Simulation of House Prices for Improved Land Valuation

Terry Bossomaier[(a)], Md Zahidul Islam[(a,b)], Rod Duncan[(a,c)], Junbin Gao[(a,d)]
[(a)]Centre for Research in Complex Systems (CRiCS)
Charles Sturt University, Australia
[(a)]Email: tbossomaier@csu.edu.au [(b)]Email: zislam@csu.edu.au
[(c)]Email:rduncan@csu.edu.au [(d)]Email: jbgao@csu.edu.au

***ABSTRACT*—The housing market is an important part of most economies in the developed world, but is also a significant factor in setting land valuation. Thus good housing models serve multiple purposes. Many studies use powerful statistical techniques to study pricing, but these are not so effective for handling large scale social or demographic shifts. Agent based modelling (ABM) is more flexible and therefore this paper describes a housing market simulation using ABM. A standard economic model is used for estimating utility, combined with a second price auction model and a decision forest for linking house features to price estimates at the beginning. Simulations are presented for a range of market temperatures, revealing different levels of price inflation.**

***Index Terms*—house price, agent based model, decision tree**

1.. INTRODUCTION

One of the challenges facing government is the accurate valuation of land parcels, for a variety of purposes, including taxation. Many factors go into valuation, but previous sale values are of paramount importance. In some ways the land may itself become more or less valuable. It may be rezoned, allowing for development, and thus increase in value. On the other hand it may become polluted or some other environmental hasard, previously not recognised may become apparent. In this case the value would go down. In some situations the value may change as a result of change of use. If a petrol station is built on top of a piece of land where a petrol station had already existed, then the land value may not change much. But if the land, were to become some other sort of retail outlet, then contamination by fuel would decrease its value relative to another adjacent parcel which had not been exposed to fuel. Sometimes external factors such as a change of flight route over a land and development of new areas close by can also change the land price.

Accurate valuation is important for equity and general satisfaction with the process. But errors may lead to costly legal challenges, making accuracy financially desirable. This project is the first stage in a project to reduce errors in land valuation in NSW, with a particular focus on residential areas. Before tackling the complexity of the major cities, such as Sydney, we focus on a medium size (by Australian standards) regional city of around 40,000 inhabitants, Bathurst.

The Financial Review ranked suburbs and towns for housing retail price growth over the last three years. A wide variation from housing stagnation or decline to substantial growth is seen, sometimes with quite large variations in areas which are close together and similar in many respects. Bathurst came in the top 20 with an aggregate growth of 14.8% and the Council has provided support for the project. The city has several interesting dynamic features:

1) It has a significant manufacturing base, with nearly one third of jobs in manufacturing, which creates several residential foci. It also adds a special link to the performance of the economy in general.
2) It is 35km from the next nearest town, thus has a lot of land for potential development and the council strategic plan includes new areas zoned for development
3) It is Australia's oldest inland city, with a lot of distinctive old housing which is attractive to some buyers
4) It has a number of big schools of state-wide significance, also creating foci for residential development.

Accurate house price prediction is, of course, still an unsolved problem, presenting numerous difficulties. It is especially difficult because it operates at several different levels and timescales. At the coarsest timescale, the Economist regularly measures house price statistics across the developed world, using the ratio of capital investment to rental returns as an indicator. By such measures some areas, such as Hong Kong, and including Australia, have grossly overvalued housing. Outside of major financial catastrophes such indicators may imply long term stagnation or even decline in house prices.

Of more immediate concern are the changes in macro-economic conditions: jobs; interest rates and financing opportunities; and trends in consumer spending and taste (such as the balance between apartments and houses). Closely coupled to these factors is the availability of land for new development.

Lastly, there are factors relating to individual attributes of the houses themselves. Some are relatively static, such as proximity to schools, busy roads, while some can change rapidly as a result of owner investment in the property, or lack of it – some houses in Sydney, for example, are bought for the land only and allowed to become derelict.

These different considerations and scales have led to a very extensive range of modelling approaches. Statistical methods of many kinds (McMillen 2008) are popular at all levels, while artificial intelligence techniques, such as fuzzy logic (Lughofer, Trawiński, Trawiński, Kempa & Lasota 2011), have proved useful at the individual house level. Simulation using *Agent Based Models* (ABMs) are less widely used, but have two big advantages: they are capable of handling many disparate factors and timescales; and with suitable changes in parametrisation can easily be applied to different towns, cities and countries. Since they are intrinsically parallel, they can be scaled to very large numbers of agents with distributed computing.

The ABM comprises two parts: the housing market, comprising land availability, macro-economic factors and buyer/seller populations; and the models for buyer preferences. Magliocca et al. (Magliocca, McConnell, Walls & Safirova 2012, Magliocca, Safirova, McConnell & Walls 2011) provide a strong framework for the first of these, but there is no agreement on the best model for buyer choice. In fact the best data for buyer choice comes from statistical models and is not readily incorporated into an agent based model.

House prices can be estimated using a Fuzzy Delphi method where a number of experts are asked separately to estimate a house price based on a set of price factors such as green areas, seashore and grave yard (Damigos & Anyfantis 2011). An average of all the estimated prices of the property can then be used to re-estimate the price in a recursive manner until all individual estimates become very close to the average. The final average can then be considered as the estimated price. In order to establish a relationship between a set of price factors and the price of a property a group of buyers can be surveyed where they estimate the price of a property and also assign scores for various price factors of the property (Kusan, Aytekin & Ozdemir 2010). Records having scores on various price factors and an estimated price can be used as a training data set in order to build a classifier which is then used to estimate prices of other properties.

Determining buyer preferences is different. Hedonic modelling has been quite successful here (Bourassa, Hoesli & Zhang 2011), but one of the most interesting findings comes from quantile statistics. Quantile regression suits property valuation for a number of reasons. It produces a number of regression models supporting a decision maker with the ability to use alternate models in order to make a more accurate valuation (Narula, Wellington & Lewis 2012). Moreover, it transpires (Zietz, Zietz & Sirmans 2008) that housing preferences are dependent upon relative house value. So the size of plot of land and the number of bathrooms are more significant in higher priced homes. But the newness of the building is more significant in lower priced homes. Factors, such as commuting distance, are less dependent on value. Furthermore the distributions seem to vary according to locality. These results pertain to houses in the State of Utah in the USA and may not generalise to other parts of the world such as Australia. Thus the first stage of the Bathurst model is the quantile modelling of house price factors.

Given the impact of these different factors on house prices, we plan a soft computing model for agent behaviour. Fuzzy logic is a good choice here, since it allows linguistic expression of rules in a natural way. It has been used in a variety of housing studies. Together with the preference factors for the house and its immediate spatial location, macro-economic factors such as interest rates and limitations on loan amounts, general economic conditions and propensity to take financial risks in such conditions produce a set of fuzzy rules with a desirability value, D_n resulting for any given house. For the present paper a simpler approach is adopted, comparing a house feature vector with desired features based on socio-economic category (section 2.4.).

Given the buyer and seller properties for existing houses, new developments now have to be considered. The Magliocca et al. (Magliocca et al. 2011) model consists of farmer, f_i, developer, d_i and consumer agents, c_i, with effectively three markets: farmers selling to developers; and developers selling to consumers; and consumers selling to each other. However, the development market may be limited not only by the willingness of farmers to sell, but also council restrictions on the rezoning of land for housing development. Thus we restrict this first model to two markets: developer and consumer.

The market then uses a Cobb-Douglas model (Magliocca et al. 2011) to determine the *utility*, U for each house, n

$$U(c_i, n) = (I_i - P_{ask})_i^{\alpha} D_i^{\beta} B_i^{\gamma} \tag{1}$$

where P is the asking price, and B the block size. From this the maximum price the consumer will pay is P_{max} is

$$P_{max} = \frac{I}{\beta_i + \gamma_i} \tag{2}$$

The offer price on any house will then be $\frac{P_{max} U}{U_{max}}$

To make the income and asking price of the same order, we use the annual payments on the property based on a fixed, simple (effective) interest rate, η, a non-variable mortgage term of M years, with M set to 20.

$$P = H(\eta + \frac{1}{M}) \tag{3}$$

where H is the actual price.

The maximum utility for each agent is set relative to *all* houses. Thus most of the time the available houses will have a lower utility and an agent will offer below their maximum price.

2.. THE FULL MODEL

In the first stage of the model our focus is on testing the underlying dynamics and thus several simplifications were made:

- Only the market for exisisng houses is modelled; there is a pool of buyers which is larger than the number of houses (representing incoming buyers and investors). The tricky issue of how to set rents is thus obviated. Thus each agent in the population may be either a buyer, seller or both.
- All housing factors (e.g. interior, proximity to non-housing spatial entities, such as schools, and shopping centres) stay constant.
- Each vendor/buyer has a separate preference, and family status that do not change during the simulation.
- The model is a small town – commuting costs are negligible.

All of these extensions will be introduced later. The software written is sufficiently flexible to embrace them easily.

A. The Market Process

In a typical housing market, sales occur by either private treaty or by auction. There is a considerable body of theory, on such processes, which allows some simplification.

Auctions have existed in many possible formats in different industries, countries and eras. Auctions may be first-price or second-price, where the winner of the auction pays respectively the highest bid or the second-highest bid submitted. Auctions may be sealed bid, where the buyer's bid is kept private, or open bid, where bids are public information. Open auctions may be English auctions, where bidders make successively higher bids, or Dutch auctions, where the auctioneer announces successively lower prices until a bidder enters a bid. For a general set of assumptions, the Theory of Revenue Equivalence stated first by Vickrey (Vickrey 1961) and then developed more formally by Myerson (Myerson 1981) establishes that the seller's expected revenue, and so also the buyer's expected price, will be independent of the design of the auction.

The following assumptions would satisfy the Theory of Revenue Equivalence.

1) The houses are auctioned individually.
2) The winner of the auction is the highest bidder. Only the winning bidder need contribute to the purchase.
3) The buyers are risk-neutral, and buyers' valuations are independently distributed and are known only to themselves.

If these assumptions are satisfied, then the expected sales price for the house will be the second-highest valuation of the buyers participating in the auction. This result holds true for all of the forms of auctions listed above.

Thus we adopt a single mechanism here which is a blind, second price auction. Each buyer makes an unseen bid. The buyer who makes the greatest offer gets the house, but pays the price of the *second-highest* bidder.

B. Overarching Framework

Every house put up for sale at each timestep gets auctioned once. If it does not sell, it is held over to the next week and re-auctioned.

1) A random number of vendors and buyers are added to the market at each timestep. Two parameters, the *market temperature,* ξ, and the *housing demand,* ζ control the number of sellers and buyers respectively. Thus there can be a glut of houses with no buyers, (high ξ, low ζ), a housing shortage (low ξ, high ζ) or a boom or

bust where both are high or low respectively. For the first simulations we set $\xi = \zeta$.
Vendors set a selling price. A buyer can also be a vendor. We make the implicit assumption that if someone buys another house without selling the first, they are acting as an investor.

2) Each buyer finds the house on the market with maximum utility, U, which determines their willingness to pay, W.
3) A second-price auction is run on each house, the ownership changes hands and vendor, buyer and house disappear from the market.

C. Houses

Houses will in the full model have a set of, say, H parameters, number of bedrooms, proximity to school etc, which will be determined from real data. To begin with we set up a random distribution of some kind. All houses have their price initialised at the start, as discussed in section 2.7.

Each parameter has a desirability value, d_i, between 0 and 1, formed as either a ratio of, say, the number of bedrooms, to the maximum number, or a qualitative value representing view, proximity to school and so on. Hedonic studies such as the one presented by Bourassa et al. (Bourassa et al. 2011) provide a starting point for the scope of such parameters.

D. Owners

All owners decide whether to buy or sell at random according to the market temperature. There is no refractory period on how soon an agent may re-enter the housing market after a sale.

Each buyer has their own values of α, β, and γ. But the calculation of the desirability of a house is formed as a dot product between the desirability vector, d_n and the weight vector w_c applied to it based upon the buyers social category. Current categories in use: families; single income, no kids (SINK); double income no kids (DINK); and retiree. For a family the weighting vector would emphasise number of bathrooms over say number of garages. The weight vector we use in our simulation is shown in Table I. A buyer belonging to the Family category has weight on House Size equal to 0.7 and weight on Mountain View equal to 0.1 meaning that they give more emphasise on house size than mountain view.

E. Vendors

Vendors set a reserve price based on what they paid for on the house to begin with (or its value at the start of the simulation). We could set a fixed increase percentage (and watch what happens as we change this), or we could measure the current inflation/deflation rate, or median value increase and produce an estimate from this. They have to sell if the reserve price is met.

F. Buyers

At each timestep each buyer bids on one house (the one with the maximum utility *of the houses on the market*). The buyer weighting vector and the house properties are combined to produce a desirability value (maybe just a weighted average).

G. Decision Tree Estimation of Initial Asking Price (Reserve Price)

In order to determine the utility function of a property a buyer uses the Cobb-Douglas equation as mentioned before in Eqn. 1. The buyer needs to know the asking price $P_{ask|n}$ of a property n in order to estimate the utility of the property for him/her. The vendors set a reserve price of the property based on what they paid. They also estimate the asking price based on their experience on the current market situation. That is, they make an estimate of their property price based on the recent sale prices of similar properties. In this study we consider that the reserve price and asking price of a property are same. We also assume that the reserve price of a property is a public knowledge as people know the recent sale prices of the similar properties.

In order to simulate the experience based asking price (reserve price) estimation by a vendor agent, we consider that we have access to a data set having various information on the properties and their recent sale prices. The data set can be considered as a two dimensional table where the records represent the properties and the columns (attributes) represent various information of the properties. One of the attributes is the Sale Price of a property and the other attributes are on various information such as Lot Size, Floor Size and Number of Bed Rooms. The attribute representing the sale price is considered as the label or class attribute, and all other attributes are considered as non-class attributes.

We then can build a decision tree (Quinlan 1993, Quinlan 1996, Islam 2012) or a decision forest (i.e. a set of decision trees) (Islam & Giggins 2011, Abellan & Masegosa 2009) from the data set in order to explore various logic rules for predicting/estimating the price of a property that we are interested in.

In this study we first generate a synthetic data set and then build a decision forest from the data set in order to learn various logic rules. The data set has 2000 records and 10 attributes, out of which nine are non-class attributes and the remaining one

Category	House Size	Year	Bedroom (number)	Bathroom (number)	Garage (number)	Landscape	Sprinkler	Mountain View
1	2	3	4	5	6	7	8	9
Family	0.7	0.3	0.9	0.9	0.5	0.1	0.2	0.1
Single Income	0.3	0.5	0.5	0.1	0.2	0.1	0.1	0.1
Double Income (no kids)	0.5	0.6	0.6	0.6	0.5	0.3	0.2	0.1
Retiree	0.1	0.9	0.3	0.1	0.3	0.7	0.9	0.3

TABLE I
WEIGHT VECTORS FOR DIFFERENT BUYER CATEGORIES

House Price ($1000)	Dist. (%)	House Size (sqm)	Lot Size (sqm)	Year	Bed	Bath	Garage	Land-scape	Sprink-ler	Mount. View
1	2	3	4	5	6	7	8	9	10	11
50-100	5	45-90	90-120	1960-2010	1	1	0	0	0	0
100-200	10	80-170	100-400	1960-2010	1-2	1	0	0-1	0	0
200-300	35	150-220	350-650	1980-2012	2-4	1-2	1	0-1	0	0
300-400	25	200-280	500-1000	1995-2012	3-5	1-3	1-2	1	0-1 (40%)	0-1 (40%)
400-500	15	250-390	700-1000	1990-2012	4-6	2-3	2-3	1	0-1 (70%)	0-1 (40%)
500-600	5	350-450	800-1000	1985-2012	4-6	2-3	2-3	1	1	0-1 (60%)
600-700	2	420-640	700-1200	1980-2012	5-7	3-4	2-3	1	1	0-1 (70%)
700-800	1	600-700	800-1000	1980-2012	5-7	3-4	2-4	1	1	0-1 (80%)
800-900	1	650-750	900-1200	1980-2012	5-7	4-5	2-4	1	1	0-1 (80%)
900-1000	.5	700-900	800-1200	1975-2012	6-7	4-5	2-4	1	1	0-1 (80%)
1000-1100	.2	700-950	800-1400	1975-2012	6-8	4-6	2-4	1	1	0-1 (80%)
1100-1200	.2	700-950	800-1500	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)
1200-1300	.034	700-1000	900-1600	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)
1300-1400	.033	800-1000	1000-1700	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)
1400-1500	.033	800-1100	1200-1800	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)

TABLE II
SYNTHETIC DATA SET GENERATION RULES

is the class attribute (“Sale Price”). The non-class attributes are “House Size (sqm)”, “Lot Size (sqm)”, “Year of building”, “Number of Bed Room”, “Number of Bath Room”, “Number of Garage”, “Landscape”, “Sprinkler”, and “Mountain View”.

Table II presents the rules that we use to generate the synthetic data set. There are altogether 11 columns and 17 rows in the table. The first column shows ranges of prices, while the second column shows the distribution of the prices in the data set. For example, in the synthetic data set there are 5% records within the price range of $50K - $100K, and 35% records within the price range of $200K - $300K. Column 3 to 11 show the ranges of various other attribute values corresponding to a price range. For example, for the price range of $50K - $100K the lot size varies between 90 sqm to 120 sqm.

We generate a record of the synthetic data set using the following steps.

- Step 1: Generate Property Price.
  We first generate a property price following the distribution of price ranges.

- Step 2: Generate Lot Size.
  Considering the high correlation between a property price and lot size (Zietz et al. 2008) we next generate the lot size within a range as shown in Table II. The probability distribution within a range of lot size is uniform.
- Step 3: Calculate House Size.
  We next calculate the House Size (i.e. Floor Size) using $P = 1000 \times h + 150 \times l$, where P, h and l are the house price, house size and lot size, respectively. If the calculated house size falls outside the upper limit of the range of House Size as shown in Table II then we consider the value equal to the upper limit of the range. We do the same for the lower limit of the range as well.
- Step 4: Generate other attribute values.
  We also generate other attribute values based on the rules as shown in Table II following a price range. For example, the number of bed rooms for a property having price within the range of \$500K and \$600K can be anything between 4 and 6 with uniform probability distribution.
  The domain of the attributes Landscape, Sprinkler and Mountain View is $\{0,1\}$, where 0 means no landscaping/sprinkler and 1 means the existence of landscaping/sprinklers. A percentage value within parenthesis indicates the probability of having the feature. For example, there is a 40% probability of having sprinklers for a property within the price range of \$300K and \$400K (see Table II).
- Step 5: Check Floor Size.
  Considering an average bed room size is 10 sqm, bath room size is 4.5 sqm, and garage size is 20 sqm we calculate the house size (using the generated number of bedroom, bathroom etc.) in order to make sure that the calculated house size is not greater than the house size estimated in Step 3. If that is not the case then we reduce the number of a bedroom, and/or a garage as necessary.

From the synthetic data set we build a decision forest having four trees using SysFor algorithm (Islam & Giggins 2011). Figure 1 shows the first tree of the forest. In this study we use the first tree to estimate the initial reserve price of a property. For example, the logic rule for Leaf 1 of the tree indicates that if there is no landscaping, no garage and the lot size is $\leq$ 120 sqm then the property price should be within \$50K to \$100K range. As the starting reserve price of a property we estimate the lower limit of the range.

H. Implementation

The population of houses was initialised with attributes randomly generated according to the distributions from the synthetic data (i.e. 5% with prices from 50-100, size 45-90 and son on). The houses are then randomly assigned among the agents. Agents are randomly assigned a type from the four categories SINK, DINK, Family (Fam) and Retiree (Ret). Incomes are determined using a base of 20 with a random exponent from 1.0 to 1.75, yielding incomes in a range from 20 to 189 - DINKs and Families are given two incomes.

A fixed number of buyers (10) and sellers (5) is used at the start of the simulation from a total of 1000 agents – these are changeable by parameters. There are parameters for the market temperature for both buying (i.e. housing demand) and selling (5 each). These numbers are divided by the scaling factor (1000) to yield the chance that an agent enters the market. With the current numbers, at every time step, an agent has a 5/1000 (0.5%) chance of becoming a buyer and a 5/1000 chance of becoming a seller.

The second-price auction model was implemented making the assumption that if there is no second price (i.e. only one bidder), or the second price is below the reserve, then the price paid is the reserve price. If the reserve price isn't met by buyers at the auction, the reserve price is reduced by 10% for the next auction. If a house attracts no bidders, no auction is deemed to have taken place, so it remains in the available pool for the next time step.

Values for the utility function were set assuming that people spend at least half their income on housing, so α, β and γ were each set to a random value from 0.6 to 1.0 (yields P_{max} equal to 50-83% of income, I).

For desirability, the weightings are given in Table I.

3.. ILLUSTRATIVE SIMULATIONS

The model was implemented in RePast Symphony Version 2. Simulations were run for 1500 time steps, approximately 5 years with a timestep being one day. All houses being put up for sale are auctioned once and only once during a timestep. Figure 2 shows the price as a function of time. A natural inflation effect is observed. The growth in price is more rapid at the beginning of the simulation, suggesting a settling period where houses undervalued at the beginning rapidly approach a more market driven value.

Figure 3 shows the size of the market as judged by number of agents for different market temperatures (2,5,8 as before). The lower market temperature produces more of an "equilibrium" with a matched numbers of buyer and sellers. The higher market

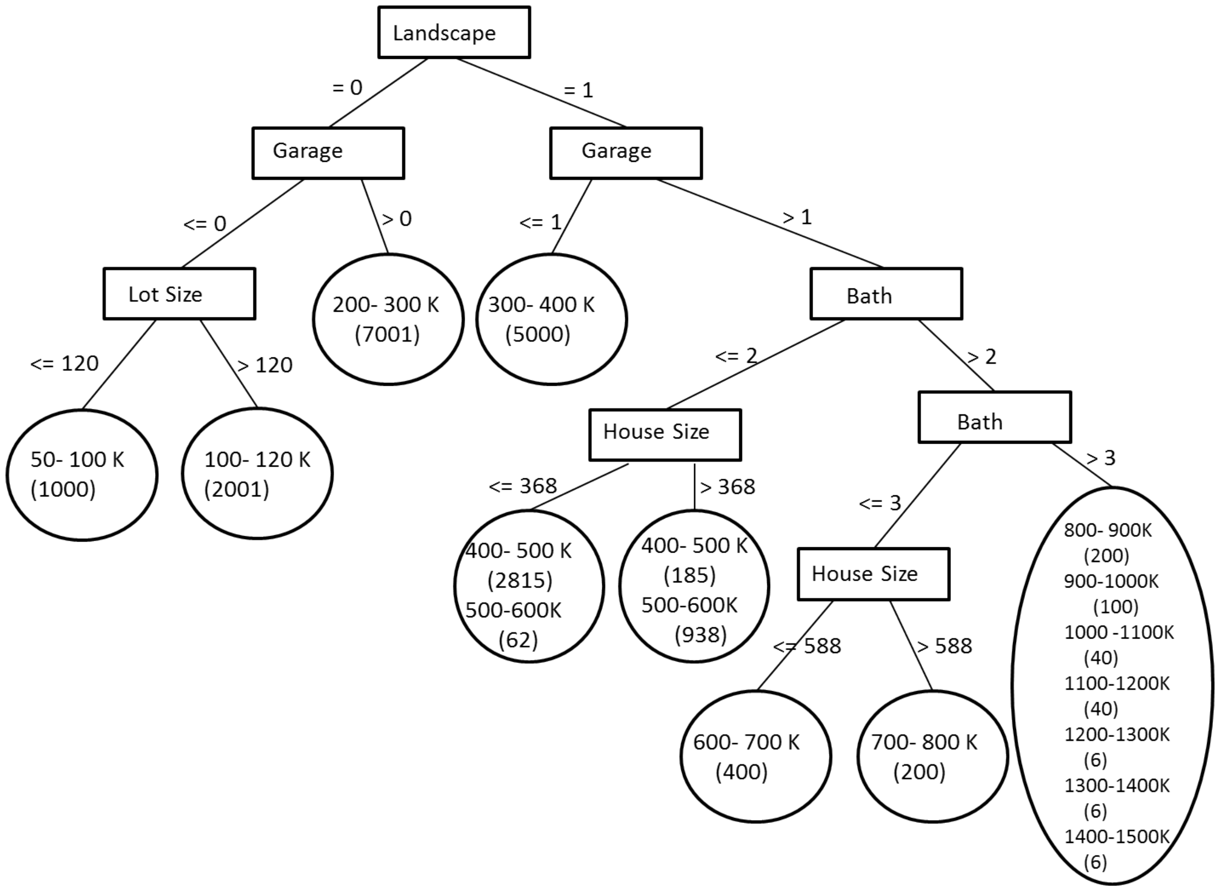

Fig. 1. A Decision Tree Obtained from the Synthetic Data Set

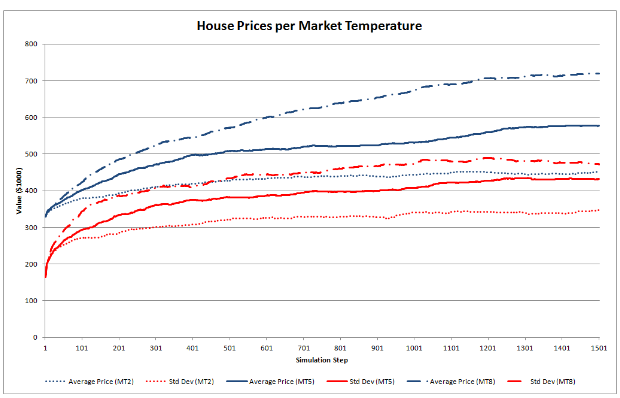

Fig. 2. Growth in house prices with time for three different market temperatures, MT2, MT5 and MT8, corresponding to temperature values of 2,5 and 8 (0.2% chance of selling, see section 2.8.

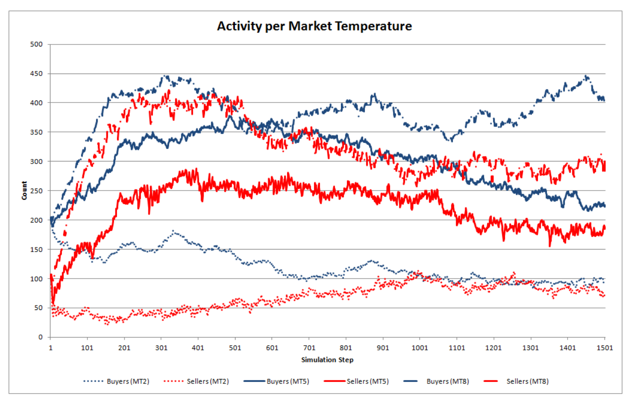

Fig. 3. Number of buyers and sellers in the market for three market temperatures MT2, MT5 and MT8, corresponding to temperature values of 2,5 and 8 (0.2% chance of selling, see section 2.8.

temperatures create an excess demand, with more buyers than sellers, consistent with the inflationary trend observed in figure 2.

4.. CONCLUSIONS AND FURTHER WORK

Agent based modelling provides a way to project beyond hedonic modelling and study house price growth under various scenarios. From house price growth, land valuation can be predicted and adjusted.

A decision tree approach has been introduced to estimate house price ranges. The next stage of the project is to determine the decision tree from real data, use census data to determine the actual probability distributions for socio-economic categories to parameterise the agents.

An interesting change in Bathurst demographics has been foreshadowed. One of the mining companies is set to dramatically increase its gold mining activity, with the expectation of around 3,000 new jobs. As almost 10% of the current population, this will present challenges to the housing market. The challenge for our simulation will be to see if we can successfully predict the outcome.

REFERENCES

Abellan, J. & Masegosa, A. R. 2009. An ensemble method using credal decision trees, *European Journal of Operational Research* .

Bourassa, S., Hoesli, M. Scognamiglio, D. & Zhang, S. 2011. Land leverage and house prices, *Regional Science and Urban Economics* 41: 134–144.

Damigos, D. & Anyfantis, F. 2011. The value of view through the eyes of real estate experts: A fuzzy delphi approach, *Landscape and Urban Planning* 101(2): 171–178.

Islam, M. Z. 2012. Explore: A novel decision tree classification algorithm, *LNCS* 6121: 55–71.

Islam, M. Z. & Giggins, H. 2011. Knowledge discovery through sysfor: A systematically developed forest of multiple trees, *in* 'Proc. of the Ninth Australasian Data Mining Conference (AusDM 11', ACS, pp. 195–204.
URL: *http://crpit.com/vol121.html*

Kusan, H., Aytekin, O. & Ozdemir, L. 2010. The use of fuzzy logic in predicting house selling price, *Expert Systems with Applications* 37(3): 1808–1813.

Lughofer, E., Trawiński, B., Trawiński, K., Kempa, O. & Lasota, T. 2011. On employing fuzzy modeling algorithms for the valuation of residential premises, *Information Sciences* 181: 5123–5142.

Magliocca, N., McConnell, V., Walls, M. & Safirova, E. 2012. Zoning on the urban fringe: Results from a new approach to modeling and housing markets, *Regional Science and Urban Economics* 42: 198–210.

Magliocca, N., Safirova, E., McConnell, V. & Walls, M. 2011. An economic agent-based model of coupled housing and land markets (chalms), *Computers, Environment and Urban Systems* 35: 183–191.

McMillen, D. 2008. Changes in the distribution of house prices over time: Structural characteristics, neighborhood, or coefficients?, *J. Urban Economics* 64: 573–589.

Myerson, R. 1981. Optimal auction design, *Mathematics of Operations Research* 6: 58–73.

Narula, S. C., Wellington, J. F. & Lewis, S. A. 2012. Valuating residential real estate using parametric programming, *European Journal of Operational Research* 217: 120–128.

Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, California, USA.

Quinlan, J. R. 1996. Improved use of continuous attributes in c4.5, *Journal of Artificial Intelligence Research* 4: 77–90.

Vickrey, W. 1961. Counterspeculation, auctions and competitive sealed tenders, *Journal of Finance* 16: 8–37.

Zietz, J., Zietz, E. N. & Sirmans, G. S. 2008. Determinants of house prices: A quantile regression approach, *Journal of Real Estate Finance Economics* 37: 317–333.