

Mitigating housing market shocks: an agent-based reinforcement learning approach with implications for real-time decision support

Sedar Olmez, Alison Heppenstall, Jiaqi Ge, Corinna Elsenbroich & Dan Birks

To cite this article: Sedar Olmez, Alison Heppenstall, Jiaqi Ge, Corinna Elsenbroich & Dan Birks (09 Jul 2024): Mitigating housing market shocks: an agent-based reinforcement learning approach with implications for real-time decision support, *Journal of Simulation*, DOI: [10.1080/17477778.2024.2375446](https://doi.org/10.1080/17477778.2024.2375446)

To link to this article: <https://doi.org/10.1080/17477778.2024.2375446>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 09 Jul 2024.



[Submit your article to this journal](#)



Article views: 124



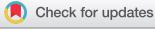
[View related articles](#)



[View Crossmark data](#)

CrossMark

RESEARCH ARTICLE

 OPEN ACCESS 

Mitigating housing market shocks: an agent-based reinforcement learning approach with implications for real-time decision support

Sedar Olmez^{a,b}, Alison Heppenstall^c, Jiaqi Ge^a, Corinna Elsenbroich^d and Dan Birks^{b,e}

^aSchool of Geography, University of Leeds, Leeds, UK; ^bThe Alan Turing Institute, British Library, London, UK; ^cSchool of Social Political Sciences, University of Glasgow, Glasgow, Scotland; ^dSchool of Health and Wellbeing, University of Glasgow, Glasgow, Scotland; ^eSchool of Law, University of Leeds, Leeds, UK

ABSTRACT

Research in modelling housing market dynamics using agent-based models (ABMs) has grown due to the rise of accessible individual-level data. This research involves forecasting house prices, analysing urban regeneration, and the impact of economic shocks. There is a trend towards using machine learning (ML) algorithms to enhance ABM decision-making frameworks. This study investigates exogenous shocks to the UK housing market and integrates reinforcement learning (RL) to adapt housing market dynamics in an ABM. Results show agents can learn real-time trends and make decisions to manage shocks, achieving goals like adjusting the median house price without pre-determined rules. This model is transferable to other housing markets with similar complexities. The RL agent adjusts mortgage interest rates based on market conditions. Importantly, our model shows how a central bank agent learned conservative behaviours in sensitive scenarios, aligning with a 2009 study, demonstrating emergent behavioural patterns.

ARTICLE HISTORY

Compiled April 28, 2024

KEYWORDS

UK housing market; agent-based model; reinforcement learning; crisis management; decision-making

1. Introduction

¹Agent-based models (ABMs) have been adopted in various research areas since their inception in the late '90s to early 2000s (Filatova, 2015; Ge, 2017; Groff Elizabeth et al., 2018; Heppenstall et al., 2006; Kothari et al., 2014; Tang & Bennett, 2010). ABMs enable researchers to simulate a complex system with autonomous agents interacting with each other within an environment. The main strength of ABMs over mathematical models is that they simulate, validate, and verify behavioural characteristics at granular spatio-temporal resolutions (Olmez, Thompson, et al., 2022; Secchi, 2015; Todd et al., 2017). This allows researchers to analyse complexity and investigate how a studied phenomenon develops at the individual level (Epstein & Axtell, 1997). This article focuses on housing markets and investigates market shocks, which are unanticipated changes to economic variables that impact the market's health (Ramey, 2016).

ABM has been used in housing market research. Researchers investigated the emergence of housing bubbles (Axtell, 2014; Erlingsson et al., 2014; Ge, 2014, 2017), the dynamics of urban regeneration (Jordan et al., 2011, 2012; Piscopia et al., 2014), and how real-world shocks such as the 2008 financial crash affected the housing market (Gilbert et al., 2009; Hamill & Gilbert, 2015). The number of ABMs for studying housing and financial markets is growing (Bae et al., 2019;

Baptista et al., 2016; Carstensen, 2015; Geanakoplos et al., 2012). These models generally allow agents to make decisions in volatile scenarios, either to hedge against volatility or profit from it (Fischer & Riedler, 2014, Todd et al., 2017; Westerhoff, 2010).

A research area less explored is applications of machine learning (ML) algorithms supporting decision-making in alleviating shocks once they have occurred, which central-bank policymakers can use to inform policy. Most models cited earlier examined how, when, and why shocks occur. However, developing techniques to counteract these shocks can reduce the impact on the economy and people's health (Oguibenine, 2011). This article proposes a hybrid model that integrates reinforcement learning (RL), with a housing market ABM. Conducting a series of experiments, we investigate if an intelligent adaptive central bank agent (Almahamid & Grolinger, 2021; Littman, 2015; Mehta, 2020) can learn trends from a housing market in real-time. During learning, this central bank agent makes decisions to fulfil a goal, for example, decreasing homelessness. In this article, "intelligent adaptive agent" is defined as: "systems or machines that utilise inferential or complex computational algorithms to modify or change control parameters, knowledge-bases, problem-solving methodologies, course of actions, or other objects in order to

CONTACT Sedar Olmez  olmez49@gmail.com  School of Geography, University of Leeds, Seminary St, Woodhouse, Leeds LS2 9JT, UK
 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/1747778.2024.2375446>.

accomplish a set of tasks required by the user" (Imam & Kerschberg, 1997).

We identified several benefits of utilising RL in the housing market domain (1) researchers can test macroeconomic policies in a safe "sandbox" environment without real-world consequences. (2) researchers can adopt various RL goal criteria to test policy interventions in the housing market. (3) researchers can test various interventions in their housing markets and document the steps to counteract these interventions. (4) shocks (crashes) can artificially be induced to speed up learning, whereas market shocks are rare events in the real world.

This article replicates an ABM of the UK housing market (Gilbert et al., 2009). Other notable housing market ABMs exist (Baptista et al., 2016; Filatova, 2015; Ge, 2017; Rosenfield et al., 2013; Yun & Moon, 2020). However, we found that either these articles were not open access and did not include download links to the models (Filatova, 2015; Ge, 2017) or the articles were open access. However, no documentation was provided to access the ABMs (Baptista et al., 2016; Rosenfield et al., 2013; Yun & Moon, 2020). We chose (Gilbert et al., 2009) as it was well received by researchers (61 citations as of May 20 2022 on Google Scholar), for its documentation. Furthermore, (Gilbert et al., 2009), strikes a good balance between simplicity (where results are tractable) and realism (simulating important processes unique to the UK housing market, such as chain trade and can replicate empirical patterns).

To investigate whether RL can manipulate the housing market, this article reproduces two identical experiments conducted in (Gilbert et al., 2009) as a comparator. Where exogenous shock events occur, and the decisions made by the central bank RL agent are observed. These results are compared to baseline scenarios where the RL intervention is removed. The model outputs reflect the consequences of RL decisions, and findings are compared with the original assertions made in (Gilbert et al., 2009).

The motivation behind the investigation lies in the potential applications of RL for policymakers. By assessing the adaptability and decision-making prowess of the RL agent, particularly in adjusting mortgage interest rates based on market conditions, our study offers insights into novel strategies for mitigating the consequences of unforeseen events in the housing market. These insights aim to empower policymakers with a proactive tool-set, allowing them to navigate and respond effectively to the challenges posed by housing market dynamics.

To summarise, this article will investigate whether (i) an RL agent can be integrated with a housing market ABM and (ii) can an RL agent be trained using input data from the housing market ABM and make decisions to counteract shocks when they occur during run-time.

Section 2 reviews pre-existing studies with **Section 3** describing the ABM developed for this article, including the RL application of the central bank agent. The results **Section 5** defines the experiments conducted and the subsequent outcomes. Lastly, a discussion and conclusion **Section 6** discusses the findings from the experiments, limitations and strengths and concludes with future avenues to be explored.

2. Literature review

Economic crises sometimes take the form of debt crisis (where a government's debt increases while repayments decrease), banking crisis (when a large swathe of people withdraw their savings as confidence in the banks depletes), asset bubble burst (i.e., housing bubble bursts which leads to a sudden devaluation of houses, an example of this was the subprime mortgage credit crisis in 2007–2008 (Dou & Wang, 2014)) and balance of payment crisis (when a country cannot afford the price of imports or services). Regulatory policy is vital when a country tries to prevent or counteract an economic crisis (Malyshev, 2015), such as a central bank's monetary policy. Martin et al. (2022, p. 3) researched whether central banks can stabilise housing markets via interest rates. Researchers found that the ability of central banks to manage housing markets by increasing interest rates, which softens the demand pressure on house prices, is limited. However, they note that "central banks can significantly improve the stability of housing markets by dynamically adjusting interest rates". Researchers agree that ML can be used to support decision-makers in alleviating economic crises (Chirita, 2011; Ho, 2020; Loukis et al., 2020; Maghdid & Ghafoor, 2020; Nik et al., 2016).

RL algorithms are a subset of ML approaches which enable artificial agents to learn. An agent tries to complete a task and, in doing so, maximises its internal rewards (Sutton & Barto, 2018). Typically, these agents learn how to complete a task through trial-and-error by interacting with their environment (Kaelbling et al., 1996). RL theory was derived from empirical observations of the psychological and neuroscientific studies in animal behaviours (Mnih et al., 2016). RL has successfully demonstrated the ability of an agent to learn how to achieve long or short-term goals through interactions with the immediate environment, the reflection of one's past knowledge and decisions influenced by rewards and penalties. Many applications of RL exist, including but not limited to (Liu et al., 2020) where researchers optimise the choice of medications identifying the correct drug dosing and timing of interventions. Spatharis et al. (2019) developed a model where air traffic is managed through an RL agent that observes millions of data points and

makes optimal decisions as to when and where planes should land.

Most RL applications in the housing market domain are related to “house price forecasting” and prediction techniques (Chen et al., 2017; Zhan et al., 2020). Some studies have integrated deep neural networks to investigate housing markets, given the recent growth in data from websites like Craigslist, Rightmove, and Gumtree. Researchers trained a neural network using textual data to identify how the rental market dynamics were changing (Zhou et al., 2019). Similarly, researchers implemented neural networks, to classify physical and socio-demographic characteristics, to assess how interrelated these factors are in the housing market of Budapest, Hungary (Norwegian, 2007). An article developed an early warning system that identified market volatility from house price training data (Park & Ryu, 2021). A drawback of this approach was that rich data sources are usually placed behind paywalls, and the neural network would have to be trained every time new data was accessible. In our research, the ABM of the housing market acts as a continuous data stream. Most importantly, in our approach, we can artificially introduce shocks (crashes) to the system to speed up learning, whereas market shocks are rare events in the real world.

Researcher articles such as Yamaguchi et al. (2018) show how RL identifies specific behaviours worms possess pre and post-feeding. Sali et al. (2021) used RL to deal with the feature selection problem, where researchers identified the most accurate and optimal features for reducing computation costs. As evidenced by the limited yet critical studies above, RL can learn to identify a particular phenomenon/pattern in data and develop effective interventions using neural networks to achieve a particular goal. Such as identifying the correct dosage for a patient’s medication (Jalalimanesh et al., 2017). Compared to the above studies, examples of applied ABM and RL in housing market research are rare (only four articles with the terms “housing market”, “reinforcement learning”, and “agent-based”, source Web of Science). The articles (Cincotti et al., 2005; Suzuki et al., 2014; Zhou, Wu, et al., 2017) utilise RL as an optimisation method to identify the most efficient strategies in power-to-power (P2P) sharing of energy between households and companies. Kang et al. (2019), on the other hand, uses data assimilation and RL to fit real-world Korean housing market data to an ABM. In the light of these advances, this article contributes to the literature by integrating an RL decision-making algorithm in a housing market ABM focusing on shocks. It is worth noting that the work proposed here is purely experimental at this stage and acts as a proof of concept.

In this article, the artificial “central bank” agent observes data streams from the housing market ABM (Olmez, 2022) and makes dynamic decisions that impact the market (such as raising, holding or reducing interest rates), demonstrating how RL can be used to stabilise a market effectively in real-time in simulation. The opportunities for using RL and ABM are considerable. For example, this article demonstrates how RL can support decision-making in stabilising the housing market during volatile times. However, in future studies, it may be used to identify early signals of a recession or a financial crisis and alleviate the negative impact of exogenous shocks such as pandemics.

3. The housing market model

The housing-market model simulates the characteristics of the UK housing market. The model contains agents that are either buyers, sellers, estate agents or houses. An aggregate distribution of these agents interact in the environment where agent-environment and agent-agent interactions grow micro and macro emergent properties. The model simulates the interactions between buyers and sellers, who utilise information from local estate agents Figure 1. Buyers make offers depending on budget and successful acquisition of mortgages, while sellers depend on valuations from estate agents, who evaluate a property’s price depending on past sales and a markup.

The proposed ABM in Figure 2 is a reproduced version of (Gilbert et al., 2009). The purpose of this reproduction in the Python programming language (Olmez, 2022) was to access a broader set of novel ML algorithms and tools which the Netlogo framework was not able to harness. The model described in the following paragraphs was inspired by the works from (Gilbert et al., 2009).

3.1. Model environment

The environment generates a 60×60 grid, which can be changed depending on computing power – producing 3600 cells that can either be a house, occupied house, unoccupied space or an estate agent. Houses are randomly distributed and, depending on density, in the case of Figure 2 70% of the space is occupied. The initialVacancyRate sets the proportion of unoccupied houses at the start, making these houses available to buy. The price of these unoccupied houses follows the same rules outlined earlier. Estate agents find the highest valuation from previous sale records. The house prices are randomly distributed using a uniform random distribution, as is the case in Figure 2. Each house has a quality index calculated upon initialisation. This measure is a ratio of the average price of other houses within the locality of the constructed house’s sale price. The process mentioned above adheres to Tobler’s first law of geography (Miller, 2004) which states that nearer things

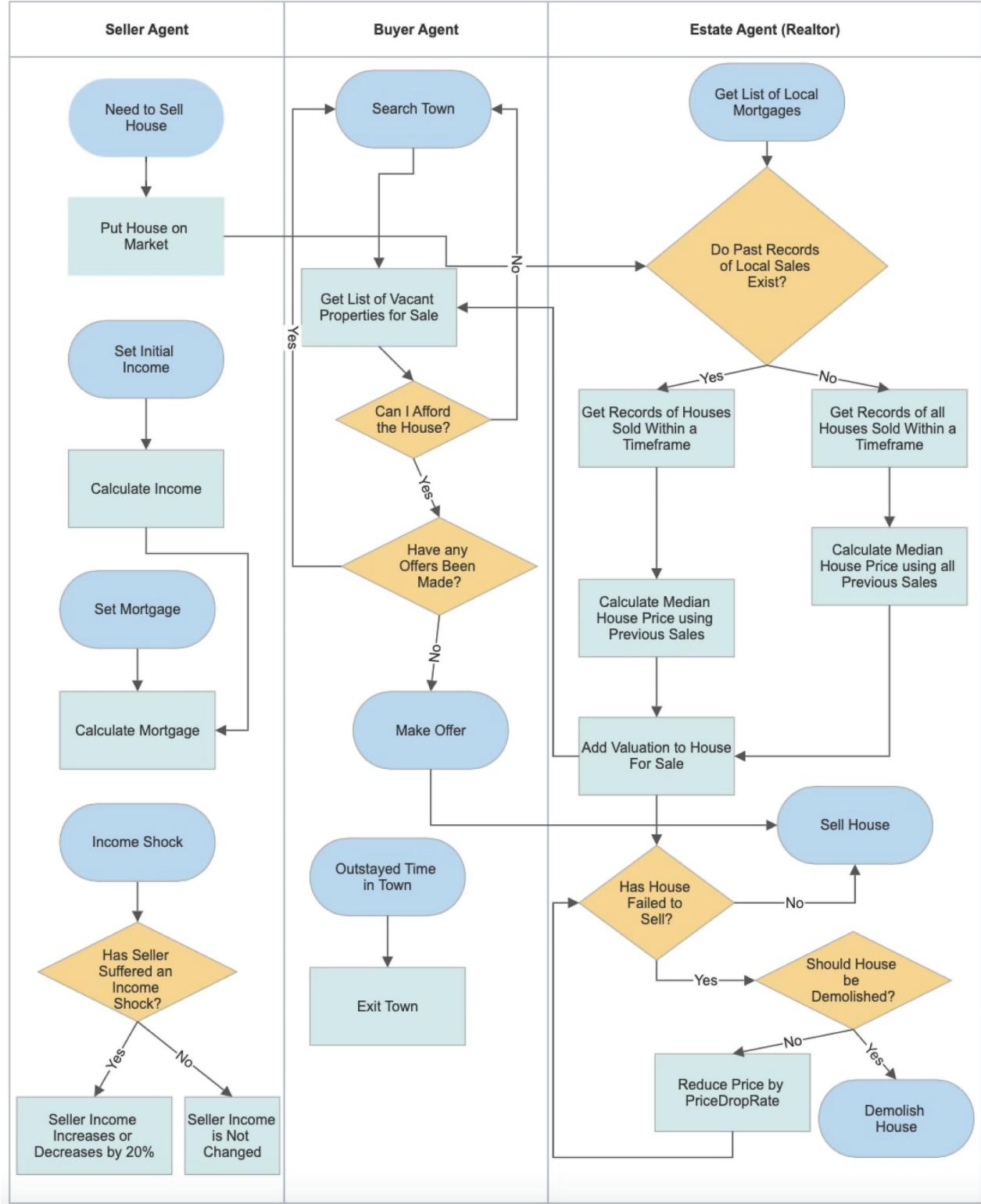


Figure 1. Flowchart presenting decisions the Seller, Buyer and Realtor (Estate) agents undertake.

are more likely to be similar than those farther apart. Every output parameter is described in Appendix A, Table A1.

3.2. Seller agents

Every step, the model moves forward in time; a step is 3 months defined by the TicksPerYear parameter Figure 2.

A percentage of homeowners (ExitRate parameter) vacate and try to sell the house at a price set by the estate agent valuation. If the house does not sell at the current timestep, it remains on the market for the next period. Every homeowner agent has an initial income determined randomly using a gamma distribution from parameters 1.3 and 5×10^{-5} multiplied by the MeanIncome parameter. Furthermore, a mortgage is calculated by the

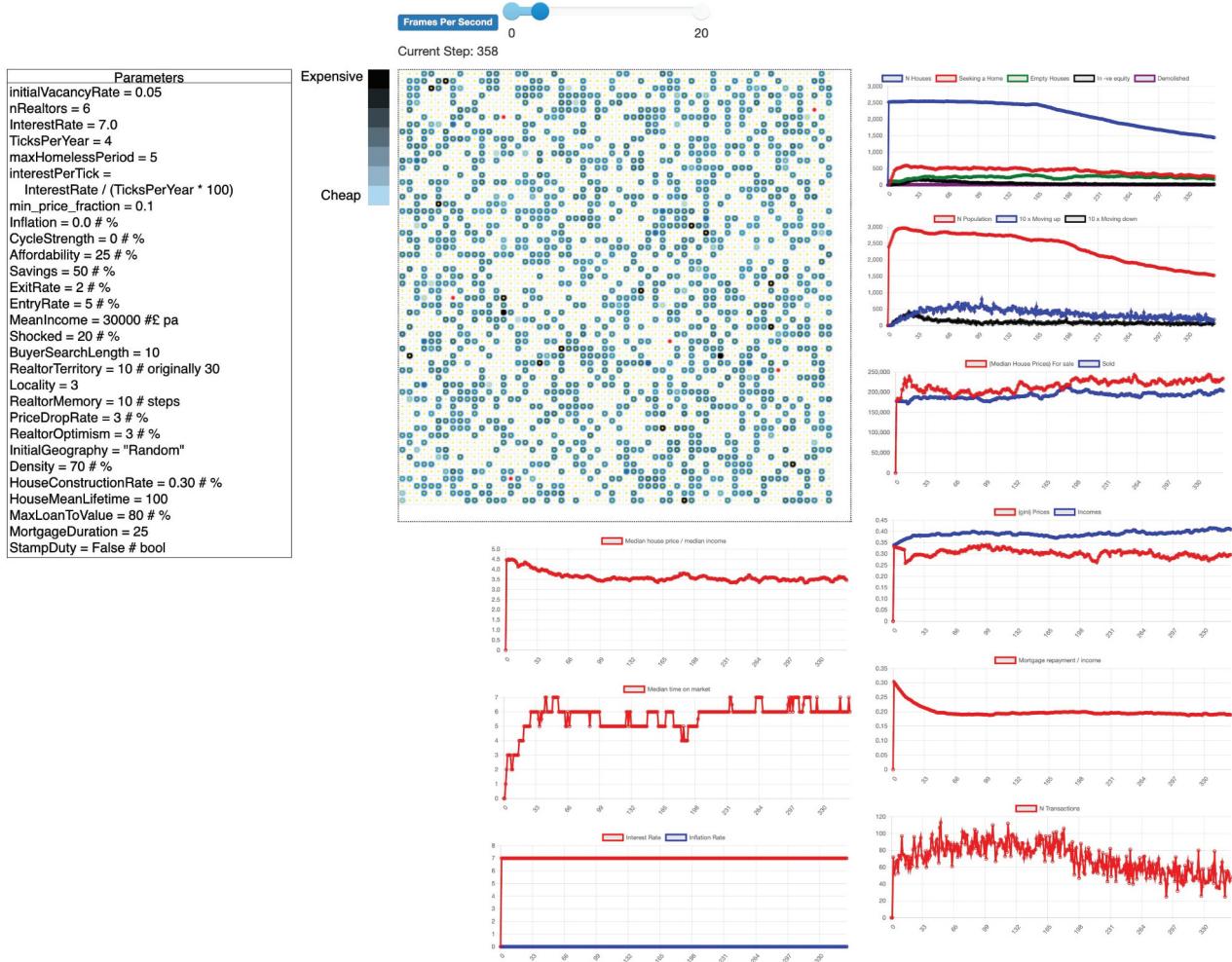


Figure 2. The user interface of the model, the parameters that can be changed on the left, the visual representation of the ABM in the centre where the small squares represent houses. Yellow dots are occupants, red dots are estate agents, white grid cells represent free space. Output plots are on the right.

ratio of the Affordability parameter, divided by interestPerTick multiplied by the owner's income. Initially, the mortgage duration is 25 years. However, this can be changed. People borrowing money must have some deposit from their capital determined by MaxLoanToValue parameter and their mortgage. At every step, a percentage of homeowners suffer income shocks determined by the Shocked parameter, which is +20%, and the same percentage suffers a shock of -20%. This leads to some homeowners having their income increase or decrease by this percentage permanently. When the ratio of the mortgage repayment is higher than twice the affordability, the homeowner trades down. Conversely, they trade up when the ratio is less than half the affordability.

3.3. Estate agents

The term estate agents is used interchangeably with realtors. Every realtor agent has a coverage radius called the RealtorTerritory. Any house outside a realtor's territory is assigned the closest realtor

calculated by the Euclidean distance. Each realtor keeps records of the previous sale. These records contain the following information: record ID, the house sold, selling price and date of sale. At the start, the mortgage value of each house is sent to one local realtor, providing realtors with a starting point for their valuations (Gilbert et al., 2009). When a seller asks for a valuation, the realtor looks through their records within the last RealtorMemory timesteps and gathers all the house prices of houses sold locally multiplied by the quality index of these houses. It then calculates the median house price of these previous sales as a valuation. If no sales have been made within the locality and period, any past sales made within the locality are considered regardless of time. Every valuation made is increased by the RealtorOptimism percentage, allowing realtors to try to sell a house more than the going rate. Lastly, if a house fails to sell at timestep N, the selling price of this house is reduced by PriceDropRate %, and it remains on the market for N + 1 until it is sold or demolished.

3.4. Buyer agents

At each timestep, people arrive in town. The amount depends on the EntryRate parameter, which is a percentage of the current population [Figure 2](#). New entrants and sellers who remain search the whole town for several timesteps defined by BuyerSearchLength parameter. Looking for vacant properties for sale that they can afford and which no offers have yet been made. Any accumulated capital from buying and selling can be put towards the costs of the new property. Subsequently, buyers choose the nearest property in price to their maximum budget and make an offer at a price set by the seller. The first buyer to make an offer has their offer accepted.

3.5. Sales

A sale is only successful if the chain of buyers and sellers remains intact. A successful chain can only occur if the house being bought is either empty or the seller succeeds in purchasing a new house and moving to that house. The people leaving town move out, and potential buyers move into these vacant properties if their offers are successful. Once all sales down the chain are complete, the model moves forward one step in time. When a house is sold, the seller receives the sale price and uses as much of it as necessary to pay off any remaining mortgage. If money is in excess, this is added to capital and can be used as a partial or complete payment of the house being bought. Conversely, if the sale price is less than the amount remaining to pay off the mortgage, the seller is in negative equity and withdraws the house from the market. The estate agent records successful sales and uses these records (as discussed above) to value houses within the same area. Finally, if an offer falls through, it lapses ([Gilbert et al., 2009](#)).

3.6. Building new houses

New houses are constructed at random empty grid cells at every timestep. The number of houses depends on the HouseConstructionRate [Figure 2](#) % of the total number of constructed houses unless there are no empty cells.

3.7. Demolition of houses

Every house has a lifetime set when it is created. This is drawn from a random exponential distribution with a mean of HouseMeanLifeTime. When a house reaches its lifetime, it is demolished, and the cell becomes vacant and available for new construction. If a house's sale price falls below one-tenth of the median price of all houses, it is demolished. If

someone occupies a house that is being demolished, they attempt to purchase a new home, and if they fail after MaxHomelessPeriod, they leave town.

3.8. Model outputs

The data produced from the model are presented below. Visually, houses are assigned a colour that reflects their current value. The lighter the shade, the cheaper the house and vice versa [Figure 2](#). The quantitative model outputs are the following:

- Number of houses, empty houses, and demolished houses. Number of people searching for a home, the number of people occupying a home in negative equity, and the number of transactions.
- The number of people in the model.
- The median house price of houses for sale and sold.
- The Gini index of the median house prices and median incomes.
- The ratios between median house price to median income, and mortgage repayments to median income.
- The mortgage interest rate, inflation rate and median time houses have spent on the market.

3.9. Reinforcement learning agent

RL allows agents to learn without explicitly telling the agent what the task is or how it is completed. A feedback reward allows the agent to learn through trial-and-error by performing actions for each state in the environment. If the reward is positive, the agent has enacted a desirable action. If the reward is negative, the action is undesirable ([Sutton & Barto, 2018](#)).

Given how well policy-gradient methods have performed ([Agarwal et al., 2020](#); [Schulze et al., 2017](#)), this was an applicable approach. Put simply, we denote a policy as π , where $\pi\theta(a|s)$ is the probability of taking action a in state s and θ are the parameters of our policy. Our goal is to update θ_t to θ_{t+1} such that we reach the optimal policy. In our model, the optimal policy would be the state where the “healthy housing market” criteria (described below) are met. If we assume a^* is the optimal action, i.e., raise interest rates by 0.01 at time t , then we want to perform gradient ascent on $\pi\theta(a^*|s)$ (ascent as we want to increase our cumulative reward). Therefore, at each iteration, we update θ in the following way $\theta_{t+1} = \theta_t + \alpha \pi_{\theta_t}(a^*|s)$ this can be described as we keep “pushing” towards more of action a^* in our policy, which is indeed what we want as raising the interest rate by 0.01 will mean we are closer to achieving our “healthy housing market” criteria.

This article proposes a novel application of RL to identify and counteract market shocks in the housing market in real-time in simulation. Several steps were taken to integrate RL with the housing market ABM:

- (1) Re-producing the well-known housing market (Gilbert et al., 2009) in a new framework (Olmez, 2022) to use as our experimental sandbox.
- (2) Replicating two experiments originally described and subsequently investigated in (Gilbert et al., 2009, p. 5) known as the loan-to-value experiments where, the MaxLoanToValue parameter is set to 80% and 100% respectively. During these experiments, an exogenous shock known as “ratefall” where a sharp increase 7% to 10% in interest rates is triggered. These two experiments were selected, as in the original article, Gilbert et al. demonstrated how the impact varying interest rates on the market were prone to being less sensitive when loan-to-value was reduced compared to loan-to-value being 100%. These experiments are poised to test RL’s ability to adapt its behaviour in two similar initial conditions but with very different outputs.
- (3) Training the RL agent on the housing market scenarios over 100 episodes, this process can be observed in Figure 3.

Due to computational complexity of training RL agents with neural networks (Baker et al., 2019; Juliani et al., 2018; Olmez, Birks, et al., 2022; Sutton & Barto, 2018), the experiments are kept concise. It is worth noting that the proposed model is a proof-of-concept used to demonstrate how the adaptive

qualities of cognitive models such as RL are suited to modelling housing market dynamics and supporting decision-making to counteract shock events.

This article models simplistic behaviours of a central bank agent. As discussed in Section 2. In reality, a central bank has more policy tools and goals to achieve beyond housing market stability. This simplicity is necessary to demonstrate a proof-of-concept. In future research, these behaviours will become more advanced.

To train the RL agent, we first identify the healthy housing market indicators. This way, the agent can learn to differentiate between an undesirable state and a desirable one. A healthy housing market is characterised by stability, affordability, and accessibility Maliene2008SustainableRelations (Cai & Lu, 2015; Maliene et al., 2008). Stability is indicated by small fluctuations in house prices, suggesting a balance between supply and demand (Zhu, Betzinger, et al., 2017). Affordability is often affected by the house price to income ratio, with a lower ratio reflecting better affordability for the average household. Accessibility can be indirectly measured through the prevalence of negative equity, as excessive negative equity rates can imply barriers to entry or exit from the housing market.

In line with these principles, we have set the following indicators to define a healthy housing market within our model:

- Stable median house prices for sale with small fluctuations up to $\leq 400,000$.
- Median house price to income ratio ≤ 7 .
- Number of people in negative equity is $\leq 5\%$
(123) where $N_people = 2466$.

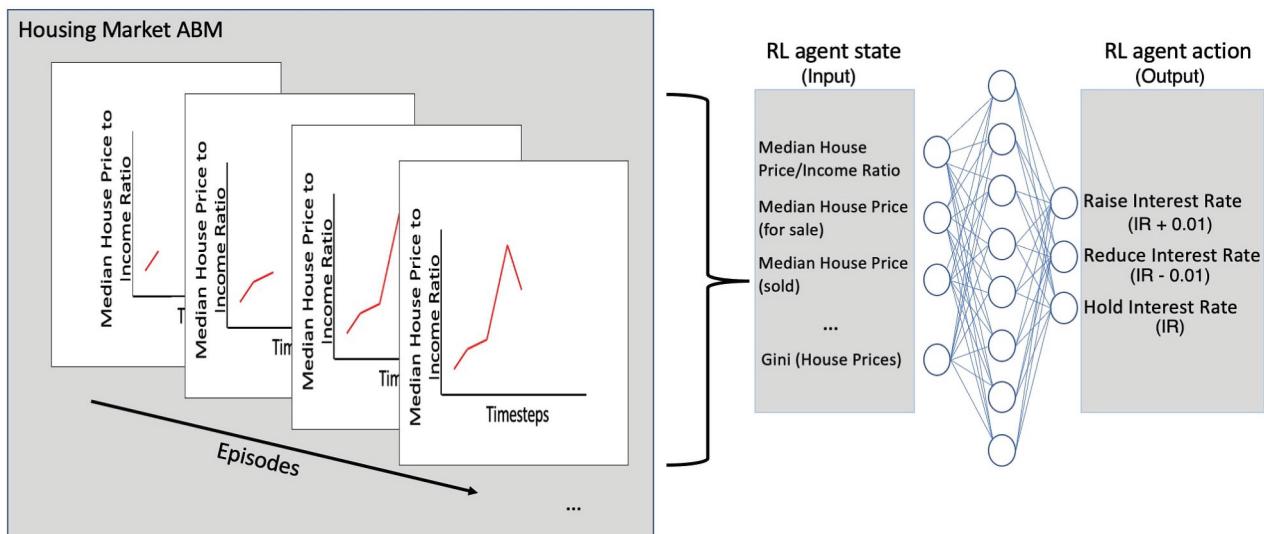


Figure 3. RL central bank agent neural network, that determines the central bank agent’s decision regarding interest rates, in the current instance, an action with high probability may be to raise interest rates as a sharp increase in house price to income ratio is observed.

If the above conditions are met, we have a desirable state (reward returned to the central bank agent, 0). If all but one of these conditions are not met, we are in an undesirable state (reward returned, -1).

The results from the RL process are presented and compared to base case scenarios in the following section. The RL outputs illustrate the central bank agent's learning process during training and how the RL agent adapts to the LTV scenarios. We compare findings to those discussed in the original article to demonstrate how RL has or has not benefitted the housing market in alleviating shocks and fulfilling goal criteria.

3.10. Training methodology

The RL agent was trained using an offline training methodology, as this allows for controlled and repeatable experiments essential for rigorous scientific evaluation. Training involved the following steps:

- (1) The agent observed state-action pairs from a pre-simulated dataset generated by the ABM representing the housing market dynamics over multiple historical scenarios.
- (2) Based on the feedback (rewards) calculated from these scenarios, the agent learned optimal policies using the Proximal Policy Optimisation (PPO) algorithm, aimed at maximising long-term rewards.
- (3) The training phase consisted of 100 episodes, with each episode representing a complete simulation run from start to finish of the housing market model.
- (4) Convergence of the agent's learning was determined by a consistent increase in rewards and stabilisation of policy outputs across episodes.

Post-training, the agent was tested in a simulated environment that replicated emerging housing market trends to evaluate its real-time decision-making capabilities.

4. Model validation

The proposed model (Olmez, 2022) was reproduced by interpreting (Gilbert et al., 2009) source code, refer to Appendix B, Figure B1 for class diagram. The behaviour of our model must be compared and deemed similar to the original, whereby model outputs produce similar trends in data. If the model outputs differ, we have deviated from the original at some point in the development. Suppose the model produces similar trends in the data outputs. We can be confident in our model's behaviour in producing

realistic trends of the UK housing market. Model replication is an important topic in the ABM literature, as discussed by (Donkin et al., 2017, p. 1) "model replication remains rare, yet is vital to assessing the repeatability of existing ABMs".

Two housing market scenarios were simulated for both models using input parameters in Appendix A, Table A2. In scenario one, no shock is introduced, and in scenario two, a shock is introduced mid-simulation run. This allows us to compare behaviours in two unrelated scenarios to quantify two completely different outcomes.

To compare both models, we adopt a visual statistical approach known as quantile-quantile (Q-Q) plot, the benefits of which have been thoroughly discussed in the following literature (Dhar et al., 2014; Oldford, 2016). Where, two probability distributions are compared using their quantiles. In our case, one variable in (Olmez, 2022) is compared to the same variable in (Gilbert et al., 2009). Furthermore, over 100 model runs over 100 simulation years are drawn for each scenario providing a large sample size to quantify the stochasticity produced and output-variability (Bogdolla et al., 2012; Lelei & McCalla, 2019). A one-degree gradient (45°) reference line is plotted to compare variables. If $x = y$, each point sits on the reference line, then both variables compared are identical and vice versa.

Due to the large volume of output variables (18) Figure 2. We select a sub-set of Q-Q plots to use in the results. These include parameters that capture the housing market's health, for example, the median house price to income ratio and the median price of houses for sale.

Figure 4 presents (Gilbert et al., 2009) on the x-axis and (Olmez, 2022) on the y-axis. Each column and row represents the scenario and variable respectively. These results show a normal distribution and results are correlated which quantitatively replicate similar trends, that sometimes deviate due to stochasticity, such as (Olmez, 2022) overestimating (Figures 4(c,f)) or underestimating (Figures 4(g,h)). Note, due to model architectures, frameworks and other factors, models cannot be replicated perfectly as highlighted by (Donkin et al., 2017; Yingfei, 2009).

The descriptive statistics Appendix B, Figure B2 show that 79% of the output variables returned a positively correlated Pearson's correlation coefficient, where 11 of these from scenario two had $r > 0.60$ which were all statistically significant $p < 0.01$. In scenario one, eight variables had $r > 0.60$ and $p < 0.01$.

In scenario two, all variables returned an $0.659 \leq r \leq 0.992$, all statistically significant, Figure 4. In scenario one, $0.235 \leq r \leq 0.839$ four variables were statistically significant. The only variable that was not statistically significant was the mean

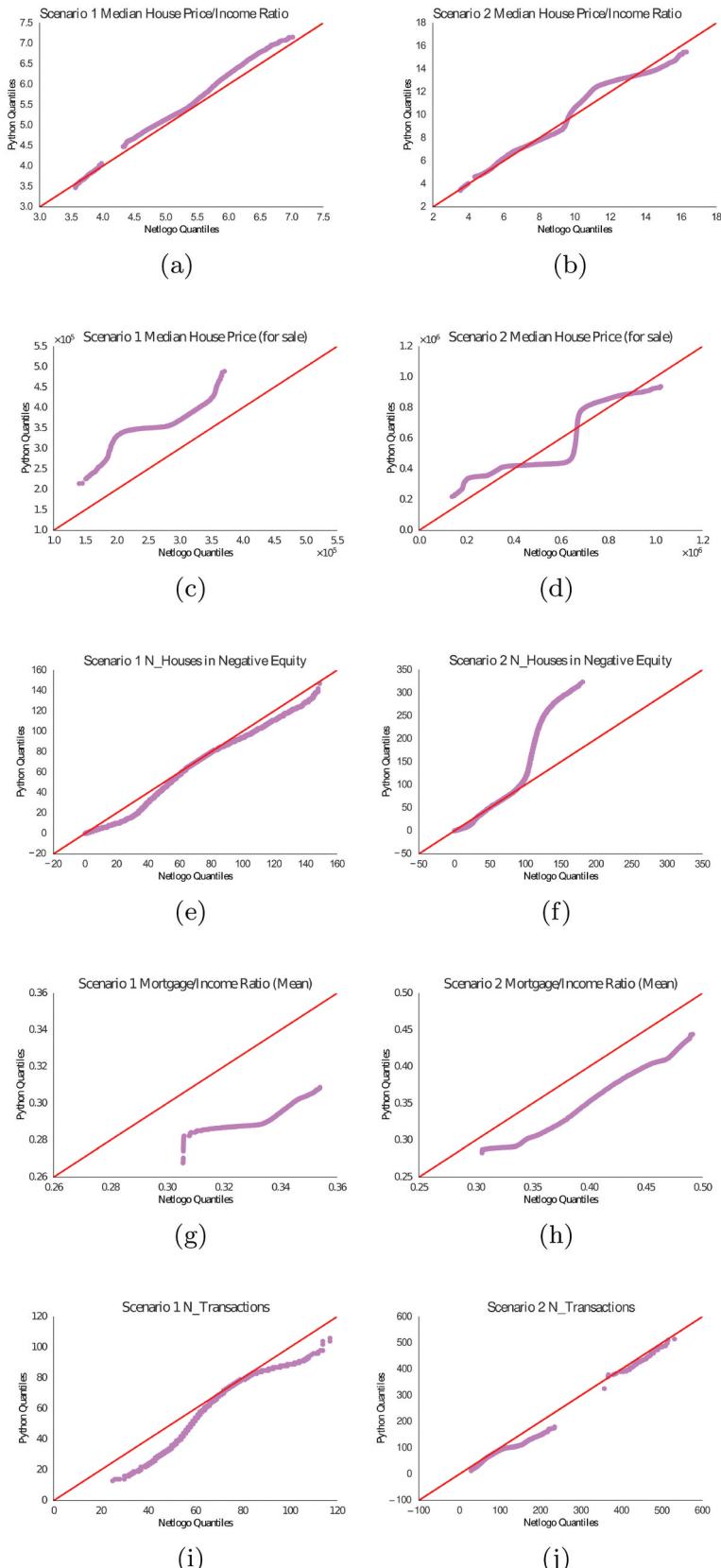


Figure 4. Q-Q plots comparing the distributions of model output variables. The solid line indicates $x = y$ for reference. Where (a-b): Median house price to income ratio (Scenarios 1–2), (c-d): Median house price for sale (Scenarios 1–2), (e-f): Number of households in negative equity (Scenarios 1–2), (g-h): Mean mortgage to income ratio (Scenarios 1–2), (i-j): Number of transactions (Scenarios 1–2).

mortgage-to-income ratio for scenario one. The most likely reason could be stochasticity, where the algorithm has more steps to process (Donkin et al., 2017).

Overall, the proposed model demonstrates the UK housing market characteristics observed in the original (Gilbert et al., 2009). Trends develop when external tweaks to the market are made, showing that indicators are sensitive to these changes in both models. The next stage is to integrate RL, to test whether an intelligent observer agent can learn to identify shocks to the market and deploy countermeasures to minimise their effects in real-time in simulation.

5. Results

This section describes the market shock, provides an overview of central bank decisions to deal with shocks, and outlines the healthy housing market conditions. This is followed by a detailed analysis of findings from original housing market research conducted in 2008 (Gilbert et al., 2009) emphasising the experiments we aim to conduct in this article to gauge the strengths and weaknesses of RL. Lastly, experiments and subsequent results show how intelligent RL agents can learn and make decisions in real-time in simulation to counteract induced shocks. Note that actions the RL agent can undertake simplify how decisions are made in the real world. We aim to explore the strengths and weaknesses of this methodology in a simplified housing market, hoping to identify the potential for future applications.

Shocks come in various forms, as described in Section 2. To manage shocks, central banks often enforce a monetary policy which stabilises and counteracts the aftermath of the shock (Martin et al., 2022). From 2007 to 2009, the world endured

a financial shock resulting in a crash in the UK housing market (Whitehead & Williams, 2011). In Figure 5, the house price to income ratio in England and Wales dropped from 7.17 in 2007 to 6.35 in 2009. Furthermore, a drop of 18.7% in house prices (Munro, 2018), from Q3 2007 to Q1 2009. To counteract the pressures, the central bank reduced interest rates from 5.25 in Jan 2007 to 1.50 in Jan 2009 (Tse et al., 2014). In contrast, a healthy housing market trend may look like Figure 2, where mortgage repayment to income ratio is 20%, and the median house price to income ratio oscillates between 3.5 and 4.0. House prices increase gradually as people move in and out of the market. The number of people in negative equity is as small as possible (Been et al., 2021; Melzer, 2010; Morescalchi et al., 2018).

Gilbert et al. (2009) conducted several experiments using their UK housing market model in 2008, where the model showed how properties of the UK housing market are emergent. Some crucial findings that aligned with empirically observed behaviours of the UK housing market were:

- House price to income ratio showed a stable relationship given mortgage interest rates and the loan-to-value ratio. For example, if interest rates are reduced or loan to value increased, house price to income rises in response.
- When the loan-to-value ratio is 100%, and the market experiences an exogenous interest rate hike from 7% to 10%, a sharp drop in the house price-to-income ratio is observed. However, if loan-to-value is set at 80% and the same increase in the interest rate is observed, the effect is much weaker (Gilbert et al., 2009, p. 5).

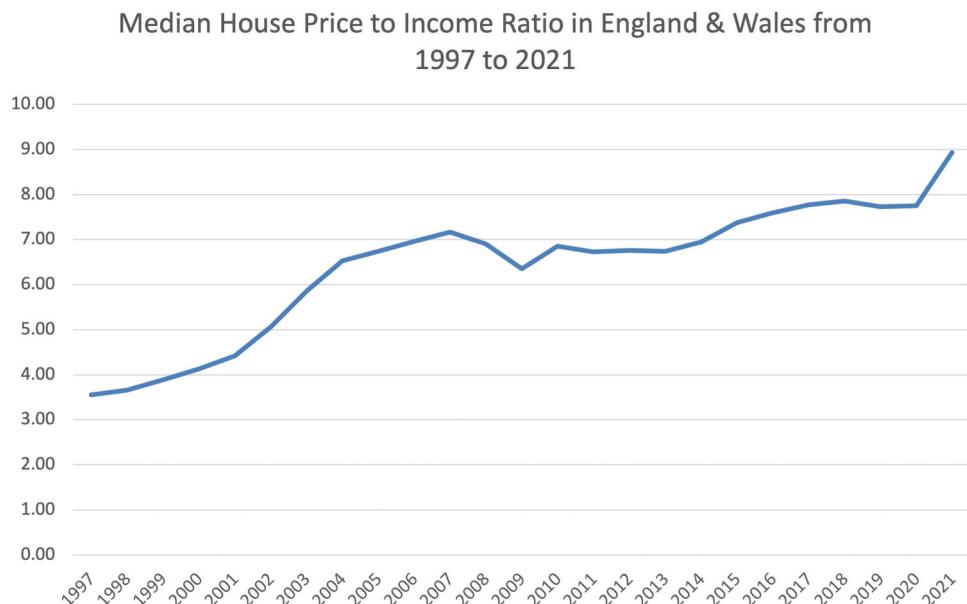


Figure 5. The median house price to income ratio in England and Wales from 1997 to 2021 (source: (Housing Team, 2022)).

These findings were also explored in other housing market research, such as (Narayan & Narayan, 2011; Tse et al., 2014; White, 2015). The experiments (Gilbert et al., 2009) make for a well-documented comparator for this article, where the strengths and weaknesses of RL can be tested in relation to the earlier assertions. Given these original experiments and results, in this article, we expect RL to behave in a certain way when adjusting interest rates in the 100% LTV scenario compared to the 80% scenario. The goal is to test if the RL central bank agent can adapt to these scenarios and fulfil its goals. In the next paragraph, we describe the experiments in detail.

We replicate two experiments; these are “loan-to-value A and B”. In the A experiment, the MaxLoanToValue parameter (refer to Table A1) is 100, and in the B experiment, it is set to 80. In both experiments, an exogenous shock occurs at timestep 200, where mortgage interest rates suddenly increase from 7% to 10%. In the base case experiments (no RL),

we observe findings from the (Gilbert et al., 2009). In the RL experiments, we observe behavioural differences and consequences of actions taken by the RL agent in achieving the “healthy housing market” criteria. As our results show, we believe there is value in utilising these contemporary methods to support future research in modelling housing markets.

Given the computational complexity in these models and the nature of RL training, we run the experiments for 100 iterations to capture a distribution of results quantifying model stochasticity. Moreover, during training, we found that the “healthy housing market” criteria were met.

To recap, the experiments were run for 400 simulation timesteps. At 200th timestep, a shock impacts key indicators to varying extents as observed in (Gilbert et al., 2009, p. 4) and presented in Figure 6. The ratefall shock severity is greater in 100% LTV compared to 80% LTV; this can be observed in Figures 6(g,h). Similar behaviour is observed for median house prices

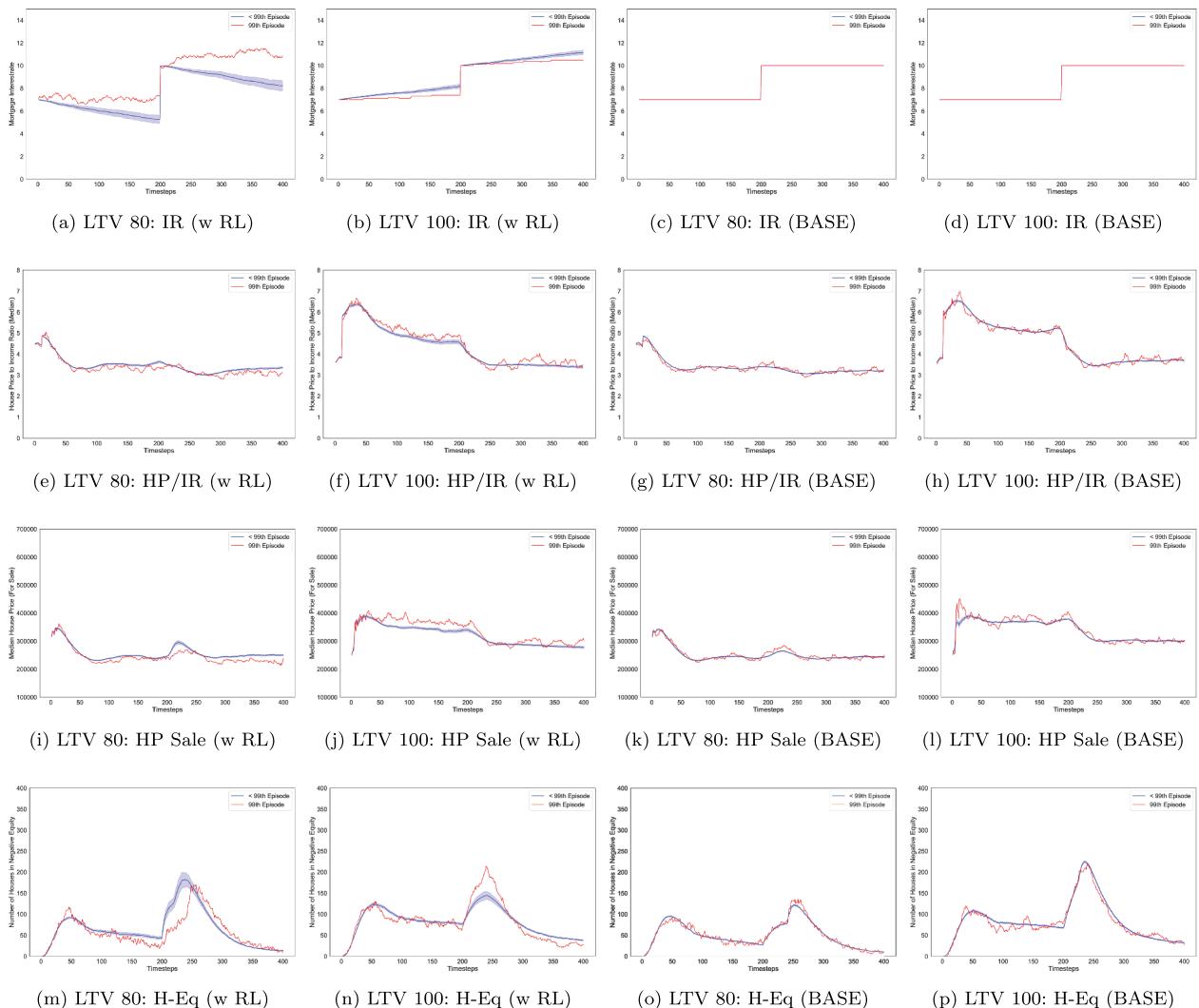


Figure 6. Line graphs showing the last model run (99th due to index starting at 0) and the average with a confidence interval for all previous runs (< 99) aggregated for each experiment condition, including base case conditions. Each row is a tracked variable, and the column is the experiment, where IR = Interest Rates, HP/IR = House Price to Income ratio and H-Eq = Houses in Negative Equity.

for sale and the number of households in negative equity ([Figures 6\(k,l,o,p\)](#)).

Interestingly, the RL agent learns to approach the 100% LTV experiment conservatively with slight adjustments but mainly holding interest rates evidenced in [Figure 6\(b\)](#) compared to the 80% experiment [Figure 6\(a\)](#). This is a response to the sensitivity of indicators of the housing market to prevailing interest rates. This particular finding demonstrates the adaptive capabilities of RL, where slight environmental changes in a model can be responded to by learning and experiences. Furthermore, on average the RL agent reduces interest rates for the LTV 80% scenario (see [Figure 6\(a\)](#)), conversely, in the 100% LTV scenario, it increases interest rates on average (see [Figure 6\(b\)](#)).

Another finding from [Figure 6](#) when analysing RL behaviours, in particular, is that leading up to and right after the shock at timestep 200, the confidence intervals are much wider (further from the mean). This means the RL agent has explored more state space at these crucial points during training. This behaviour can be observed most clearly in [Figures 6\(a,b\)](#). In the 100% LTV experiment, the variance statistic for interest rates is 2.691 (mean 9.095, std 1.640) compared to the 80% experiment 4.521 (mean 7.607, std 2.126), where RL agent is exploring more and subsequently adjusting interest rates more often in the 80% compared to the 100% LTV experiment.

RL trains iteratively across several simulation runs, also known as episodes. The latest episode in the training phase is the most recent behavioural output, usually representing the stage at which RL has learned the most optimal set of behaviours to achieve some goal. Conversely, the most recent episodes are those in which RL is yet to learn effective behaviours (Sutton & Barto, 2018). The data observed at the 99th model-run (episode) are the outputs for when the RL was most trained. Therefore, we compare these to our goal criteria for a healthy housing market state. For some indicators, the RL agent was better at achieving the healthy housing market goals than others, which would be expected as some indicators are more sensitive to interest rates than others. For both loan-to-value experiments, RL successfully ensured the house price-to-income ratio was below seven even after the ratefall shock, refer to [Figures 6\(e,f\)](#). The median-house-price-for-sale indicator shows that RL was more effective in achieving the $\leq 400,000$ goal in the 80% experiment ([Figure 6\(i\)](#)) compared to the 100% experiment ([Figure 6\(j\)](#)) which was only above 400,000 for a short time at the earlier timesteps from 0–100. A similar outcome was also observed in the base case [Figure 6\(l\)](#). The data show that the houses-in-negative-equity was a more complex indicator for RL to achieve the goal of ≤ 123 where the shock exacerbated the complexity as presented by the wide

confidence intervals post-shock [Figures 6\(m,n\)](#). However, pre and post-shock, we observe a downward trend where the number of households in negative equity is less than 123, even achieving less than 50 near the end of the simulation. It is worth noting that while differences exist between the RL and base case scenarios, these can be considered small. However, this can result from the chosen healthy housing market goal conditions, and differences may be more significant if other goal conditions or a combination of conditions were chosen. These results also demonstrate the housing market's ability to settle after a shock.

Overall, these experiments show that RL can achieve healthy housing market goal criteria and alleviate the shock effect on the market, which varies in effectiveness across the different indicators. Given these results, we presume that as the number of goal indicators increases from 3 to $3 + i$ for some i , complexity in achieving these goals also increases. Given this complexity, we believe that RL's housing market goals may be unachievable at some point. This may be due to equilibrium whereby increasing one indicator, the goal is met, but another indicator is reduced; thus, the goal is not met and vice-versa. This study has demonstrated that RL is a valuable technique that should be welcomed by housing-market and macroeconomic researchers interested in utilising autonomous decision-making methods to aid policy making in dealing with uncertainty like economic shocks. In the next section, we break down the learning process of RL and describe the strengths and weaknesses in utilising RL within this domain.

6. Discussion and conclusion

This article reproduces a well-known ABM of the UK housing market to integrate a novel RL algorithm that learns to counteract housing market shocks in real-time. We answer the following research questions: can RL be integrated with housing market ABMs? Moreover, can agents learn trends from the housing market and adapt to economic shocks by counteracting the impact of these shocks in real time? Findings show RL can be integrated with housing market ABMs, as evidenced in sub-Section 3.9 and does well, in learning to counteract shocks through monetary policies such as interest rate adjustments.

This article shows how RL could adapt its behaviours and, over time, through training, learn behaviours that enable it to achieve the goal state. Furthermore, the RL agent portrayed characteristics of the original model (Gilbert et al., 2009). One example is the effects of interest rates in the 80% loan-to-value (LTV) compared to the 100% loan-to-value environment. Responding to the impact of interest rates being more sensitive in the 100% LTV case

compared to 80%. The RL agent learned to explore a greater range of interest rates in the less-sensitive scenario (80%) compared to the more sensitive scenario (100%) intended to counteract the market's sensitivity in these two conditions, which was purely learnt and not hard-coded.

A drawback of our approach is that the RL agent's tools are limited. This is not indicative of a real-world central bank, which has more policies to counteract crises, such as regulatory, monetary, and fiscal policies. However, macroeconomic policies such as adjusting interest rates is a critical intervention central banks make (Martin et al., 2022; Popescu, 2014; Valadkhani et al., 2019) with the most recent example tackling inflation in the UK (Inman, 2022; Luhnow & Colchester, 2022; Lynch & Adam, 2022). Another caveat is that our model is, a simplified version of the real world. This would ensure computational tractability. Thus, this article only focuses on a single policy tool to demonstrate the application of RL in the housing market research domain and is exploratory in nature.

There are several weaknesses in RL methods, including overfitting, the exploration-exploitation trade-off (Sledge & Principe, 2017), and computational demand. To address the exploration-exploitation issue, we used an objective function that was not greedy but balanced both aspects (Silver et al., 2014; Sutton & Barto, 2018). While computational demand was not a problem for our model, it could become an issue for more complex environments with more agents and action spaces. In these cases, advanced computational resources may be required.

There are several exciting directions for future research based on this work. For example, the findings can support housing market modelling, where researchers forecast the potential for exogenous shocks and identify policy decisions to alleviate economic downturns. The technique can also be adapted to simulate a realistic case where the goal is to optimise the current state of the housing market through RL. Additionally, given the recent release of the 2021 UK census, researchers can enhance the model with this data to study the dynamics of the housing market. This is the first example in the literature that uses novel RL algorithms within housing market agent-based models to develop a methodology for autonomously counteracting exogenous shocks to the market. There is also value in exploring this application in macroeconomics, where artificial intelligence-assisted policy-making and signal detection can have a significant impact. For example, it may help a central bank to detect a recession or financial crisis and take action early.

Notes

The RL model and associated agent-based model (ABM) code can be freely accessed at the following repositories

respectively: (Barhate, 2021) and (Olmez, 2022), which includes detailed documentation and instructions for setup and execution. The datasets utilised for training and validating the model are publicly available at <https://doi.org/10.6084/m9.figshare.21719879.v1>, which includes comprehensive metadata and access guidelines.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This document is the result of research funded by the Economic and Social Research Council (ESRC), grant numbers: ES/P000401/1 and ES/R007918/1, UK Prevention Research Partnership (UKPRP) MR/S037578/2, Medical Research Council MC_UU_00022/5 and Scottish Government Chief Scientist Office SPHSU20.

References

- Agarwal, A., Kakade, S. M., Lee, J., & Mahajan, G. (2020). Optimality and approximation with policy gradient methods in markov decision processes. *Conference on learning theory*, Graz, Austria.
- Almahamid, F., & Grolinger, K. (2021, 9). Reinforcement learning algorithms: An overview and classification. *Canadian Conference on Electrical and Computer Engineering*, 2021- September, Toronto, Canada.
- Axtell, R. L. (2014). *An agent-based model of the housing market bubble in metropolitan*. SSRN.
- Bae, J. W., Paik, E., Dongoh, K., Jung, J., & Lee, C. H. (2019, 1). Simulation framework for self-evolving agent-based models: A case study of housing market model. *Proceedings - Winter Simulation Conference*, Gothenburg, Sweden, 2018–December, (pp. 1120–1131).
- Baker, B., Kanitscheider, I., Markov, T., Wu, Y., Powell, G., McGrew, B., & Mordatch, I. (2019). Emergent tool use from multi-agent autocurricula. *arXiv*. <https://doi.org/10.48550/arXiv.1909.07528>
- Baptista, R., Farmer, J. D., Hinterschweiger, M., Low, K., Tang, D., & Uluc, A. (2016, 10). Macroprudential policy in an agent-based model of the UK Housing Market. *SSRN Electronic Journal*. <https://papers.ssrn.com/abstract=2850414>
- Barhate, N. (2021). *Minimal PyTorch implementation of proximal policy optimization*. GitHub. <https://github.com/nikhilbarhate99/PPO-PyTorch>.
- Been, V., Ellen, I., Figlio, D. N., Nelson, A., Ross, S., Schwartz, A. E., & Ellen, A. (2021). *The effects of negative equity on children's educational outcomes*. NBER. <http://www.nber.org/papers/w28428>
- Bogdoll, J., Hartmanns, A., & Hermanns, H. (2012). Simulation and statistical model checking for modestly nondeterministic models. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7201, pp. 249–252). https://link.springer.com/chapter/10.1007/978-3-642-28540-0_2
- Cai, W., & Lu, X. (2015). Housing affordability: Beyond the income and price terms, using China as a case study. *Habitat International*, 47, 169–175. <https://doi.org/10.1016/j.habitatint.2015.01.021>

- Carstensen, C. L. (2015). An agent-based model of the housing market: Steps toward a computational tool for policy analysis. *Analysis & Policy Observation*.
- Chen, J. H., Ong, C. F., Zheng, L., & Hsu, S. C. (2017, 7). Forecasting spatial dynamics of the housing market using support vector machine. *International Journal of Strategic Property Management*, 21(3), 273–283. <https://doi.org/10.3846/1648715X.2016.1259190>
- Chiri, t'a, M. (2011). Usefulness of artificial neural networks for predicting financial and economic crisis. *Economics and Applied Informatics* 2, 61–66. <https://ideas.repec.org/a/djifseesiay2012i2p61-66.html>
- Cincotti, S., Guerci, E., & Raberto, M. (2005, 5). Price dynamics and market power in an agent-based power exchange. *Noise and Fluctuations in Econophysics and Finance*, 5848, 233. https://www.researchgate.net/publication/228675480_p_rice_dynamics_and_market_power_in_a_gent-based_power_exchange
- Dhar, S. S., Chakraborty, B., & Chaudhuri, P. (2014, 8). Comparison of multivariate distributions using quantile–quantile plots and related tests. *Bernoulli*, 20(3), 1484–1506. <https://doi.org/10.3150/13-BEJ530>
- Donkin, E., Dennis, P., Ustalakov, A., Warren, J., & Clare, A. (2017, 6). Replicating complex agent based models, a formidable task. *Environmental Modelling & Software*, 92, 142–151. <https://doi.org/10.1016/j.envsoft.2017.01.020>
- Dou, X., & Wang, J. (2014, 2). Asset securitization and bubbles: An illustration of subprime mortgage default crisis. *Advances in Economics and Business*, 2(2), 112–119. <http://www.hrpublish.org>
- Epstein, J. M., & Axtell, R. (1997, 3). Artificial societies and generative social science. *Artificial Life and Robotics*, 1(1), 33–34. <https://doi.org/10.1007/BF02471109>
- Erlingsson, E. J., Teglio, A., Cincotti, S., Stefansson, H., Sturluson, J. T., & Raberto, M. (2014, 12). Housing market bubbles and business cycles in an agent-based credit economy. *Economics*, 8(1), 2014–2022. <https://www.degruyter.com/document/doi/10.5018/economics-ejournal.ja.2014-8/html>
- Filatova, T. (2015, 11). Empirical agent-based land market: Integrating adaptive economic behavior in urban land-use models. *Computers, Environment and Urban Systems*, 54, 397–413. <https://doi.org/10.1016/j.compenvurbsys.2014.06.007>
- Fischer, T., & Riedler, J. (2014, 11). Prices, debt and market structure in an agent-based model of the financial market. *Journal of Economic Dynamics and Control*, 48, 95–120. <https://doi.org/10.1016/j.jedc.2014.08.013>
- Ge, J. (2014). Who creates housing bubbles? An agent-based study. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 8235, pp. 143–150). https://link.springer.com/chapter/10.1007/978-3-642-54783-6_1
- Ge, J. (2017, 3). Endogenous rise and collapse of housing price: An agent-based model of the housing market. *Computers, Environment and Urban Systems*, 62, 182–198. <https://doi.org/10.1016/j.compenvurbsys.2016.11.005>
- Geanakoplos, J., Axtell, R., Farmer, J. D., Howitt, P., Conlee, B., Goldstein, J. & Yang, C. Y. (2012). Getting at systemic risk via an agent-based model of the housing market. *American Economic Review*, 102(3), 53–58. <https://doi.org/10.1257/aer.102.3.53>
- Gilbert, N., Hawksworth, J. C., & Swinney, P. A. (2009). An agent-based model of the english housing market. In *Technosocial predictive analytics, papers from the 2009 aaai springsymposium, technical report ss\09\09* (pp. 30–35). AAAI. <http://www.aaai.org/Library/Symposia/Spring/2009/ss09-09-007.php>
- Groff Elizabeth, R., Johnson, S. D., & Amy, T. (2018, 2). State of the art in agent-based modeling of urban crime: An overview. *Journal of Quantitative Criminology*, 35(1), 155–193. <https://doi.org/10.1007/s10940-018-9376-y>
- Hamill, L., & Gilbert, N. (2015). *Agent-Based Modelling in Economics* (1st ed.). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118945520>
- Heppenstall, A., Evans, A., & Birkin, M. (2006). Using hybrid agent-based systems to model spatially-influenced retail markets. *JASSS*, 9(3). <https://www.jasss.org/9/3/2.html>
- Ho, N. (2020, 8). How AI can help build resiliency for small businesses in a global economic crisis. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 3606–3606).
- Housing Team. (2022). *House price to workplace-based earnings ratio office for national statistics*. <https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/ratioofhousepricetoworkp>
- Imam, I., & Kerschberg, L. (1997). Adaptive intelligent agents. *Journal of Intelligent Information Systems*, 9(3), 211–213. <https://link.springer.com/article/10.1023/A:1008672326807>
- Inman, P. (2022). *What is the Bank of England doing in bid to stabilise UK economy?* Bank of England. *The Guardian*. <https://www.theguardian.com/business/2022/sep/28/what-bank-of-england-doing-pound-dollar-uk-e>
- Jalalimanesh, A., Shahabi Haghighi, H., Ahmadi, A., & Soltani, M. (2017). Simulation-based optimization of radiotherapy: Agent-based modeling and reinforcement learning. *Mathematics and Computers in Simulation*, 133, 235–248. <https://doi.org/10.1016/j.matcom.2016.05.008>
- Jordan, R., Birkin, M., & Evans, A. (2011). Agent-based simulation modelling of housing choice and urban regeneration policy. *Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6532, pp. 152–166). https://link.springer.com/chapter/10.1007/978-3-642-18345-4_1
- Jordan, R., Birkin, M., & Evans, A. (2012, 1). Agent-based modelling of residential mobility, housing choice and regeneration. *Agent-Based Models of Geographical Systems*, 511–524. https://doi.org/10.1007/978-90-481-8927-4_2
- Juliani, A., Berges, V.-P., Vckay, E., Gao, Y., Henry, H., Mattar, M., & Lange, D. (2018). Unity: A general platform for intelligent agents. *arXiv preprint arXiv:1809.02627*.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4, 237–285. <https://doi.org/10.1613/jair.301>
- Kang, D. O., Bae, J. W., Lee, C., Jung, J. Y., & Paik, E. (2019, 1). Data assimilation technique for social agent-based simulation by using reinforcement learning. *Proceedings of the 2018 IEEE/ACM 22nd International Symposium on Distributed Simulation and Real Time Applications, DS-RT 2018*, Madrid, Spain (pp. 220–221).
- Kothari, V., Blythe, J., Smith, S., & Koppel, R. (2014). Agent-based modeling of user circumvention of security. *ACM international conference proceeding series*, Sanibel Island Florida USA.
- Lelei, D. E. K., & McCalla, G. (2019). How many times should a pedagogical agent simulation model be run? In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 11625, pp. 182–193). Springer.

- https://link.springer.com/chapter/10.1007/978-3-030-23204-7_1
- Littman, M. L. (2015). Reinforcement learning improves behaviour from evaluative feedback. *Nature*, 521(7553), 445–451. <https://www.nature.com/articles/nature14540>
- Liu, S., See, K. C., Ngiam, K. Y., Celi, L. A., Sun, X., & Feng, M. (2020, 7). Reinforcement learning for clinical decision support in critical care: comprehensive review. *Journal of Medical Internet Research*, 22(7), e18477. <https://www.jmir.org/2020/7/e18477>
- Loukis, E., Kyriakou, N., & Maragoudakis, M. (2020). Using government data and machine learning for predicting firms' vulnerability to economic crisis. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 12219, pp. 345–358). Springer. https://link.springer.com/chapter/10.1007/978-3-030-57599-1_2
- Luhnow, D., & Colchester, M. (2022). U.K.'s central banker struggles with inflation, a financial crisis and his own government - WSJ. <https://www.wsj.com/articles/u-k-central-banker-andrew-bailey-inflation-financial-crisis-1166>
- Lynch, J., & Adam, K. (2022). Bank of England intervenes to stabilize UK finances after Liz Truss budget. *The Washington Post*. <https://www.washingtonpost.com/world/2022/09/28/boe-uk-pound-intervention/>
- Maghdid, H. S., & Ghafoor, K. Z. (2020, 9). A smartphone enabled approach to manage COVID-19 lockdown and economic crisis. *SN Computer Science*, 1(5), 1–9. <https://link.springer.com/article/10.1007/s42979-020-00290-0>
- Maliene, V., Howe, J., & Malys, N. (2008). Sustainable communities: Affordable housing and socio-economic relations. *Local Economy: The Journal of the Local Economy Policy Unit*, 23(4), 267–276. <https://doi.org/10.1080/02690940802407989>
- Malyshev, N. A. (2015, 12). The importance of regulatory policy. *SSRN Electronic Journal*. <https://papers.ssrn.com/abstract=3323598>
- Martin, C., Schmitt, N., & Westerhoff, F. (2022, 3). Housing markets, expectation formation and interest rates. *Macroeconomic Dynamics*, 26(2), 491–532. <https://www.cambridge.org/core/journals/macroeconomic-dynamics/article/abs/housing-markets-ex>
- Mehta, D. (2020). State-of-the-art reinforcement learning algorithms. *International Journal of Engineering Research Technology*, 8, 717–722. <https://doi.org/10.17577/IJERTV8IS120332>
- Melzer, B. T. (2010). Mortgage debt overhang: Reduced investment by homeowners with negative equity. *Journal of Finance*, 72(2), 575–612. <https://doi.org/10.1111/jofi.12482>
- Miller, H. J. (2004, 6). Tobler's first law and spatial analysis. *Annals of the Association of American Geographers*, 94(2), 284–289. <https://doi.org/10.1111/j.1467-8306.2004.09402005.x>
- Mnih, V., Badia, A. P., Mirza, L., Graves, A., Harley, T., Lillicrap, T. P., & Kavukcuoglu, K. (2016). Asynchronous methods for deep reinforcement learning. *33rd international conference on machine learning, ICML 2016*, New York, USA (pp. 1–19).
- Morescalchi, A., van Veldhuizen, S., Voogt, B., & Vogt, B. (2018). Negative home equity and job mobility. *Data-driven policy impact evaluation: how access to microdata is transforming policy design*, 183–202. https://link.springer.com/chapter/10.1007/978-3-319-78461-8_1
- Munro, M. (2018, 10). House price inflation in the news: A critical discourse analysis of newspaper coverage in the UK. *Housing Studies*, 33(7), 1085–1105. <https://doi.org/10.1080/02673037.2017.1421911>
- Narayan, P. K., & Narayan, S. (2011, 5). The importance of real and nominal shocks on the UK Housing Market. *SSRN Electronic Journal*. <https://papers.ssrn.com/abstract=2052122>
- Nik, P. A., Jusoh, M. A., Shaari, A. H., & Sarmdi, T. (2016). Predicting the probability of financial crisis in emerging countries using an early warning system: Artificial neural network. *Journal of Economic Cooperation and Development*, 37, 25–40. <https://api.semanticscholar.org/CorpusID:157881366>
- Norwegian, T. (2007). A heterodox economic analysis of the housing market structure in budapest using neural network classification. *Journal of Real Estate Literature*, 15 (1). <https://doi.org/10.1080/10835547.2006.12090194>
- Oguibenine, B. (2011). Economic recession and mental health: An overview. *Neuropsychiatrie: Klinik, Diagnostik, Therapie und Rehabilitation: Organ der Gesellschaft Österreichischer Nervenärzte und Psychiater*, 25(3), 113–117.
- Oldford, R. W. (2016, 1). Self-calibrating quantile–quantile plots. *American Statistician*, 70(1), 74–90. <https://doi.org/10.1080/00031305.2015.1090338>
- Olmez, S. (2022, 3). *SedarOlmez94/pythonic UK housing market ABM 2022: UK Housing Market Model 2022*. Zenodo. <https://zenodo.org/record/6362146>
- Olmez, S., Birks, D., & Heppenstall, A. (2022). Learning complex spatial behaviours in ABM: an experimental observational study. *arXiv*. <https://arxiv.org/abs/2201.0109v1>
- Olmez, S., Thompson, J., Marfleet, E., Suchak, K., Heppenstall, A., Manley, E., & Vidanaarachchi, R. (2022). An agent-based model of heterogeneous driver behaviour and its impact on energy consumption and costs in urban space. *Energies*, 15(11), 4031. <https://doi.org/10.3390/en15114031>
- Park, D., & Ryu, D. (2021). A machine learning-based early warning system for the housing and stock markets. *Institute of Electrical and Electronics Engineers Access*, 9, 85566–85572. <https://doi.org/10.1109/ACCESS.2021.3077962>
- Picascia, S., Camarca, A., Picascia, S., DiLuccia, A., & Gianfrani, C. (2014). Cereal-based gluten-free food: How to reconcile nutritional and technological properties of wheat proteins with safety for celiac disease patients. *Nutrients*, 6(2), 575–590. <https://doi.org/10.3390/nu6020575>
- Popescu, I. V. (2014, 1). Analysis of the behavior of central banks in setting interest rates. The case of central and eastern european countries. *Procedia Economics and Finance*, 15, 1113–1121. [https://doi.org/10.1016/S2212-5671\(14\)00565-6](https://doi.org/10.1016/S2212-5671(14)00565-6)
- Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. *Handbook of Macroeconomics*, 2, 71–162. <https://doi.org/10.1016/bs.hesmac.2016.03.003>
- Rosenfield, A., Chingcuanco, F., & Miller, E. J. (2013, 1). Agent-based housing market microsimulation for integrated land use, transportation, environment model system. *Procedia Computer Science*, 19, 841–846. <https://doi.org/10.1016/j.procs.2013.06.112>
- Sali, R., Adewole, S., & Akakpo, A. (2021). Feature selection using reinforcement learning. *arXiv*. <https://doi.org/10.48550/arXiv.2101.09460>
- Schulze, J., Müller, B., Groeneweld, J., & Grimm, V. (2017). Agent-based modelling of social-ecological systems: Achievements, challenges, and a way forward. *Journal of*

- Artificial Societies and Social Simulation*, 20(2). <https://doi.org/10.18564/jasss.3423>
- Secchi, D. (2015, 3). A case for agent-based models in organizational behavior and team research. *Team Performance Management*, 21(1/2), 37–50. <https://doi.org/10.1108/TPM-12-2014-0063>
- Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., & Riedmiller, M. (2014, 6). Deterministic policy gradient algorithms. *Proceedings of the 31st international conference on machine learning* (Vol. 32, pp. 387–395). PMLR, Beijing, China. <https://proceedings.mlr.press/v32/silver14.html>. In E. P. Xing & T. Jebara (Eds.).
- Sledge, I. J., & Principe, J. C. (2017, 6). Balancing exploration and exploitation in reinforcement learning using a value of information criterion. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, New Orleans, USA (pp. 2816–2820).
- Spatharis, C., Blekas, K., Bastas, A., Kravaris, T., & Vouros, G. A. (2019). Collaborative multiagent reinforcement learning schemes for air traffic management. *10th international conference on information, intelligence, systems and applications, IISA 2019*, Patras, Greece.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
- Suzuki, M., Nakatani, R., & Nishikawa, I. (2014). A mechanism design of solar power trading by autonomous agent based on reinforcement learning. *Frontiers in Artificial Intelligence and Applications*, 262, 392–401. <https://ebooks.iospress.nl/doi/10.3233/978-1-61499-405-3-392>
- Tang, W., & Bennett, D. A. (2010). *Agent-based modeling of animal movement: A review*. *Geography Compass*, 4(7), 682–700. <https://doi.org/10.1111/j.1749-8198.2010.00337.x>
- Todd A., Beling P., Scherer W., & Yang S. Y. (2017, 2). Agent-based financial markets: A review of the methodology and domain. *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016*.
- Tse, C. B., Rodgers, T., & Niklewski, J. (2014, 2). The 2007 financial crisis and the UK residential housing market: Did the relationship between interest rates and house prices change? *Economic Modelling*, 37, 518–530. <https://doi.org/10.1016/j.econmod.2013.08.013>
- Valadkhani, A., Nguyen, J., & O'Brien, M. (2019). Asymmetric responses of house prices to changes in the mortgage interest rate: Evidence from the Australian capital cities. *Applied Economics*, 51(53), 5781–5792.
- <https://www.tandfonline.com/doi/abs/10.1080/00036846.2019.1619026>
- Westerhoff, F. (2010). A simple agent-based financial market model: Direct interactions and comparisons of trading profits. *Nonlinear Dynamics in Economics, Finance and Social Sciences: Essays in Honour of John Barkley Rosser Jr*, 313–332. https://link.springer.com/chapter/10.1007/978-3-642-04023-8_1
- White, M. (2015, 5). Cyclical and structural change in the UK housing market. *Journal of European Real Estate Research*, 8(1), 85–103. <https://doi.org/10.1108/JERER-02-2014-0011>
- Whitehead, C., & Williams, P. (2011, 10). Causes and consequences? Exploring the shape and direction of the housing system in the UK post the financial crisis. *Housing Studies*, 26(7–8), 1157–1169. <https://doi.org/10.1080/02673037.2011.618974>
- Yamaguchi, S., Naoki, H., Ikeda, M., Tsukada, Y., Nakano, S., Mori, I., Ishii, S., & Brown, A. (2018, 5). Identification of animal behavioral strategies by inverse reinforcement learning. *PLOS Computational Biology*, 14(5), e1006122. <https://doi.org/10.1371/journal.pcbi.1006122>
- Yingfei, X. (2009). A language-based approach to model synchronization in software engineering.
- Yun, T. S., & Moon, I. C. (2020). Housing market agent-based simulation with loan-to-value and debt-to-income. *Journal of Artificial Societies and Social Simulation*, 23(4), 1–19. <https://www.jasss.org/23/4/5.html>
- Zhan, C., Wu, Z., Liu, Y., Xie, Z., & Chen, W. (2020, 7). Housing prices prediction with deep learning: An application for the real estate market in Taiwan. *IEEE International Conference on Industrial Informatics (INDIN)*, Beijing, China, 2020–July, (pp. 719–724).
- Zhou, X., Tong, W., & Li, D. (2019). Modeling housing rent in the atlanta metropolitan area using textual information and deep learning. *ISPRS International Journal of Geo-Information*, 8(8), 349. <https://www.mdpi.com/2220-9964/8/8/349>
- Zhou, Y., Wu, J., Long, C., Cheng, M., & Zhang, C. (2017, 12). Performance evaluation of peer-to-peer energy sharing models. *Energy Procedia*, 143, 817–822. <https://doi.org/10.1016/j.egypro.2017.12.768>
- Zhu, B., Betzinger, M., & Sebastian, S. (2017). Housing market stability, mortgage market structure, and monetary policy: Evidence from the euro area. *Journal of Housing Economics*, 37, 1–21. <https://doi.org/10.1016/j.jhe.2017.04.001>

Appendices

Appendix A.

Table A1. Model input parameters and description (source (Gilbert et al., 2009)).

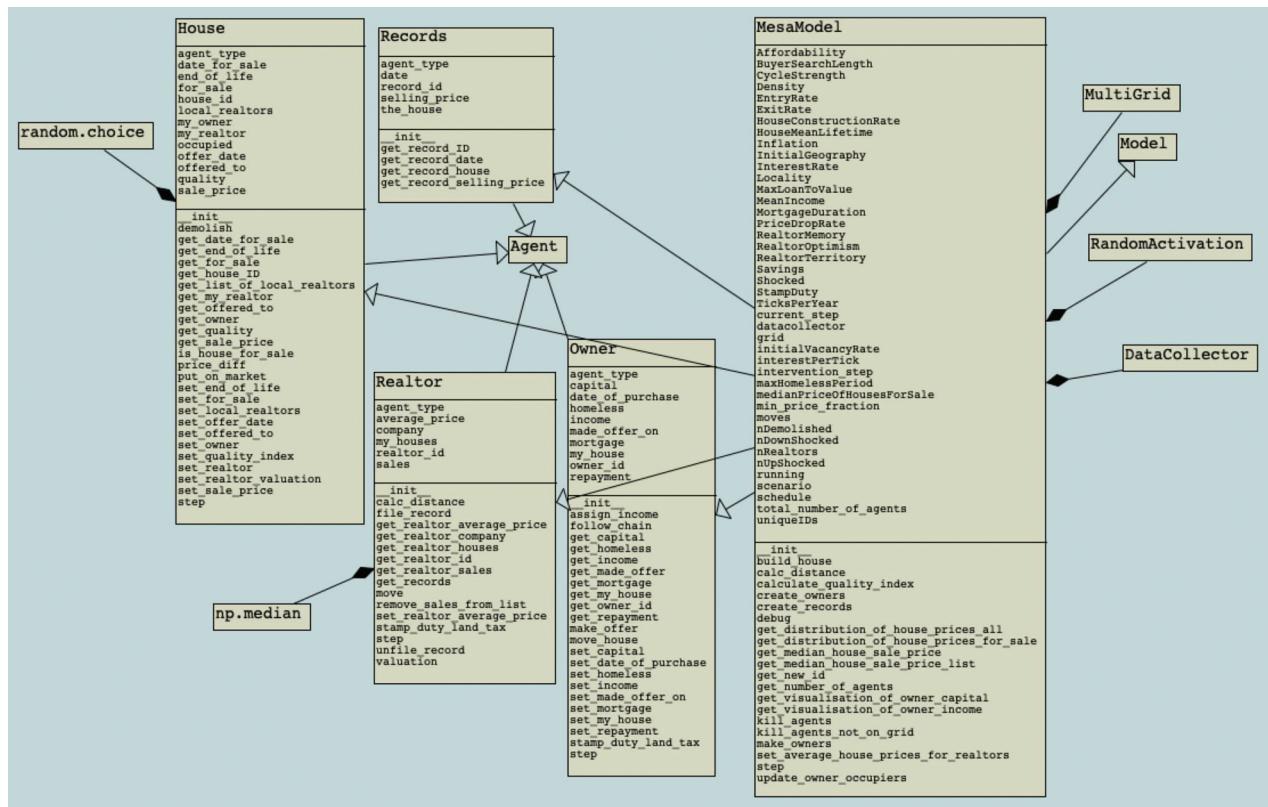
Parameter	Description
InitialVacancyRate	Proportion of empty houses.
nRealtors	Number of realtors.
InterestRate	Mortgage interest rates.
TicksPerYear	Number of timesteps per year.
MaxHomelessPeriod	Maximum number of timesteps a person can be homeless.
InterestPerTick	Interest rate, after a cyclical variation has been applied.
MinPriceFraction	If a house price falls below this fraction of the median price, it is demolished.
Inflation	Inflation percentage per year.
CycleStrength	Percentage of interest rate change periodically.
Affordability	Owners trade down or up depending on the affordability ratio.
Savings	The amount of capital a person has.
ExitRate	Percentage of people exiting the market.
EntryRate	Percentage of people entering the market.
MeanIncome	average annual income.
Shocked	percentage of person's income increased or decreased.
BuyerSearchLength	Length of time a buyer searches for a home.
RealtorTerritory	The locality a realtor covers.
Locality	Distance between houses used for valuations.
RealtorMemory	Historical collection of records used for valuations.
PriceDropRate	Percentage a house drops in price.
RealtorOptimism	A fee added to the valuation.
InitialGeography	Type of geography.
Density	How clustered the agents are spatially.
HouseConstructionRate	The rate at which houses are constructed.
HouseMeanLifetime	The average lifetime of a house.
MaxLoanToValue	The amount of deposit.
MortgageDuration	The timeframe for mortgage repayment.
Stampduty	A tax levied on house sales (UK only).

Table A2. Model input parameters for similarity testing.

Parameter	Input Value
NumberOfRuns	100
Timesteps	400
CycleStrength	0%
MeanIncome	£30000 pa
ExitRate	2%
Density	70%
StampDuty	False
TicksPerYear	4
HouseMeanLifetime	101
MortgageDuration	25
HouseConstructionRate	0.30%
MaxLoanToValue	100%
Shocked	20%
BuyerSearchLength	10
EntryRate	5%
MaxHomelessPeriod	5
InitialGeography	"Random"
PriceDropRate	3%
InterestRate	7% ^a
RealtorOptimism	3%
Inflation	0%
Affordability	25%
Savings	50%
RealtorTerritory	30
RealtorMemory	10
Locality	3

^aFor scenario two, the interest rate will drop from 7% to 3% initiating a shock at timestep 200.

Appendix B.



	Simulation	Python 2	Netlogo 2	Python 1	Netlogo 1
Houseprice/Income.Ratio(Median)	mean	8.575	8.236	5.244	5.094
	std	3.385	3.204	0.494	0.466
	min	3.444	3.552	3.474	3.569
	25%	5.305	5.265	5.004	4.841
	50%	6.682	6.401	5.139	5.001
	75%	11.655	10.700	5.354	5.264
	max	15.483	16.314	7.156	7.017
Median.Houseprice.for.sale	pearson (corr)	0.947 (<i>p</i> < 0.01)		0.839 (<i>p</i> < 0.01)	
	mean	540784.707	492637.040	376547.125	303944.850
	std	205852.457	209494.168	21553.608	28655.851
	min	220618.472	140373.576	214807.564	140800.961
	25%	370983.716	306518.909	366592.465	297064.457
	50%	406323.037	344237.304	377485.583	308279.053
	75%	793104.838	694714.404	388691.637	319404.545
Median.Houseprice.for.sold	max	939038.900	1024133.000	489329.873	371461.380
	pearson (corr)	0.809 (<i>p</i> < 0.01)		0.323 (<i>p</i> < 0.01)	
	mean	468766.092	572591.395	280877.291	227158.909
	std	200050.365	159258.916	26848.436	20349.625
	min	138245.853	118400.328	138660.344	118881.129
	25%	286752.743	231172.403	285557.479	225151.212
	50%	314505.776	252806.798	294130.012	231070.367
Gini.index.house.prices	75%	674358.036	515474.713	302448.809	236661.114
	max	811625.267	688856.713	340738.414	263754.427
	pearson (corr)	0.961 (<i>p</i> < 0.01)		0.837 (<i>p</i> < 0.01)	
	mean	0.307	0.283	0.290	0.279
	std	0.037	0.037	0.034	0.035
	min	0.193	0.197	0.205	0.199
	25%	0.278	0.255	0.265	0.252
Gini.index.incomes	50%	0.304	0.280	0.282	0.271
	75%	0.337	0.314	0.314	0.304
	max	0.432	0.387	0.394	0.382
	pearson (corr)	0.641 (<i>p</i> < 0.01)		0.832 (<i>p</i> < 0.01)	
	mean	0.394	0.374	0.388	0.373
	std	0.014	0.016	0.013	0.015
	min	0.331	0.301	0.329	0.302
Mean.Mortgage.repayment.to.income.ratio	25%	0.386	0.364	0.379	0.364
	50%	0.395	0.376	0.388	0.373
	75%	0.403	0.385	0.397	0.383
	max	0.445	0.412	0.456	0.416
	pearson (corr)	0.534 (<i>p</i> < 0.01)		0.667 (<i>p</i> < 0.01)	
	mean	0.353	0.401	0.295	0.340
	std	0.059	0.065	0.004	0.005
Number.of.transactions	min	0.283	0.306	0.268	0.306
	25%	0.396	0.376	0.293	0.338
	50%	0.306	0.353	0.295	0.340
	75%	0.417	0.473	0.298	0.342
	max	0.444	0.491	0.309	0.354
	pearson (corr)	0.992 (<i>p</i> < 0.01)		-0.150 (<i>p</i> = 0.2026)	
	mean	62.960	70.694	53.087	60.676
Median.time.on.market	std	24.733	26.553	14.230	9.463
	min	13.000	9.000	13.000	9.000
	25%	52.000	59.000	42.000	54.000
	50%	61.000	67.000	53.000	60.000
	75%	70.000	76.000	64.000	67.000
	max	515.000	531.000	106.000	117.000
	pearson (corr)	0.659 (<i>p</i> < 0.01)		0.235 (<i>p</i> < 0.01)	
N.houses	mean	8.733	9.287	9.007	9.566
	std	1.871	2.316	1.588	1.796
	min	1.000	1.000	1.000	1.000
	25%	9.000	9.000	9.000	9.000
	50%	9.000	10.000	9.000	10.000
	75%	10.000	11.000	10.000	10.000
	max	13.000	15.000	12.000	14.000
N.people.seeking.home	pearson (corr)	0.865 (<i>p</i> < 0.01)		0.824 (<i>p</i> < 0.01)	
	mean	2034.573	2508.608	2085.355	2475.416
	std	460.394	80.058	397.827	44.420
	min	1074.000	2196.000	1267.000	2354.000
	25%	1600.000	2373.000	1712.000	2442.000
	50%	2219.000	2423.000	2223.000	2473.000
	75%	2438.000	2481.000	2436.000	2511.000
N.empty.houses	max	2568.000	2587.000	2567.000	2615.000
	pearson (corr)	0.746 (<i>p</i> < 0.01)		-0.038 (<i>p</i> = 1.868)	
	mean	380.377	536.058	369.355	370.709
	std	122.045	109.796	108.300	79.474
	min	23.000	0.000	21.000	2.000
	25%	275.000	286.000	284.000	330.000
	50%	393.000	367.000	365.000	357.000
N.houses.in.negative.equity	75%	458.000	438.000	426.000	407.000
	max	701.000	628.000	694.000	634.000
	pearson (corr)	0.907 (<i>p</i> < 0.01)		0.882 (<i>p</i> < 0.01)	
	mean	307.697	536.020	378.740	558.276
	std	152.558	226.519	104.894	180.824
	min	2.000	3.000	12.000	5.000
	25%	222.000	418.000	328.000	482.000
N.demolished.houses	50%	299.000	565.000	383.000	605.000
	75%	439.000	720.000	448.000	683.000
	max	632.000	1044.000	626.000	913.000
	pearson (corr)	0.329 (<i>p</i> < 0.01)		0.423 (<i>p</i> < 0.01)	
	mean	111.932	86.326	63.344	66.777
	std	66.820	27.779	22.645	22.442
	min	0.000	0.000	0.000	0.000
N.people	25%	72.000	71.000	49.000	53.000
	50%	88.000	86.000	63.000	64.000
	75%	133.000	104.000	77.000	77.000
	max	324.000	181.000	147.000	149.000
	pearson (corr)	0.674 (<i>p</i> < 0.01)		0.776 (<i>p</i> < 0.01)	
	mean	6.518	7.242	6.228	7.004
	std	3.073	2.936	3.016	2.774
N.demolished.houses	min	0.000	0.000	0.000	0.000
	25%	4.000	5.000	4.000	5.000
	50%	6.000	7.000	6.000	7.000
	75%	8.000	10.000	8.000	9.000
	max	23.000	31.000	21.000	23.000
	pearson (corr)	0.117 (<i>p</i> = 1.406)		0.109 (<i>p</i> = 1.364)	
	mean	2107.253	2256.194	2073.900	2294.336
N.people	std	544.515	362.434	485.047	261.922
	min	934.000	1522.000	1142.000	1859.000
	25%	1596.000	1920.000	1645.000	2119.000
	50%	2253.000	2240.000	146.000	2200.000
	75%	2404.000	2494.000	2408.000	2398.000
	max	3149.000	3092.000	3136.000	3092.000
	pearson (corr)	0.897 (<i>p</i> < 0.01)		0.830 (<i>p</i> < 0.01)	