Analysing Urban Systems using Agent-Based Modelling

Scott Heckbert, **Alex Smajgl, CSIRO Sustainable Ecosystems**
Email: scott.heckbert@csiro.au

Keywords: ***Agent-based modelling, urban systems, residential decision making***

EXTENDED ABSTRACT

This paper outlines the application of an agent-based model (ABM) within the context of urban systems. The model is developed with the goal of exploring residential market dynamics under various scenarios, and addresses two issues of concern to Australian urban systems, namely interest rates and population growth.

Simulated agents are faced with the selection of a residence, and use market and non-market factors in decision making. From the field of discrete choice theory, a preference based utility calculation is linked to the attributes of properties, as provided from GIS data and populated from a variety of secondary literature sources. In a second decision making step, the affordability of residences is calculated based on prevailing interest rates, and compared to income based thresholds for housing expenditure.

The model highlights the benefits of micro scale simulation in the ability to represent and report not only the average outcomes of policy impacts, but how the impacts are distributed within the society, addressing the issue of 'social gap'.

Impacts of two sets of scenarios are examined, comparing different possible interest rate changes, and increases in population growth through immigration. It is shown that the ABM methodology is uniquely able to identify thresholds of change across varying scenarios, and offers particular value in the ability to communicate the distributive effects of impacts.

A generic urban landscape and agent population is simulated using data from ABS, Brisbane DCDB and Sydney community profiles, with discussion on techniques and data sources to calibrate specific case studies.

1. INTRODUCTION

This paper outlines the application of an agent-based model (ABM) within the context of urban systems. The model is developed with the goal of exploring residential market dynamics under various scenarios. Although urban landscapes have begun to be simulated using ABM techniques in other countries, the application to Australian cities is as yet very limited. As such, ABM research stands to make significant contributions to issues of concern to Australian cities.

ABM systems consist of a set of 'micro level entities' that interact with each other and an 'environment' in prescribed ways (Lane 1993). The interaction over time of micro level entities, or agents, produces a history of the changing states of the overall system.

The strength of using an ABM methodology over other techniques lies in the ability to represent the decision making process at the individual scale. As such, systematic processes that yield the macro-level outcomes receive adequate representation and are not 'lost' in overall population aggregation. This allows the model to specifically identify who is affected, how impacts are distributed across the population, and what outcomes emerge on a higher scale of measurement. These are questions that cannot be addressed using aggregated scale techniques. Urban systems are emergent from the multitude of individual actors within, guided by top-down policy controls. Thus analysis of the system requires an adequate representation of these individual entities and constraints.

In the context of urban systems, this technique can be particularly useful for exploring issues of 'social gap', which is concerned with the distribution of benefits within a society. Although traditional methods of analysis deal with aggregate measures of wellbeing, this is limited by the assumptions in utilitarian theory that do not adequately explore the distribution across all individuals in the group. Even in developed modern societies, there are bound to be some people poorer than others, who are then relatively deprived (Hammond 1997). Policy makers are concerned with issues across different segments of society, and therefore pay particular attention to marginalised members of the community. These may not be adequately reflected in aggregated analysis.

Two current issues that receive a great deal of attention in Australia are the impact of interest rate changes and population growth (immigration). In a highly urbanised society such as Australia, the impact of various changes to these rates is of particular importance. The model outlined addresses the connection between interest and population growth rates, effects on the housing market, and the resulting impacts on affordability issues for households. A variety of other factors influence household affordability issues, however here we single out the impacts associated with housing affordability.

To explore this issue, the model described here simulates agents involved in residential decision making, who use market and non-market criteria to inform their choices. Agents represent a household, which is taken to be the base unit of residential decision making. The land base is represented as a set of GIS polygons representing plots of residential land. Agents are represented as objects (discrete bundles of data) which include perception of model variables and cognitive operations based on these variables (methods). The agents retain 'ownership' of landscape polygons. Thus, agents in the model represent an individual decision making unit who decide on a residential location. Following from Doran (2001), agents are software entities that are autonomous loci of decision making that sense, decide and act.

Decision making occurs according to preference (utility) functions and affordability calculations. The model therefore takes into account the market (affordability) and non-market (utility) influences that affect residential decision making. Thus, the demand side of the land use market is represented in a fashion that is consistent with processes involved in real world decision making. Impact assessment can therefore be undertaken, focussing on the effects of interest rate and population growth rate changes. Using ABM methods allows for examination of the urban system at a disaggregated level, and incorporates indicators that examine impacts on this scale.

A 'generic' urban landscape and agent population is simulated using data from ABS, Brisbane digital cadastre database (DCDB) and Sydney community profiles, with discussion on techniques and data sources to calibrate specific case studies.

2. RESIDENTIAL DECISION MAKING

Agents in the model described perform a two-step calculation which separates market and non-market decisions, both of which are integral to selecting a residence. The first question an agent asks when selecting a home is "Do I like this location?". Secondly, agents ask the question "Can I afford to live here?". Because decision making is not driven solely by either financial conditions or by preferences of the individual, both factors are considered in the decision making criteria outlined here.

2.1. Representation of Preference-Based Decision Making

Agents representing households face the decision of where to select their residence. Residence locations all have the common characteristic of consisting of a set of unique attributes. Decision making is a function of an individual's perception of, and preferences for these attributes, and also their expectations of outcomes of their decisions. Thus, following from the area of discrete choice theory, a model of decision making is presented according to Louviere, Hensher and Swait (2000) and Ben-Akiva and Lerman (1985). This involves:

- A decision maker facing a choice problem
- Perception of a set of possible alternatives
- Perception of attributes of each alternative
- The ability to evaluate the outcomes of each alternative by some decision rule
- Action based on the decision making rule and the feasible options.

For residential location decisions, 'buyers' are faced with a set of mutually exclusive options from which to choose, from a finite set of discrete bundles of attributes. Agents ultimately select the residence which is expected to generate the highest utility within the feasible choice set.

Within discrete choice theory, utility is seen to be derived from the attributes, characteristics, or properties of a good, and not directly from the good itself. Thus, individuals hold a utility function such that:

$$U_{ij} = U(X_j^k) \quad (1)$$

where X_j^k is a good, in this case a residence, with attributes k, as considered by individual i, at location j.

The utility function is dependant on the relative preferences, β_i^k, for location attributes. Preference coefficients are utility function weights such that:

$$U_{ij} = \sum_{k=1}^{K} \beta_i^k X_j^k \quad (2)$$

Dealing firstly with the question "Do I like this location?", agents in the model are initialised with an individual set of preference parameters, β_i^k, normally distributed according to a given variance. Each preference parameter is associated with a specific X_j^k attribute, as is described later. Using this set of β_i^k parameters, a calculation can be performed on residential plots to determine the amount of utility derived from that location.

Agents maintain a low-end threshold for acceptable utility levels, which is set and varied for each agent in the same fashion that the preference parameters are set. Hence, some agents will accept a very low level of utility, and some will be more 'picky' and be initialised with a higher threshold. For a property to be considered as a prospective residence, the utility calculation must fall above this threshold.

For the purposes of calibrating a case study specific model, elicitation of β_i^k parameters for individuals' preference structures for location attributes X_j^k is possible through random utility modelling (RUM) using stated preference (SP) and revealed preference (RP) techniques (See Adamowicz et al. 1997, Adamowicz et al 2001, Louviere, et al. 2000), and also by multi criteria analysis (MCA) (see Hajkowicz et al. 2000). In this generic application, preference weights are assigned to individual agents from a population level mean of 1 with a variance of 0.3 for all k attributes.

2.2. Residential Affordability Calculation

The second consideration when purchasing a new home is the affordability of the property, which answers the question "Can I afford to live here?". Agents have an income and residences have a purchasing price. Home buyers in the real world typically secure financing from a financial institution in the form of home loans/mortgages, and hence interest rates play a role in determining the affordability of houses, and therefore which location will inevitably be selected. To accommodate this, the model uses a 'housing affordability calculator' which informs the agent of what monthly payments they could expect to make when considering purchasing a specific property.

The specific equation used is an amortisation function such that:

$$MP_j = P_j \times (\gamma - 1)/(1 - \gamma^{-n}) \quad (3)$$

Where MP_j is the monthly payment for home j, as a function of the purchase price of the home, P_j, and a monthly interest amortisation factor, γ, as applied to n repayment periods. The monthly amortisation factor is in turn a function of interest rates, r, such that:

$$\gamma = 1 + (r/100/12) \quad (4)$$

The resulting equation returns the expected monthly mortgage payment an agent can expect on the home they are considering for purchase. This value is in turn used to calculate the affordability, A, of the home using the equation:

$$A = I \times (1/3) \times (1/12) - MP_j \quad (5)$$

Where one third of (monthly) income, I, is taken to be the amount spent on housing before being in a condition of 'housing stress' (Landt and Bray, 1997).

However, in the real world some households may choose to spend slightly more or slightly less than this 1/3 value. To account for this, the final set of equations applies a stepwise function, such that:

$$\text{If } A \begin{cases} >0 \Rightarrow \alpha = A - A^{1.05} \\ <0 \Rightarrow \alpha = |A|^{1/\sqrt{A}} \times A \end{cases} \quad (6)$$

Where α represents an 'adjusted affordability' measure. Hence, if A is negative, the affect of 'overspending' on housing is magnified. Likewise, if the household 'underspends' on housing (such as a high income family living in an inexpensive house) they will become increasingly inclined to move into a home suitable to their earnings. As with the utility function described in the previous section, each agent has a unique lowest acceptable threshold for α.

3. METHODS

The model is implemented in C#, using GIS .dbf files from the Brisbane DCDB as landscape input files. Data are organised upon input into the model into two interacting linked list data structures, one representing the landscape data, and one representing household agents, the latter generated within the model according to the specified population (adjusted for population growth in each time step). The number of landscape nodes is determined by the GIS data loaded into the model. The DCDB data is the primary GIS layer needed to build the landscape linked list, and identifies unique plots of land according to contour, area, tenure, location, the type of land parcel, and other relevant details of land ownership.

Data layers for various landscape attributes are added to the DCDB polygons to represent attributes of the houses, neighbourhood, and greater city and constitute the X_j^k attributes described in section 2.1, including:

City attributes from Sydney community profiles:[1]

- IndustryBase
- MajorCorporates
- BusinessPark
- CityCommunityServices

Neighbourhood attributes from Sydney community profiles, ABS and Brisbane DCDB data:

- Accessibility
- Education
- Recreation
- AverageHousePrice
- AverageLotSize
- NeighbourhoodCommunityServices

Residence attributes from Brisbane DCDB data and proxy assumptions:

- HouseSize
- LotSize
- NumberBedrooms

Agent attributes from ABS data and proxy assumptions:

- Income
- Preference values for each landscape attribute
- Utility and affordability thresholds

In this generic application, landscape and agent attribute data are drawn from a variety of secondary literature sources, as per availability. Calibrated simulations can be populated from a variety of sources, including ABS census data, DCDB data, property valuation databases, and community profiles for cities and neighbourhoods. Agent utility functions can be calibrated as discussed in section 2.1.

4. SCENARIOS AND INDICATORS

Model simulations are used to examine the effect of various scenarios of population growth (immigration) and interest rate levels. Indicators for the current application include:

- Monthly payments made on housing
- Adjusted affordability
- Gini coefficients based on adjusted affordability

Monthly payments and adjusted affordability are calculated as outlined in section 2.2.

The gini coefficients are calculated according to the technique outlined in Liao (2005). Here, the population is clustered into percentiles based on individual adjusted affordability, with the gini coefficient reporting on the distribution of this indicator across the population. A gini coefficient of 1 would describe a population whose indicator

[1] See www.gws.org.au

is uniformly distributed across all individuals; a perfectly equitable outcome. A coefficient of 0 would report that one individual in the population holds a 'monopoly' on the indicator. As an example, the gini coefficient for distribution of income in Australia, as reported by ABS (2005a) is currently 0.315, having increased consistently in the last ten years. This is interpreted to mean that the income of Australians in becoming more equitably distributed over time. The gini coefficient applied here reports on adjusted affordability, which is a measure of dissatisfaction, and hence a negative number. Thus, the gini coefficient ranges from -1 to 0, with 0 representing a uniformly distributed effect throughout the population.

Indicators are presented for four possible interest rate conditions and three population growth scenarios, communicated as a graphical trajectory of outcomes. Ten simulation runs were performed for each indicator, per scenario. Population level trends are reported in averages, and use the gini index as a measure of how outcomes are distributed across the agent population.

5. RESULTS

Results for variable population growth and interest rate levels are compared to a baseline case. The baseline interest rate is assumed to be 7.5%, which is the mean value between the prime and real prime interest rates as of March 2005(APL 2005). Loan term is agent-specific, with a mean value of 20 years, and a variance of 8 years. In terms of population growth, the baseline growth rate for Brisbane (statistical division) is used, and set at 2.3% per annum (ABS 2005b). Each simulation was run for 30 time steps, each representing a three month period, the time frame by which it is assumed that a residential relocation decision might take place. Results are generated *ceteris paribus*, such that only the scenarios (interest and population growth rates) under question are altered for the simulation, and are done so at model initialisation. The following figures depict the mean outcome for 10 simulation runs, per scenario.

5.1. Interest Rate Changes

Focussing first on interest rate scenarios, indicators are presented for interest rate levels of 7.5%, 7.75%, 8% and 8.25%. Figure 1 describes the average monthly payment that households can expect to make under different interest rate levels. As would be expected, there is an equal increase in monthly payments for each interest rate increase of 0.25%.

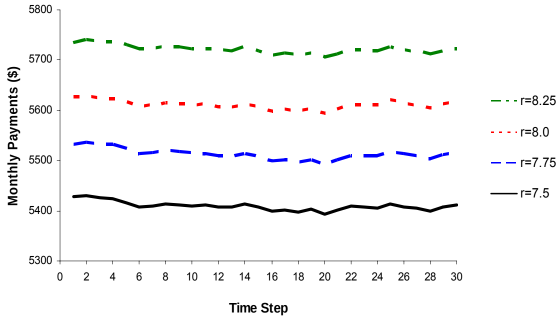

Figure 1: Average monthly payments for four interest rate scenarios. A unitary increase in loan payments is seen.

However, the effects of an increase in interest rates in not 'felt' in a unitary fashion as might be interpreted from the above figure. Hence, the adjusted affordability measurement described in section 2.2 accounts for this. Figure 2 describes the fashion by which further 'squeezing' of the household's budget constraint has a magnifying effect.

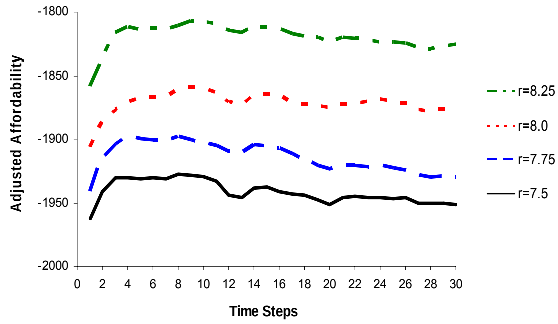

Figure 2: Average adjusted affordability for four interest rate scenarios. A magnifying effect of interest rate increases is seen.

As such, the higher the interest rate, the more household agents find their threshold for affordability not being met. How this is distributed across the population is described in figure 3, reporting the gini coefficient for adjusted affordability in each time step.

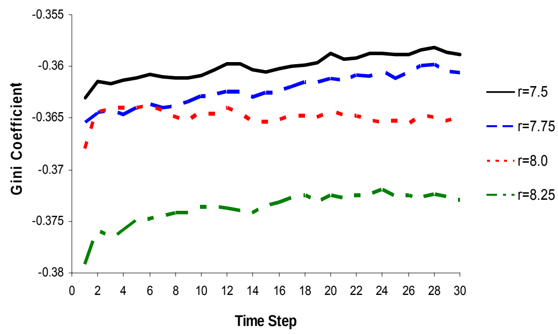

Figure 3: Gini coefficient for four interest rate scenarios. The distribution of adjusted affordability becomes less uniform as interest rates increase.

Here we see that the effect on adjusted affordability is not equally felt across all segments of the population. As interest rate increases, the impact of adjusted affordability in the population becomes less uniformly distributed. In other words, the number of 'relatively deprived' individuals in the society increases, indicating a widening 'social gap'. The greatest impact is seen to occur between the 8.0% and 8.25% increases, with the change in the first two scenarios of 7.5% and 7.75% to be relatively small. This suggests that there exists some threshold for maintaining housing affordability, where households are generally able to be resilient to changes up to an increase in the range of 0.5%. After this, further increases have a sharpened effect.

5.2. Population Growth Changes

Focussing now on the effects of various population growth (immigration) possibilities, we examine the same set of indicators for three scenarios: the baseline population increase of 2.3%, and two possible increases to 3.0% and 3.7%. We see in figure 4 that increases in population for the different growth scenarios eventually begins to differentiate the three trajectories in terms of monthly payments, albeit the effect is minimal. Although very similar in the first half of the simulation, the effect of compounding growth has an effect over time that begins to differentiate the trajectories.

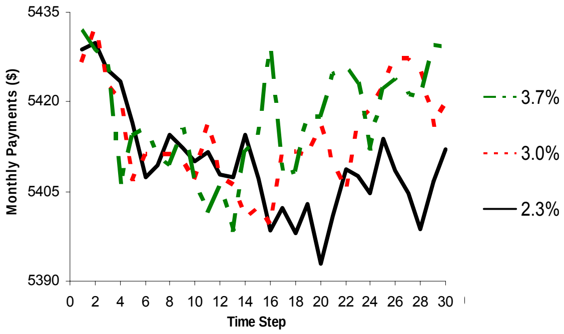

Figure 4: Average monthly payment for three population growth scenarios. Increases in population slowly place upward force on market price.

The outcomes for adjusted affordability are shown in figure 5. Magnification of effects is seen in a similar fashion described for the previous indicator, again however, the trend is minimal, and there is not a clear delineation between the two higher population growth rates, with these trajectories overlapping at several points.

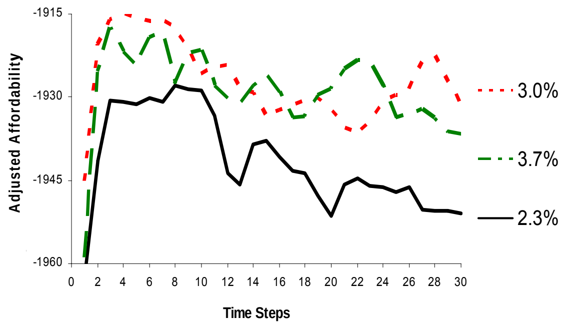

Figure 5: Average adjusted affordability for three population growth scenarios. Trajectories for higher rates differ from the baseline but trends are not clearly seen between the two**.**

The distribution of adjusted affordability across the population is described in figure 6, reporting the gini coefficient for each time step.

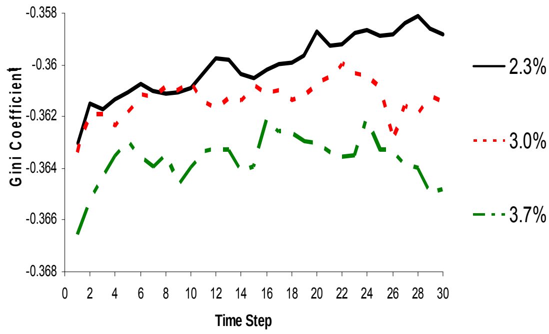

Figure 6: Gini coefficients for distribution of adjusted affordability for three population growth scenarios. Less equitable distribution is seen with higher growth rates.

As population growth rates increases, the impact of adjusted affordability in the population becomes less uniformly distributed. This trend is more observable than the effects on monthly payments and adjusted affordability across the population, suggesting that the number of 'relatively deprived' individuals in the society increases, even while other indicators do not show marked trends. This serves to show that distributional considerations may be changing, even while overall average indicators may not.

6. DISCUSSION

The aim of simulating the dynamics of urban environments is to improve understanding of impacts within these systems. Using the ABM approach allows for analysis not only of average trends, but also how impacts are distributed across the population. Hence, the model presents an improvement on traditional aggregated techniques, and allows consideration of 'social gap' concerns.

The model presented outlines a decision making criteria that uses market (affordability) and non-market (preference for attributes) drivers. Although a 'generic' urban landscape and agent population is simulated, calibration of the model to specific case studies is possible using secondary literature, including ABS census data, DCDB data, property valuation databases, and community profiles for cities and neighbourhoods. Agents' preference-based decision making can be calibrated from RUM (using SP /RP) and MCA.

The specific value of results discussed here lies in the identification of potential thresholds and distributive effects. From the scenarios involving interest rate changes, the distribution of housing affordability across the population, as reported in the gini coefficient, decreases as interest rates increase. Furthermore, the scenario involving the highest increase in interest rates shows a magnified effect of 'inequity' under this policy, identifying potential thresholds for interest rate rises of 0.5% and above. In the scenarios involving population growth, it was seen that although indicators which report population averages may not show large changes in trajectories, the equity of distribution within the society can however be changing. Hence, support for the use of disaggregated scale methodologies is given, in order to track distributional impacts that may be lost in average indicators and aggregated techniques.

7. ACKNOWLEDGMENTS

The authors would like to thank Geoff Carlin and Colin Mayocchi for programming contributions. Funding for this project was provided by the CSIRO Sustainable Ecosystems Internal Venture Capital program.

8. REFERENCES

Adamowicz W., Boxall P, Williams M., Louviere J., Swait J. (1997): Perceptions Versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. *Journal of Environmental Economics and Management* 32: 65-84.

Adamowicz W., Bennett J. (2001): Some Fundamentals of Environmental Choice Modelling. In: J. Bennett and R. Blamey (eds.) The Choice Modelling Approach to Environmental Evaluation. Edward Elgar. Northampton. 37-69.

Australian Bureau of Statistics (2005a): Household Income and Income Distribution, Australia. Available online: www.abs.gov.au/Ausstats/abs@.nsf

Australian Bureau of Statistics (2005b): Media Release 3218.0 Melbourne and Sydney experience largest population growth, but Brisbane records the fastest growth. Available online: www.abs.gov.au/Ausstats/abs@.nsf

Australian Parliamentary Library (2005): Monthly Economic and Social Indicators 2003-2004. Available online: www.aph.gov.au/library/pubs/mesi/mesi51.htm

Ben-Akiva, M. and Lerman, S. (1985): Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, MA: MIT Press.

Doran, J. (2001): Intervening to Achieve Co-operative Ecosystem Management: Towards an Agent Based Model. *JASSS.* (4)2.

Hajkowicz S., McDonald G., Smith P. (2000): An evaluation of multiple objective decision support weighting techniques in natural resource management. *Journal of Environmental Planning and Management.* 43(4). 505-518.

Hammond, P. (1997): Progress in the Theory of Social Choice and Distributive Justice; in Inequality and Taxation: Research on Economic Inequality, vol. 7; edited by Sourushe Zandvakili; JAI Press; Greenwich, Conn:87-106.

Landt, J. and Bray, R. (1997): Alternative Approached to Measuring Rental Housing Affordability in Australia. National Centre for Social and Economic Modelling, Faculty of Management, University of Canberra. Discussion Paper No. 16.

Lane, D. A. Artificial Worlds and Economies, part II. *Journal of Evolutionary Economics*. (3): p. 177- 197.1993.

Liao, T. (2005): The Gini Unbound: Analyzing Class Inequality with Model-Based Clustering. International Conference in Memory of Two Eminent Social Scientists: C. Gini and M. O. Lorenz. University of Siena.

Louviere, J. D. Hensher and J. Swait. Stated Choice Methods - Analysis and Application. Cambridge: Cambridge University Press. 2000.