 2), while the NN shows a concavity in this value. The remaining explanatory variables are held constant (in this case, with average values over the sample). An additional square metre in the dwelling size increases its value by over a thousand euros, increasing with size in the network. The conclusions are in line of the finding of García Rubio (2004) in a smaller city. The adaptability of the networks to the situations at both extremes of the size range is not possible with classical hedonic models.

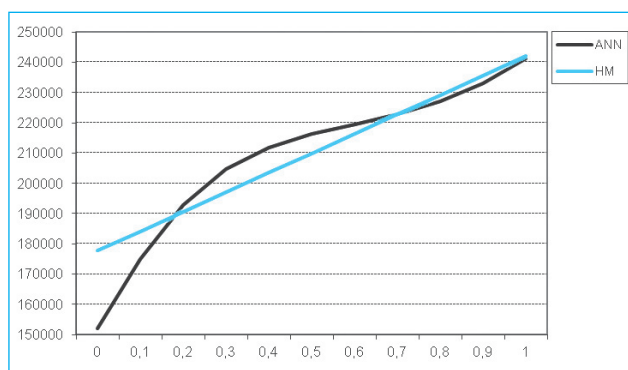
Figure 2. Implicit prices (Euros) associated to the dwelling surface (meters)





The localisation index is obtained taking into account the urban zone and the income level. Upper values correspond to better surroundings and higher income areas. In the hedonic model, an increase in 0.1 units in this index is valued at 6429.73€. Being constant is not very realistic, so it is necessary to estimate the corresponding implicit price using the NN, as can be seen in the Figure 3.

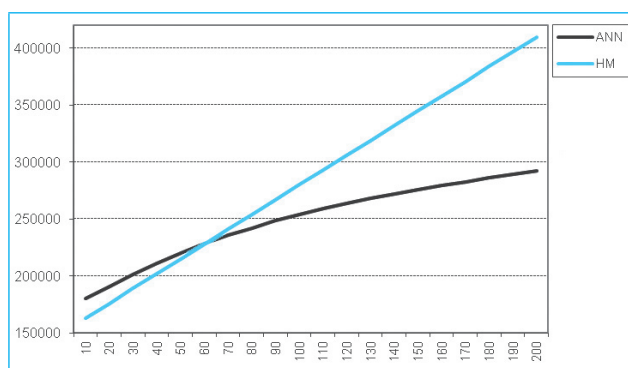
**Figure 3.** Implicit prices (Euros) associated to the location index



The influence of the localisation of the property show a decreasing rate of increase at lower values of the localisation index, and then a stabilisation until reaching an almost constant value for the best urban areas.

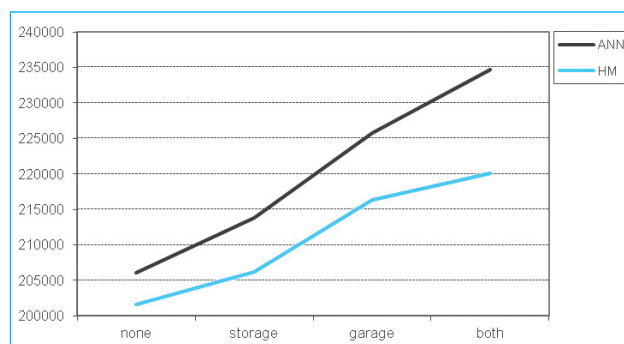
The common expenses influence the valuation of each dwelling; an additional euro in these expenses is linked to nearly 1.3 thousands euros in the value of the property. The NN shows a stabilising influence. When the common expenses increase over a certain level, the implicit price is almost constant (Figure 4).

**Figure 4.** Implicit prices (Euros) associated to the common expenses (Euros)



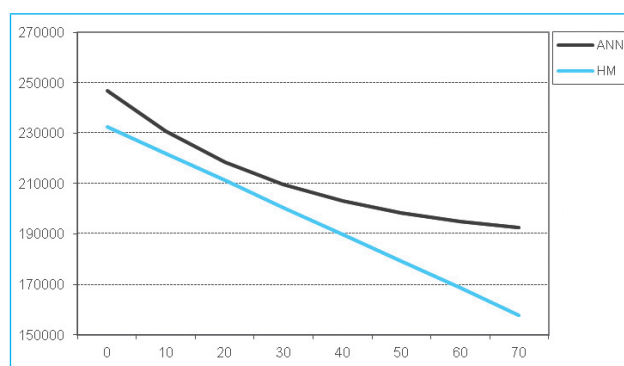
The additional facilities to the property, as an additional storage room and a parking place, are valued in a different way with the NN than with the hedonic model; with the former, 28552€, more realistically than with the latter, 18459€, according to the market price of these characteristics, when sold independently (Figure 5).

**Figure 5.** Implicit prices (Euros) associated to the annexes



When the number of years, since the property was built, is considered, the decrease in the implicit price is a decreasing concave function, with the NN, while the hedonic price do not show the real depreciation perceived by the buyers, showing a linear decrease over time (Figure 6).

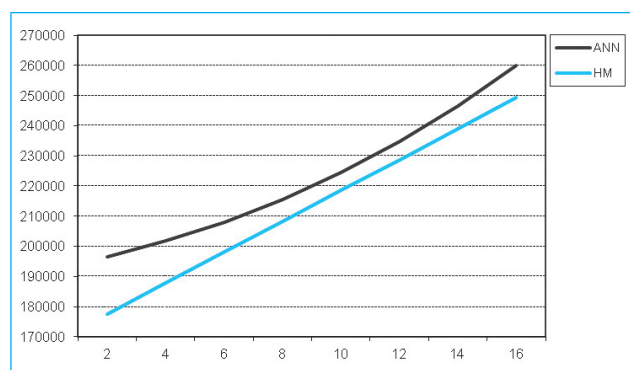
**Figure 6.** Implicit prices (Euros) associated to the age of the building (years)



The last input variable is categorical, with 16 possible values, as it represents the combination between of

two interacting variables, each of them with four levels (Figure 7). The hedonic model undervalues the situation with quality at both extremes, while the NN adapts better to all the range of this interaction. The differences take into account other aspects of the maintenance and quality of the building, that are positively correlated with both variables included in this interaction.

**Figure 7.** *Implicit prices (Euros) associated to the interaction between the state of the pavement and of the carpentry*



## 6. Conclusions

NN seem well suited to estimate price functions in the real estate market. This option provides several advantages over classical hedonic modelling. For example, it is easier to introduce original information avoiding, partly, the always subjective index building process. The non linear relations between the internal and external attributes of a property and its sales prices are directly introduced

in the network. It is possible to build parsimonious models using MLP with a small amount of hidden units that produce smaller absolute error in price estimation. The minimum relevant information to produce a precise estimate is better processed than in an hedonic model, so NN are a useful statistical tool to be applied in the real estate market, with the advantage of not needing a special knowledge of the situation. It can also be used to detect the relevant price determination variables that are influencing the sale price, at different ranges and types of properties. These variables can change depending of each individual pattern. NN methodology cannot avoid the identification problems associated to the building process of offer and demand function estimation, or the economic difficulties with the hedonic prices interpretation.

NN produce better estimates (i.e. lower mean absolute errors) than the hedonic models using the same set of variables; it should be expected as there are nonlinearities in the real estate market that are more fit within are difficult to include in a hedonic model with many medium and low value dwellings, while NN introduce a better fit in all the price range.

Also, it has to be stated that some statistical problems linked to the categorical attributes that partially define a property are dealt smoothly by NN.

To estimate a NN it is necessary to have a quite large sample, and this can be a problem, as there are periods that the real estate market is very volatile and, others with not many transactions.

The price to be paid for this more efficient black box approach in the lack of interpretation for the NN weights.

## References

- Aguiló-Segura, P.M. (2002). El método de valoración de los precios hedónicos. Una aplicación al sector residencial de las Islas Baleares. Tesis Doctoral. University of The Isles Baleares (Spain).
- Azqueta, D. (1994). La Función de Precios Hedónicos en la Práctica. Valoración económica de la Calidad Ambiental. Mc Graw-Hill. Madrid.
- Bengochea Morancho, A. (2003). A hedonic valuation of urban green areas. *Landscape and Urban Planning*, 66, 35-41.
- Bilbao Terol, C. (2000). Relación entre el precio de venta de una vivienda y sus características: un análisis empírico para Asturias. *Revista Asturiana de Economía*, 18, 141-150.
- Bover, O. and Velilla, P. (2001). Precios hedónicos de la vivienda sin características: el caso de las promociones de viviendas nuevas. *Estudios Económicos*, 73. Banco de España. Madrid.
- Caridad, J. M. and Brañas, P. (1996). Demanda de características en Córdoba: un modelo de precios hedónico. *Revista de Estudios Regionales*, 46, 139-153.
- Ceular, N. and Caridad, J. M. (2001). Sistemas neuronales en valoración urbana. *Estudios de Economía Aplicada*, 18, 67-81.
- Court, A.T. (1939). Hedonic Price Indexes with Automotive Examples. *Dynamics of Automobile Demand* (General Motors Corporation), 99-117. New York.

- Freeman, A. M. (1979). The hedonic approach to measuring demand for neighborhood Characteristics, The Economics of Neighborhoods. Academic Press.
- Freeman, J. and Skapura, D.M. (1993). Redes neuronales algoritmos, aplicaciones y técnicas de programación. Ed.Wilmington Addison-Wesley Díaz de Santos, 26-241.
- Fuentes Jiménez, A. M. (2004). Métodos estadísticos y económicos para la determinación del precio de la vivienda. Phd Thesis (in spanish). University of Granada (Spain).
- Gallego Mora-Esperanza, J. (2004). La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid. CT: Catastro, nº 50, abril, 51-67.
- García Rubio, N. (2004). Desarrollo y aplicación de redes neuronales artificiales al mercado inmobiliario: aplicación a la ciudad de Albacete. Phd Thesis (in spanish). Universidad de Castilla, La Mancha.
- Greene, W.H. (1993). The econometric approach to Efficiency Analysis. H.O. fried, CAK Lovell y S. S. Schmidt (editor) The Measurement of Productive Efficiency, Oxford, University Press, New York, 68-119.
- Griliches, Z. (1971). Introduction: Hedonic Price Indexes Revisited. In Price Indexes and Quality Changes: Studies in New Methods of Measurement, Griliches (Eds.), Cambridge: Harward University Press.
- Hamzaoui, Y. E. & Hernández Pérez, J.A. (2011). Application of Artificial Neural Networks to predict the selling Price in the real estate valuation. 10th Mexican International Conference on Artificial Intelligence, 175-181. Puebla.
- Haykin, S. (1999). Neural networks: A comprehensive foundation. Ed. Prentice - Hall.
- Jaén, M. and Molina, A. (1995). Modelos econométricos de tenencia y demanda de vivienda. Universidad de Almería.
- Khalafallah, A. (2008). Neural Network Based Model for predicting housing market performance. Tsinghua Science and Technology, 13 (S1) 325-328.
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. Journal of Political Economy, 74, 132-157.
- Lara Cabeza, J. (2005). Aplicación de las redes neuronales artificiales al campo de la valoración inmobiliaria. Mapping, 104, 64-71.
- Liu, J., Zhang, G. X. L., & Wu, W. P. (2006). Application of fuzzy neural network for real estate prediction, Lecture Notes in Computer Science (LNCS) 3973, 1187-1191.
- Martín del Brío, A. & Sanz de Molina, A. (1997). Redes neuronales y sistemas borrosos. Ra-Ma. Madrid.
- Mohamed Amar, R. (2002). Estimación del precio de la vivienda urbana mediante redes neuronales artificiales: Estudio de un caso. XII, Jornadas Hispano-Lusas. Covilhã.
- Muth, R. F. (1969). Cities and Housing. Chicago. University Press, Chicago.
- Richardson, H. W. (1978). Urban economics. Hinsdale, Dryden Press.
- Roseblatt, F. (1958). The Perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65, 304-408.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy, 82, 34-55.
- Saura P. (1995). Demanda de características de la vivienda en Murcia. Secretariado de Publicaciones, University of Murcia.
- Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. Expert Systems with Applications, 36, 2843-2852.
- Sirmans, G.S. (2005). The composition of Hedonic Pricing Models. Journal of Real Estate Literature. 13 (1) 3-43.
- Solow, R. (1972). Congestion, density and the use of land in transportation. Journal of Economics, 74, 161-173.
- Tiebout, C. (1956). A pure theory of local expenditures. Journal of Political Economy, 64, 416-535.
- Wallace, H. A. (1926). Comparative farmland values in Iowa. The Journal of Land and & Public Utility Economics, 2(4), 385-392.
- Waight, F.V. (1929). Quality as a Determinant of Vegetable Prices. Columbia University Press. New York.