



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*



New Zealand Agricultural and
Resource Economics Society (Inc.)

**House Price Prediction:
Hedonic Price Model vs. Artificial Neural Network**

Visit Limsombunchai

Commerce Division, Lincoln University, Canterbury 8150, New Zealand.
e-mail: limsombv@lincoln.ac.nz

**Paper presented at the 2004 NZARES Conference
Blenheim Country Hotel, Blenheim, New Zealand. June 25-26, 2004.**

Copyright by author(s). Readers may make copies of this document for non-commercial purposes only, provided that this copyright notice appears on all such copies.

House Price Prediction: Hedonic Price Model vs. Artificial Neural Network

Visit Limsombunchai¹
Commerce Division, Lincoln University,
Canterbury 8150, New Zealand.
Email: limsombv@lincoln.ac.nz

Abstract

The objective of this paper is to empirically compare the predictive power of the hedonic model with an artificial neural network model on house price prediction. A sample of 200 houses in Christchurch, New Zealand is randomly selected from the Harcourt website. Factors including house size, house age, house type, number of bedrooms, number of bathrooms, number of garages, amenities around the house and geographical location are considered. Empirical results support the potential of artificial neural network on house price prediction, although previous studies have commented on its black box nature and achieved different conclusions.

Key Words: Hedonic Model, Artificial Neural Network (ANN), House Price.

JEL Classification: C53, L74.

1. Introduction

An accurate prediction on the house price is important to prospective homeowners, developers, investors, appraisers, tax assessors and other real estate market participants, such as, mortgage lenders and insurers (Frew and Jud, 2003). Traditional house price prediction is based on cost and sale price comparison lacking of an accepted standard and a certification process. Therefore, the availability of a house price prediction model helps fill up an important information gap and improve the efficiency of the real estate market (Calhoun, 2003).

In New Zealand, most people know the benefit of owning a house, because buying a house is considered the most utilised and profitable investment. New Zealand has one of the highest ratios of people owning their houses in western world with over 70% of its citizens living in their own houses. As house market in New Zealand is thriving, house price becomes a crucial factor for house seekers.

Over the last two decades there has been a proliferation of empirical studies analysing residential property values, with Ball (1973) being last major study. Each succeeding research has generally improved the predictive power of the models by emphasising attributes of property value such as housing site, housing quality, geographical location and the environment. More recent studies have focused on location externalities, transaction costs and factors affecting the future expected cost in homeownership (Norman, 1982).

¹ The author is a Ph.D. student and would especially like to thank Dr.Christopher Gan and Dr.Minsoo Lee for invaluable comments and suggestions.

The hedonic price models have been commonly used to estimate house prices and property values. Most of the models include housing attributes such as location, neighbourhood, and house size. However, there is a limited number of studies in this area using an artificial neural network technique. This paper uses the hedonic method and artificial neural network to empirically determine the house prices in Christchurch, New Zealand. Secondary data from 200 houses in Christchurch is used in a hedonic price framework and artificial neural network to empirically compare the predictive power of both techniques and to suggest an appropriate technique for the house price prediction.

This paper is divided into the following sections. Section 2 provides an overview of the hedonic price model and artificial neural network. Section 3 presents the models and section 4 discusses the data, variables and methodology that used in this paper. Section 5 reports the empirical results, and section 6 concludes the findings.

2. Hedonic Price and Artificial Neural Network Theory

2.1 Hedonic Price Theory

Hedonic price theory assumes that a commodity such as a house can be viewed as an aggregation of individual components or attributes (Griliches, 1971). Consumers are assumed to purchase goods embodying bundles of attributes that maximize their underlying utility functions (Rosen, 1974). Rosen (1974) describes the process in which prices reveal quality variations as relying on producers who "tailor their goods to embody final characteristics described by customers and receive returns for serving economic functions as mediaries". Hedonic price theory originates from Lancaster's (1966) proposal that goods are inputs in the activity of consumption, with an end product of a set of characteristics.

Bundles of characteristics rather than bundles of goods are ranked according to their utility bearing abilities. Attributes (for example, characteristics of a house such as number of bedrooms, number of bathrooms, number of fireplaces, parking facilities, living area and lot size) are implicitly embodied in goods and their observed market prices. The amount or presence of attributes associated with the commodities defines a set of implicit or "hedonic" prices (Rosen, 1974). The marginal implicit values of the attributes are obtained by differentiating the hedonic price function with respect to each attribute (McMillan et al., 1980). The advantage of the hedonic methods is that they control for the characteristics of properties, thus allowing the analyst to distinguish the impact of changing sample composition from actual property appreciation (Calhoun, 2001).

Hedonic price theory has been applied to valuation of agricultural commodities (Brorsen et al., 1984; Ethridge and Davis, 1982; Wilson, 1984), residential amenities (Blomquist and Worley, 1981; McMillan et al., 1980; Witte et al., 1979; and Milon et al., 1984,) and wildlife related recreation resources (Pope and Stoll, 1985; Livengood, 1983; Pope et al., 1984; and Messonnier and Luzar, 1990). Other applications have involved the estimation of the benefits of environmental improvements (Freeman, 1979; Blomquist and Worley, 1981; Harrison and Rubinfeld, 1978; and McMillan et al., 1980).

While the hedonic technique is an acceptable method for accommodating attribute differences in a house price determination model, it is generally unrealistic to deal with the housing market in any geographical area as a single unit. Therefore, it seems more reasonable to

introduce geographical information or location factor into a model that allows shifts in the house price level. Frew and Wilson (2000) employ the hedonic price model to examine the relationship between location and property value, in Portland, Oregon, and the authors found that there was a significant relationship between location and property value.

Fletcher et al. (2000) examine whether it is more appropriate to use aggregate or disaggregate data in forecasting house price using the hedonic analysis. It is found that the hedonic price coefficients of some attributes are not stable between locations, property types and age. However, it is argued that this can be effectively modelled with an aggregate method. The hedonic price model has also been used to estimate individual external effects (e.g. environmental attribute) on house prices. For example, there is a number of studies that have applied the hedonic price model in quantifying the effects of noise (Mieszkowski and Saper, 1978; Damm et al., 1980; Uyeno et al. 1993) and air pollution on house prices (Ridker and Henning, 1982; Graves et al, 1988).

Even though the hedonic price model has been widely recognized, issues such as model specification procedures, multicollinearity, independent variable interactions, heteroscedasticity, non-linearity and outlier data points can seriously hinder the performance of hedonic price model in real estate valuations. The artificial neural network model has been offered as a possible solution to many of these problems, especially when the data patterns show non-linearity (Lenk et al. 1997; Owen and Howard, 1998).

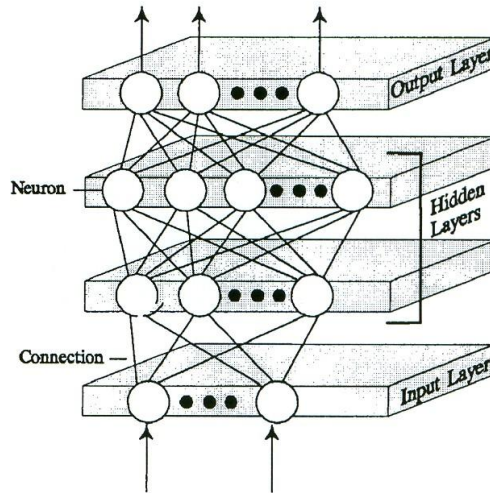
Tay and Ho (1991), using a large sample of data from the apartment sector in Singapore, found that a neural network model performs better than a multiple regression model for estimating value. The authors concluded that the neural network can learn valuation patterns for “true” open market sales in the presence of some “noise” as a way of establishing a robust estimator. Similar results can be found in Do and Grudnitski (1992) and McCluskey (1996) studies.

Worzala et al. (1995), on the other hand, take on a contrary position and cast some doubt upon the role of neural networks compared to the traditional regression models. The authors argued that even when the same data is used, results from models prepared by different neural network software package could be inconsistent and did not always outperform regression models. Lenk et al. (1997) also reached the similar conclusions. Their study documented very similar performance between the hedonic model and the neural network models.

2.2 Artificial Neural Network Theory

Neural network is an artificial intelligence model originally designed to replicate the human brain’s learning process. The model consists of three main layers: input data layer (example the property attributes), hidden layer(s) (commonly referred as “black box”), and output layer (estimated house price). Neural network is an interconnected network of artificial neurons with a rule to adjust the strength or weight of the connections between the units in response to externally supplied data (see Figure 1) (Stanley et al., 1998).

Figure 1: Feed-forward neural network structure with two hidden layers.



Source: James and Carol, 2000.

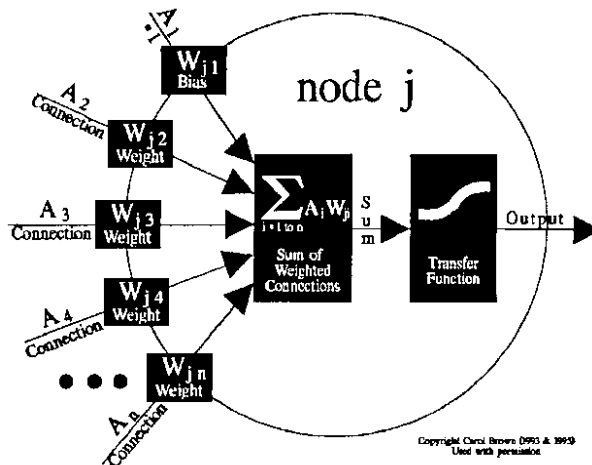
Each artificial neuron (or computational unit) has a set of input connections that receive signals from other computational units and a bias adjustment, a set of weights for input connection and bias adjustment, and transfer function that transforms the sum of the weighted inputs and bias to decide the value of the output from computational unit (see Figure 2). The output for the computation unit (node j) is the result of applying a transfer function ϕ to the summation of all signals from each connection (A_i) times the value of the connection weight between node j and connection i (W_{ji}) (refer to equations 1 and 2).

$$\text{Sum}_j = \sum_j (W_{ji}A_i) \quad [1]$$

$$O_j = \phi (\text{Sum}_j) \quad [2]$$

where O_j is output for node j and ϕ is transfer function which can take many different forms: linear functions, linear threshold functions, step linear functions, sigmoid function or Gaussian functions (James and Carol, 2000).

Figure 2: Structure of a Computational Unit (node j)



Source: James and Carol, 2000.

3. Models

3.1 Hedonic Price Model

The hedonic model involves regressing observed asking-prices for the house against those attributes of a house hypothesized to be determinants of the asking-price. Attributes hypothesized to contribute to the price of a house include land size (in square meters), house age (in years), number of bedrooms, number of bathrooms, number of toilets, and number of garages, binary variables representing the type of house (with garden, or without garden) and amenities around the residential areas (such public facilities). In addition, the geographical location of the house also plays an important factor in influencing the house price. In this paper, the Christchurch area is divided into six different geographical locations. They are Inner Christchurch, North Christchurch, South Christchurch, East Christchurch, West Christchurch, and Northwest Christchurch. The location dummy variables equal to 1 if a particular property is situated in the identified location, 0 otherwise.

Implicitly, the model for the hedonic price function is specified as:

$$\text{PRICE} = f(\text{LAND, AGE, TYPE, BEDROOMS, BATHROOMS, GARAGES, AMENITIES, INNER CHRISTCHURCH, NORTH CHRISTCHURCH, SOUTH CHRISTCHURCH, EAST CHRISTCHURCH, WEST CHRISTCHURCH, NORTHWEST CHRISTCHURCH, } \epsilon) \quad [3]$$

Variables in the model are defined as:

| | |
|-----------------|--|
| PRICE = | Price of house in Christchurch in NZD |
| LAND (+) = | Land size (in square meters) |
| AGE (-) = | Age of the house (in years) |
| TYPE (+) = | Type of house; 1 if the house has a garden, 0 otherwise |
| BEDROOMS (+) = | Number of bedrooms |
| BATHROOMS (+) = | Number of bathrooms |
| GARAGES (+) = | Number of garages |
| AMENITIES (+) = | Amenities around the house; 1 if the house is close to two or more public facilities (i.e. bus stop, school, public park and so on), 0 otherwise |
| ϵ = | Error term |

A priori hypotheses are indicated by (+) and (-) in the above specification. Based on previous literature, it is hypothesised that most of the variables have a positive relationship with the house price, except age of the house. For example, a house with garden is more expensive than a house without garden. A small house should be cheaper than a large house. A house that has multiple bedrooms, bathrooms, garages and close to public amenities (such as public parks, public libraries, etc) is expected to command a higher price than a house that has less number of bedrooms, bathrooms, toilets, garages and no public amenities nearby. Conversely, the age of a house would have a negative relationship with house price since an old house commands a lower price compared to a newly built house.

3.2 Artificial Neural Network Model

The use of the neural network model is similar to the process utilized in building the hedonic price model. However, the neural network must first be trained from a set of data. For a particular input, an output (estimated house price) is produced from the model. Then, the model compares the model output to the actual output (actual house price). The accuracy of this value is determined by the total mean square error and then back propagation is used in an attempt to reduce prediction errors, which is done through the adjusting of the connection weights.

The performance of the network can be influenced by the number of hidden layers and the number of nodes that are included in each hidden layer. Unfortunately, there exists little theory to support the process for the determination of the optimal number of hidden layers and nodes, and also the optimal internal error threshold (Lenk et al., 1997). Therefore, a trial-and-error process is applied to find the optimal artificial neural network model. A feed-forward/back-propagation neural network software package, NeuroShell, was used to construct the artificial neural network model.

There are no assumptions about functional form, or about the distributions of the variables and errors of the model, neural network model is more flexible than the standard statistical technique (Mester, 1997). It allows for nonlinear relationship and complex classificatory equations. The user does not need to specify as much detail about the functional form before estimating the classification equations but, instead, it lets the data determine the appropriate functional form.

In accordance to standard analytical practice, the sample size was divided on a random basis into 2 sets, namely the “training set” and the “production set” (as known in neural network literature), or the “estimation set” and the “forecasting set” (as know in regression analysis literature). The training set and the production set contain 80% and 20% of the total sample, respectively. To evaluate the forecasting accuracy of both models, an out-of sample forecasting is operated, subsequently, the R^2 and the Root Mean Square Error (RMSE) were calculated and compared (refer to equations 4 and 5). The model with a higher R^2 and lower RMSE was considered to be a relatively superior model.

$$R^2 = 1 - \frac{\sum_{i=1}^n (P_i - \hat{P}_i)^2}{\sum_{i=1}^n (P_i - \bar{P})^2} \quad [4]$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - \hat{P}_i)^2} \quad [5]$$

where P_i is actual house price, \hat{P}_i is estimated house price and n is the number of observations.

4. Data and Procedures

A sample of 200 housing information in the Christchurch area is randomly selected from the largest real estate agent, Harcourt. The data set is retrieved from Harcourt's website (www.bluebook.co.nz) in May 2003.

Since most business offices, restaurants and shops are located in the inner city centre, the proportion of residential houses is quite small. Only 15 housing data is collected from the inner city, 25 housing information is from North Christchurch, and 40 housing information for the remaining four identified locations. There are 200 observations utilized in this study.

Economic theory offers little guidance with respect to the choice of functional form for the hedonic model as the hedonic price function represents an equilibrium relationship derived from individuals' preferences and suppliers' cost functions (Freeman, 1993). While earlier hedonic studies used linear specifications, recent investigations aimed at identifying more appropriate functional specifications have indicated the superiority of flexible forms (Cooper et al., 1987; Milon et al., 1984). Coefficients resulting from linear specifications identify the relative contribution of their respective attributes to the price of the product. Linear specifications, however, imply constant marginal willingness-to-pay for all households consuming the good (Freeman, 1979). This does not allow for the identification of the demand schedule for the attribute in question and also ignores the possibility that demand for the attribute may be a function of its level as well as the level of other attributes. In the case of non-linear specifications, the first derivative of the hedonic price function with respect to the specified attribute yields the implicit marginal price of the attribute (McMillan et al., 1980).

As economic theory provides no clear guidance regarding the choice of functional form to be used in hedonic regression, this paper employed the semi-log model because price is a very sensitive and volatile component (Shonkwiler and Reynolds, 1986).

5. Empirical Results

The estimated coefficients of equation 1 are shown in Model 1 (see Table 1). The weighted least square (WLS) technique and the White (1982) adjustment for estimating a heteroscedasticity consistent covariance matrix are applied to equation 1 instead of the ordinary least square technique because of heteroscedasticity. The number of toilets was dropped from equation 1 to avoid multicollinearity problem since the number of toilets (TO) was found to have a high correlation with the number of bathrooms (BA) (see Table 2).

Model 1 shows that all of coefficients have correct hypothesised signs and most of the coefficients are statistically significant. It should be noted that White heteroscedasticity test still indicate the heteroscedasticity problem, even if the weighted least square (WLS) and the White adjustment techniques are utilized. The estimated results demonstrate that houses with more bedrooms and bathrooms are priced higher. A relatively new house is more expensive than an old house, and a house with garden is priced higher than one without garden. Location variables play a significant role on house prices. For example, houses in the Northwest of Christchurch (such as Burnside, Fendalton, Ilam, and Merivale) are priced higher since they have access to good public and private high schools in those area due to the school-zone policy and the University of Canterbury. Furthermore, Fendalton has traditionally

been known as an upper income area. On the contrary, properties in East of Christchurch (such as Linwood, Phillipstown, Aranui, and Bexley) are priced lower than the rest of areas since it is relatively a poor neighbourhood and most of houses are relatively older than those in other areas.

Table 1: Hedonic Price Models

| Variable ^{1/} | Coefficient | | |
|-------------------------------|-----------------------|-----------------------|-----------------------|
| | Model 1 ^{2/} | Model 2 ^{2/} | Model 3 ^{2/} |
| C | 11.1763 ^{**} | 11.2526 ^{**} | 10.3075 ^{**} |
| LAND (L) | 0.0001 | 0.0004 | -0.0020 |
| AGE (A) | -0.0028 [*] | -0.0029 ^{**} | -0.0118 |
| TYPE (TY) | 0.3599 [*] | | |
| BEDROOMS (BD) | 0.0788 | 0.0622 | 0.6687 ^{**} |
| BATHROOMS (BA) | 0.2411 ^{**} | 0.3517 ^{**} | 0.3314 |
| GARAGES (G) | 0.1826 [*] | 0.0506 | 0.0321 |
| AMENITIES (AM) | 0.0366 | 0.0941 [*] | 0.1997 |
| North Christchurch (NC) | -0.1955 [*] | -0.0281 | 0.2436 |
| South Christchurch (SC) | -0.2759 ^{**} | -0.0170 | na. |
| East Christchurch (EC) | -0.4521 ^{**} | -0.2483 ^{**} | -0.1040 |
| West Christchurch (WC) | -0.2250 [*] | -0.0001 | 0.1947 |
| Northwest Christchurch (NWC) | 1.9423 ^{**} | 2.3804 ^{**} | 1.5601 ^{**} |
| R ² | 0.7833 | 0.8780 | 0.7817 |
| Adj. R ² | 0.7657 | 0.8660 | 0.6944 |
| F-stat. | 44.2913 ^{**} | 73.2795 ^{**} | 8.9516 ^{**} |
| White Heteroscedasticity Test | 7.1970 ^{**} | 2.3153 ^{**} | 5.3612 ^{**} |
| | n = 160 | n = 124 | n = 36 |

Note: 1/ Dependent Variable is Log(P).

2/ WLS and White adjustment for estimating a heteroscedasticity consistent covariance matrix.

*, ** represent 10% and 5% significant level, respectively.

Model 1 is hedonic price model for both house with garden and without garden.

Model 2 is hedonic price model for house with garden.

Model 3 is hedonic price model for house without garden.

In general, houses with gardens are usually located away from the city or shopping mall areas, while the houses without garden are located closer to the business district centre, town, and university. Thus, houses with gardens versus houses without gardens reflect different market segment and different pricing strategy. For example, Model 1 shows the average price of a house with garden is higher than a house without garden in every location (see Table 1). Therefore, it can be concluded that house prices are determined differently according to its type.

Table 2: Correlation Coefficient Matrix

| | P | L | A | BD | BA | TO | G |
|----|-------|------|-------|------|-------------|------|------|
| P | 1.00 | | | | | | |
| L | 0.09 | 1.00 | | | | | |
| A | -0.19 | 0.30 | 1.00 | | | | |
| BD | 0.18 | 0.50 | -0.12 | 1.00 | | | |
| BA | 0.35 | 0.22 | -0.33 | 0.61 | 1.00 | | |
| TO | 0.47 | 0.20 | -0.35 | 0.58 | 0.85 | 1.00 | |
| G | 0.33 | 0.35 | -0.32 | 0.52 | 0.52 | 0.53 | 1.00 |

The hedonic price models (Models 2 and 3) are segregated according to property type, that is, houses with gardens and houses without gardens respectively (see Table 1). The R^2 in both models is relatively high but the coefficients in both models, such as land size, garages and some geographical locations, are statistically not significant. Furthermore, the null hypothesis of White heteroscedasticity test is rejected at 5 the percent level in both models. The results indicate that the segregation model improves the explanatory power of the model but cannot overcome the problem of heteroscedasticity. The insignificant of the variables may be caused by the reduction of the sample size since there are only 36 observations on house without garden model.

The back propagation training process is always regarded as a black box in the neural network model, thus the internal characteristics of a trained network is simply a set of numbers which prove to be difficult in relating back to the application in a meaningful fashion. For that reason, the learned output (weights or coefficients) cannot be interpreted or utilized as price adjustments.

The relative contribution factors of the best artificial neural networks (the relative importance of inputs) are shown in Table 3. All three networks employ the same variables for the input layer nodes that are used as the independent variables to create the hedonic price models. Ward networks (multiple hidden slabs with different activation functions), which use Gaussian, Tanh, and Gaussian Complement (Ward System Group Inc., 1993) as the activation functions for 3 hidden slabs and each slab contains 6 hidden nodes, are considered as the best networks in this study. Although neural networks with 1 and 2 hidden layers are examined and their results are slightly better than the hedonic price models, the results are not presented here because they do not outperform Ward networks.

The relative contribution factor in Table 3 shows that land size and number of garages, respectively, are important factors that determine the house price for house with garden while amenities near the house area is the less important factor (see model 2). Generally, houses with gardens are located in the outskirts of the business district centers since they require large land sizes. Thus amenities around the house area may not be an important factor impact the house price. However, a larger land size means a higher price of the house.

Table 3: Neural Networks' Relative Contribution Factor

| Factor | Relative Contribution | | |
|------------------------------|-----------------------|---------|---------|
| | Model 1 | Model 2 | Model 3 |
| LAND (L) | 0.0879 | 0.1724 | 0.0608 |
| AGE (A) | 0.2231 | 0.0936 | 0.1804 |
| TYPE (TY) | 0.0766 | | |
| BEDROOMS (BD) | 0.0649 | 0.0598 | 0.1749 |
| BATHROOMS (BA) | 0.0621 | 0.1206 | 0.0517 |
| GARAGES (G) | 0.1700 | 0.1615 | 0.1824 |
| AMENITIES (AM) | 0.0675 | 0.0355 | 0.1160 |
| North Christchurch (NC) | 0.0299 | 0.0747 | 0.0463 |
| South Christchurch (SC) | 0.0306 | 0.0639 | na. |
| East Christchurch (EC) | 0.0391 | 0.0453 | 0.0375 |
| West Christchurch (WC) | 0.0493 | 0.0788 | 0.0319 |
| Northwest Christchurch (NWC) | 0.0990 | 0.0940 | 0.1181 |
| R^2 | 0.9450 | 0.9942 | 0.9378 |
| | n = 160 | n = 124 | n = 36 |

Note: Ward network is utilized with learning rate = 0.1, momentum = 0.1 and initial weight = 0.3.
 Model 1: house with and without garden.
 Model 2: house with garden.
 Model 3: house without garden.

For house without garden, age of house and the number of garages are factors that have strong impact on the house price (see Model 3 in Table 3). Land size for house without garden is less important compared to house with garden. On the other hand, age of the house, the number of bedrooms, the number of garages and amenities around the house areas do impact the house price for the house without garden when compared to the house with garden.

On the aggregate model (see Model 1 in Table 3), the neural networks' relative contribution factor demonstrates that the age of the house and the number of garages, respectively, have contributed to the predictive power of model than the other variables. Geographical location such as Northwest of Christchurch has a relatively high impact to the house price compared to land size, house type, number of bedrooms, number of bathrooms and amenities around the house area. The result indicates that geographical location plays an important role on the house price determination.

The R^2 from neural network models are higher than the R^2 from hedonic price models (see Table 3). The results imply that the neural network model can estimate the house price more accurately than the hedonic price model in both aggregate and disaggregate models (see Figure 3). However, the results do not provide strong and conclusive evidence of superiority in term of prediction capability between both models, as shown by the sample results.

Table 4 shows the out-of-sample forecast evaluation results for hedonic price models and neural network models. Again, the R^2 of neural network models are higher than the R^2 of hedonic price models, and the RMSE of neural network models are lower than hedonic price models. Therefore, it can be concluded that the neural network model is relatively superior model for house price prediction (see Figure 4).

Figure 3: Actual and estimated house prices in log form (in sample forecast)

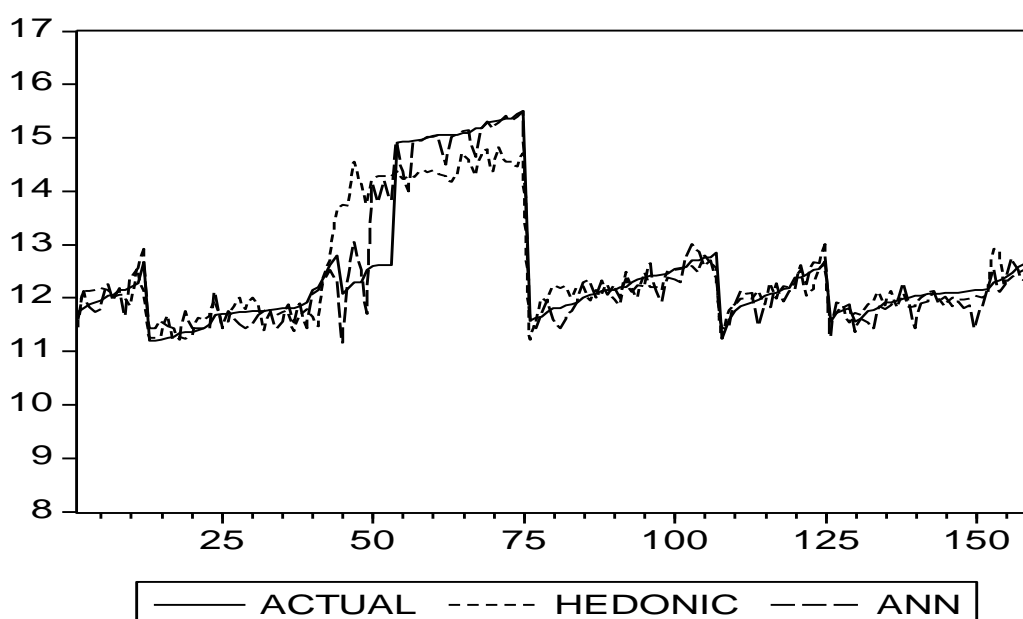


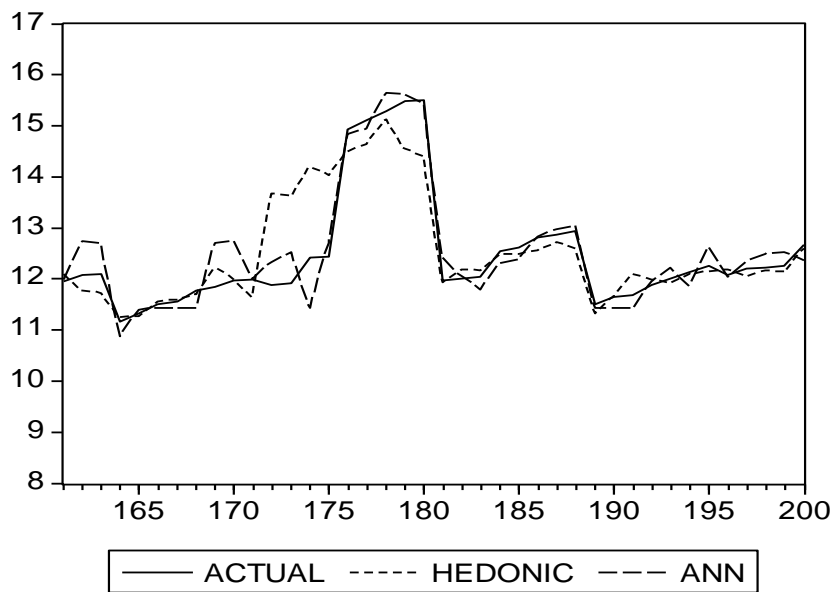
Table 4: Comparing the Out-of-Sample Forecast Evaluation Results for Hedonic Price Model and Neural Network Model

| | Model 1 | Model 2 | Model 3 |
|-----------------------------|------------|------------|--------------|
| Hedonic price model | | | |
| - R^2 | 0.6192 | 0.7499 | 0.3807 |
| - RMSE | 876,215.63 | 642,580.05 | 1,435,810.81 |
| Neural network model | | | |
| - R^2 | 0.9000 | 0.8408 | 0.6907 |
| - RMSE | 449,111.46 | 512,614.99 | 1,014,721.92 |
| | n = 40 | n = 31 | n = 9 |

Note: Model 1: house with and without garden.
 Model 2: house with garden.
 Model 3: house without garden.

The results from Table 4 also suggest that the better model for house price prediction should be the aggregate neural network model rather than the disaggregate models, as it has the highest R^2 and the lowest RMSE. Even though the neural network models for house with and without garden have relatively high R^2 in the case of in sample forecast (0.9942 and 0.9378, respectively), their performances are not good compared to the out-of-sample forecast, especially houses without garden. The low number of observations may be one of the possible explanation for the poor performance of the model since the aggregate model has higher number of observations than the disaggregate models.

Figure 4: Actual and estimated house prices in log form (out-of-sample forecast)



6. Conclusion

This paper empirically compares the predictive power of the hedonic price model with an artificial neural network model on house price prediction. Artificial neural network models and hedonic price models are tested for their predictive power using 200 houses information in Christchurch, New Zealand.

The results from hedonic price models support the previous findings. Even if the R^2 of hedonic price models are high (higher than 75%) for in sample forecast, the hedonic price models do not outperform neural network models. Moreover, the hedonic price models show poorer results on out-of-sample forecast, especially when comparing with the neural network models. Thus, the empirical evidence presented in this paper supports the potential of neural network on house price prediction, although previous literatures have commented upon its black box nature and reached different conclusions.

The non-linear relationship between house attributes and house price, the lack of some environmental attributes, and inadequate number of sample size could be the cause of the poor performance of the hedonic price models. However, it should be noted that the optimal artificial neural network model is created by a trial-and error strategy. Without this strategy, the results may not indicate superiority of the neural network model (Lenk et al., 1997).

There are, however, some limitations in this paper. Firstly, the house price used is not the actual sale price but the estimated price due to the difficulty in obtaining the real data from the market. Secondly, this paper considered only the current year's information of the houses. The time effect of the house price, which could potentially impact the estimated results was ignored (the same house should have different price in different years, assuming that age factor is constant). Finally, the house price could be affected by some other economic factors (such as exchange rate and interest rate) are not included in the estimation.

References:

- Allen C. G., 1977, "Hedonic Prices, Price Indices and Housing Markets", Center for metropolitan planning and research and department of political economy, the Johns Hopkins University, Baltimore, Maryland.
- Blomquist, G. and L. Worley, 1981, "Hedonic Prices, Demand for Urban Housing Amenities and Benefit Estimates", *Journal of Urban Economics*, 9: 212 - 221.
- Brorsen, B. W., W. R. Grant and M. E. Rister, 1984, "A Hedonic Price Model for Rough Rice Bid/Acceptance Markets", *American Journal of Agricultural Economics*, 66(2): 156 - 163.
- Calhoun C. A., 2001, "Property Valuation Methods and Data in The United States", *Housing Finance International*, 16(2): 12 - 23.
- Calhoun C. A., 2003, "Property Valuation Models and House Price Indexes for The Provinces of Thailand: 1992 – 2000", *Housing Finance International*, 17(3): 31 – 41.
- Cooper, M. L., L. B. Deck and K. E. McConnell, 1987, "On the Choice of Functional Form for Hedonic Price Functions", Resources for the Future Discussion Paper QE87-08, Washington, DC: Resources for the Future.
- Damm D., Lerman S. R., Lemer-Lam E. and J. Young, 1980, "The Response of Urban Real Estate Values in Anticipation of The Washington Metro", *Journal of Transport Economics and Policy*, September, 315 – 336.
- Do A. Q. and G. Grudnitski, 1992, "A Neural Network Approach to Residential Property Appraisal", *The Real Estate Appraiser*, 58(3): 38 – 45.
- Ethridge, D. E. and B. Davis, 1982, "Hedonic Price Estimation for Commodities: An Application to Cotton", *Western Journal of Agricultural Economics*, 7(2): 293 - 300.
- Fletcher M., Gallimore P. and J. Mangan, 2000, "The Modeling of Housing Sub-markets", *Journal of Property Investment & Finance*, 18(4).
- Freeman, M. A., 1993, *The Measurement of Environmental and Resource Values*, Washington, D.C.: Resources for the Future.
- Freeman, M.A., 1979, *The Benefits of Environmental Improvement*, Washington, D.C.: Resources for the Future.
- Frew J. and B. Wilson, 2000, "Estimation The Connection Between Location and Property Value", Essay in Honor of James A. Graaskamp, Boston, MA: Kluwer Academic Publishers.
- Frew J., and G. D. Jud, 2003, "Estimating The Value of Apartment Buildings", *The Journal of Real Estate Research*, 25(1): 77 - 86.
- Graves P., Murdoch J. C., Thayer M.A. and D. Waldman, 1988, "The Robustness of Hedonic Price Estimation – Urban Air Quality", *Land Economics*, 64: 220 – 233.
- Griliches, Z., ed., 1971, *Price Indexes and Quality Changes: Studies in New Methods of Measurement*, Cambridge, Mass: Harvard U. Press.
- Harrison, D. and D. L. Rubinfeld, 1978, "Hedonic Housing Prices and the Demand for Clean Air", *Journal of Environmental Economics and Management*, 5: 81 - 102.
- James R. C. and E. B. Carol, 2000, "Artificial Neural Networks in Accounting and Finance: Modeling Issues", *International Journal of Intelligent Systems in Accounting, Finance & Management*, 9: 119 – 144.
- Lancaster K. J., 1966, "A New Approach to Consumer Theory", *Journal of Political Economy*, 74: 132 – 157.
- Lenk M. M., Worzala E. M. and A. Silva, 1997, "High-tech Valuation: Should Artificial Neural Networks Bypass The Human Valuer?", *Journal of Property Valuation & Investment*, 15(1): 8 – 26.
- Livengood, K. R., 1983, "Value of Big Game from Markets for Hunting Leases: The Hedonic Approach", *Land Economics*, 59:3: 287-291.

- McCluskey W., 1996, "Predictive Accuracy of Machine Learning Models for The Mass Appraisal of Residential Property", *New Zealand Valuers' Journal*, July: 41 – 47.
- McMillan, M. L., B. G. Reid and D. W. Gillen, 1980, "An Extension of the Hedonic Approach for Estimating the Value of Quiet", *Land Economics*, 56(3): 315 - 328.
- Messonnier, Mark L. and E. Jane Luzar, 1990, "A Hedonic Analysis of Private Hunting Land Attributes Using an Alternative Functional Form", *Southern Journal of Agricultural Economics*, 232: 129 - 135.
- Mester, L. J., 1997, "What's The Point of Credit Scoring?", *Business Review*, Federal Reserve Bank of Philadelphia, Sep./Oct. 1997: 3 – 16.
- Mieszkowski P. and A. M. Saper, 1978, "An Estimate of the Effects of Airport Noise on Property Values", *Journal of Urban Economics*, 5: 425 – 440.
- Milon, J. Wally, Jonathan Gressel and David Mulkey, 1984, "Hedonic Amenity Valuation and Functional Form Specification", *Land Economics*, 60(4): 378 - 387.
- Owen C. and J. Howard, 1998, "Estimation Realisation Price (ERP) by Neural Networks: Forecasting Commercial Property Values", *Journal of Property Valuation & Investment*, 16(1): 71 – 86.
- Pope, C. A. and J. R. Stoll, 1985, "The Market Value of Ingress Rights for White-Tailed Deer Hunting in Texas", *Southern Journal of Agricultural Economics*, 17(1): 177 - 182.
- Pope, C. A., C. E. Adams and J. K. Thomas, 1984, "The Recreational and Aesthetic Value of Wildlife in Texas", *Journal of Leisure Research*, 16(1): 51 - 60.
- Ridker R. and J. Henning, 1982, "The Determinants of Residential Property Values with Special Reference to Air Pollution", *Review of Economics and statistics*, 46: 246 – 257.
- Rosen S., 1974, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", *Journal of Political Economics*, 82: 34 – 55.
- Shonkwiler J. S. and J. E. Reynolds, 1986, "A Note on The Use of Hedonic Price Models in The Analysis of Land Prices at The Urban Fringe", *Land Economics*, 62(1).
- Stanley M., Alastair A., Dylan M. and D. Patterson, 1998, "Neural Networks: The Prediction of Residential Values", *Journal of Property Valuation & Investment*, 16(1): 57 – 70.
- Tay D. P. H. and D. K. H. Ho, 1991, "Artificial Intelligence and The Mass Appraisal of Residential Apartments", *Journal of Property Valuation & Investment*, 10(2): 525 – 539.
- Uyeno D., Hamilton S. W. and A. J. G. Biggs, 1993, "The Density of Residential Land Use and The Impact of Airport Noise", *Journal of Transport Economics and Policy*, January, XXVII(1).
- Ward System Group Inc., 1993, NeuroShell 2, release 4.0.
- White, H., 1982, "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, 50: 1 – 26.
- Wilson, W. W., 1984, "Hedonic Prices in the Malting Barley Market", *Western Journal of Agricultural Economics*, 9(1): 29 - 40.
- Witte, A. D., H. J. Sumka and H. Erekson, 1979, "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets", *Econometrica*, 47(5): 1151 - 1173.
- Worzala E., M. Lenk and A.Silva, 1995, "An Exploration of Neural Networks and Its Application to Real Estate Valuation", *The Journal of Real Estate Research*, 10(2): 185 – 201.