Agent-Based Modelling of House Price Evolution

Siti Amri
School of Information Technology,
Charles Sturt University
Bathurst, NSW, Australia
samri@csu.edu.au

Terry Bossomaier
School of Information Technology,
Charles Sturt University
Bathurst, NSW, Australia
tbossomaier@csu.edu.au

Abstract
Housing prices result from many factors, from government control to individual needs. Yet they seemingly take on a will of their own, booming rather than crashing and so defying all analysis and expectation. We discuss a comprehensive agent-based model of housing prices, in which real data is embedded from a variety of sources. Much of the data is spatial collected from Geographical Information Systems (GIS). The variety of different stakeholders, costs, formats and infrastructure presents considerable challenges in developing a general model. This paper describes the preliminary design work towards such a general model of the way house prices evolve. A number of initial simulations are developed in MATLAB to examine the impact of real-estate agents and fuzzy inference of the vendor's agent on house price evolution.

INTRODUCTION

House Prices are determined by the interaction of supply and demand factors [8]. These factors have a different temporal lag and may have positive or negative effects on house prices. For example, an increase in population will result in increasing demand on housing consumption and consequently will increase housing price. However, an increased construction of new houses will add to the "for sale" stock and will decrease price. These housing factors have a positive or negative feedback loop and are dynamic in nature. Historical observation of house prices indicates a cyclical trend with distinctive regional variations.

The housing market is in a perpetual state of disequilibrium due to the inelasticity of supply [8] [18]. The construction rate of new houses is relatively low proportional to the overall stock. Durability of houses lead to a housing market which is determined by existing homes. Harvey argued that house prices are less affected by supply but more demand driven. The housing market also functions inefficiently due to, amongst other factors, imperfect knowledge of buyers and sellers, the uniqueness of each site and building, and high legal and transaction costs, which prevent the asset from transferring smoothly to its most profitable use.

These market inefficiencies allow speculation of houses for short term gain [2] [13] [9]. Other researchers have also identified the speculation sentiment on house price and term it the "frenzy effect" [3] [13] [20]. This speculation sentiment explains the expectation that once house prices rise, the rise is continuing thus reinforcing the upward trend of house prices. These studies confirmed that house prices are generally too volatile to be driven by economic fundamentals alone. The varying sentiment of buyers and sellers' behaviour contribute to the volatility as well. Balchin et. al [18] summarised the housing market as "disorganised with no central buying or selling place, comprised of a vast number of housing transactions involving heterogeneous houses".

Knowledge of how house prices evolve is useful for investors, banks, institutions, and private purchasers as it allows them to make informed decisions about a purchasing strategy. This is particularly important in Australia, where real estate is the greatest component of household wealth. For most households, the family home constitutes the single most important investment and the most expensive. Being able to model the evolution of house prices has the potential to have a significant impact.

The general model of house prices is that they have been on the rise over time. However, the pattern of the rise has not been continual and includes a high number of fluctuations [4]. In general, house prices are determined by the interaction of demand and supply factors within the housing market. There are both global and local factors that impact on pricing. In addition, the housing market is driven by the expectation of buyers and vendors about future price growth [5]. Varying levels of market knowledge and cycles of buyer-vendor sentiment add further complexity to the way house prices evolve.

Because of the number of factors at different levels, producing a model of the way house prices evolve has proved extremely difficult. Using an agent-based model is an alternative approach to modelling such complex systems. Such systems have proved successful in many other diverse and complex domains such as stock market trading and petroleum exploration. In this paper we describe an agent-based model to incorporate the various temporal, spatial and human factors that impact on the evolution of house prices.

The paper reviews methodological issues on house price estimation. This paper also proposes a simulation model to capture the adapting cognitive processes of vendors, buyers, and real-estate agents and the impact of their changing sentiments upon prices. A simulation has been developed in MATLAB [14] to further investigate this aspect of the model. The implementation and results from this simulation are described before the paper concludes with a summary of outcomes and directions for future work.

METHODOLOGICAL ISSUES

There are 3 distinct groups of methodologies to calculate house prices:

- Statistical Methods
- Preference Methods
- Artificial Intelligence Methods

Statistical Methods

Of the statistical techniques, Multiple Regression Analysis is by far the most commonly used method to forecast house prices. In this method, the dependent variable, house price, is regressed against a set of predictors (or independent variables), as illustrated in the equation below. These predictors are comprised of variables such the housing structure (size and features) and location.

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \ldots + \beta_n X_n + \mu$$

where:

Y = dependent variable

$\beta_1, \beta_2, ..., \beta_n$ = unknown parameters to be estimated

$X_1, X_2, ..., X_n$ = independent variables

μ = error term

In many studies, past values of house prices are used to predict its future price [7] [19]. A further extension of this model created market segments based on spatial factors [17][19]. These spatial factors adjusted prices for different areas. This allows different regression functions to be computed for each of the market segments.

Another extension of this method is to apply factor analysis, principal component analysis (PCA) or multidimensional scaling (MDS) to segment the data. The data with reduced dimensions are then grouped into submarkets using cluster analysis [19].

Critics of these statistical methods argue against the generalisation of aggregate price indices from regression results [7] [10]. This model assumed a uniform appreciation rate for all housing units in a given locality and they claimed that this contradicted with their own findings. Seward et. al found that higher price housing appreciates at a more rapid rate than low and medium price housing during a growth phase yet there is no difference in a down trend. The difference in the market behaviour of the housing unit is lost in the regression approach due to the averaging effects over available data. The question of sample bias has been raised in the way regression techniques use previous sales prices [7]. Jud and Seales determined that the sample included only homes that were sold and not all the houses in a given area. Thus, in an economic downturn, bias is largest when fewer houses are sold [7]. The dependence of this technique on secondary datasets is also a problem because variable selection tends to be based on what is available rather than on what is sensible [19].

Kauko further questioned the use of a partitioning technique, explaining that "heterogeneity of dwellings and households can imply not necessarily the existence of several submarkets, but rather the same diversified markets". This argument raises doubt about whether a locally partitioned dataset with multiple regression functions or a single hedonic price function is more appropriate for forming house price estimation.

The strongest counterargument against this technique is that it ignores the importance of household and institutional preferences [6] [19]. Parker et. al [6] criticised this technique as it failed to represent feedbacks and capture a dynamic process. They argued that this top-down modelling approach is only suitable when analysing with a very coarse temporal and spatial resolution. If there is strong local heterogeneity and interactions then this type of modelling is not considered appropriate.

Preference Methods

The second group of modelling techniques on house forecasting is focusing not on prices but on housing preference. This stream of research activity includes conventional stated preferences (interview surveys, experimental choice design, contingent valuation) and analytic decision tools. These studies generate house price estimates associated with value differences from information collected through interview. The argument against this method is that it is based on hypothetical rather than actual behaviour and reflects consumer tastes and not house price formation [19].

Artificial Intelligence Methods

The final group of methods, from Artificial Intelligence (AI), has recently emerged offering a simple approach to simulating market behaviour and house price estimation. These approaches have qualitative underpinnings and simulate market behaviour with the aim not to generate perfect accuracy but to simulate judgemen [11] [19].

In the house price application, Rossini [15] [16] and Kauko [19] both employed an artificial neural networks (ANN) together in comparison with multiple regression analysis (MRA). Kauko provides an extensive review of ANN use in house price research and discusses the issues surrounding the relative poor "take up" of this technique. The consensus from these researchers is the results from ANN are superior to the MRA. However, Rossini, who applies ANN to develop a mass appraisal valuation of property in South Australia, finds that ANN produces superior results when the data set is small.

In summary, the MRA technique requires the exclusion of outliers and homogeneous data to prevent large error margins and to trace average dependencies respectively. It assumes equilibrium and linear data trend. The ANN technique used together with the genetic algorithm provides an excellent data grouping but assumes rational choice with perfect knowledge.

Existing methodologies of calculating house prices assumes rational choice, complete knowledge, and optimal solution. These models often fail to predict and explain the reality of the housing system. The nature of housing is influenced not only by its location and associated structural attributes but also by its relative value reflected by individual or institutional preference. These individuals and institutions have incomplete knowledge of the market and are boundedly rational i.e. they have evolving choices that move toward achieving their goals.

THE AGENT-BASED MODEL

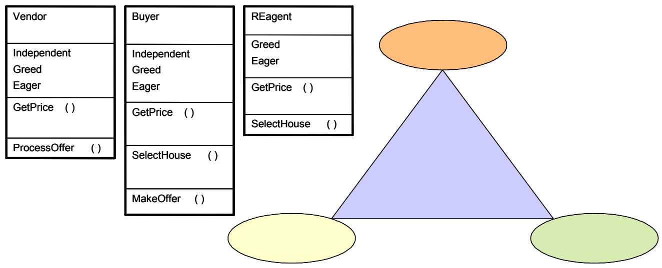
The housing system can be viewed as an evolving complex system. Supply and demand factors may have linear or non-linearity tendency. Furthermore, in the urban areas the housing system can generate surprising complexity and often ordered pattern in aggregate form. These emergence and self-organisation phenomena are reflected in the pricing cycle and in the spatial clustering of socio-economic groups by housing markets in the urban landscape. In the complex system framework, house price distribution and dynamics are the result of interactions from multiple independent entities or agents, which may be individuals (buyers, sellers, real-estate agents) or institutions (banks, property developers, government).

The agent-based modelling approach is a technique derived from complexity. It captures non-linearities and emergent behaviour from cycle to chaos. It represents the housing economics as a decentralised collection of interacting, adaptive, autonomous agents with no global control. It focuses on the simulation of agents' behaviour and how they interact. This research proposes the construction of the computational model of the housing system. This model will be used as a test bed to examine alternative agents' traits eg. pricing strategies. This computational model enables simulation of the house price evolution and the testing of its processes. Validation of model results will be compared with existing house price data.

1.1 The Agents

There are 3 types of agents considered in this study; vendor, buyers and real-estate agents (Figure 1). These agents have imperfect knowledge of the housing market and may rely on information from past decisions, their own and those of other agents to update decision-making strategies. This process leads to temporal interdependencies among agents. These decisions will likely have temporal dynamic impacts on the house price distribution and market environment.

Figure 1: Agents Relationship



In this study, the agents have simple adaptive traits for speculating house price strategies. Each agent constructs a probability measure of house price based on their knowledge of housing factors and level of access to sales history information. In this study, the varying level of knowledge is represented by the different methods of estimating house prices employed by the agents. The vendor and buyer agents adapt their pricing and transaction strategies according to their personal attributes. In this initial set-up the agent strategy is reflected by how it speculates the house price estimate. Also, house price is determined by the cost of acquiring the house plus the speculation of its appreciation over time. The price function of the agents has to ensure that the vendors or real estate agent never sells a house for less than it paid to acquire it.

The vendor agent perceives the house estimate through a limited scope of comparing recent property sales in their neighbourhood. The vendor deduces its house projection based on a subset of previous sale records. In this example, it is a list comprising of 20 houses which make up the recent sales. The vendor then locates the 2 nearest houses and calculates the sale mean value of the prices. This value will be the vendor's perception of its house value and constitutes the asking price.

$$Dist(x) = \sum_{i=1}^{n} (r_1 - r_n)^2 + (c_1 - c_n)^2$$

$$P_v = \frac{P_{\min_1} + P_{\min_2}}{2}$$

$where:$

P_v = Vendor's price

$P_{\min_1}$ = Price at 1st closest

$P_{\min_2}$ = Price at 2nd closest

$r_1, c_1 \ldots r_n, c_n$ = House locations

The real estate agent is assumed to have a better knowledge of the sale price variations of a given area [12]. The real estate agent constructs an estimate of house prices in a current year based on a comprehensive list of sold prices using the following equation:

$$P_{y_c} = \sum_{i=1}^{n} P_{y_s}(1+r)^{y_c - y_s}$$

$where:$

P = house price

y_c = the year current

y_s = the year sold

r = inflation rate

When a vendor approaches a real estate agent for a valuation on a property, the following steps are executed

1. Start with vendor's house location and search recent sales record;
2. Calculate the distance of sold houses to vendor's location, sort the distance table;
3. Take the first n values(n=8); The estimate is sum of the new list divided by n

The buyer agent is by far the simplest in terms of estimating their offer price. Buyer uses the same recent sale record as the vendor agent but employs the median price to set its price range. Buyer bid should be as low as possible without failing to win the house. If a bid fails, the buyer agent increases the price it will bid at the next offer. Buyer preference to particular neighbourhood and associated amenities constrain their housing search behaviour. This behaviour is encapsulated in buyer's eagerness index randomly assigned to each of the buyer agent. Buyer agent behaviour is characterised by an upgrading of their housing situation, thus stimulating the higher quality housing market. This behaviour is set by its choice of houses with high desirability index. The second attribute of the buyer is confidence.

$$P_b = P_{median}(b_{eager}(i)+1)$$

$where:$

P_b = Buyer Price

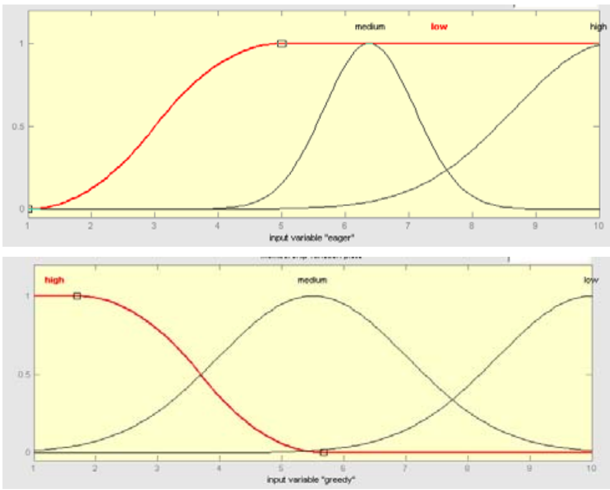
P_{median} = Median Price

b_{eager} = Buyer's eagerness index between 0 and 1

The vendor is given 3 personal attributes to determine their behaviour and interactions with other agents (independent, greed, and eager). The vendor with high independent values will conduct the sale of their house directly to the buyer agent. If the vendor independent attribute is less than the given threshold values than the vendor will employ the real estate agent to provide a valuation of the vendor's property.

The vendor's decision to accept or reject buyer's offer employs the fuzzy logic to map input of the vendor's attribute (*eager* and *greed*) to the decision space (*accept* or *reject*). Eager has a positive relationship with accept while greed has an inverse relationship with accept. Membership function of the vendor's attribute follows the asymmetrical polynomial function for greed input and sigmoidal function for eager input (Figure 2) [14].

Figure 2: Membership Functions



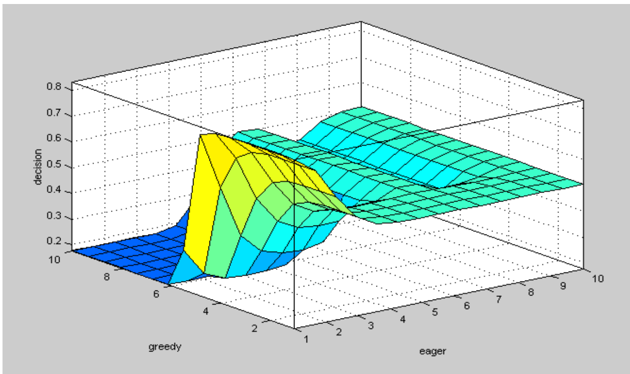
The fuzzy relationship is summarised in Figure 3. The decision rules for determining the vendor's decision is as follows:

If (*eager* is low) or (*greed* is high) then decision is reject

if *(eager* is medium) the decision is accept

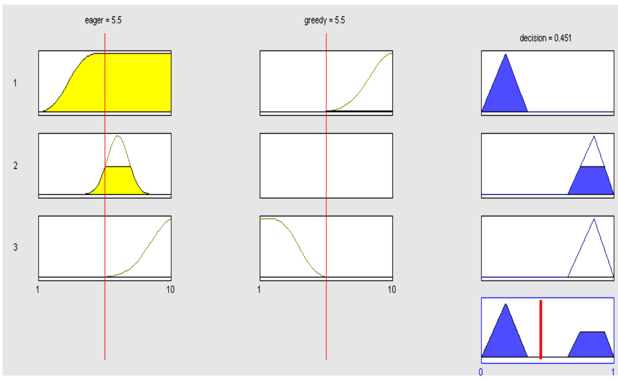
if (*eager* is high) or (*greed* is low) the decision is accept

Figure 3: Fuzzy Logic Decision Space



The defuzzification method is a decision set represented by a number between 0 and 1 (Figure 4). A threshold value is set to determine the final decision to accept or reject.

Figure 4: Defuzzification Method



SIMULATING HOUSE PRICES

In this study an artificial environment of 10x10 grid is selected to provide the model framework. In future developments, this grid will be replaced with the actual cadastre data of the study area. Each grid will correspond to the parcel boundary and will have specific structural and locational attributes.

In this simulation, each cell is assigned initial price and a random desirability value.

$$P = 1000(d^2 + 1)$$

where:

P = House Price

d = desirability index between 0 and 1

The desirability index encapsulates the neighbourhood quality or some special features of the house such as view or aspect. The total number of houses in the simulation is 100. The number of houses for sale is fixed at 10 and the same number is set for buyer and vendor agents. These initial settings provide a framework to test dynamics of house price through agents' interactions.

The problem in this study is considered as a dynamic housing simulation. Given a set of buyers-vendors, all agents fixed with initial house estimates, the task is to find speculation strategies for each buyer/vendor to optimize their housing options.

The model works as follows

Initialization:

- each cell is given a random value of house price
- each vendor agent sets an initial price
- each buyer agent sets an initial price
- set number of buyer, vendors, house for sale
- set speculation rate, r

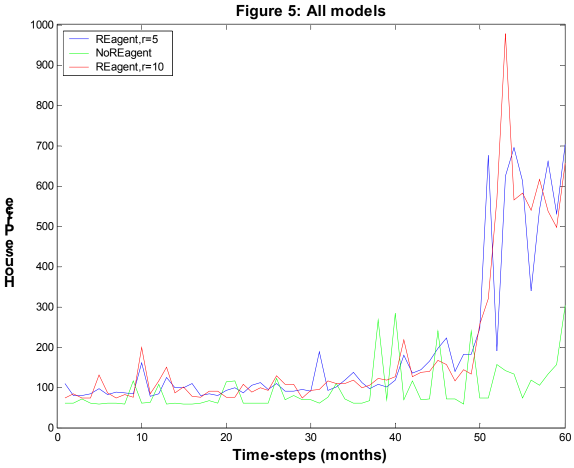
one time step:

Run the simulation for 12 time steps and until all buyer list is exhausted. If house is sold, another house is randomly added to the list for sale. Each buyer will have a fixed set of 10 houses to select from. Buyer and vendor agents are given an independent value which determines whether they will negotiate directly with one another or use the real estate agent to mediate. In the current setting the agents have varying levels of memory, to simplify the model implementation. Later developments will increase agent intelligence.

DISCUSSION

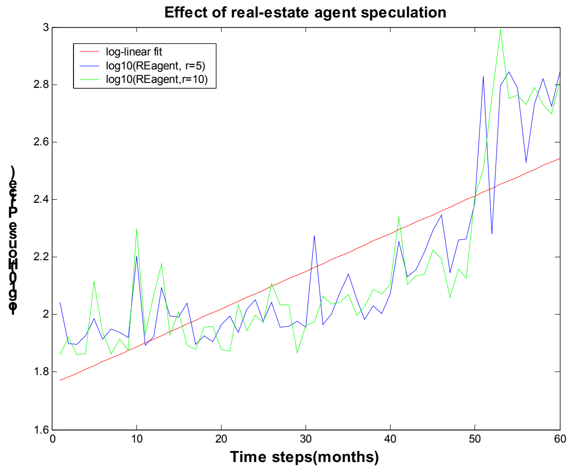
In this study, 3 models of agent pricing strategies are examined. In model 1, the housing simulation has no real-estate agent and buyers and vendors agents conduct direct transaction in the sale process. Buyers and vendors have no speculation strategy in their house price estimate. In model 2, the vendors and buyers' independence attributes are integrated in the pricing and house selection decisions. In this model, the vendors and buyers are assigned a confidence attribute randomly between 0 and 1. In model 2, vendors and buyers with confidence less than .50 per cent will rely on the real-estate agent to calculate the selling price and select house respectively. Model 2 inclusion of the real-estate agent adds a simple speculation element in the house price evolution. In model 2, the real-estate agent sets a uniform price increase of 5 per cent. Model 3 is similar to model 2 but with a speculation rate of 10 per cent. Figure 5 summarizes the results of the 3 models

Figure 5: All models



Using a log-linear plot, the speculative inflation attempted by the real-estate agent shows a straight line. The simulation shows that inflation occurs at 5 per cent level but does not follow the 10 per cent speculation so well.

Figure 6: Logarithm Analysis



OUTCOMES AND FUTURE WORK

There are various applications of agent-based simulation ranging from environmental system, financial markets to social systems. Agent-based modelling is unique because of the focus on behaviour of adaptive agents. In this project, research questions on how agents form their house price speculations and whether or not asset prices are predictable are explored. Initially, the agents deduce house price estimates based on their information processing of housing factors. The agents adapt their property pricing and transaction strategies based on their personal traits. In this model, the interpretation of the emergent behaviour may lead to new strategies and tactics that will benefit household in making housing investment decision.

The housing market has been characterized by strong cyclical and regional price variations. Previous housing studies confirmed that the housing market is not only driven by economic fundamentals but also by individual speculation. The motivation behind this study is to understand house price dynamics and its associated spatial distribution in the study area. This research has constructed buyer-seller behavior to investigate agent-based modeling approach in simulation of house price evolution. Comparative analysis of existing housing models and the agent-based approach will be carried out. Finally, extension of this study will integrate the model results with Geographic Information System (GIS) for the purpose of capturing the spatial dynamics of house price variations.

ACKNOWLEDGMENTS

This project was sponsored by Australian Research Council Linkage Grant no. LP0212023

REFERENCES

[1] A. Karantonis, *An Empirical Study of Demand Factors Affecting the Segmented Sydney Housing Market*, 7 th Pacific Rim Real Estate Society Conference, Brisbane 2003

[2] A.M.Zarestsky, *Asset Prices and Market Speculation*, Regional Economist, http://www.stls.frb.org/publications/re/1999/b/re1999b3.html, 1999

[3] B. Badcock, *Snakes or Ladders? The Housing Market and Wealth Distribution in Australia,* International Journal of Urban and Regional Research, 18(4):609-627

[4] Bized, *House Price Simulation*,

http://www.bized.ac.uk/virtual/vla/house_prrices/resource_pack, 2003

[5] Charter Keck Cramer, *Housing Gets That Cyclical Feeling*, http://www.charterkc.com.au, 2003

[6] D. C. Parker, S.T. Manson, M.A. Janssen, M.J. Hoffmann, and P. Deadman, *Multi-Agent Systems for the Simulations of Land-Use and Land-Cover Change: A Review*, Annals of the Association of American Geographers,2002

[7] G.D. Jud and T.G. Seaks, *Sample Selection Bias in Estimating Housing Sales Prices*, The Journal of Real Estate Research,9(3):289-298,1994

[8] J. Harvey, Urban Land Economics 4th Edition, Macmillan Press Ltd., 1996

[9] J. Muellbauer and A. Murphy, *Booms and Busts in the UK Housing Market,* Economic Journal, 107(445):1701-1727,1997

[10] J.A. Seward, C.J. Delaney, and M.T. Smith, *An Empirical Analysis of Housing Price Appreciation in a Market Stratified by Size and Value of the Housing Stock*, The Journal of Real Estate Research,7(2):195-205,1992

[11] K. Nagel, *Truly Agent-Based Strategy Selection For Transportation Simulations* http://www.sim.inf.ethz.ch/papers/agdb/html/agdb.html , 2002

[12] L. V. Zumpano, K.H.Johnson and R.T. Anderson, *Internet Use and Real Estate Brokerage Market Intermediaries*, Journal of Housing Economics,12:134-150,2003

[13] M. Riddel, *Fundamental, Feedback Trading, and Housing Market Speculation: Evidence from California*, Journal of Housing Economics, 8:272-284,1999

[14] MATLAB, http://www.mathworks.com

[15] P. Kershaw and P. A. Rossini, *Using Neural Networks to Estimate Constant Quality House Price Indices* ,5th Pacific Rim Real Estate Society Conference, Kuala Lumpur,1999

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

[17] P.A. Waddell, A. Borning, M. Noth, N. Freir, M. Becke, and G. Ulfarsson, *UrbanSim Reference Guide*, version 0.9, http://www.urbansim.org, 2003

[18] P.N. Balchin, G.H. Bull and J.L. Kieve, Urban Land Economics and Public Policy 5th edition, Macmillan Building and Surveying Series, 1995

[19] T.J. Kauko, *Modelling the Locational Determinants of House Prices: Neural Networks and Value Tree Approaches*, PhD thesis, University Utrecht, http://www.library.uu.nl/digiarchief/dip/diss/2002-1204-091756/inhoud.htm

[20] Yates, J. *A Spatial Analysis of Trends in Housing Markets and Changing Patterns of Household Structure and Income*, Australian Housing and Urban Research Institute (AHURI), http://www.ahuri.gov.au,2002