Fifth Annual Pacific-Rim Real Estate Society Conference
Kuala Lumpur, Malaysia, 26-30 January 1999

Using Neural Networks to Estimate Constant Quality House Price Indices

Paul Kershaw and Peter Rossini

School of International Business,
Centre for Land Economics and Real Estate Research (CLEARER),
University of South Australia

Keywords: Constant Quality Indexing, Neural Networks, Genetic Algorithms, Time Series Analysis, Property Cycles

Abstract: This paper extends work in the field of Constant Quality Price Indexing by using Neural Networks to estimate the model. A series of housing data sets is used to develop constant quality price indices using traditional econometric techniques and using neural networks incorporating genetic algorithms. The analysis indicates that neural networks are a real alternative to the econometric methods.

Introduction

In a previous study the Local Government areas of Port Pirie, Stirling and Unley were used for some basic Constant Quality Indexing using an adjusted time series approach. In this paper the same locations are chosen for study. Cross sectional data are used involving all probable market transactions from July 1980 to June 1998. This data is analysed using a standard hedonic method utilising dummy variables to indicate time periods and also using a Neural Network model. The neural network model uses the same data, in the same format and uses Genetic Algorithms to optimize the model structure. The aim of this research is to determine is neural networks are a useful inclusion to the tools used for property price indexing.

Literature Review

The analysis of property data over time has always been fraught with difficulty. Invariably the market evidence (from transactions) involves the sale of different properties for each time period. This has always made analysis difficult because the sample of properties varies for each time period. Early analysis was based on simple statistics such as the mean or median. However this was improved with the consideration of methods to adjust for the variations in quality. Works by Bailey et al.(1963), Goodman (1978), Palmquist (1980), Nourse (1982) and Hendershott & Thibodeau (1990) helped to develop a body of methodology to consider constant quality indexes.

This came to a head at the start of the 1990's with a paper by Mankiw & Weil.(1989). In this paper the authors challenged estimates on future housing estimates by using constant quality methodologies. This brought about a flood of papers in response. Works such as those by Case et.al.(1991) Engelhardt & Poterba. (1991), Hamilton (1991), Hendershott (1991), Holland (1991) Meese & Wallace (1991). Woodward (1991) helped to develop the methodological arguments. Various studies have investigated the use of constant quality price index methodologies in the Australasian region. Research by Rossini, et al (1995), Costello & Schwann (1996) Rossini (1996a), Rossini (1996b) and Schwann (1997) have all contributed to the regional knowledge.

The opportunity for the use of Artificial Neural Networks (ANN) to estimate price indices has been investigated in recent years.

For example Borst (1991) reported the use of ANN to data sets of family residences in New England. Tay and Ho (1992, 1994) examined sets in Singapore using 833 residential apartment properties for training and tested this against 222 case set of similar apartment properties. Do and Grudnitiski (1992) used data from a

multiple listing service in California while Evans (1993) worked with residential housing in the United Kingdom. The most recent work comes from Worzala (1995), Borst (1995, 1996), McCluskey (1996a, 1996b) and Rossini (1997a, 1997b, 1997c). Rossini's research was based on data from South Australia and demonstrated that the results from artificial neural networks could potentially produce superior results to more traditional econometric models in certain circumstances.

Methodology

The research used three regions (at the local government area level) to compare the results of constant quality indexing using ordinary least squares and neural networks. The three locations chosen were those used by Rossini (1996a) which would allow for some checks of consistency of results. The three areas chosen exhibit some significant differences which makes their use attractive. Broadly they can be described as

1. Port Pirie LGA – this is the centre of Australia's major lead smelter and has been the target of a major lead decontamination project over the last decade. It also stands out as a location where there have been dramatic changes to house quality over time and a substantial amount of new housing. (This is highlighted in studies by Rossini (1996a, 1996b) and Rossini et al.(1995). Port Pirie is a regional centre of some ten thousand people.
2. Unley LGA - this is an inner city local government area bordering the City of Adelaide to the South. There is a wide variety of detached housing from small single fronted cottages to mansions on large allotments. There was a perception that parts of this area have been highly gentrified and that much of the smaller old houses, close to Adelaide's CBD, have been upgraded, renovated and extended. Any new development is limited to redevelopment and minor in-filling. The 1996 study (Rossini 1996a) showed that over time there had been minor changes to housing over the last decade and a half. This was evidenced by increases in both the average building area and building condition.
3. Stirling LGA - this location takes in parts of the Adelaide Hills. Development in this area has been limited due to water catchment and other regulations, although a substantial amount of vacant land has been built on, over the last decade. There is been very limited redevelopment and renovation in this area, and the perception of local real estate agents is that prices have "stood still or got worse" over the last few years. The past studies (Rossini 1996a) show that while Port Pirie has had a dramatic increase in house quality that Unley a smaller (but significant) change, housing quality in Stirling has remained much the same.

For each location all sales occurring in the locations from January 1985 to December 1998 which were recorded by the Lands Title Office in conjunction with the Department of Environment and Natural Resources (DENR) were examined. In each location the following sales were removed from the sample to minimise the affect of non-market transactions.

- all sales which included other land, other items with sale or were conditional transfers. (as indicated by the DENR records)
- all sales tagged as Not considered to be Market Value by DENR
- sales involving an apparent family transaction
- sales involving a Government Department or Instrumentality where an apparent non-market consideration is paid.

As a result of this analysis the following samples were obtained

Location	Total Sample	Average per quarter
Port Pirie LGA	2717	50.3
Unley LGA	7941	147.1
Stirling LGA	3927	72.7

The data was collected as individual property transactions to allow for cross sectional analysis and was then aggregated as quarterly time series. In each case the following information was collected for each quarter

- Mean and Median Price (transaction price in dollars)
- Mean Condition Code (Code 0 poor - 9 good)
- Mean Equivalent Area (weighted building area estimate)

The choice of condition code and equivalent was based results of models across a broad number of Adelaide suburbs which showed that these were consistently good hedonic indicators (Rossini,1997c).

Model Specifications

Hedonic Estimate using Dummy Variables

This is a standard hedonic approach to constant quality price indexing. The model is based on cross sectional data. The model estimates log of price as the dependent variable against an array of dummy variables for each quarter and an array of property attributes.

The property attributes used are listed in Table 1.

Table 1 - Property Characteristics used in Models

Variable	Description	Code
Land Area	Area in Hectares	Continuous
Equivalent Building Area	Calculated Equivalent Area of Buildings based on weighted average formula for main buildings and other buildings	Continuous
Condition	Scaled code from 1 - demolition to 9 - high quality new condition	Scale
Year built	Date of Construction of the Main Building	Continuous
Zone - Dummies	Series of Interactive dummy variables based on the Government Code for Zone which determines Development Controls - zero if not in the zone and land area if in the zone	Dummy
Wall Type - Dummies	Series of Interactive dummy variables - based on the external wall cladding - zero if not the wall type and building area if it is	Dummy
Building Style - Dummies	Series of Interactive dummy variables - based on building style - zero if not that style and building area if it is	Dummy

The model allows the estimate of the quarterly dummies while holding the property characteristics constant.

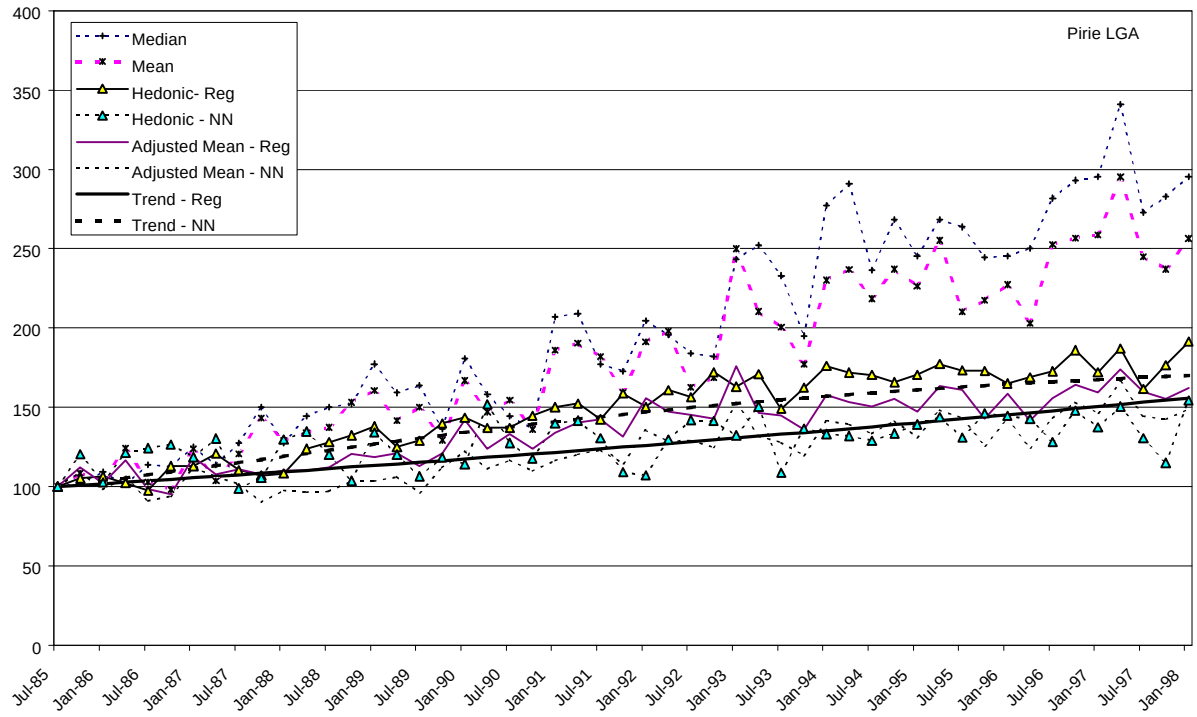
The regression model is estimated using ordinary least squares. The neural network model is estimated using a three layer model. The input layer uses the variables, which were found to be significant in the OLS model. The hidden lay uses one node for each input node. This proved to produce reasonably consistent results. The use of genetic algorithms to develop the network structure is also considered possible.

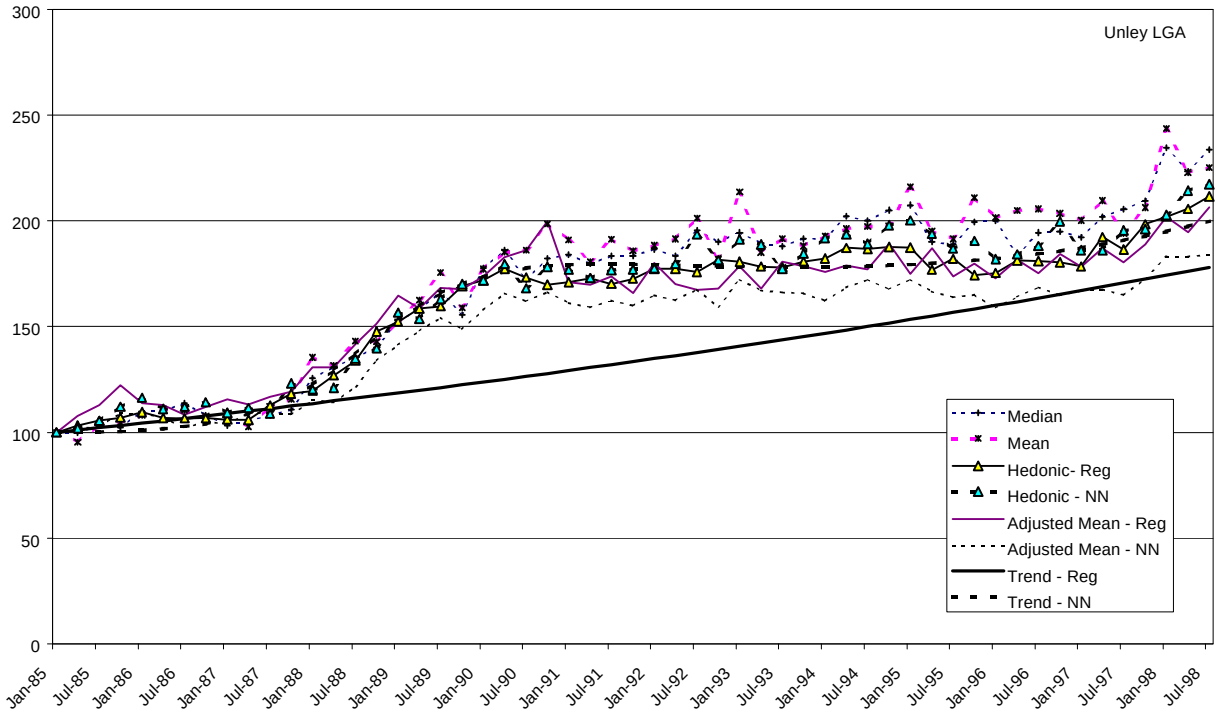
Adjusted Mean Price based on Time Series Model

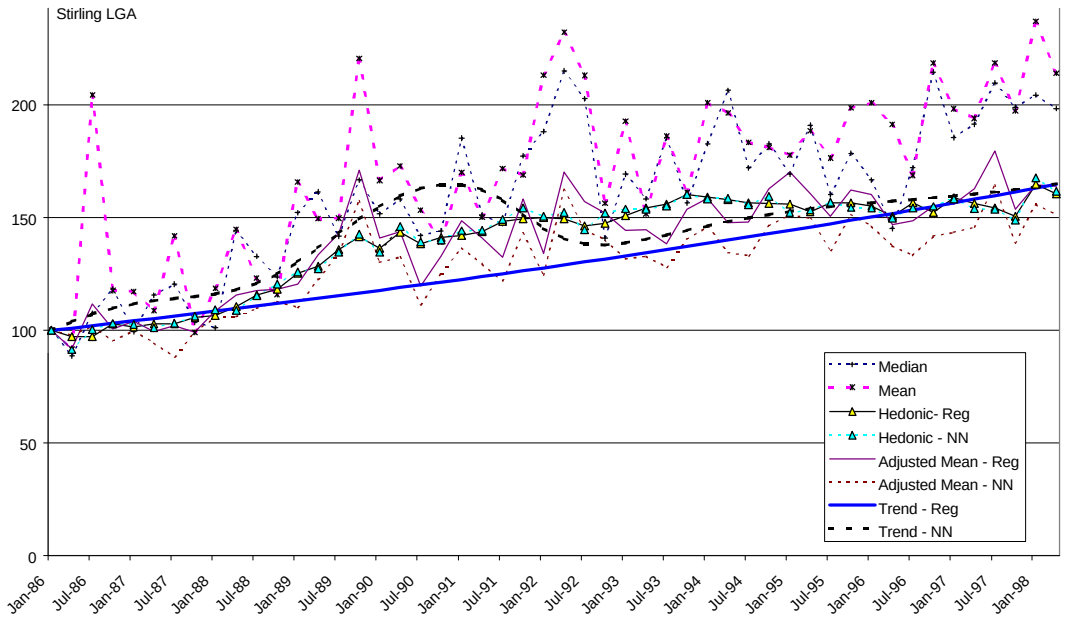
This model uses the price adjustment method adopted by Rossini (1996a). The method is based on the Grid Adjustment approach commonly used in real estate valuation. The model used is from the basic trends shown above. However in this case rather than estimate the trend, the original mean price is adjusted for differences in building area and condition producing an adjusted mean value. This ensures that the error term is incorporated into the final estimate and only the systematic component is removed.

Basic Trends

These are simple time series trends. In each case the dependent variable is Log of mean price with the independent variables, time, mean building area and mean condition. The model is estimated using ordinary least Squares regression and using neural networks with 5 layers, 3 nodes per hidden layer.


Pirie LGA
Median
Mean
Hedonic- Reg
Hedonic - NN
Adjusted Mean - Reg
Adjusted Mean - NN
Trend - Reg
Trend - NN
400
350
300
250
200
150
100
50
0
Jul-85
Jan-86
Jul-86
Jan-87
Jul-87
Jan-88
Jul-88
Jan-89
Jul-89
Jan-90
Jul-90
Jan-91
Jul-91
Jan-92
Jul-92
Jan-93
Jul-93
Jan-94
Jul-94
Jan-95
Jul-95
Jan-96
Jul-96
Jan-97
Jul-97
Jan-98


Unley LGA
Median
Mean
Hedonic- Reg
Hedonic - NN
Adjusted Mean - Reg
Adjusted Mean - NN
Trend - Reg
Trend - NN
300
250
200
150
100
50
0
Jan-85
Jul-85
Jan-86
Jul-86
Jan-87
Jul-87
Jan-88
Jul-88
Jan-89
Jul-89
Jan-90
Jul-90
Jan-91
Jul-91
Jan-92
Jul-92
Jan-93
Jul-93
Jan-94
Jul-94
Jan-95
Jul-95
Jan-96
Jul-96
Jan-97
Jul-97
Jan-98
Jul-98



Results

The unadjusted mean and median for each of the three study areas consistently over estimate the change in property values over time and have larger inter period fluctuations. This is the expected result because no account is taken of changes in building size and condition in particular. This is most clearly demonstrated in the Port Pirie data where considerable improvement to the housing stock has taken place over the period of study. This is fully described in Rossini et al. (1995).

The hedonic regression and the hedonic Neural Network analyses produce similar results in the three study areas but with a wider divergence in the Port Pirie study area where there is greater inter temporal variation in property characteristics and a smaller sample size. The inter period fluctuations are less extreme and the results indicate the superiority of the hedonic approach over the unadjusted method in removing the variation attributable to inconsistent quality attributes.

The grid adjusted mean price analyses using the regression and neural network approaches produce similar results across the three study areas. In each case the neural network approach indicates a lower indexed price. The regression method utilises a linear adjustment whilst the neural network method uses a non-linear adjustment which produces a larger adjustment and therefore a lower indexed price.

The Basic Trend analyses produces contrasting results. The regression trend shows the typical logarithmic curve whilst the neural network trend line is more undulating. This disparity is accentuated by the use of a single order relationship in the regression trend and a multi-order trend with the neural network approach. The neural network results for the Stirling LGA, where the trend line appears unrepresentative, graphically illustrates one of the inherent difficulties in using neural networks, that of over training and confirms the conclusions reached by Worzala et al. (1995).

Conclusions

The results clearly indicate the superiority of the hedonic models over the unadjusted and adjusted models. In addition it has been demonstrated that, for the chosen data sets at least, neural network analysis of hedonic models shows no significant advance on the results obtainable from hedonic regression models.

Implementation of the neural network trend using time series data provides a satisfactory proxy for the hedonic models which use cross sectional data thus reducing the data requirements and processing complexity considerably. In the case of the Unley LGA study the data requirements are reduced from 3927 cases to 52. These results justify further work being carried out on additional datasets to verify that this conclusion may be extrapolated to a general case.

References

Bailey M.J., R.F. Muth & H.O. Nourse.(1963) A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association*, 58: 933-942, December, 1963.

Borst, R.A and McCluskey (1996) The Role of Artificial Neural Networks in the Mass Appraisal of Real Estate, paper presented to the *Third European Real Estate Society Conference*, Belfast, June 26-28

Borst, R.A. (1991) Artificial Neural Networks: The Next Modelling/Calibration Technology for the Assessment Community? *Property Tax Journal*, IAAO, 10(1):69-94

Case B. and. Quigley J.M. (1991) The Dynamics of Real Estate Prices. *Review of Economics and Statistics*, 73(3), 1991.

Case B., Pollakowski H.O &. Wachter S.M.(1991) On Choosing Among House Price Index Methodologies. *AREUEA Journal*, 19(3):287-308, 1991.

Costello, G.J. & Schwann, G.M. (1996) "The Impact of Holding Periods on the Measurement of Real Estate Prices" *Pacific Rim Real Estate Society Conference*, Brisbane ,1996

Do, A.Q. and Grudnitiski, G. (1992), A Neural Network Approach to Residential Property Appraisal, *The Real Estate Appraiser,* Dec 1992:38-45

Engelhardt G.V. & J.M. Poterba. (1991) House prices and demographic change - Canadian evidence. *Regional Science and Urban Economics* 21: 539-546, 1991.

Evans, A. James,H. And Collins, A. (1993), Artificial Neural Networks: an Application to Residential Valuation in the UK, *Journal of Property Valuation & Investment:* 11:195-204

Goodman A.C.. Hedonic Prices, Price Indices and Housing Markets. (1978) *Journal of Urban Economics*, 5: 471-484, 1978.

Hamilton B.W.(1991) The baby boom, the baby bust, and the housing market - a second look. *Regional Science and Urban Economics* 21: 547-552, 1991.

Hendershott P.H. & T.G. Thibodeau.(1990) The Relationship between Median and Constant Quality House Prices: Implications for Setting FHA Loan Limits. *AREUEA Journal*, 18(3): 323-334, 1990.

Hendershott P.H. (1991) Are real house prices likely to decline by 47 percent? *Regional Science and Urban Economics* 21: 553-563, 1991.

Holland A.S. (1991) The Baby Boom and the housing market - Another look at the evidence. *Regional Science and Urban Economics* 21: 565-571, 1991.

James, H. And Lam, E, (1996) The Reliability of Artificial Neural Networks for Property Data Analysis, paper presented to the *Third European Real Estate Society Conference*, Belfast, June 26-28

Kershaw, P.J. & Rossini, P.A. (1997a) Residential Time Series Analysis - An Automated Approach, *Pacific Rim Real Estate Society Conference*, New Zealand ,1997

Mankiw N.G. & D.N. Weil.(1989) The Baby Boom, The Baby Bust, and the Housing Market. *Regional Science and Urban Economics* 19: 235-258, 1989.

McCluskey, W.(1996a) Predictive Accuracy of Machine Learning Models for Mass Appraisal of Residential Property, *New Zealand Valuer's Journal*, July:41-47

McCluskey, W., Dyson, K., McFall, D. & Anand,S. (1996b) Mass Appraisal for Property Taxation: An Artificial Intelligence Approach, *Land Economics Review*, Vol 2, No 1, 25-32

Meese R. & Wallace N.,(1991). Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices. *AREUEA Journal*, 19(3):309-331, 1991.

Nourse H.O.(1982) Single-Family House Values: A Comparative Analysis of National Indices. In C.F. Sirmans, editor, *Urban Housing Markets and Property Valuation (Research in Real Estate*, Vol.2). JAI Press, 1982

Palmquist R.B. (1980) Alternative Techniques for Developing Real Estate Price Indexes. *Review of Economics and Statistics*, 62(3): 442-448, 1980.

Rossini P.A. (1997a) Application of Artificial Neural Networks to the Valuation of Residential Property *3rd Pacific Rim Real Estate Society Conference*, New Zealand ,1997

Rossini P.A.(1997b) Artificial Neural Networks versus Multiple Regression in the Valuation of Residential Property *Australian Land Economics Review*, November 1997 Vol 3 No 1

Rossini, P.A. (1996a) "Using Constant Quality House Prices to Assess Property Market Performance" *The Valuer and Land Economist* , August 1996

Rossini, P.A. (1997c) Improving the Results of Artificial Neural Network Models for Residential Valuation, *4th Pacific Rim Real Estate Society Conference*, Perth 1998

Rossini, P.A.(1996b) "Constant Quality House Prices Are They a Better Way to Gauge House Price Movements *Pacific Rim Real Estate Society Conference*, Brisbane ,1996

Rossini, P.A., Kooymans, R.R. and Kershaw, P.J. (1995)" Constant Quality House Prices in an Australian Context - A Case Study of Port Pirie, South Australia ", *Pacific Rim Real Estate Society Conference*, Melbourne,1995

Schwann, G.M. (1997) "A Real Estate Price Index for Thin Markets" *3rd Pacific Rim Real Estate Society Conference*, New Zealand, 1997

Woodward S.E..(1991) Economists' Prejudices: Why the Mankiw-Weil story is not credible. *Regional Science and Urban Economics* 21: 531-537, 1991.

Worzala, E., Lenk, M. And Silva, (1995). An Exploration of Neural Networks and Its Application to Real Estate Valuation. *The Journal of Real Estate Research,* Vol. 10 No. 2

Paul Kershaw, Principal Lecturer - University of South Australia
School of International Business
North Terrace, Adelaide, Australia, 5000
Phone (61-8) 83020352
Fax (61-8) 83020512
E-mail paul.kershaw@unisa.edu.au