Strategic decision-making in multi-agent markets: The emergence of endogenous crises and volatility

BENJAMIN PATRICK EVANS

Supervisor: Mikhail Prokopenko
Auxiliary Supervisor: Michael S. Harré

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Centre for Complex Systems
School of Computer Science
Faculty of Engineering
The University of Sydney
Australia

2023

To my partner Lily, and my family.

Abstract

Traditional economic frameworks are built upon perfectly rational agents and equilibrium outcomes. However, during times of crises, these frameworks prove insufficient. In this thesis, we take an alternative perspective based on "Complexity Economics", relaxing the assumption of perfectly rational agents and allowing for out-of-equilibrium dynamics. While many contemporary approaches explain crises and non-equilibrium market phenomena as the rational reaction to external news, the emergence of endogenous crises remains an open question.

We begin addressing this question by demonstrating how a multi-agent model of heterogeneous boundedly rational agents acting according to heuristics can reproduce and forecast key non-linear price movements in the Australian housing market, during boom and bust cycles. In order to provide foundations for such heuristic-based reasoning, we then propose a novel information-theoretic approach, *Quantal Hierarchy*, for modelling limitations in strategic reasoning, demonstrating how this convincingly and generically captures the decision-making of interacting agents in competitive markets outperforming existing approaches. In addition, we demonstrate how a concise generalised market model can generate important stylised facts, such as fat-tails and volatility clustering, and allow for the emergence of crises, purely endogenously. This thesis provides support to the interacting agent hypothesis, addressing a crucial question of whether crisis emergence and various stylised facts can be seen as endogenous phenomena, and provides a generic method for representing strategic agent reasoning.

Authorship attribution statement

This is to certify that the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes (other than the publications outlined below). I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

The findings of this thesis have appeared in the following peer-reviewed publications:

- Chapter 3 of this thesis is published as: Evans, B.P., Glavatskiy, K., Harré, M.S., Prokopenko, M. The impact of social influence in Australian real estate: market forecasting with a spatial agent-based model. *J Econ Interact Coord* **18**, 5–57 (2023).
- Chapter 4 of this thesis is published as: Evans, B.P., Prokopenko, M. A Maximum Entropy Model of Bounded Rational Decision-Making with Prior Beliefs and Market Feedback. *Entropy* 23, 669 (2021).
- Chapter 5 of this thesis is published as: Evans, B.P., Prokopenko, M. Bounded rationality for relaxing best response and mutual consistency: the quantal hierarchy model of decision-making. *Theory and Decision* (2023).
- Chapter 6 of this thesis is published as: Evans, B.P., Prokopenko, M. Bounded strategic reasoning explains crisis emergence in multi-agent market games. *Royal Society Open Science* 10(2), 221164 (2023).

Benjamin Patrick Evans

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Mikhail Prokopenko

Acknowledgements

I would like to thank my supervisor, Professor Mikhail Prokopenko, for his dedication and relentless work ethic, from the very first email to the final submission. Without your support, extensive knowledge, and insights, this thesis would not have been possible. I continue to learn something new from you at every meeting. Your deep curiosity is truly inspiring, both inside and outside of research. This curiosity and exploration lead to the best discoveries, even if, at times, this curiosity has led us off the well-defined walkways to scaling cliff faces in national parks around Australia (following heuristics such as "at the second set of spider webs, turn right").

I would also like to thank my auxiliary supervisor, Dr Michael Harré, and collaborator, Dr Kirill Glavatskiy, for many invaluable discussions and contributions, particularly with Chapter 3, and the wider Centre for Complex Systems for being so welcoming throughout my studies.

Additionally, I have benefited from several research visits during my candidature. I want to thank Dr. ir. Qingyi Feng and Prof. dr. ir. Vincent Buskens from the Centre for Complex Systems Studies at Utrecht University, who I was fortunate enough to visit under the Swaantje Mondt travel fund, supported by grantors Martine and Jaap Mondt. I would also like to thank the Center for Adaptive Rationality, at the Max Planck Institute for Human Development, for organising the 20th Summer Institute on Bounded Rationality, and the University of Oxford for organising the 2022 Summer Institute in Computational Social Science (SICSS-Oxford).

Finally, this thesis would not have been possible without the financial support of the Australian government under their generous Research Training Program scholarship.

"The very complexity that has made a theory of the decision-making process essential has made its construction exceedingly difficult... It seemed almost utopian to suppose that we could put together a model of adaptive man that would compare in completeness with the simple model of classical economic man... [Computational modelling] provides us with a tool of research-for formulating and testing theories-whose power is commensurate with the complexity of the phenomena we seek to understand... As economics finds it more and more necessary to understand and explain disequilibrium as well as equilibrium, it will find an increasing use for this new tool" - Visionary Computer Scientist and Economist, Herbert Simon, 1959.

Contents

CHAPTER 1

Introduction

1.1 Scope

The conventional theory for pricing in economic markets is the efficient-market hypothesis (EMH), based on rational expectations and market equilibrium. EMH states that pricing reflects all known information about an asset and, as such, reflects the fair value, e.g. "the price is always right". Based on the EMH, it should be impossible to beat the market, as pricing follows a purely "random walk" driven only by unpredictable external news (the link between EMH and random walks is itself controversial). However, experimentally, substantial deviations from this efficiency are observed, for example, with firms such as Renaissance Technologies consistently outperforming the market, where the founder Jim Simons is quoted as saying "[EMH] is just not true", leading Econophysicist Jean-Philippe Bouchaud remarking "In reality, markets are not efficient... free markets are wild markets" (Bouchaud, 2008), and with much empirical evidence "[eroding] the trust in a theory which denies the existence of any systemic deviations of stock prices from their fundamental values" (Lux, 1995). The occurrence of significant asset bubbles aid in eroding this trust, where pricing sees substantial deviations from the fundamental asset value, for example, with housing market bubbles (Axtell et al., 2014), casting doubts on the efficiency of such markets.

These substantial price deviations from the underlying value bring market modelling back to the original beliefs of Keynes, who believed "animal spirits" (Nyman et al., 2021) played a significant role in market dynamics. A similar opinion is shared by Shiller, who believes many not necessarily perfectly rational factors (such as structural, cultural, and psychological factors) help shape price dynamics (Shiller, 2015b), for example, irrational exuberance being the psychological basis of a speculative bubble (Shiller, 2015a), which the ex-Federal Reserve Board chairman also supported as an explanation for the dot-com bubble. Such considerations motivate the relaxation of rationality and efficiency when modelling various economic markets, allowing for alternate concepts, such as the interactions among bounded rational agents (e.g. market participants), to potentially explain market dynamics and understand the resulting "wildness" (Bouchaud, 2008).

Under the EMH, such bubbles would only burst under significant external news. The 2008 Global Financial Crisis can be seen as an inflection point that helped trigger wider spread doubts about the EMH, with the Federal Reserve chairman reflecting "it should be clear that among the causes of the recent financial crisis was an unjustified faith in rational expectations, market efficiencies, and the techniques of modern finance" (Volcker, 2011), the US Treasury Secretary remarking the “the [traditional] economic models were worthless” during times of crises, and the president of the European Central Bank stating they "felt abandoned by the conventional tools" (Arthur et al., 2020). Under these rational expectations approaches, crises can only arise due to the arrival of external news – meaning endogenous crises formation can not be explained under such a hypothesis. These concerns highlighted the need for a "revolution" in economic modelling (Bouchaud, 2008). These limitations in conventional approaches for explaining out-of-equilibria economic phenomena were one of the early motivations of work at the Santa Fe Institute, around a simple question: "What would economics look like if we allowed nonequilibrium?" (Arthur et al., 2020).

In addressing this question, an alternate view on economics and financial markets began to emerge, which did away with some of the traditional notions of perfect rationality, market efficiency, and equilibrium, which would go on to be called "Complexity Economics", following Arthur, 1999. Under Complexity Economics, markets are seen as nonlinear complex adaptive systems, evolving dynamically (Hommes, 2001). Rather than nonequilibrium states being a temporary result of the rational reaction to external news, "from the complexity economics perspective, change is largely an endogenous phenomenon, not simply the result of unexplained shocks from outside the system" (Arthur et al., 2020). Throughout this thesis, we take a similar perspective – and look to explain various market phenomena as endogenously occurring.

The interacting agent hypothesis (IAH) (Lux and Marchesi, 1999) is one approach leading this revolution, driven by the fact many market phenomena are incompatible with the EMH, instead appearing to arise due to "intrinsic dynamic forces of speculative markets not related to fundamental factors" (Lux, 1995). Rather than various stylised facts arising due to a news arrival process, this phenomenon can be explained as a result of interactions among agents with various (not necessarily perfectly rational) factors influencing the decision-making (such as psychological fears or processing constraints). For example, the IAH explains market volatility and speculative bubbles and crashes based on crowd dynamics and herding among agents (Gontis et al., 2016; Kaizoji, 2000). In contrast, under the EMH, such behaviour would only occur as a reaction to external news arrival.

Throughout this work, we look to develop upon this idea, providing support to the IAH. During "general" market times, in well-established professional markets, the EMH may serve as a fair assumption, but crucially, for important turning points of markets, the EMH begins to diminish in applicability. Throughout this thesis, we look to explain phenomena unexplainable by the EMH, where rather than market dynamics being driven purely by underlying fundamental changes, the endogenous interaction between market participants can drive various market phenomena. Importantly, this provides an alternative tool to understand economic markets both in and out of equilibrium, with an early focus on the Australian housing market, before moving into more general economic markets.

1.2 Objectives

This thesis aims to explain pricing dynamics and out-of-equilibrium market phenomena (for example, endogenous crises formation) as a result of the interaction among bounded rational agents. The main objectives of this thesis are detailed below.

I Demonstrate how interactions among inductive reasoning agents can recreate key price dynamics

Human decision-making in economic markets is often comprised of heuristics, and the resulting decisions effected by various structural, cultural, and psychological factors (Shiller, 2015a), such as emotion and social forces (Ackert and Deaves, 2009; Strahilevitz et al., 2011).

Therefore, when attempting to model economic decision-making, such observations motivate the relaxation of *homo economicus*, the perfectly rational representative economic actor, who can perform exact deductive reasoning (De Bondt and Thaler, 1995). Instead, suggesting decision-makers perform inductive reasoning (Arthur, 1994), where such a reasoning process may be heterogeneous among the market participants (Agliari et al., 2018), and influenced by various social and psychological factors such as the choices of others (Garnier-Brun et al., 2022). Representing these various factors in terms of behavioural heuristics and/or information processing resources becomes a key task of this thesis.

Specifically, the first objective of this work is to demonstrate how interacting inductive-reasoning agents, influenced by various social pressures and search processing constraints, can generate key market phenomena, such as pricing dynamics, in accordance with actual market conditions. In the first instance, we wish to *a.)* demonstrate this with hand-crafted behavioural heuristics, to see how well they perform, before *b.)* generalising these heuristics into information-processing terms to represent a more generic class of economic phenomena. In both cases, we aim to place a particular focus on real data from the Australian housing market.

II Model strategic higher-order reasoning to capture decision-making under constraints

Rather than asset pricing being strictly driven by the underlying fundamentals (as is the case with the EMH), it has often been said that instead, beliefs about other investors' beliefs drive the price dynamics (Keynes, 1937; Casti, 1996). Representing this type of strategic decision-making, where an agent reasons about another agent who themselves is reasoning about the first agent, and so on, is difficult due to the self-referential nature of the recursive reasoning process (e.g. the Holmes–Moriarty problem, Koppl and Barkley Rosser Jr, 2002). Through relaxing this *homo economicus* assumption, and admitting agents may display bounded rational behaviour, this self-reference can be seen as partial self reference (Löfgren, 1979), due to the ability to break at various points of recursion when processing resources deplete. Developing a generic model of this partial self-reference, becomes essential for capturing agents' strategic higher-order reasoning in economic markets. The key question here is: How can we capture this higher-order reasoning in a generic (information-theoretic) sense?

Hence, our second objective is to develop an information-theoretic model quantifying information-processing resources of strategically reasoning agents, and verify this model against a set of representative canonical games and market settings.

III Demonstrate how interactions among bounded strategic reasoning agents can lead to endogenous crisis emergence

To reiterate, the endogenous formation of crises and periods of heightened volatility can not be explained by the efficient market hypothesis, without the arrival of external news. However, frequently, such dynamics occur in actual economic markets, displaying endogenous crisis formation, clustered volatility, and "fat-tails" (Hommes, 2002), which contradict the random-walk explanation of the EMH (in the absence of news). Despite much effort, a simple and concise explanation of crises emergence and these stylised facts (such as fat-tails and volatility clustering) in terms of strategic reasoning between interacting agents has yet to be given. A key question is, can endogenous crisis form out of the interaction among bounded rational agents that reason recursively (and thus, strategically) about other agents. The specific objective here is to show under a concise and general market settings how agent interactions and bounded rational reasoning can lead to endogenous crises formation, in the absence of any fundamental market changes. We aim to demonstrate applicability using a canonical market entrance problem, and compare to existing solutions.

1.3 Contributions and Main Findings

To address these objectives, we have produced four self-contained studies, each forming a chapter in this thesis. In the first study, we focus specifically on housing markets before moving into more general economic markets in the later studies.

1.3.1 Thesis Overview

The first study presents a fine-grained computational model of the Australian housing market. The second study is more theoretically grounded, and helps infer decision-making factors that had been hand-crafted in the first model. The third study moves into strategic reasoning between interacting agents, developing a novel information-theoretic approach for quantifying inductive strategic reasoning. Finally, the fourth study brings these key concepts together, where we develop a concise and general multi-agent model of economic markets, where interacting agents using inductive strategic reasoning recreate crucial market phenomena.

The impact of social influence in Australian real estate: market forecasting with a spatial agent-based model

This study (Evans et al., 2023) presents a computational model of the Greater Sydney housing market. Agent decision-making is driven by buy and sell heuristics, with decisions influenced by psychological and economic factors, such as social influences and pressures, connecting agent decision-making to the "Animal spirits" of Keynes. The bottom-up simulation provides well-performing forecasting for overall market pricing and area-specific breakdowns arising from these individual agent decisions (addressing Objective I.a). In addition, the simulation results elucidate agent preferences in submarkets, highlighting differences in agent behaviour and the importance of agent heterogeneity, for example, between first-time home buyers and investors, and between local and overseas investors.

This study was published in the Journal of Economic Interaction and Coordination (Evans et al., 2023), and is presented in Chapter 3.

A Maximum Entropy Model of Bounded Rational Decision-Making with Prior Beliefs and Market Feedback

This study (Evans and Prokopenko, 2021) proposes a principled approach to infer agent decisions in competitive markets based on prior beliefs and observed macroeconomic outcomes. We show how such a model can explain and reconstruct the distribution of price fluctuations in the Greater Sydney housing market, based on the interactions between these inductive reasoning agents (driven by the observed profit rates), addressing Objective I.b. Incorporating prior beliefs into such a framework provided insights into potential leading indicators of market state changes, allowed for analysis into the evolution of agent beliefs over time, and improved the accuracy of the model in reproducing actual market data.

This study was published in Entropy (Evans and Prokopenko, 2021), in the special issue "Three Risky Decades: A Time for Econophysics?", and is presented in Chapter 4.

Bounded rationality for relaxing best response and mutual consistency: The Quantal Hierarchy model of decision-making

In this study (Evans and Prokopenko, 2023a), we provided a principled and generic way of quantifying limitations in higher-order strategic reasoning, addressing Objective II. Bounds in player processing abilities are expressed as information costs, where future chains of reasoning are discounted, implying a hierarchy of players where lower-level players have fewer processing resources. Specifically, the proposed model is based on a recursive form of the variational free energy principle, representing higher-order reasoning as (pseudo) sequential bottom-up decisions. We demonstrate the applicability of the proposed model to several canonical economic games, showing the proposed model is a good fit for out-of-sample human behaviour, outperforming existing state-of-the-art approaches.

This study was published in Theory and Decision (Evans and Prokopenko, 2023a), and is presented in Chapter 5.

Bounded strategic reasoning explains crisis emergence in multi-agent market games

In this study (Evans and Prokopenko, 2023b), we propose a concise model that explains the endogenous emergence of punctuated out-of-equilibrium dynamics based on the interactions between bounded rational agents (addressing Objective III). In a market entrance game, we show how inductive strategic reasoning can lead to endogenously emerging crises, exhibiting fat tails in "returns". We also show how other common stylised facts of markets, such as clustered volatility, arise due to agent diversity (or lack thereof) and the varying learning updates across the agents (addressing Objective I). This work explains various stylised facts and crisis emergence in economic markets, in the absence of any external news, based on agent interactions and bounded rational reasoning. This provides an alternative perspective to the efficient market hypothesis, which can model periods of out-of-equilibrium dynamics without external news arrival.

This study was published in Royal Society Open Science (Evans and Prokopenko, 2023b), and is presented in Chapter 6.

1.3.2 Mapping Findings to Objectives

When analysing market outcomes, it seems improbable that market participants (agents) are perfectly rational, which is also supported by various experimental economic studies. Instead, market participants are subject to various factors influencing decision-making and may exhibit bounded rational behaviour. Throughout this thesis, we have developed novel techniques for modelling human behaviour and computational techniques for analysing the resulting out-of-equilibria phenomena in economic markets composed of these agents. Thus, we addressed our objectives as follows

I Demonstrate how interactions among inductive reasoning agents can recreate key price dynamics

Using Australian housing market data, we demonstrate how models of boundedly rational agents acting according to heuristic rules can recreate and forecast key pricing movements across the Greater Sydney region, capturing boom and bust cycles (Chapter 3), and how a more generalised model can reconstruct the distribution of price fluctuations (Chapter 4). Additionally, in a generic setting, we show how a concise model can generate key stylised facts of economic markets such as volatility clustering (Chapter 6). This shows how accounting for interactions among agents can explain key market phenomena and price dynamics from a bottom-up perspective.

II Model strategic higher-order reasoning to capture decision-making in competitive markets

We formally represent limitations in higher-order reasoning in an information-theoretic sense by representing reasoning as an extensive form game tree and introducing processing costs for exploring such a tree (Chapter 5). This provides a principled and generic approach for modelling recursive agent reasoning in economic markets.

III Demonstrate how interactions among bounded strategic reasoning agents can lead to endogenous crisis emergence

In a concise market entrance game, we show that boundedly rational strategic reasoning can lead to endogenously emerging crises, exhibiting fat tails in “returns”, following actual market conditions (Chapter 6). This confirms that crises can emerge endogenously in economic markets without external news based on agent interactions and bounded rational reasoning.

1.3.3 Closing the loop

We begin this thesis by looking at how a population of inductive reasoning agents, acting according to predefined heuristic rules, can recreate key price dynamics of an actual economic market (the Greater Sydney housing market). We then generalise some of these heuristics, using the concept of Smithian competition to capture decision-making and the market feedback loop, and formalise bounded reasoning through an information processing constraint. From this foundation, we then move from basic inductive reasoning to inductive strategic reasoning, where agents are again subject to information processing constraints but now characterising more generalised strategic reasoning in competitive markets. Finally, we apply this generalised strategic reasoning in a multi-agent system, "closing the loop", providing support to the findings of the original computational model with behavioural heuristics. This thesis goes from specific fine-grained examples to building a more generalised understanding at each step.

This thesis demonstrates a fundamental idea: Market dynamics are endogenously driven by the interactions among bounded rational agents. This bounded rational reasoning can be captured with the usage of hand-crafted behavioural heuristics, but we also show how more generally, an information-theoretic approach, the Quantal Hierarchy (QH) model, can generically quantify limitations in strategic reasoning. By providing this principled QH approach of modelling bounded rational strategic reasoning, we support the use of heuristics in fine-grained ABMs for approximating such decision-making under processing costs. The developed methods are verified with actual market data and underlying economic decision-making data, outperforming a range of canonical approaches. This thesis supports the interacting agent hypothesis, explaining the endogenous emergence of out-of-equilibria phenomena through the interaction among bounded rational agents.

1.3.4 Thesis Structure

The remainder of the thesis is presented as follows. Chapter 2 provides a background to the primary concepts used in this thesis, such as complexity economics, information theory, and decision-making. Chapter 3-Chapter 6 are the main contribution chapters, each addressing the outlined objectives and working towards a more generalised representation. Finally, conclusions and future work are detailed in Chapter 7.

CHAPTER 2

Background

2.1 Complexity Economics

Complexity economics was pioneered at the Santa Fe Institute in the late 1980s, with an interdisciplinary conference titled "The Economy as an Evolving Complex System" (Arthur et al., 2020), with many notable attendees and organisers such as Phil Anderson (Nobel Prize in Physics), Kenneth Arrow (Nobel Prize in Economics), Lawrence Summers (Secretary of US Treasury), and John Holland (leading computer scientist) (Farmer, 2000).

While there have been many accomplishments in complexity economics (see Arthur et al., 2020), here, we provide a brief background of some of the most relevant work for this thesis, splitting the work into two broad (overlapping) categories: Computational Economics and Econophysics. Due to the breadth of literature, we focus on one specific aspect from each which is directly relevant to this thesis: agent-based modelling from computational economics and statistical equilibrium models from econophysics. ABM are not necessarily separate from econophysics (Abergel et al., 2013; Chakraborti et al., 2011b), however, we have made this distinction for presentation and clarity, with "Complexity Economics" being the overarching umbrella term.

2.1.1 Computational Economics

Computational economics makes use of the increasing availability of computing power to address problems in economics that were previously considered intractable under traditional mathematical examination. Here, we focus on one technique, agent-based computational economics (ACE), and more specifically, ACE approaches to market modelling.

2.1.1.1 Agent-based modelling

Well-documented failures in traditional rational expectation models highlighted the need for alternative approaches for market modelling (Farmer and Foley, 2009; Bouchaud, 2008). One tool in particular, agent-based modelling (ABM) has shown much promise in the modelling of economic markets (Chen et al., 2012) due to the ability to represent realistic individual behaviour, such as bounded rational decision-making, social influences, and heterogeneous expectations among agents.

Rather than the top-down approach of assuming efficiency and rational expectations, ABM's view markets from the "bottom-up" (Tesfatsion, 2002). This bottom-up view explains market phenomena as arising from individual (not necessarily perfectly rational) decision-making. In this sense, ABMs can model macroeconomic dynamics from individual micro-level behaviour (LeBaron and Tesfatsion, 2008), creating a bottom-up adaptive approach to macroeconomics (Gatti et al., 2011), helping to address limitations of neoclassical approaches (Cincotti et al., 2022).

Thus far, ABM have seen a variety of successes in modelling economic markets. Early work notably showed how ABMs could create an artificial stock market (Palmer et al., 1999; LeBaron, 2002; LeBaron et al., 1999). More recently, various methods have been proposed to capture the dynamics of housing markets (Geanakoplos et al., 2012; Axtell et al., 2014), as well as various other use cases, such as modelling risks and financial stability (Bookstaber, 2017), scenario-based exploration of economic recovery patterns (Sharma et al., 2021), and beginning to show use for economic forecasting (Poledna et al., 2022). These works show the usefulness and flexibility of ABM, and an extensive review of ABM in economics is presented in Axtell and Farmer, 2022.

2.1.2 Econophysics

Driven by the empirically observed statistical properties of financial time series (Chakraborti et al., 2011a), during the 1990s, an influx of physicists began to enter the world of finance (both in academia and industry) (Farmer, 2000), and research at the interplay of these two areas (physics and markets) became more common. The term "econophysics" was coined by Eugene Stanley in 1995 (Stanley et al., 1996), and sees economic and financial phenomena through the "lens" of statistical physics (Schinckus, 2012). There have been many notable works in this area, for example, explaining distributions of stock price variations (Cont and Bouchaud, 2000), price fluctuations (Laloux et al., 1999), and risk management (Bouchaud, Potters et al., 2003). A general background of econophysics is presented in Kutner et al., 2019; Kutner et al., 2022. Similar to ABM, this area aims to provide an alternate perspective on economic and financial markets, relaxing these traditional economic assumptions.

For this thesis, work involving statistical mechanics, specifically the concept of statistical equilibrium, is most significant. A review of the influence of statistical mechanics and the usefulness of entropy in economics is presented in (Rosser Jr, 2021).

2.1.2.1 Statistical Equilibrium

Statistical equilibrium approaches to economics (Foley, 2003; Kaizoji, 2006; Scharfenaker and Yang, 2020b; Scharfenaker and Yang, 2020a) are most relevant to our work here. While it may seem counterintuitive to consider statistical "equilibrium" approaches, statistical mechanics provides an alternative equilibrium to the traditional (Walrasian) equilibrium approaches to economics. Statistical equilibrium offers a more flexible definition that instead represents a probability distribution of all possible states of the system (Ömer, 2020). This line of thinking may be considered "post-Walrasian" (Foley et al., 2017).

For modelling markets, statistical equilibrium provides an interesting new perspective, which can help relax some of the unrealistic ad-hoc assumptions, such as the perfect rationality of particpants or the efficiency of market clearance (Ömer, 2020). The most relevant model here is the Quantal Response Statistical Equilibrium (QRSE) model of Scharfenaker and

Foley, 2017, which utilises the maximum entropy principle (detailed in later sections) to infer individual decisions and statistical fluctuations of macroeconomic outcomes. Likewise, Kaizoji, 2006 uses the maximum entropy principle to infer traders' decision-making rules.

The statistical equilibrium models outlined utilising the principle of maximum entropy follow from Jaynes' work that shows the fundamental relationship between information theory and statistical mechanics (Jaynes, 1957a; Jaynes, 1957b). Crucially, Jaynes showed that "the maximisation of entropy is not an application of a law of physics, but merely a method of reasoning which ensures that no unconscious arbitrary assumptions have been introduced", which allows for the modelling of abstract phenomena not necessarily governed by physical processes (Rosser Jr, 2021). We adopt this perspective throughout this thesis for the modelling of decision-making and social phenomena in economic markets.

2.2 Information Theory

Agents are often limited in the amount of information they can process, for example, in a market, it may be impossible for a human to consider all available information. Throughout this thesis, we use information theory to formalise this information processing constraint.

Shannon pioneered information theory in his paper "A Mathematical Theory of Communication" (Shannon, 1948), as a method of representing information communication subject to a noisy channel, building upon prior fundamental work of Nyquist, 1924; Hartley, 1928. Due to the generality of the representation, information theory has found widespread usage, e.g., as a method for constraining agent decision-making (Sims, 2003; Ortega and Braun, 2013; Foley, 2020; Harré, 2021), as an alternative to conventional equilibrium models in economics (Yang, 2018; Yang, 2022), and as a general explanatory framework for complex systems analysis (Prokopenko et al., 2009).

This thesis adopts this perspective. In this section, we review some of the fundamental information-theoretic concepts utilised throughout this thesis.

2.2.1 Information-theoretic measures

2.2.1.1 Entropy

The central measure in information theory is the (Shannon) entropy H of a random variable X, which quantifies the amount of "information" contained within X, or the amount of "uncertainty" associated with the observations. Assuming a discrete distribution, the entropy is defined as

$$\mathrm{H}(X) = -\sum_{x \in X} p(x) \log p(x) \tag{2.1}$$

The choice of the logarithm base is problem specific, however, throughout this thesis, we use base 2 to quantify entropy in terms of bits (other common examples are e or 10).

As a simple example of the entropy measure, if we assume X is uniformly distributed, we get $H(\mathcal{U}) = \log k$, i.e. the maximally uncertain distribution, where k is the number of items in X, $k = |X|$. When observing an outcome from $\mathcal{U}$, our expected "surprise" is maximised. To see this connection with surprisal s, we can consider the logarithm of the inverse of the probability, e.g.

$$s(x) = \log \frac{1}{p(x)} \tag{2.2}$$

We can rewrite entropy in terms of the expected surprisal as

$$\mathrm{H}(X) = \sum_{x \in X} p(x) s(x) \tag{2.3}$$

which shows the connection between entropy, uncertainty, and surprise.

Enforcing a constraint on the entropy restricts the amount of "information" one can process and is utilised to represent informationally-constrained decision-makers in later sections.

2.2.1.2 Kullback Leibler Divergence (Relative Entropy)

The Relative Entropy, or the Kullback Leibler (KL) Divergence, $D_{\mathrm{KL}}(P \parallel Q)$ quantifies the amount of information gain from using probability distribution P over probability distribution Q, and is given by:

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_{x \in X} P(x) \log \left(\frac{P(x)}{Q(x)} \right) \tag{2.4}$$

We can see when $P = Q$, we get $D_{\mathrm{KL}}(P \parallel Q) = 0$. That is, we receive no additional information from using P. Alternatively, if we assume that Q is the uniform distribution $\mathcal{U}$, we recover the definition of Shannon entropy from Eq. (2.1) (up to a constant term), as this is the expected excess surprise from using $\mathcal{U}$ when the actual distribution is P. More generally, KL divergence can serve as a measure of the *divergence* between two probability distributions. In a similar manner to the entropy constraint, enforcing a constraint on $D_{\mathrm{KL}}(P \parallel Q)$ limits the divergence from a base "reference" distribution Q. We utilise this in later sections to limit the divergence from an agent's prior beliefs, i.e., to limit the amount of information acquisition based on information processing abilities.

2.2.2 Principle of maximum entropy

One helpful concept building upon the information-theoretic measures outlined in the previous section is the principle of maximum entropy (Jaynes, 1957a; Jaynes, 1957b) (MaxEnt), which has seen success in economic applications (Scharfenaker and Yang, 2020b; Scharfenaker and Yang, 2020a). When choosing among a set of probability distributions, MaxEnt speculates one should choose the distribution with the maximum entropy, as this makes the fewest assumptions. This specification is in accordance with Occam's razor, which (overly simplified) specifies, amongst all else being equal, the simplest explanation is the correct one. MaxEnt provides a systematic information-theoretic approach to modelling inference (Golan and Foley, 2022).

For example, in the simplest case, we have the natural constraints that the probability distribution must sum to one $\sum_{x \in X} p(x) = 1$, and since it is probability distribution, all values must be non-negative $p(x) \geq 0, \forall x$. Given this constraint, the distribution which should be chosen is the uniform distribution, as this has the maximal entropy among all possible options.

However, what distribution would arise if we were to enforce additional constraints on the first and second-order moments (mean and standard deviation)? MaxEnt can be used for answering this question by maximising

$$\max_p \mathrm{H(p)} = \max_p \left(-\sum_{x \in X} p(x) \log p(x) \right) \tag{2.5}$$

subject to the additional constraints that

$$\sum_{x \in X} p(x) x = \mu \tag{2.6}$$

$$\sqrt{\sum_{x \in X} p(x)(x-\mu)^2} = \sigma \tag{2.7}$$

which corresponds to a constrained optimisation problem. To treat this as an unconstrained optimisation problem, we convert this to a Lagrangian function, where we get

$$\begin{aligned} \mathcal{L} = -\sum_{x \in X} p(x) \log p(x) + \lambda_0 \left(\sum_{x \in X} p(x) - 1 \right) + \lambda_1 \left(\sum_{x \in X} p(x) x - \mu \right) \\ + \lambda_2 \left(\sqrt{\sum_{x \in X} p(x)(x-\mu)^2} - \sigma \right) \end{aligned} \tag{2.8}$$

where λ_i are Lagrange multipliers. Taking the first order conditions of $\mathcal{L}$ with respect to $p(x)$ yields:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{2.9}$$

i.e., the Gaussian distribution. This result shows that the distribution which maximises the entropy subject to constraints on the first and second-order moments is Gaussian. The usefulness of MaxEnt is that any arbitrary constraints can be added:

$$\sum_{x \in X} p(x) f_k(x) \geq F_k \qquad k = 1, \ldots, m. \tag{2.10}$$

which take the form of additional Lagrange multipliers, and the appropriate distribution (i.e. the maximally non-committal one) is of the form

$$p(x) = \frac{1}{Z(\lambda_1, \ldots, \lambda_m)} e^{\lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)} \tag{2.11}$$

i.e., the Boltzman distribution, where Z is the partition function. MaxEnt is very flexible, and provides a general framework for rational inferences under incomplete information (Scharfenaker and Yang, 2020b). MaxEnt is utilised in this thesis for representing agent decisions based on information processing constraints, and also for inferring resulting statistical equilibrium's based on market feedback loops.

2.3 Decision Making

Throughout this thesis, we follow the expected utility hypothesis based upon the Von Neumann–Morgenstern (VNM) utility theorem (Von Neumann and Morgenstern, 2007). The expected utility hypothesis states that when faced with a decision, a perfectly rational agent wishes to maximise their expected utility, and serves as a foundational assumption throughout economics.

To define this utility, VNM-rationality requires four fundamental underlying axioms of the decision-maker's preferences: completeness, transitivity, continuity, and independence. If all axioms are satisfied, one can assign a numeric utility function U to measure these preferences. The expected utility $\mathrm{E}[U]$ is then the weighted likelihood of achieving each utility, with agents trying to maximise this expected utility:

$$\max \mathrm{E}[U] = \max \sum_{a \in A} p(a)U(a) \tag{2.12}$$

This framework represents a perfectly rational decision-maker who can consider all outcomes and choose the option which maximises their expected utility. In experimental settings, however, this is frequently violated (Kahneman, 2003).

2.3.1 Bounded Rationality

It is instead plausible that agents are attempting to maximise their expected utility, however, they are subject to constraints when reasoning (for example, due to cognition, time, or money), that is, they display *bounded rationality* (Simon, 1957) (also referred to as "computational rationality", Gershman et al., 2015; Lewis et al., 2014). Following the information-theoretic discussion in the previous section, these constraints can be quantified using an information-processing constraint. Rather than an agent attempting to maximise their expected utility, e.g. from Eq. (2.12), agents are instead maximising their expected utility subject to an information processing (entropic) constraint:

$$\max \sum_{a \in A} p(a)U(a) \tag{2.13}$$

$$\text{subject to} - \sum_{a \in A} p(a) \log p(a) = B \tag{2.14}$$

and the usual constraint that the probabilities must sum to 1. Solving for $p(a)$, again using the method of Lagrangian multipliers, leads to the softmax (or logit) function:

$$p(a) = \frac{1}{Z} e^{\beta U(a)} \tag{2.15}$$

which restricts information-processing abilities through the Lagrange multiplier β (which serves as the resource parameter), and reconstructs the maximum expected utility principle when $\beta \to \infty$. This information-processing constraint has many benefits for modelling human decision-making, most notably, it abstracts away any specific type of constraint, meaning bounded rational reasoning can still be treated as an optimisation problem (Sims, 2003).

Ortega and Braun, 2013 extended this framework to account for a prior belief distribution p_0 over actions, representing decision-making as state changes, showing that this follows from the (negative) free-energy difference in thermodynamics:

$$-\Delta F(p) = \sum_{a \in A} p(a)U(a) - \frac{1}{\beta} \sum_{a \in A} p(x) \log \frac{p(x)}{p_0(x)} \tag{2.16}$$

where now, rather than the entropy constraint seen previously, this information processing constraint naturally takes the form of the KL divergence. This is maximised by the equilibrium distribution:

$$p(a) = \frac{1}{Z} p_0(x) e^{\beta U(a)} \tag{2.17}$$

which allows agents to trade off their expected utility gain with information-processing costs from their prior beliefs. This foundational framework of Ortega and Braun, 2013 is utilised and extended throughout this thesis for representing informationally-constrained reasoning.

2.3.2 Strategic Decision-making

Until now, we have considered a decision-maker choosing from a set of options with pre-defined utilities and a fixed environment. Frequently, however, the decisions to be made depend on the choices of other agents. This configuration brings us to strategic decision-making, where the decisions are made by considering the choices of other agents. Throughout this thesis, we focus on non-cooperative games, where the agents compete with each other (such as in economic markets).

2.3.2.1 Game Theory

In competitive (non-cooperative) games of two or more players, a fundamental solution concept is Nash Equilibrium, where no player has any incentive to deviate from their chosen strategy. For all finite games, a Nash Equilibrium solution exists (Nash, 1951).

These solutions can be split into pure strategy equilibria and mixed strategy equilibria. In pure strategy equilibria, all players play according to a pure strategy. With mixed strategy equilibria, at least one player is using a mixed strategy.

Pure Strategies. A pure strategy details an exact strategy a player will play. For example, if faced with three strategies A_1, A_2, A_3, a pure strategy is playing any one of these with certainty, e.g. $p(A_1) = 1$.

Mixed Strategies. Rather than playing one strategy with certainty, mixed strategies assign probabilities to each strategy. With this in mind, pure strategies can be seen as a special case where the selected strategy has $p = 1$, and all others $p = 0$. To see why mixed strategies are required, consider the game of Rock Paper Scissors. If any player were to play a pure strategy (A_1=rock, A_2=paper, A_3=scissors), this would be easily exploitable. For an equilibrium to arise, players must randomise amongst each strategy. If players weight each strategy equally $p(A_1) = p(A_2) = p(A_3) = \frac{1}{3}$, there is no incentive to deviate, as there there is no possibility to exploit the move. This weighting is the unique mixed Nash equilibrium for this game. While this is a simplistic example, it shows the need and use of mixed strategies.

Mixed strategies are used more generally throughout this thesis to represent uncertainty in play and allow erroneous play from information-constrained decision-makers (e.g. an agent may not choose the best option with certainty). This representation better models human behaviour and leads to behavioural game theory (Camerer, 2010).

2.3.2.2 Behavioural Game theory

While the game-theoretic concepts outlined above are based on perfect rationality, frequently deviations from this perfect rationality are observed. Below we review the two most relevant approaches for incorporating bounded rational play into game theory. A more comprehensive analysis of behavioural game theory approaches is presented in (Camerer, 2003; Camerer et al., 2004a; Camerer, 2010).

Quantal Response Equilibrium. Quantal Response Equilibrium (QRE) (McKelvey and Palfrey, 1995; McKelvey and Palfrey, 1998) relaxes the notion of best-response, allowing for erroneous play. Above, we have seen how enforcing an entropic constraint on the maximum expected utility principle leads to an information-constrained decision-maker, which can account for these errors in play. This same specification is frequently used in QRE:

$$P_{i,j} = \frac{\exp(\beta\,\mathrm{E}[U_{i,j}(P_{-i})])}{\sum_k \exp(\beta\,\mathrm{E}[U_{i,k}(P_{-i})])} \tag{2.18}$$

where P is the probability of player i choosing strategy j, and P_{-i} is the probability distribution of the other players taking each strategy. When $\beta \to \infty$, the original Nash behaviour is recovered, and $\beta < \infty$ represents bounded rational play. The key here is that to maintain an equilibrium, we must calculate the fixed point solution (since P_{-i} also depends on $P_{i,j}$). For example, if we have a two-player (a and b) game with two outcomes (H or T), we get the following

$$P_{a,H} = \frac{\exp(\beta\,\mathrm{E}[U_{a,H}(P_b)])}{\exp(\beta\,\mathrm{E}[U_{a,H}(P_b)]) + \exp(\beta\,\mathrm{E}[U_{a,T}(P_b)])}, P_{a,T} = 1 - (P_{a,H}) \tag{2.19}$$

$$P_{b,H} = \frac{\exp(\beta\,\mathrm{E}[U_{b,H}(P_a)])}{\exp(\beta\,\mathrm{E}[U_{b,H}(P_a)]) + \exp(\beta\,\mathrm{E}[U_{b,T}(P_a)])}, P_{b,T} = 1 - (P_{b,H}) \tag{2.20}$$

which can be solved with various approaches to find the resulting equilibrium (McKelvey et al., 2016).

Level-k models. Level-k type models (Stahl and Wilson, 1995) relax the notion of mutual-consistency, but maintain the notion of best-response. Unlike QRE, level-k type models do not necessarily give rise to equilibrium. Instead, closer to ABM type rules, they progress "bottom-up". A level-0 player acts according to some basic predetermined rule, a level-1 player then exploits this fact and best responds based on the other player being level-0. Likewise a level-2 player for level-1, and so on, with level-k best responding to level-$k-1$. This representation has a number of benefits, for example, when computing equilibrium is impractical, or when the resulting dynamics are out-of-equilibrium.

There are various extensions of level-k type models, most notably, Cognitive Hierarchy (Camerer et al., 2004b), which fits a distribution of $< k$ thinkers, rather than assuming all opponents are at $k-1$. Most commonly, a one-parameter Poisson distribution is utilised (Chong et al., 2005).

This thesis builds upon these ideas, proposing an alternative integrated approach for modelling bounded rational reasoning in strategic decision-making.

2.4 Summary

This section covered key relevant topics from complexity economics, information theory, and behavioural game theory. We have outlined how techniques from complexity economics can provide an alternate view of economic markets and how information theory can provide a generalised framework for quantifying processing constraints and performing inference. Additionally, we have shown that agent-based computational economics can be utilised to analyse nonequilibrium emergent phenomena from a "bottom-up" perspective. This background gives an important foundation for the thesis, which we build upon to model the endogenous emergence of stylised facts and crises through information-constrained decision-making.

CHAPTER 3

The impact of social influence in an agent-based model of the Australian housing market

We begin with developing a computational agent-based model of the Australian housing market able to explain and forecast key price movements. Importantly, the model makes no assumptions about equilibrium, efficiency, or rationality. Instead, the interactions between heterogeneous bounded rational agents acting according to simple behavioural heuristics are enough to explain key market phenomena. We also demonstrate how various factors such as social pressures better capture actual market dynamics than stricter assumptions, and how agent heterogeneity can explain differences between buyer behaviour (e.g. investors vs first-time-home-buyers).

Journal of Economic Interaction and Coordination (2023) 18:5–57
https://doi.org/10.1007/s11403-021-00324-7

REGULAR ARTICLE

The impact of social influence in Australian real estate: market forecasting with a spatial agent-based model

Benjamin Patrick Evans[1] · **Kirill Glavatskiy**[1] · **Michael S. Harré**[1] · **Mikhail Prokopenko**[1]

Received: 10 September 2020 / Accepted: 19 February 2021 / Published online: 22 March 2021

Abstract

Housing markets are inherently spatial, yet many existing models fail to capture this spatial dimension. Here, we introduce a new graph-based approach for incorporating a spatial component in a large-scale urban housing agent-based model (ABM). The model explicitly captures several social and economic factors that influence the agents' decision-making behaviour (such as fear of missing out, their trend-following aptitude, and the strength of their submarket outreach), and interprets these factors in spatial terms. The proposed model is calibrated and validated with the housing market data for the Greater Sydney region. The ABM simulation results not only include predictions for the overall market, but also produce area-specific forecasting at the level of local government areas within Sydney as arising from individual buy and sell decisions. In addition, the simulation results elucidate agent preferences in submarkets, highlighting differences in agent behaviour, for example, between first-time home buyers and investors, and between both local and overseas investors.

Keywords Agent-based modelling · Spatial models · Housing markets · Decision-making · Social influence · Trend following

JEL Classification H31 · R31 · C63 · O18 · D91

The authors are thankful to Paul Ormerod, Adrián Carro and Markus Brede for many helpful discussions of the baseline model. The authors acknowledge the HPC service at The University of Sydney for providing HPC resources that have contributed to the research results reported within this paper. The authors would also like to acknowledge the Securities Industry Research Centre of Asia-Pacific (SIRCA) and CoreLogic, Inc. (Sydney, Australia) for their data on Greater Sydney housing transactions.

Benjamin Patrick Evans
benjamin.evans@sydney.edu.au

[1] Centre for Complex Systems, The University of Sydney, Sydney, NSW, Australia

1 Introduction

Within economic markets, housing markets are unique for a variety of reasons. The combination of durability, heterogeneity, and spatial fixity amplifies the role of the dwellings' perceived value and the buyers and sellers' expectations (Alhashimi and Dwyer 2004). The extremely high cost of entry and exit into the market (with moving fees, agent fees, etc.) further complicates the decision-making of participating households (Huang and Ge 2009). There are long time delays in the market response as houses cannot be erected instantaneously to accommodate an increase in demand (Bahadir and Mykhaylova 2014). The fact that real estate can be seen as both an investment asset and a consumption good (Piazzesi et al. 2007) [and even a status good, Wei et al. (2012)] magnifies the impact of social influence on both decision-making and resultant market dynamics and structure.

Consequently, housing markets are notoriously difficult to model as the ensuing market dynamics generates volatility, with nonlinear responses, and "boom–bust" cycles (Sinai 2012; Miles 2008; Burnside et al. 2016), making traditional time series analysis insufficient. Nonlinear dynamics of housing markets are ubiquitous, being observed throughout the world, from Tokyo (Shimizu et al. 2010) to Los Angeles (Cheng et al. 2014).

Traditional economic modelling methods, such as dynamic stochastic general equilibrium models (DSGE), typically use representative aggregated agents while making strong assumptions about the behaviour of the markets (rational and perfect competition). Such representative agents may be limiting for economic models (Gallegati and Kirman 1999). Furthermore, these assumptions (and many other traditional economic assumptions) are known to be inadequate in housing markets, motivating a well-recognised need for change in housing market modelling (McMaster and Watkins 1999). In addressing this need, a specific type of models, called agent-based models (ABM) has been applied. ABMs aim to capture markets from the "bottom up" (Tesfatsion 2002), i.e. by focusing on the decision-making of individual agents in the market, possibly influenced by non-economic factors. In this sense, ABMs are capable of modelling macroeconomies from micro (i.e. agent-specific) behaviour (LeBaron and Tesfatsion 2008) and analysing the economic decision-making in counterfactual settings. While ABMs have shown promise in housing market modelling (Geanakoplos et al. 2012) [and wider economic modelling, Poledna et al. (2019)], current ABMs themselves are not exempt from some limitations. Many of the existing housing ABMs tend to introduce at least one of the following constraints: the spatial structure of markets is neglected (meaning submarkets are not considered), perfect information is still assumed, and/or the impact of social influence on decision-making of individual agents is underestimated.

A fine-resolution model of spatiotemporal patterns within such markets is desirable: it would give an understanding of how market dynamics shape within local areas, explaining how the pricing structure directly affects the agents' mobility over time (i.e. by forcing households out of certain regions due to gentrification and higher cost of living). Furthermore, within such housing markets [and in fact, many economic markets, Conlisk (1996)], it is also known that agents do not act perfectly rational (Wang et al. 2018, instead following bounded rationality Simon 1955, 1957). Firstly, humans are often influenced by social pressure (e.g. herd mentality), with the decisions being made purely based on social pressure rather than a perfectly rational choice. Secondly, it is difficult to process all the relevant information in the market (i.e. it is impractical for an agent to be able to view every dwelling listing within the housing market). Thus, it is unreasonable to assume that agents act in a perfectly rational manner, yet this is what many current housing market models assume (despite ABMs not intrinsically requiring these assumptions to be made).

To address these limitations, we introduce a spatial agent-based model, in which the constraints imposed by various search and mobility costs create effective spatial submarkets. These submarkets are modelled graph-theoretically, with a graph-based component used in both representing imperfect information and modulating social influences. The spatial ABM is then capable of capturing the "boom–bust" cycles observed in Australia (in particular, Greater Sydney) over the last 15 years as arising from individual agent decisions. That is, we show complex nonlinear market dynamics as arising through individual buy and sell decisions in the market, with introduced factors which affect and explain the decision-making behaviour. The model succeeds in forecasting nonlinear pricing and mobility trends within specific submarkets and local areas. In exploring the pricing dynamics, we focus on the influence of imperfect spatial information and the role of social influence on agent decision-making. In doing so, we identify the salient parameters which drive the overall dynamics and decision-making, and pinpoint the parameter thresholds, beyond which the resultant dynamics exhibit strong nonlinear responses. These thresholds allow us to distinguish between different configurations of the market (e.g. markets with supply or demand dominating) and differences in agent behaviour. In addition, we identify and trace specific interactions of parameters, in particular the interplay of social influences, such as the fear of missing out and the trend-following aptitude, in the presence of imperfect spatial information.

The remainder of the paper is organised as follows. In Sect. 2, we provide an overview of agent-based models of housing markets. In Sect. 3, we outline the baseline model, while in Sect. 4 we outline the proposed spatial extensions and new parameters. In Sect. 5, we analyse the sensitivity and parameters of the model, before presenting the results and discussion in Sect. 6. In Sect. 7, we provide conclusions and highlight future work.

2 Background

2.1 Agent-based models of housing markets

One of the pioneering works for agent-based modelling (ABM) of the housing markets was by Geanakoplos et al. (2012) [and extended further in Axtell et al. (2014), Goldstein (2017)], where the Washington DC market was modelled from 1997 to 2009 in an attempt to understand the housing boom and crash. Macroeconomic experiments were then conducted to see how changing underlying factors, such as interest rates or leverage rates, would affect this pricing trend.

Baptista et al. (2016) model the UK housing market to see the effects that various macroprudential policies have on price cycles and price volatility. Gilbert et al. (2009)

also look at the English housing market, varying exogenous parameters and policies, and tracking the effect these have on median house prices. Likewise, Carstensen (2015) explores the Danish housing market and macroprudential regulations, such as income and mortgage rate shocks.

Ge (2017) analyses how housing market bubbles can form (and bust) purely endogenously without external shocks, due to leniency and speculation of agents. Kouwenberg and Zwinkels (2015) also show an ABM can "endogenously produce boom-and-bust cycles even in the absence of fundamental news".

A recent ABM of the Australian housing market proposed by Glavatskiy et al. (2020) explained the volatility of prices over three distinct historic periods, characterised by either steady trends or trend reversals and price corrections. This model highlighted the role of the agents' trend-following aptitude in accurately generating distinct price dynamics, as detailed in Sect. 3. In this paper, we further develop this model by introducing several features directly capturing social influences and bounded rationality in agent decision-making, as elaborated in Sect. 4.

Traditionally, the modelling goal is to explain the housing market pricing (rather than predict its trajectory) and trace how possible macroeconomic policy changes may have affected the dynamics. Here, instead, we focus on predicting the pricing dynamics beyond the period covered by the current datasets (i.e. presenting out-of-sample forecasting) as arising from individual buy and sell decision-making behaviour in the market (as opposed to a black box technical analysis machine learning-based approach), as well as illuminating agent preferences within specific housing submarkets. This motivation is aligned with the growing suggestions that agent-based models should be predictive (Polhill 2018) [which has admittedly been met with some resistance, Edmonds and ní Aodha (2018)]. Agent-based models have recently been shown to outperform traditional economic models, such as vector autoregressive models and DSGE models for out-of-sample forecasting of macrovariables (GDP, inflation, interest rates, etc.) (Poledna et al. 2019). For example, it was demonstrated that an ABM can outperform standard benchmarks for out-of-sample forecasting in the US housing market (Kouwenberg and Zwinkels 2014), while successful out-of-sample forecasting was carried out by Geanakoplos et al. (2012) as well.

2.1.1 Spatial models

Spatial distribution of houses and dependencies between market trends on spatial patterns have been recognised as important and desirable features (Goldstein 2017). For example, Baptista et al. (2016) describe the spatial component as one that is "highly desirable", yet "this approach greatly increases the complexity of the models and hence most spatial ABMs in the field listed below make use of a highly simplified representation of the environment, often in the shape of small grids".

Spatial agent-based models have also shown to be useful in a variety of other areas such as epidemic modelling (Chang et al. 2020; Cliff et al. 2018), cooperative behaviour (Power 2009), and symbiotic processes (Raimbault et al. 2020). Despite the promise shown by housing ABMs, there are currently only relatively few spatial housing market models with the capacity to accurately forecast nonlinear price dynamics.

The seminal works in spatial housing ABMs are by Ge (2013) and Ustvedt (2016). Both use a matrix-based approach, with the region being arranged on a two-dimensional grid. In Ge (2013), each cell (row/column) in the grid is assigned a neighbourhood quality (endogenous) and a nature quality (exogenous). The neighbourhood quality is a measure of attractiveness which aims to capture concepts such as safety, and is dependent on agents that live in that region (which can change in the model, thus endogenous). In contrast, nature quality is based on outside factors not changed by the model, such as distance to a beach or weather (thus exogenous). Data used in this work are abstract; that is, it is not calibrated to a particular city, but rather used to trace how these factors affect the trends.

Ustvedt (2016) also uses a two-dimensional grid for a NetLogo model. However, an important additional spatial step is made: district borders are incorporated using GIS data (somewhat similar to what we propose in our model with the graph-based approach, however, there are important differences which we outline below), and the model is calibrated based on Oslo, Norway.

Another work is Pangallo et al. (2019), which models theoretical (i.e. not calibrated to any specific region) income segregation and inequality, using a spatial agent-based model, and the effect this may have on house prices. Again, this approach uses a two-dimensional grid for the spatial component. This model assumes a monocentric city, and measures the "attractiveness" of a location, based on the distance to the (generic) city centre. The main contribution of this work is a mathematically tractable spatial model for capturing income segregation. The effects are related to the house prices, where unequal income is shown to lower the house price globally.

Unexplored Extensions The spatial work outlined above provide strong models which achieve their purpose of policy understanding and effect of market shocks in Ge (2013) and Ustvedt (2016). However, the spatial component is often considered a secondary point of the models, which means there are some key additional insights which have yet been unexplored. Particularly, the presence of area-specific submarkets have not been explicitly modelled, and analysis into the spatial preferences of agents within the market (and submarkets) have not been explored.

For example, in studies of Ge (2013) and Ustvedt (2016) the spatial contribution affects the initial price and the supply of dwellings for a given grid cell (location and population). However, at every simulated step the effects of the spatial component do not extend to varying search costs or probability of listing for agents. Furthermore, the spatial component is not utilised to initialise agent characteristics (such as income, wealth, etc.), based on the areas in which they reside. These area-specific agent characteristics and behaviours become particularly important for capturing submarkets across various areas, especially when real-world data are available.

Another assumption of existing spatial models is that of a monocentric city, meaning a distance metric such as "distance to centre" is used for measuring attractiveness, which becomes problematic for polycentric cities or with agents who have no desire to live within the "centre" when analysing agent preferences. Furthermore, computing these distances in a 2d grid can often be misinformative, as moving across a region border (i.e. into a new zone) often incurs a far larger cost than moving a cell inward into the same zone. To address this, we explicitly capture this feature in the proposed graph-based spatial extension. Distances are measured as the shortest path through the graph, with nodes representing various regions (contained within boundaries). The graph-based approach is particularly useful, as no monocentric assumption is made.

Furthermore, as detailed area-specific analysis remains largely unexplored, the distributions of prices within individual areas is often not considered, instead, using a representative mean or median rather than sampling from the actual underlying distributions for each area, which may fail to capture certain area trends. This is often caused by the lack of underlying data. In our work, we use several contemporary datasets, such as SIRCA–CoreLogic and the Australian Census datasets, constructing the relevant pricing probability density functions for each area.

In summary, in contrast to the grid-based approach commonly used, in this work, we propose an extensible graph-based approach which is described in Sect. 4. Such an approach allows us to further exploit the spatial component by introducing submarkets with graph-based search costs and initialisation of agent characteristics based on areas, giving insights into area-specific submarkets which have not yet been explored. In addition, the graph-based approach does not assume a monocentric region, allowing for polycentric cities [which Greater Sydney is developing towards Greater Sydney Commission (2018), Crosato et al. (2021)] to be modelled more effectively.

3 The Baseline model

This work extends the work of Glavatskiy et al. (2020), which we will refer to as the "Baseline" method. In this section, we describe basic features of the baseline model, which the present work carries over.

3.1 Agents

There are three key agent types in the model: dwellings, households, and the bank.

A dwelling is a "physical" property, e.g. a house, apartment or condo. Each dwelling has an intrinsic quality, which reflects its hedonic value (e.g. large house or existence of a pool). The quality is fixed during the simulation. The quality of a dwelling is used as a reference for determining its listing price and dwelling payments. All dwellings have an owner. Dwellings can be rented or sold. A rental contract is a binding agreement between the owner and renting household. Dwellings may be vacant at any period (which means that they are not rented out).

A household represents a person or group of people (i.e. a family), which reside within Greater Sydney. Additionally, the model contains overseas agents, which can participate in the market but do not reside in the Greater Sydney region. Households have heterogeneous monthly incomes and liquid cash levels. Households can own several dwellings, but can only reside in one (overseas agents do not reside in any dwelling). When purchasing a dwelling, households always choose the most expensive dwelling they can afford. If they can afford to buy a dwelling, they always attempt to do so, putting a market bid (see below). Households that own more than one dwelling attempt to rent the additional dwellings out, and, if successful, receive rental payments as a contribution to their liquid cash. Households that do not own a dwelling rent one. Households pay tax based on their income and ownership.

The bank combines the functions of a commercial bank and the regulatory body, controlling various financial characteristics, such as income tax rates, mortgage rates, overseas approval rates, mortgage approvals, and mortgage amounts (how much can be lent to a particular household).

3.2 Behavioural rules

The agents' behaviour is governed by the price they are willing to sell their dwelling for, the listing price, and the price they can afford to buy a new dwelling, the bid price. The basis for the pricing equations comes from the pioneering housing market ABM of Axtell et al. (2014).

The household bid price, i.e. their desired expenditure, is modulated by the household's monthly income $I[t]$ according to Eq. (3.1). Following Axtell et al. (2014), the desired expenditure formulation is the result of an analysis into income dynamics, motivated by the $\frac{1}{3}$ of income on housing expenses heuristic, with modifications to capture a wider range of heterogeneity and expenditures.

$$B[t] = H\frac{U_{\mathrm{b}}[t]\phi_{\mathrm{b}}I[t]^{\phi_{\mathrm{I}}}}{\phi_{\mathrm{M}}[t] + \phi_{\mathrm{H}} - h * \Delta_{\mathrm{HPI}}[t]} \tag{3.1}$$

H is a uniformly random value between $1 \pm b_{\mathrm{h}}/2$ for each bid, where $b_{\mathrm{h}} = 0.1$ is the listing heterogeneity parameter. This allows variation in bid prices around a central value (i.e. uniform prices within a range around the centre, in this case, $\pm 5\%$). $U_{\mathrm{b}}[t]$ is the urgency of a household to buy a dwelling, which is equal to 1 if the household has recently not sold any dwelling, and is larger than 1 by a term proportional to the number of months since the last sale otherwise. Furthermore, ϕ_{I} and ϕ_{b} are the income modulating parameters, which are calibrated from the mortgage-income regression for that period of interest. The resulting expenditures are sublinear with income, meaning high-income agents spend a lower percentage of their income on housing than their lower-income counterparts. In addition, $\phi_{\mathrm{M}}[t]$ is the mortgage rate at time t, while ϕ_{H} is the annual household maintenance costs. Finally, h is the trend-following aptitude and $\Delta_{\mathrm{HPI}}[t]$ is the change in house price index (HPI) over the previous year.

The bid price of the overseas investors is determined as an average, given the total volume and quantity of the approved overseas investments by the Foreign Investment Review Board. Several aspects of the overseas investments are detailed in "Appendix K".

The dwelling list price $P[t]$ at time t is modulated by the quality of the dwelling Q according to Eq. (3.2). Again, the basis for the listing price equation comes from the estimation results of Axtell et al. (2014), where the formulation arose from a comprehensive set of real estate transaction data.

$$P[t] = H\frac{b_{\ell}\overline{Q}_{\mathrm{h}}S[t]^{b_{\mathrm{s}}}(1 + D_{\mathrm{h}}[t])^{b_{\mathrm{d}}}}{{}^{31}U_{\ell}[t]} \tag{3.2}$$

H is again a uniformly random parameter that behaves the same as in Eq. (3.1). $b_\ell = 1.75$ is the listing price factor, showing the extent to which the seller tends to increase the listing price. Furthermore, $\overline{Q}_{\mathrm{h}}$ is the average sale price of the 10 dwellings with the most similar quality to the dwelling h for sale. In addition, $S[t]$ is the market average of the sold-to-list price ratio, $b_{\mathrm{s}} = 0.22$ is the sold to list exponent parameter, $D_{\mathrm{h}}[t]$ is the number of months the dwelling has been on the market, and $b_{\mathrm{d}} = -0.01$ is the number of months exponent parameter. Finally, $U_\ell[t]$ is the urgency to sell the dwelling, which is equal to 1 if the household is not in financial stress, and increases proportionally to the number of months in financial stress otherwise.

Banks approve households desired expenditure based on the bank's lending criteria. The bank uses the households liquidity and monthly income for determining an appropriate amount to lend and offers the corresponding loan to the household. If the loan amount is greater than $p_{\mathrm{d}} \times B[t]$ (with $p_{\mathrm{d}} = 0.6$), then the household accepts the loan; otherwise, the household skips this round of the market.

The parameters from the baseline method are presented in Table 2, and sensitivity analysis around the default values of these parameters is given in "Appendix B", and more in-depth explanation is given in the study of Glavatskiy et al. (2020).

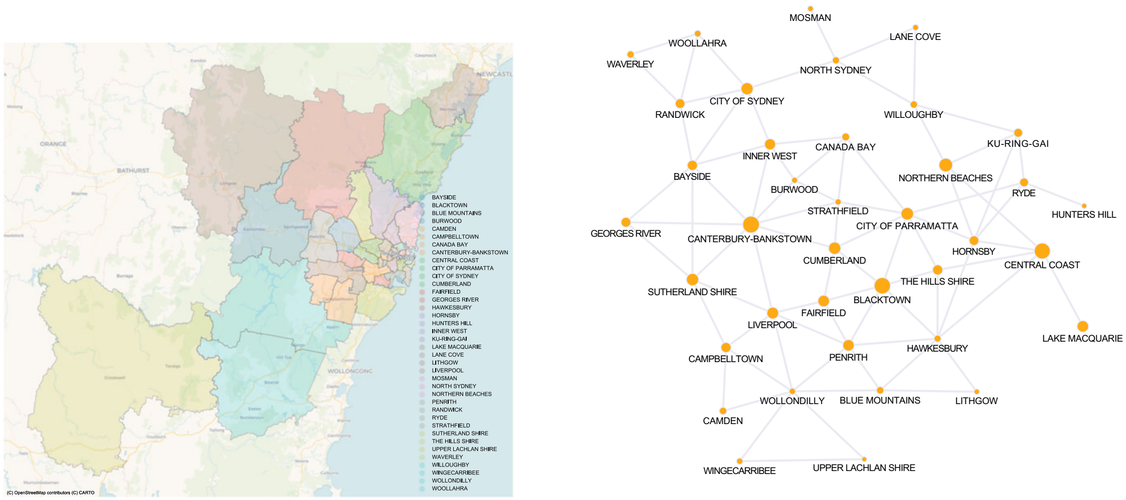
3.3 Market algorithm

An equilibration period is run at the start of the simulation to build up a "history" of the market, allowing for sales etc. to take place before the true simulation begins. Household and housing characteristics (such as income or construction) are smoothly extrapolated in the equilibration period, to match the corresponding values at the beginning of the actual simulation.

The model runs in sequential steps, with each step representing one month of actual time. That is, one sequential step (or tick) in the simulation corresponds to one month in the actual housing market. During every step, the model makes several market updates:

1. The city demographics are updated (new dwellings and households created to match the actual projected numbers for the Greater Sydney region).
2. Each household receives income and pays its living costs: non-housing expenses, maintenance fees and taxes (if owning a dwelling), rent (if renting). The balance is added to or subtracted from the household's liquid cash.
3. Expiring rental contracts are renewed.
4. Dwellings are placed on sale.
5. Households put their bids for buying.
6. The buyers and sellers are matched (described below)
7. The households receive mortgages and mortgage contracts from the bank, and the balance sheets of both buyers and sellers are updated.

To match buyers and sellers, bids and listings are sorted in descending order. Each listed dwellings is then attempted to match with the highest bid. If the bid price is higher than the list price, then the deal is made with probability $p_{\mathrm{m}} = 0.8$. Otherwise, the listing is considered unattended and the next one attempts to match. The pseudo-code for the process is given in "Appendix C"32



(a) LGA Boundary Map **(b)** LGA Network

Fig. 1 Greater Sydney LGAs. On the left, we see the raw GIS data. On the right, the processed graph (with nodes scaled based on population size)

4 Model extension

In this section, we develop the spatial component of an ABM housing market, effectively introducing submarkets into the model. Specifically, we investigate how social aspects influence selling a dwelling, as well as account for the agents' preferences to buy in a nearby neighbourhood when purchasing a dwelling. In doing so, we investigate how pricing dynamics are driven by individual agent decisions. The spatial component also affects how households are initialised (i.e. what neighbourhood they should belong to), and how the prices of nearby dwellings may affect the listing price.

4.1 Spatial component

Greater Sydney is composed of 38 local government areas (LGAs), each of which contains several suburbs (and postcode areas). The data provide sales at a postcode level and the LGA level. However, the postcode data may be too granular as the number of listings in a given time period for small areas could be low or even zero. For this reason, we analyse the data at the LGA level, but the proposed approach is general and can be used at any level of granularity (i.e. over countries, states, cities, government areas, postcodes, suburbs, or even individual streets) assuming the data are available. The LGAs are visualised in Fig. 1a.

4.1.1 Graph-based topology

To incorporate the LGA areas into the ABM, the data must be converted to an appropriate data structure. This is achieved by converting the map (from Fig. 1a) into an undirected graph G, with equal edge weights of 1 (i.e. unweighted, the weighted extensions are discussed below) shown in Fig. 1b. In doing so, the topology of the spatial relationships of the suburbs is preserved but not the exact locations, i.e. the x, y coordinates of latitude and longitude are not needed.

Each of the N LGAs (shaded polygon) in Fig. 1a is converted into a vertex v_i, $i \in [1, \dots, N]$. The cardinality of the set of all vertices V is $|V| = 38$ corresponding to the 38 LGAs. Two LGA areas v_i and v_j are adjacent to one another if they share a border (darker lines in Fig. 1a) and these borders are converted to the set of edges E all of weight one that form an adjacency matrix G for the LGAs. Formally there is a 2-D spatial region for Sydney (the map of Sydney): $\mathcal{M}$, composed of the N non-overlapping LGAs that form a complete cover of Sydney and each LGA shares at least one border with another LGA. LGA(x) associates a graph vertex x with an LGA, and Adj(l_i,l_j), defined for two LGAs l_i and l_j, is a function measuring the length of their common border in $\mathcal{M}$:

$$V = \{v_i \mid i \in [1, \dots, N],\ \mathrm{LGA}(v_i) \in \mathcal{M}\}, \tag{4.1}$$

$$E = \{e_{i,j} = 1 \mid v_i, v_j \in V,\ i \neq j,\ \mathrm{Adj}(\mathrm{LGA}(v_i), \mathrm{LGA}(v_j)) > 0\}. \tag{4.2}$$

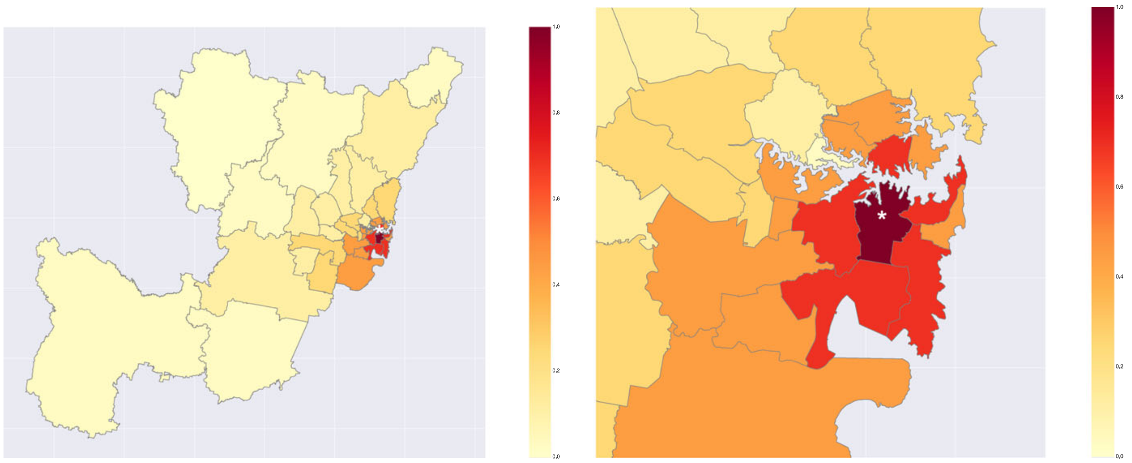
This definition implies $G(E, V)$ is a connected undirected graph, there are no disconnected subgraphs. An important edge is also added that represents the Sydney harbour bridge connecting Northern Sydney with the City of Sydney.

To calculate the distance between vertices v_i and v_j, the minimum path length (i.e. the path with the lowest number of edges) between the two vertices is used as edges are equally weighted: $\delta(v_i, v_j)$ denotes this shortest path. Because the edges have unit weighting the shortest paths are found using a simple breadth-first search. However, future extensions could consider edge weightings based on metrics such as real distance between centroids, travel time between centroids, or even adding additional edges for public transport links. In cases of weighted edges, Dijkstra's algorithm could be used to compute $\delta(v_i, v_j)$ instead.

4.1.2 Spatial submarkets

Dwellings are allocated initial prices based on the distribution of recent sales within their LGA and also populated according to the census data for dwellings in each LGA. A full description of the process is given in "Appendix I.2". Households are also distributed into LGAs based on census data, with renters then moving to LGAs in which they can afford, as described in "Appendix D".

The original dwelling list pricing equation from Eq. (3.2) is also now updated to be based on each dwelling's LGA. Rather than $\overline{Q}_\mathrm{h}$ being the average of the 10 most similar quality dwellings in the model, it is the average of the 10 most similar within the LGA. Likewise, $S[t]$ is the average sold-to-list price ratio for the dwelling's LGA (not overall). In this sense, *spatial* submarkets (LGAs) (Watkins 2001) capable of exhibiting their own dynamics are introduced. Recent research Bangura and Lee (2020) has shown the importance of submarkets in the Greater Sydney market, so capturing such microstructure is a key contribution of the proposed approach, as trends can be localised to specific submarkets (a feature not prominent in existing ABMs of housing markets).



(a) Overview **(b)** Zoomed

Fig. 2 Probability of viewing a listing based on buyers location (in this case the City of Sydney). Dark red indicates high probability, and light yellow indicates low probability (colour figure online)

4.2 Spatial outreach

In an actual housing market, a typical buyer does not review every listing in the entire city due to the high search costs and desire to live in certain areas. Rather, the buyer targets particular spatial submarkets, relating to a given area. In particular, listings immediately around the buyer's location are likely to be viewed with a higher probability than listings which are further away.

Therefore, it is unrealistic to assume perfect knowledge in an ABM of the housing market. To model this imperfect spatial information, we introduce an outreach term O, which determines the likelihood for a buyer located at v_i to view a listing located at v_j, as described by Eq. (4.3). According to this expression, the likelihood of viewing the listing decreases with the distance between the buyer and the listed dwelling. The outreach factor is illustrated in Fig. 2 for a buyer located in the LGA "City of Sydney".

$$O(v_i, v_j) = 1 - \frac{\delta(v_i, v_j)}{\max\limits_{k \in V} \delta(v_i, v_k)} \tag{4.3}$$

While the spatial outreach makes sense for first-time home buyers, for investors, the outreach becomes uniform, as they do not necessarily desire rental properties near where they reside. So for investors, we use $O(v_i, v_j) = 1, \forall i, j$.

To control the strength of the outreach, we introduce a new parameter $\alpha \in (0, 1)$, so the probability of viewing a listing $P_{\text{view}}(v_i, v_j)$ is given in Eq. (4.4).

$$P_{\text{view}}(v_i, v_j) = \alpha O(v_i, v_j)^2 \tag{4.4}$$

where α modulates the spatial information on dwelling listings: the higher the α, the more the listings are viewed by a potential buyer. That is, α adjusts the likelihood of viewing a listing, based on the distance to that listing. The effect α has on resulting decisions is further discussed in "Appendix E".

4.3 Spatial FOMO

In the baseline model, dwellings have a fixed probability of being listed, $p_{\mathrm{b}} = 0.01$. Here, we consider a spatial probability to list a dwelling located in a certain LGA l_i, which depends on the number of recent sales in l_i. For this, we introduce the "fear of missing out" (FOMO) parameter, denoted by β, which modulates the probability of listing a dwelling situated in l_i by the number of recent sales in l_i altering the agents' decision-making behaviour. In particular, if a high number of dwellings in l_i have been sold, then the owners of dwellings in l_i will be more likely to list their dwelling on the market. This will account for the possibility that if a certain LGA becomes a popular location for selling a dwelling, the owners of dwellings in this LGA would not want to miss an opportunity to sell their dwelling.

The spatial listing probability $p_{\mathrm{list}}(l_i)$ is expressed by considering the difference in the fraction of dwellings currently in l_i's submarket ($f_{l_i} = \mathrm{listings}_{l_i}/\mathrm{dwellings}_{l_i}$) with respect to the Greater Sydney average. A higher f_{l_i} means dwellings in l_i have been less likely to sell in the previous months (since all begin with a fixed probability of selling p_{b}) compared to the regions average. This spatial listing probability[1] is given by Eq. (4.5).

$$p_{\mathrm{list}}(l_i) = p_{\mathrm{b}} + p_{\mathrm{b}}\beta\left[\frac{f_{l_i}}{\sum_{\mathrm{a}\in V} f_{\mathrm{a}}/|V|} - 1\right] \tag{4.5}$$

The magnitude of β controls the strength of l_i's spatial submarket contribution to the listing probability. Rewriting Eq. (4.5) by denoting the term in the square brackets as x, i.e.

$$p_{\mathrm{b}} + p_{\mathrm{b}}\beta x$$

we see that if both β and x are negative, then the listing probability will be higher than the baseline level p_{b}. Therefore, if the number of dwellings for sale in a particular LGA is less than the average in the whole city, this means that this particular LGA's submarket has been clearing fast, and homeowners in this LGA will be more likely to list their dwelling. In contrast, if x is positive, then dwellings in the current LGA are not clearing as fast as in the other LGAs, so homeowners in this LGA will be less likely to list. Conversely, a positive β results in an opposite effect. If an LGA has comparatively few listings, the homeowner from this LGA will be less likely to list a dwelling for sale, whereas if this LGA has many listings, the probability to list a dwelling there increases. In this way, β has a direct effect on the supply of dwelling listings and alters the decision-making behaviour of sellers.

5 Optimisation and sensitivity analysis

The selection of appropriate parameters is an important step in agent-based modelling, and in most existing work, parameters are selected over the entire period of interest.

[1] p_{list} can technically be < 0 or > 1, so p_{list} is capped to be between 0 and 1, in order to be a true probability, although this is exceptionally rare and does not appear to occur in Fig. 4.

Here, we instead adopt a machine learning approach whereby we split the time series into a training and a testing portion; this ensures the model also performs well for the unseen (i.e. the testing) portion of data, and avoids biasing the selection of parameters by considering the entire time period.

We use Bayesian hyperparameter optimisation (Snoek et al. 2012) to find appropriate combinations of parameters. The training set is used for parameter selection, whereas the test set is only used to evaluate the goodness of fit of the model after the optimisation process is completed (in Sect. 6). We stress that the test set is never seen by the optimisation process. This is an important distinction from previous work (Glavatskiy et al. 2020), which constructs the models by considering all time points, and as such cannot be considered true predictions, unlike here where the model can be seen as a true predictor of future pricing trends. Optimisation details are given in "Appendix E".

5.1 Optimisation results

We run several optimisation processes in order to quantify the contribution of each component on agent decision-making and the effect these decisions have on resulting market dynamics. Each method follows the same optimisation process.

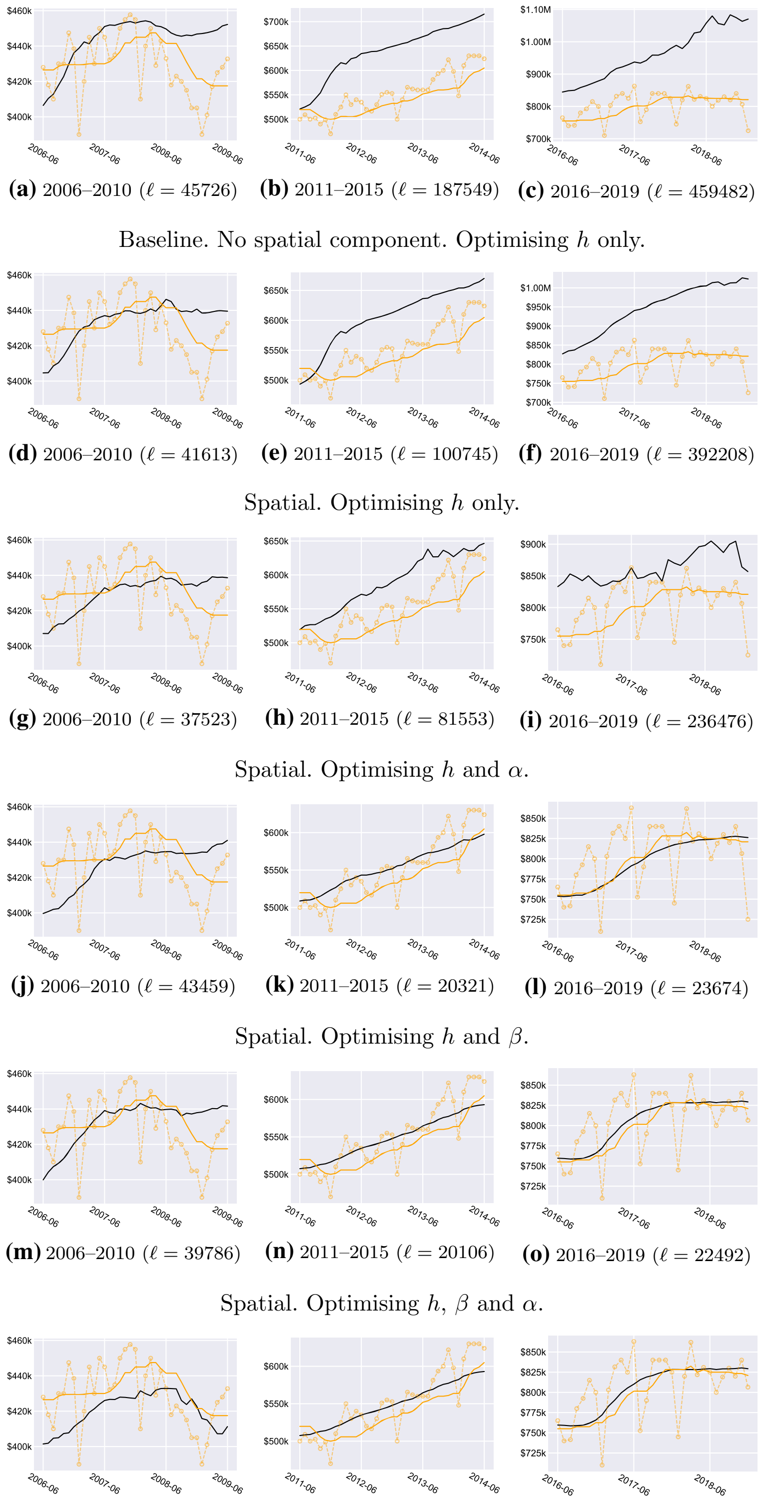
A "baseline" method is run, where there is only a single area (Greater Sydney), and only h is optimised for (with perfect knowledge[2] and no β optimisation). A spatial version of the baseline, with the 38 Greater Sydney LGAs as areas. Again, only h is optimised for (with perfect knowledge and no β). We then run pairwise combinations, so h and α, and h and β. We never run without optimising h, as this was the key decision-making parameter in the original model. Finally, we run the proposed extensions in their entirety—that is, we optimise over all three parameters (h, β, and α). We then apply a global constraint (on the result of the training optimisation) that 2006–2010 must exhibit a peak, with details outlined in "Appendix E.3". The results are presented visually in Fig. 3. We also provide additional analysis into the network architecture itself in "Appendix G".

Looking at the resulting plots, we can see that with the introduction of each new component, the resulting values of the loss function ℓ (see Eq. E.1) over the training period is reduced at each step, with the proposed extensions achieving the minimal ℓ, indicating that the predictions arising from the resulting agent decisions follow more closely to those in the actual market. From this point forward, we focus on the proposed extension in its entirety, due to the improved performance in all three periods.

5.2 Resulting parameters

Exploring the entire parameter search space would be computationally prohibitive. Bayesian optimisation intelligently explores this search space, balancing exploration and exploitation with the use of an acquisition function allowing more emphasis to be placed on well-performing or unexplored regions of the space (Shahriari et al.

[2] Perfect knowledge in this paper is assumed to mean $\alpha = 1$, $O(v_i, v_j) = 1$, i.e. ability to view every listing across all of Greater Sydney, i.e. $\mathcal{M}$ in 4.1.1. 37



(a) 2006–2010 ($\ell = 45726$) **(b)** 2011–2015 ($\ell = 187549$) **(c)** 2016–2019 ($\ell = 459482$)

Baseline. No spatial component. Optimising h only.

(d) 2006–2010 ($\ell = 41613$) **(e)** 2011–2015 ($\ell = 100745$) **(f)** 2016–2019 ($\ell = 392208$)

Spatial. Optimising h only.

(g) 2006–2010 ($\ell = 37523$) **(h)** 2011–2015 ($\ell = 81553$) **(i)** 2016–2019 ($\ell = 236476$)

Spatial. Optimising h and α.

(j) 2006–2010 ($\ell = 43459$) **(k)** 2011–2015 ($\ell = 20321$) **(l)** 2016–2019 ($\ell = 23674$)

Spatial. Optimising h and β.

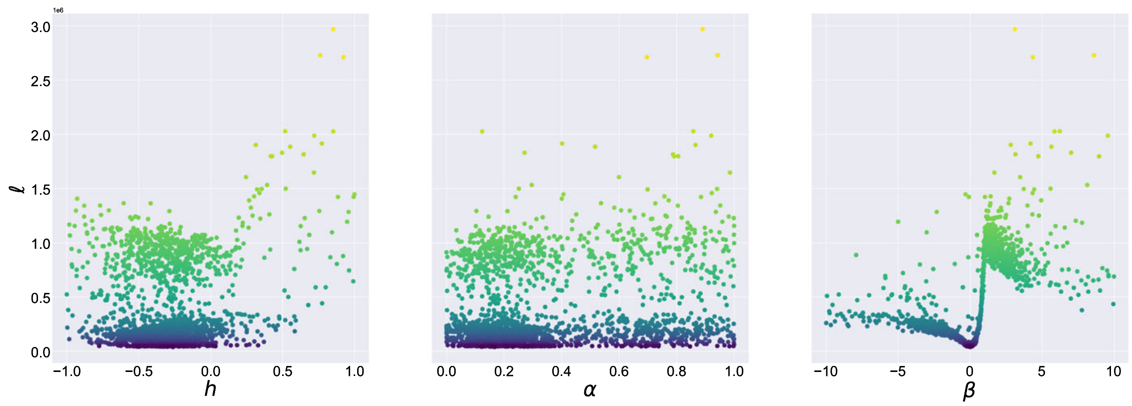
(m) 2006–2010 ($\ell = 39786$) **(n)** 2011–2015 ($\ell = 20106$) **(o)** 2016–2019 ($\ell = 22492$)

Spatial. Optimising h, β and α.

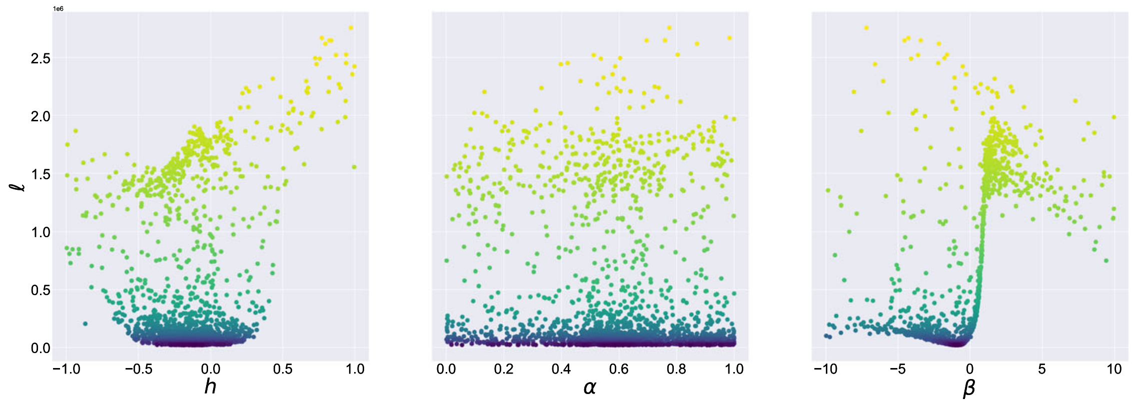
(p) 2006–2010 ($\ell = 64334$) **(q)** 2011–2015 ($\ell = 20106$) **(r)** 2016–2019 ($\ell = 22492$)

Spatial. Optimising h, β and α with global constraint in 2006–2010.

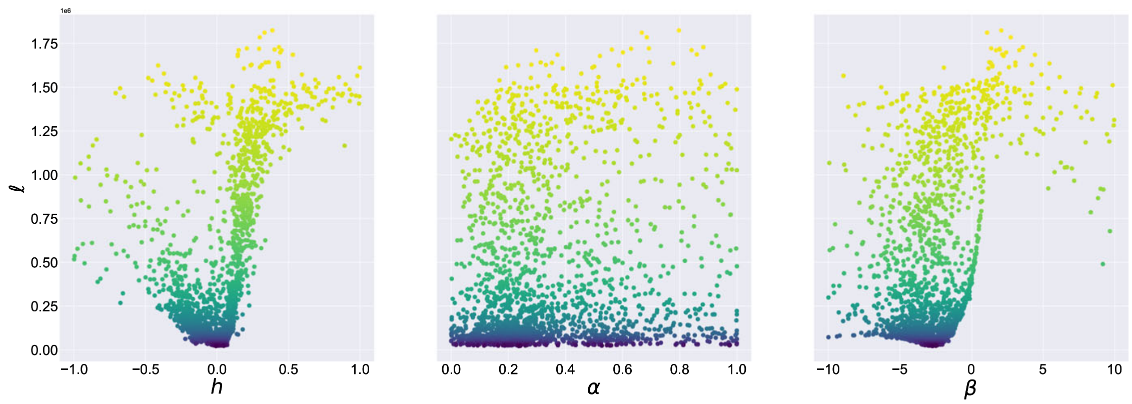
Fig. 3 Optimisation of the goodness of fit for dwelling prices across all models (for the training period). The orange lines are from the SIRCA–CoreLogic data (the solid line represents the rolling median, and the dotted line represents the month-to-month median). The black line shows the best fitted path from the model



(a) 2006–2010



(b) 2011–2015



(c) 2016–2019

Fig. 4 Univariate parameter analysis. The x-axis represents the parameter value, and the y-axis represents the loss (with logarithmic colours for consistency across the various loss plots). The other parameters are averaged over to provide the one-dimensional view (colour figure online)

2015). The resulting exploration is visualised in Fig. 12. From this, we can see the regions of interest, with dark sections indicating areas with the lowest loss ℓ (and more sample points being present in such areas). We can see the search space is fairly well explored in all cases, with obvious regions of well-performing parameter combinations (pairwise combinations are visualised in Fig. 13). While the 3D plot gives a high-level overview, it is difficult to visualise the contribution of each component. To facilitate this, we present a flattened one-dimensional view of each parameter, where the results

Table 1 Selected ABM parameters from the training period with Bayesian optimisation, rounded to 2 decimal places

	2006–2010	2011–2015	2016–2019
h	− 0.80	− 0.11	− 0.005
β	0.08	− 1.03	− 2.73
α	0.28	0.59	0.24

are averaged over the other 2 parameters to view the loss for 1 parameter at a time. This is shown in Fig. 4. From these plots, it can be seen how each parameter behaves in isolation (noting that such plots do not capture the parameter interactions). The selected ABM parameters (the ones which had the lowest ℓ) are presented in Table 1. A sensitivity analysis for the parameters is performed in "Appendix F".

The parameter search space is uniform across the ranges given in Table 3, so an uninformative parameter would be sampled uniformly as well (since there would be no Bayesian preference the sampling would be approximately uniformly random); instead, we see clear unimodal peaks in almost all cases in Fig. 13, with the h and β parameters being normally distributed around the optimal value found, and α with a clear peak but non-normally distributed. This indicates each parameter seems to have a useful range, which is further verified in the sensitivity analysis in "Appendix F".

h: Recall that the HPI aptitude h directly influences the agents' bid price [given in Eq. (3.1)] and serves as the key trend-following parameter in the original model. We can see the value of h controls the contribution of the HPI over the previous year and as such it affects the bid price based on the markets state. This relationship depends on both the current HPI and the yearly difference in Δ_{HPI}.

The absolute value of h controls the magnitude of the contribution, and we can see 2006–2010 had the highest contribution indicating the largest market effect on bidding. In the original model, h also had a large magnitude but the opposite sign. The value for 2011–2015, $h = -0.11$, is very close to that chosen in the original model $h = -0.10$, which reflects the relatively consistent price dynamics, with agents mostly ignoring the market trend.

Interestingly, in 2016–2019, the chosen h is near zero. This means the denominator of Eq. (3.1) simplifies to:

$$\phi_{\mathrm{M}}[t] + \phi_{\mathrm{h}} - h * \Delta_{\mathrm{HPI}}[t] \simeq \phi_{\mathrm{M}}[t] + \phi_{\mathrm{H}}$$

meaning the price is dependent on the mortgage rate and homeownership rate, as ϕ_{H} is set based on the current value of HPI, this means the historical values are not being used, and instead, a much more "forgetful" market based only on the previous month's HPI is used for setting the price, with agents paying less attention to historical trends.

The HPI aptitude h also has clear optimal ranges for each period, with 2006–2010 in the range $[-0.75 \ldots 0]$, 2011–2015 in the range $[-0.5 \ldots 0.25]$ (which are of similar widths), and 2016–2019 in the much narrower range $[-0.2 \ldots 0.05]$. We can see a sharp transition occurring in 2016–2019 around 0.1, in which case the loss begins increasing drastically for any higher values.

β: This parameter (the right-hand column in Fig. 4) exhibits a very sharp transition around 0 for all years. This is because β has a direct relation to the supply and demand in the model, which drastically changes the dynamics based on the availability of properties. We see increasing contributions of β throughout the years, with 2016–2019 indicating the highest levels of β. This is perhaps reflective of the market, where people are increasingly following trends when it comes to the decision to sell a dwelling (perhaps an indicator of a "bursting" bubble, with a large cascading sell-off). This indicates sharper peaks and dips are likely to occur in the future, with decision-making being made increasingly under the pressure of social influence.

α: This parameter does not present as clear of a transition, or defined range of optimal vales as h or β. This indicates the value of α has less of an impact on the results when compared to h and β (as is confirmed in the sensitivity analysis in "Appendix F"). The likely reason for this is since the buyer always attempts to purchase the most expensive viewed dwelling, the varying levels of α do not have as large of an effect on the outcome as the other parameters, since the most expensive dwellings still have a higher preference. Discussion on the effect α has on agent utilities and decisions is given in "Appendix L".

We can see a general preference towards various α ranges for each period, based on higher density of samples in these areas. The preference is never for "perfect" information, i.e. $\alpha = 1$, reinforcing that such agents often act in a boundedly rational manner (with $\alpha < 1$). In the 2006–2010 period, there is a general preference towards sampling lower levels of α (within $\approx [0.05, \ldots 0.3]$, with the optimal value being 0.28), resulting in buyers acting with less information. In 2011–2015, during the economic recovery in Australia, buyers may have been more cautious and considering a wider range of available dwellings when purchasing, reflected in higher values of α ($\approx [0.5, \ldots 0.7]$, with the optimised value being 0.59). In 2016–2019, the cautiousness of buyers appears to revert again (within $\approx [0.05, \ldots 0.3]$, with optimised value of 0.24), showing buyers making less "informed" choices perhaps due to the rapidly increasing dwelling prices and buyer's desires to partake, at the expense of making "optimal" choices with larger α values.

6 Results

In this section, we present results of ABM simulations in terms of (i) price forecasting over three historic periods, aligned with the Australian Census years (2006, 2011 and 2016), and (ii) resultant agent preferences in terms of household mobility patterns. We also identify market trends across the three periods. The considered periods 2006–2010 and 2011–2015 include 48 months, while the contemporary period, 2016–2019, covers 42 months (our SIRCA–CoreLogic dataset includes the market data until 31 December 2019). Each time period compromises a separate set of simulations (with varying optimised parameters). Discussion on the chosen time periods is given in "Appendix I.3". For each period, we run 100 Monte Carlo simulations, using the model parameters optimised for the corresponding training set, as described in Sect. 5, and then obtain predictions for the remaining (testing) part of the data. For the first two periods, the first $\frac{3}{4}$ of the time series is the training part (e.g. 36 months from 1 July

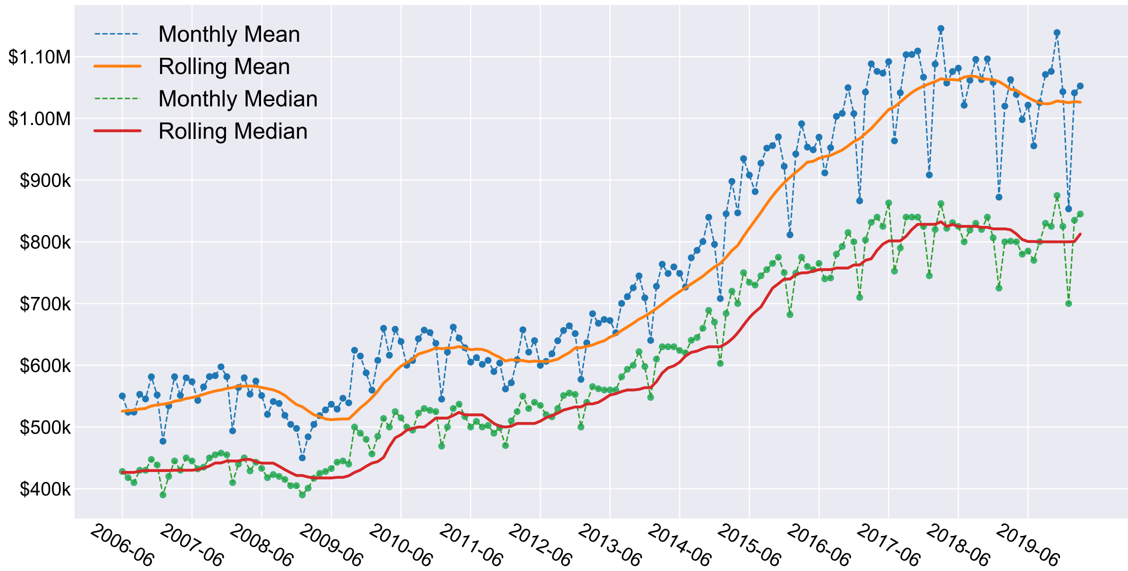


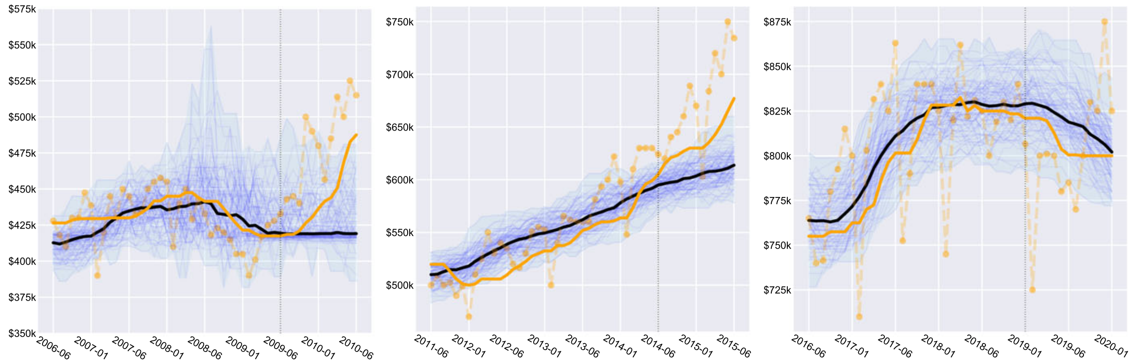
Fig. 5 Actual trends of Greater Sydney house prices from June 2006 to December 2019. Source: SIRCA–CoreLogic. The orange line represents the rolling mean, and the red line the rolling median. The raw monthly data points are also visualised (dotted lines). We use the median price trend as a more robust measure in all cases in this paper

2006 to 30 June 2009), and the remaining $\frac{1}{4}$ is the testing portion not used by any optimisation process (e.g. 12 months from 1 July 2009 to 30 June 2010). For the last period, the training part includes 30 months (from 1 July 2016 to 31 December 2018), with the remaining 12 months of 2019 used for testing.

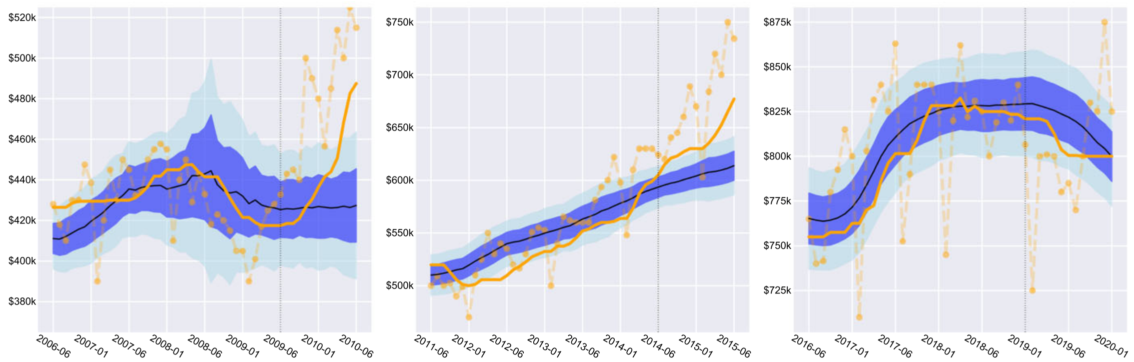
6.1 Price forecasting

We visualise the forecasting results in Fig. 6. It is evident that the model can successfully capture the key trends across the entire time series as emerging from individual buy and sell decisions, correctly identifying the peak and dip in 2006–2010, the steady growth during 2011–2015, and the growth and slow decline in 2016–2019.

The time period beginning in 2006 was a period of substantial uncertainty, triggered by the Global Financial Crisis (GFC). Tracing predictions for the last quarter of the period (i.e. the testing part of the dataset), shown in Fig. 6a, d, we observe that the average market rebound pattern is not fully followed. However, when considering the range of the simulations runs, i.e. the possibilistic regions of the simulation (as defined by Edmonds and ní Aodha (2018), and visualised by the blue boundaries in Fig. 6), we see that these contain the first segment of the fast rebound, and capture the market recovery to a good degree. In other words, the model is able to show the possibility of such a rebound. The discrepancy indicates that there was a catalyst underpinning a significant price appreciation during 2010 that was not solely driven by social influence and endogenous regulatory factors such as interest rates. It is well accepted that the market was reignited in 2009–2010 by exogenous factors, most notably by post-GFC government stimulus initiatives, such as the First Home Owners Boost (Randolph et al. 2013).



(a) Median: 2006–2010 **(b)** Median: 2011–2015 **(c)** Median: 2016–2019



(d) Mean: 2006–2010 **(e)** Mean: 2011–2015 **(f)** Mean: 2016–2019

Fig. 6 ABM simulation predictions. The top row shows the median of the simulations (black lines), every individual run (blue lines), and the minimum and maximum of every run with the light blue fill. The bottom row shows the mean of the simulations (black line), and ± 1 and 2 standard deviations (dark blue and light blue respectively). The vertical dashed lines separate training and testing parts of the time series. The orange lines show the actual trend obtained from SIRCA–CoreLogic data (colour figure online)

For the period beginning in 2011, the predicted time series correctly follows the actual trend, while slightly underestimating the slope of the growth towards the end of the period, as shown in Fig. 6b, e. It can be argued that the social influence factors, estimated during the optimisation phase, continue to affect the market dynamics during the last quarter of the period. In other words, the influence of these factors on decision-making, coupled with endogenous factors (e.g. interest rates), results in a steady growth of the market, predicted until the end of the period.

In the last considered period, starting in 2016, the model captures both the testing and training portions very well, correctly predicting the dip from 2019 onward, as shown in Fig. 6c, f. The notable decline occurs only in the testing period, and yet, the model is able to accurately predict both the peak and the correction. This indicates that the underlying reasons for the market reversal have developed during the first part of the period, and have been adequately captured by the parameter optimisation.

Overall, in all three time periods, the possibilistic output of the model contains most of the actual market dynamics. Specifically, during 2006–2010, 96% of the actual monthly prices are within the possibilistic range boundaries of the simulation (i.e. between the minimum and maximum output for each time period, the top row in Fig. 6), for 2011–2015: 84%, and 2016–2019: 100%. When using the mean ± 1 (2) standard deviations (bottom row in Fig. 6) as possibilistic boundaries, we observed the following probabilities of falling within the boundaries, during 2006–2010: 78% (92%), 2011–2015: 27% (63%), and 2016–2019: 86% (100%).

6.2 Area-specific price forecasting

The model was not directly optimised for spatial submarkets. However, in this section, we evaluate the predictive capacity of the model in terms of area-specific forecasting. During initialisation, some area-specific information for LGAs is drawn from available distributions, for example, the recent sale price of dwellings in that LGA. Likewise, households are also distributed into LGAs based on the actual population sizes of the LGAs. However, no additional optimisation is applied across different LGAs with respect to actual area-specific trends. In other words, the spatial component is used only for initialising relevant distributions, leaving the market dynamics to develop through agent-to-agent interactions.

In Fig. 7, we visualise the predicted area-specific pricing at the end of the testing period. We can see that these predictions closely follow actual data in general, despite not being directly optimised for, with all predictions characterised by high R^2 values.

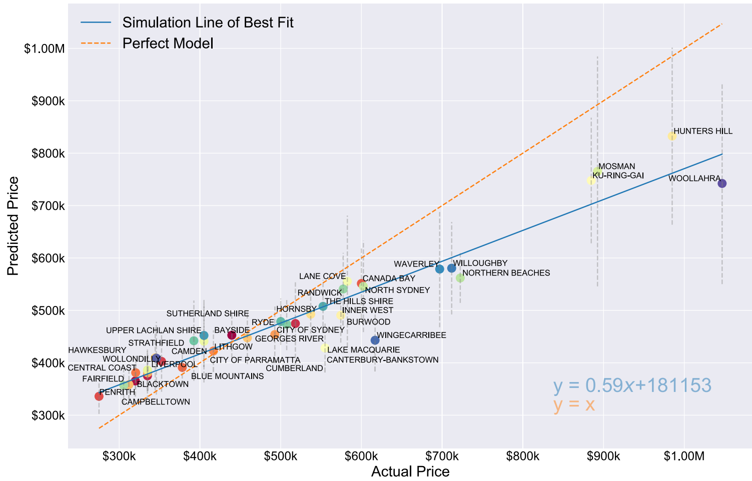
For the period 2006–2010 (with the end of the testing period mapping to June 2010), the simulations slightly overestimate the final price of the cheaper LGAs, but underestimate the resulting price of the most expensive LGAs, as shown in Fig. 7a. However, the perfect model (orange line) tends to be within the error margin (standard deviation) of the predictions of the simulation.

For 2011–2015 (with the end of the testing period being June 2015), the slopes of both actual and predicted regressions are almost identical ($m = 0.98$), as shown in Fig. 7b. However, the additive constant of the regression (i.e. y-axis intercept) for the predicted line is greater than 0 (of the perfect model): as a result, we are predicting slightly higher values across the LGAs on average. Considering the LGAs that were most overpriced with respect to the spatial trend (such as Kuringai, Waverly, Northern Beaches, and North Sydney), we can compare their resultant predicted prices in June 2015, Fig. 7b, with the actual prices depicted in Fig. 7c. The 2016–2019 plots show that this growth did eventually happen, and so the simulations produced for 2011–2015 merely predicted this appreciation for an earlier time than the actual scenario.

The predictions for the period 2016–2019 (the end of the testing period: December 2019) produce a regression strongly aligned with the actual fit, particularly for the higher-priced LGAs. Again, there is some overestimation in the cheaper LGAs, but this seems to highlight the increasing popularity of these suburbs. Analogously to the previous period, this may be indicative of some future price growth for these areas, not yet reflected in the current actual pricing.

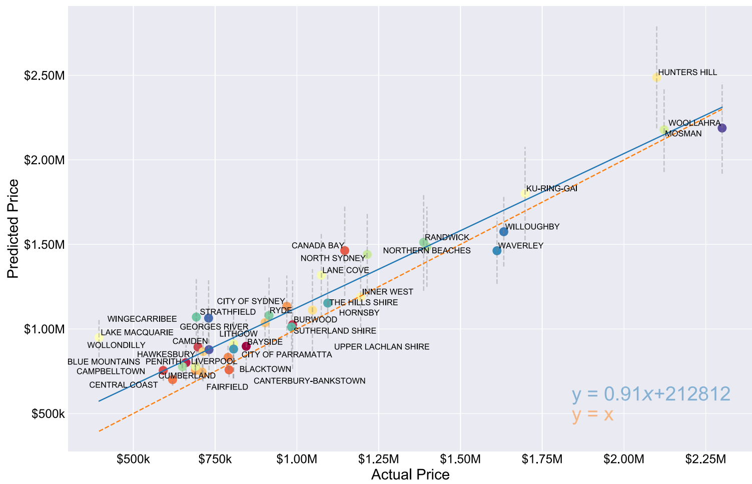
Overall, we observe that the resulting area-specific price predictions at the end of each testing period fit closely to the actual resulting prices. This indicates that the model successfully captured the spatial submarkets, despite having been optimised for the overall market dynamics of the Greater Sydney region as a whole.

Fig. 7 Predicted LGA pricing at the *end* of the testing period. The actual prices (SIRCA–CoreLogic) are shown on the x-axis, with the predicted prices on the y-axis. Error bars represent the standard deviation of the prediction across simulation runs. The orange line shows the perfect model ($y = x$), and the blue line shows the least-squares line of best fit for the predictions (equation given on plot) (colour figure online)



(a) June 2010 predictions (end of 2006–2010 test period). $R^2 = 0.92$.

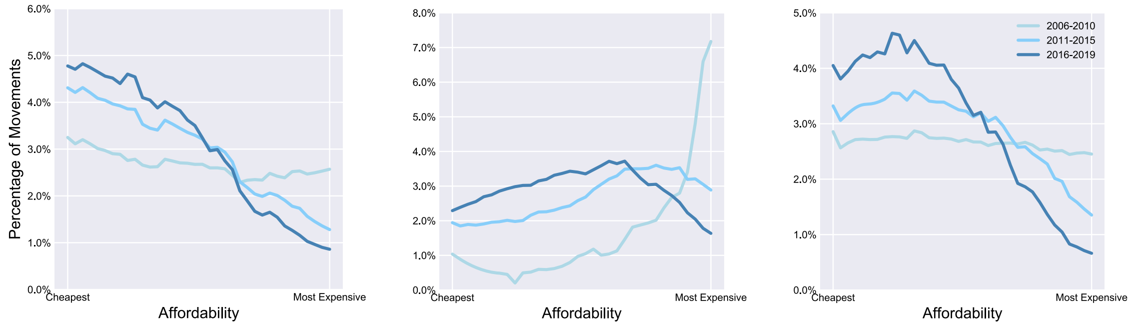
(b) June 2015 predictions (end of 2011–2015 test period). $R^2 = 0.82$.



(c) December 2019 predictions (end of 2016–2019 test period). $R^2 = 0.91$.

6.3 Agent preferences and mobility patterns

Analysing the households' movements produced by the simulation is another key insight the spatial agent-based model can provide. In this section, we consider various agent movement patterns (which we refer to as household mobility), aiming to identify the salient trends in agent preferences. There are several key areas we focus on:



(a) First-time home buyers **(b)** Overseas Investors **(c)** Local Investors

Fig. 8 Comparison of price influenced mobility between the three periods. The x-axis represents the affordability (most affordable locations on the left). The y-axis represents the (smoothed) percentage of movements to the area. Darker colours represent later years. A full breakdown of household mobility is provided in "Appendix J"

first-time home buyers, investors, and new households (i.e. migrations or households splitting). Again, no direct optimisation was applied to the movements, and so the identified preferences are intrinsic results of the model, and not attributed to some actual data. However, we show that such mobility preferences are supported by evidence, thus arguing that the model is able to produce sensible local patterns for which it was not explicitly optimised for, based only on the global calibration data and proposed spatial structure.

The mechanisms shaping the process of settlement formation and generating intra-urban mobility specifically include transitions driven by critical social dynamics, transformations of labour markets, changes in transport networks, as well as other infrastructural developments (Kim et al. 2005; Simini et al. 2012; Barthelemy et al. 2013; Louf and Barthelemy 2013; Arcaute et al. 2016; Barthelemy 2016; Crosato et al. 2018; Barbosa et al. 2018; Piovani et al. 2018; Slavko et al. 2019; Barthelemy 2019). Types of homeownership, in particular, are known to affect mobility patterns (Crosato et al. 2021). In this work, we focus solely on the movements resulting from the housing market dynamics, which in turn incorporate the imperfect spatial information and other subjective factors such as the FOMO and trend-following aptitude. We do not model any structural changes across the regions, i.e. the LGAs boundaries, transport and other infrastructure topologies, remain fixed.

6.3.1 New households/migrations

New households are those which are added throughout the simulation based on the projected household growth. New households can result from a variety of sources, such as people moving to Greater Sydney (migration), or households from Greater Sydney splitting, e.g. in the case of divorce or young adults moving out of home. We make no distinction between the two types in the simulation, and for simplicity, refer to both intercity and intra-city migration types as migrants.

The most common LGAs into which the new households move are shown in Fig. 22.[3] The simulation produces a clear trend for migration towards the cheaper

[3] All movements are scaled by the population size to allow a fair comparison, as outlined in "Appendix J".

areas, as the price begins to increase throughout the Greater Sydney region over time. This is most apparent for the 2016–2019 period, during which we can detect only a minority of the new households that purchased a dwelling in the expensive areas upon moving to the Greater Sydney. This is markedly different when in comparison with the 2006–2010 period. Furthermore, there is a clear peak in the more affordable LGAs, comprising Western Sydney (such as Campbelltown, Penrith, Blacktown, Fairfield, and Liverpool), and LGAs further away from the metropolitan area (such as Central Coast, Lake Macquarie, and Hawkesbury). A similar trend can be seen with the new renters. This agrees with the discussion in Bangura and Lee (2019), which names Western Sydney as "the first port of call for new arrivals, immigrants and refugees". This observations also agree with the study of Slavko et al. (2020a), which shows the outward sprawl from the densely populated Sydney metropolitan area.

6.3.2 First-time home buyers

Homeownership has long been a goal of many Australians (Bessant and Johnson 2013), so simulating the forecasted feasibility in Sydney—the largest and most expensive (Yetsenga and Emmett 2020), Australian city—is essential. The first-time home buyers are defined here as those who have resided in Sydney but have previously been renters, and then purchased their own dwelling. This is in contrast to the analysis in Sect. 6.3.1, where the new owners are defined as those that had just entered the simulation (by moving to Sydney).

The first home purchases are visualised in Fig. 24. The main diagonal represents an agent purchasing in the same LGA as the one where the household is currently renting. The area below the main diagonal (which we refer to as lower triangle) shows households purchasing in cheaper LGAs in comparison with those where they are renting, and the area above the main diagonal (the upper triangle) shows the agents purchasing in LGAs more expensive than those where they are currently renting. In the earlier years (i.e. the 2006–2010 period), we can see that the densities in the heatmaps are relatively evenly distributed. Over time, however, the density of the upper triangle begins to decrease, meaning that the agents are purchasing in the LGAs cheaper than those where they are renting, as the expensive LGAs become increasingly out of reach. This is also reflected in Fig. 8a: a comparison between the 2006–2010 and the 2016–2019 periods clearly shows that many suburbs are simply becoming out of reach for the first-time home buyers, with a larger percentage of them needing to purchase in the more affordable areas.

This result is in line with (Randolph et al. 2013) which shows the distribution of First Home Owner Grants within Sydney statistical districts, over the period 2000–2010, pointing out that such grants were increasingly likely in the lower-income housing markets, such as Western and Southern Sydney. Likewise, in La Cava et al. (2017), it is shown that the average distance to the CBD of dwellings that first-time home buyers can afford has been increasing from 2006 through 2016. Furthermore, La Cava et al. (2017) show that the purchasing capacity of first-time home buyers has been limited to the bottom (i.e. most affordable) 10–30% of dwellings, where in 2016 the median first-time home buyer could afford only around 10% of the available dwellings. A similar conclusion is reached in Kupke and Rossini (2011) which find key workers being able to afford fewer dwellings and commute longer distances between 2001 and 2009, a trend which seems to have been followed ever since, as shown by the mobility patterns here.

6.3.3 Investors

Investors are defined as households which own multiple dwellings, or households which live overseas yet own a property in Greater Sydney. We make an explicit distinction between local (residing in Sydney) and overseas investors, since price increases within Sydney are often attributed to the latter category (Rogers et al. 2015, 2017; Wong 2017; Guest and Rohde 2017). This distinction is visualised in Fig. 23. The overseas investment approvals are regulated by the Australian government, and details of the approval data are provided in "Appendix K".

During the 2006–2010 period, based on the overseas approvals granted, the simulation produces a clear preference for the overseas buyers towards the most expensive regions, purchasing properties almost exclusively in the highest priced regions. In later periods, we begin to observe a relatively wider range of preference, although still with a clear trend towards the mid-high-range areas. This is reflected in Fig. 8b. A driving factor behind this is the higher average government approval for overseas investment given during 2006–2007 years, in comparison with later years, reflected in our simulation (as displayed in Fig. 25). While the foreign investment data in the Greater Sydney housing market is sparse and not fine-grained, this purchasing pattern is in concordance with the recent literature. For example, the study of (Gauder et al. 2014) mentions that foreign investors tend to prefer inner-city dwellings within Sydney (which tend to correspond to higher-priced LGAs); however, recently "foreign investment has started to broaden out into other areas of Sydney".

These findings are in sharp contrast to mobility patterns of the local investors, for which the simulation produces a far wider distribution across areas. For 2011–2015, and 2016–2019 periods, the most expensive LGAs become out of reach for local investors, which is most apparent during 2016–2019 (as shown in Fig. 8c). Local investors can be seen buying properties in many of the cheaper LGAs, which also falls in line with the simulation showing many renters in these areas. Due to the affordability, these LGAs also exhibit higher population growth rates than other areas. This also agrees with existing studies, for example (Pawson and Martin 2020), which find that many high-income Australian landlords are investing in dwellings in lower socioeconomically developed regions of Sydney.

7 Conclusions and future work

In this work, we have introduced a spatial element to a model of a large, well-developed housing market (the Greater Sydney region) using an adjacency matrix based on the spatial composition of the city, and introduced several factors which affect agent decision-making within the market. We have shown the model is capable of predicting housing price dynamics as arising from individual buy and sell decisions in the market. The proposed model is capable of capturing a large variety of spatial topologies, for example, monocentric and polycentric cities. Furthermore, the graph-based approach is flexible allowing for any level of granularity, for example, over the differing scales of countries, cities, or suburbs, and extendable to weighted versions which could incorporate transport times between areas.

Using this model, we have demonstrated the usefulness of spatial analysis when it is calibrated to the Australian house price data for the Greater Sydney region. The 38 LGAs of Greater Sydney were simulated, and agents (households) were calibrated based on the LGA in which they live. We have shown that the spatial component allows an additional level of fine-tuning that results in better overall fitting of the model to data, as well as producing strong out-of-sample predictions for each individual LGA by optimising only for the overall trend. That is, spatial areas add an additional layer of predictions, while also improving the overall aggregate trend.

We investigated the agent's spatial awareness of the market, where buyers only have limited knowledge of the market, based on the area in which they reside (imperfect spatial information). We demonstrated through varying α (a parameter that controls how much of the market each agent is able to perceive) that lower values of α capture the true trend better than perfect (whole of market) knowledge, indicating the usefulness of modelling this imperfect spatial information in a housing market and showing that agents often act in a boundedly rational manner.

The spatial component also allows for the analysis of agent preferences in terms of movement patterns, where we have shown differences in mobility and purchase locations between various agent types, for example differences between first-time home buyers and investors, where first-time home buyers are limited to the more affordable locations, with investors being able to purchase higher-priced properties, and likewise between local and overseas investors, where overseas investors are shown to have a strong preference towards mid to high valued areas. We also model new migrations to the city, showing such agents becoming increasingly pushed towards cheaper areas of the city.

We have also introduced a novel fear of missing out component which alters sellers decision-making behaviour. With this parameter, we model how sellers become more likely to sell a listing if many surrounding listings have recently sold, and show a strongly localised fear of missing out occurring throughout the market. This indicates agents' decisions are often motivated by their neighbour's decisions rather than by strict optimisation of their own benefits, i.e. real households are only partially rational in this regard.

While in this work we addressed some key concerns in a housing market, there are still several areas of improvement we would like to focus on in future work. The spatial component opens up a range of additional possibilities, such as overlaying public transport maps on the network, allowing for the distance to key work areas, schools, beaches, etc., and further modelling and capturing agent mobility within the simulation. The demographics and household types could be sampled from actual data, which would allow analysis into subgroups of people (i.e. young singles vs. retirees vs. families), and allow us to model any spatial trends that arise between demographic groups. The internal optimisation functions (for example, what neighbourhood to move to, what kind of dwelling to choose) of agents could also be investigated further, as currently, agents will purchase the most expensive dwelling they can afford based on their knowledge and outreach. There are also three key equations which drive the model that could be further investigated: the bid prices, the listing prices, and the bank approval process. These are currently predefined equations but they could be treated as optimisation problems themselves, finding expressions that match most closely to the training period.

A: Implementation

The model is written from scratch in Python3, based on the C++ code from Glavatskiy et al. (2020).

B: Baseline model parameters

The work extends the model of Glavatskiy et al. (2020) (the baseline). For completeness, all parameters of the baseline method with explanations are given in Table 2.

Local sensitivity analysis of these parameters is presented in Fig. 9, where we vary the internal values around their default ranges (within a range of $\pm 20\%$) one at a time while keeping the other parameters at their default values. The resulting analysis shows the model is robust to small changes in the parameters. The only parameters that stand out from the analysis deserving additional discussion are the listing price factor, the expectation downshift ratio, and the amount quality reference.

The listing price factor b_ℓ has relatively large variation in prices, but as shown in Eq. (3.2) this is because it acts as a linear scaler on the list prices, so scales the output within the $\pm 20\%$ range too, which means the parameter is behaving as expected. The expectation downshift ratio p_{d} highlights the willingness of agents to downgrade the pre-purchase expectation based on the amount offered by the bank. For higher values, this indicates buyers wanting more from banks, and we see this increases prices overall. It is interesting to note, however, that higher values also increase the volatility of the market, with larger fluctuations seen for example in late 2017 (at the peak of the market). This confirms the discussion in the study of Glavatskiy et al. (2020) which highlighted the propensity to borrow as a key explanatory factor of volatility. For the amount quality reference $\bar{Q}_{\mathrm{h}}$, this is an integer parameter which explains why there are fewer comparison lines than the continuous parameters, but we still observe well-behaved outputs in the $\pm 20\%$ range.

For all other parameters, the outputs only result in small variations from the default value indicating the robustness of the model to variations around the internal parameters.

C: market matching

The market matching process where buyers and sellers are matched is relatively simple and given in Algorithm 1. We can see the highest bidding buyer gets preference to the listings, and every buyer attempts to purchase the most expensive listing they can

Table 2 Baseline model parameters

Parameter	Symbol	Value
External paramaters		
Mortgage rate	ϕ_{M}	2006–2010: 7.3–9.45%
		2011–2015: 5.53–7.79%
		2016–2019: 4.95–5.35%
Tax rate		2006–2010: 0%, 15%, 30%, 40%, 45%
		2011–2015: 0%, 15%, 30%, 37%, 45%
		2016–2019: 0%, 19%, 32.5%, 37%, 45%
Tax brackets		2006–2010: $0, $6k, $25k, $75k, $150k
		2011–2015: $0, $6k, $37k, $80k, $180k
		2016–2019: $0, $18.2k, $37k, $87k, $180k
Income expenditure	$\phi_{\mathrm{I}}, \phi_{\mathrm{b}}$	2006–2010: [49.37, 0.81]
		2011–2015: [76.76, 0.75]
		2016–2019: [81.75, 0.80]
Mortgage duration		30 years
Annual house tax and fees		1.7%
Stamp duty		2.5%
House care	ϕ_{H}	2.5% ± 0.5%
Purchase fees		2.5%
Loan to value ratio		80% ± 10%
Income growth rate		1.002 ± 0.001
Income consumption		60%
Liquid consumption		2.5%
Internal paramaters		
Listing probability	p_b	1%
Amount of houses for reference	$\overline{Q}_{\mathrm{h}}$	10
List price factor	b_ℓ	1.75
Sold to list power	b_{s}	0.22
Months on market power	b_{d}	− 0.01
Urgency stress		0.2
Urgency rental		0.02
Urgency cash		0.2
Probability of accepting highest bidder	p_{m}	80%
Expectation downshift ratio	p_{d}	60%

The external parameters are exogenous, sourced from, the Australian Bureau of Statistics, The Household, Income and Labour Dynamics in Australia Survey (Wilkins and Lass 2015). The internal parameters are endogenous, and the default values tend to come from the analysis of (Axtell et al. 2014). Sensitivity analysis around the internal parameters is given in Fig. 9, and a full outline is given in Glavatskiy et al. (2020)

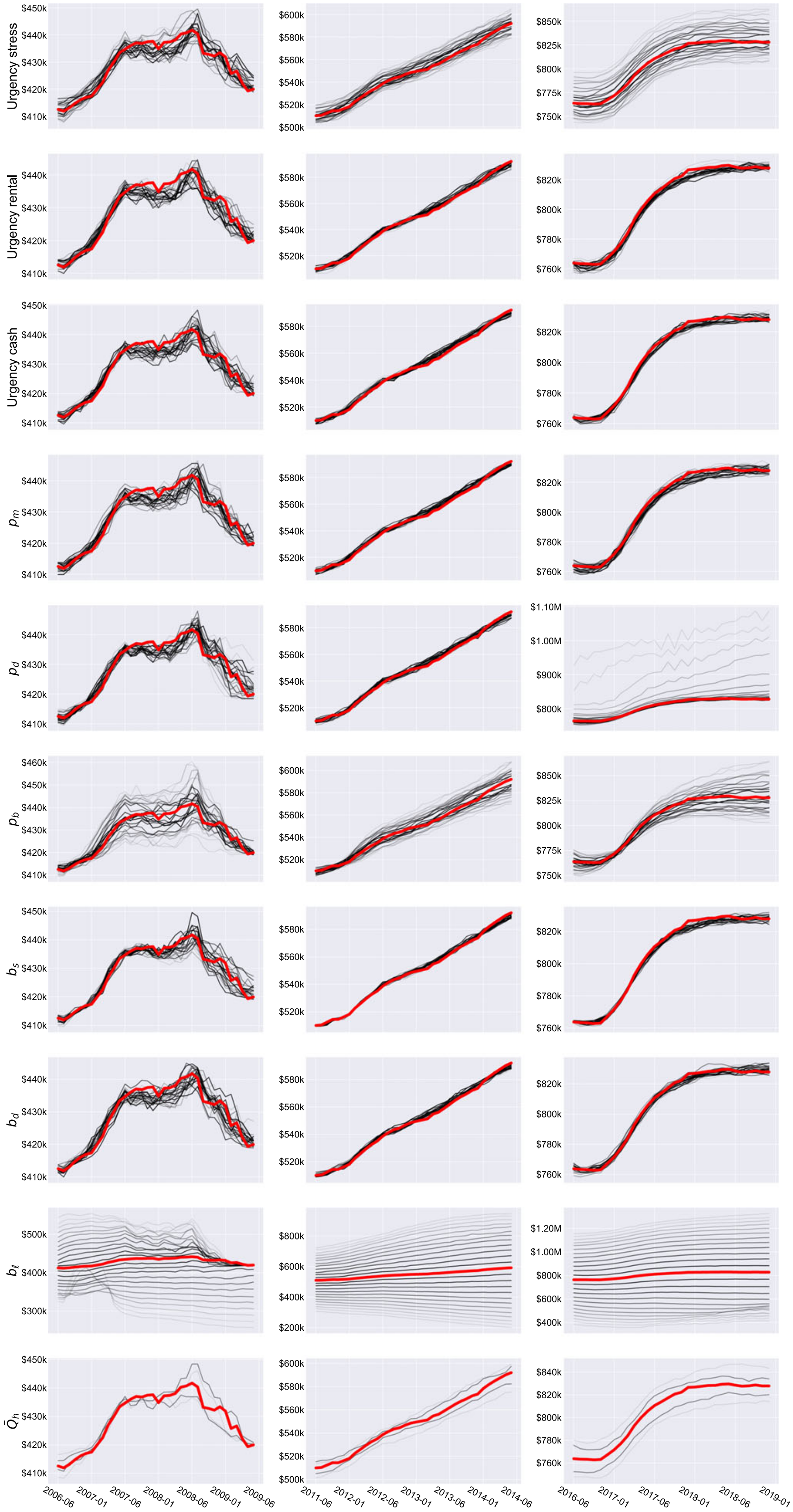


Fig. 9 Local sensitivity analysis showing variation in output based on varying the baseline model input parameters one at a time ±20% around their default value (with others fixed at their default). The red lines represent the output for the default value, and the grey lines represent for the varying values (with shade indicating the distance to the default, black meaning very near the default, and light grey being further away). All show relatively small variations in the output based on the input parameters, indicating the robustness of the parameters (colour figure online)

afford. In certain cases, deals are rejected due to external influences (modelled by a random 20% chance of rejection). This matching process is performed once every simulation step, with bids and listings that did not clear persisting into the next step.

Algorithm 1 Market Matching

```
procedure MATCH(bids, listings)
    sort bids                                  ▷ From highest bid price to lowest
    sort listings                              ▷ From highest list price to lowest
    for bidder in bids do
        best listing ← max listing buyer can afford
        if no deal breaker then
            Make deal between buyer and seller
            listings − best listing            ▷ Remove this listing from the available listings
        end if
    end for
end procedure
```

D: Rental market

In the baseline model, there was no concept of rental matching. Households were randomly assigned a rental, with no regard to the cost of the dwelling or income of the household. Here, we add in an additional matching process based on the idea that households should spend maximum 30% of their income on housing when possible to avoid housing stress (Thomas and Hall 2016; Fernald 2020).

New households are randomly assigned a "local" area (weighted by the population of each area), where they begin and have their characteristics (wealth, income, cash flow, etc.) assigned. From there, every household attempts to find a vacant rental in their price range (which will likely result in various households moving out of financial requirements). Households with extremely high incomes, where all dwellings are less than 10% of their income, get the most expensive rental available. Households with extremely low income, where all dwellings are at least 30% of their income, get the cheapest one they can afford. All other households randomly choose a rental they can afford (in the 10–30% of income range).

Households remain in their rentals for the duration of the simulation. In this work, we do not attempt to capture the rental market in its entirety and leave this for future work where we would like to model the relationship between renters and investors. The changes were made to ensure the cash flow situations of each household match closer to those seen in the real world, where in the previous model many households would be in a poor cash flow situation due to the rental price. Other work such as Mc Breen (2010) looks more in depth at modelling the rental market.

33

E: Bayesian optimisation

E.1 Details

Bayesian optimisation is performed using the Tree of Parzen Estimators approach with hyperopt from Bergstra et al. (2013). The optimisation process was run for 2000 iterations in all cases. The loss was measured as the average loss over several stochastic runs for each set of parameters, to minimise the effect of randomisation in the model and resulting loss.

E.2 Loss function

For measuring the goodness of fit, we use a loss function with two terms—a shape and temporal term to try and capture the nonlinearities overtime when predicting housing price trends. This loss function is a modification of DILATE (Vincent and Thome 2019) which was introduced as a loss function for neural networks for time series predictions, although DILATE has been simplified here (with the removal of smoothing parameters) as Bayesian optimisation does not require the loss function to be differentiable.

The loss function is given in Eq. (E.1).

$$\ell = \lambda * \text{shape} + (1 - \lambda) * \text{temporal} \tag{E.1}$$

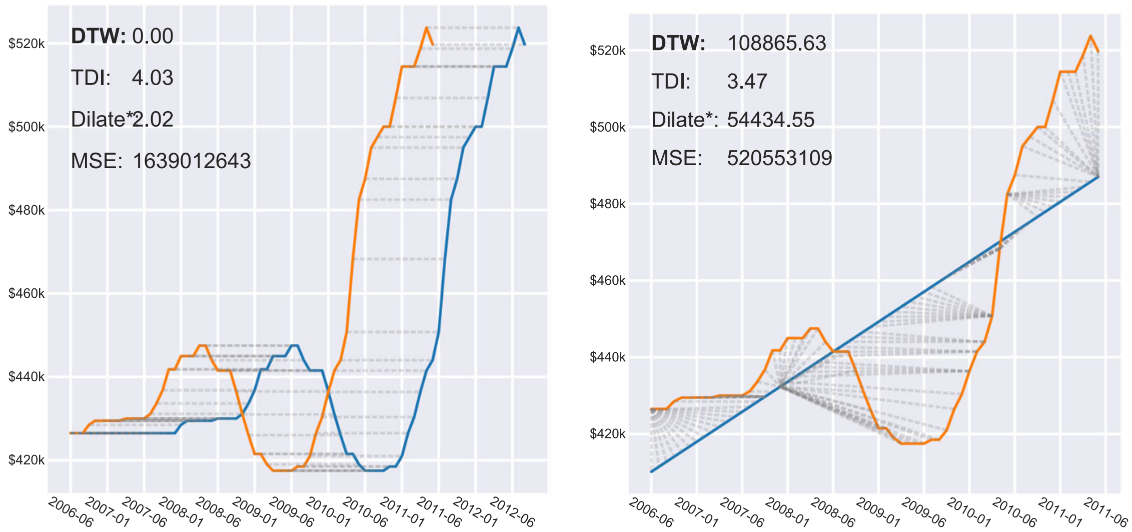
$\lambda = 0.5$ was used throughout since this was the most common in the original paper of Vincent and Thome (2019). However, as the terms are not normalised, the two do not have an equal contribution; instead, the temporal term serves more like a penalty on the shape (with λ controlling the strength of the penalty).

The shape term is based on dynamic time warping (DTW) which has commonly been used in speech recognition tasks (Sakoe and Chiba 1978; Myers et al. 1980) and, however, has a wide range of applications in time series data (Berndt and Clifford 1994). Dynamic time warping can be expressed recursively as a minimisation problem as in Eq. (E.2)

$$\text{shape} = \text{DTW} = d(x, y) + \min \begin{bmatrix} \text{DTW}(x-1, y), \\ \text{DTW}(x-1, y-1), \\ \text{DTW}(x, y-1) \end{bmatrix} \tag{E.2}$$

This can be read as minimising the cumulative distance (using distance measure d, in this case, euclidean distance) on some warped path between x and y, by taking the distance between the current elements and the minimum of the cumulative distances of neighbouring points.

Unlike the common applications in speech recognition, where words can be spoken at varying speeds (so the peaks do not necessarily match up), in financial markets, timing such peaks is important. This motivates the introduction of a temporal term, for trying to align such peaks and dips. The temporal term is based on Time Distortion



(a) Shifted Predictions **(b)** Linear Predictions

Fig. 10 Motivation of time series-based loss using a constructed example. We can see the line on the right is a very poor predictor of the true trend, failing to capture any of the peaks or dips. However, the MSE is significantly lower than the line on the left. DTW captures the shifts, and incorporating a penalty on time can penalise these shifts. The light grey lines show how DTW matches points together, even if they do not occur at the same time period

Index (TDI) (Frías-Paredes et al. 2016, 2017), which can be thought of as the normalised area between the optimal path and the identity path (where the identify path is $(1, 1), (2, 2).., (N, N)$) (Vallance et al. 2017) and aims to minimise the impact of shifting and distortion in time series forecasting (Frías-Paredes et al. 2016).

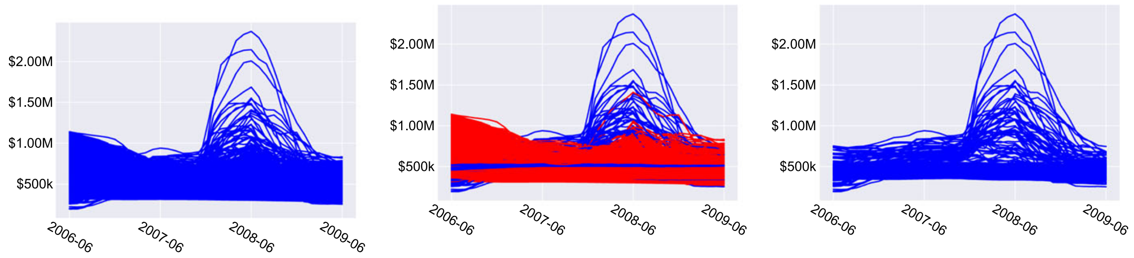
$$P_l = \int_{i_l}^{i_{l+1}} \left(x - \frac{(x - i_l)(j_{l+1} - j_l)}{(i_{l+1} - i_l)} + j_l \right) \mathrm{d}x \tag{E.3}$$

$$\text{temporal} = \text{TDI} = \frac{2 \sum |P_l}{N^2} \tag{E.4}$$

To see the usefulness over a more standard approach loss function such as MSE for time series, consider the example in Fig. 10. We can see the MSE can be a problematic approach, and in some cases (as in the example where the linear line Fig. 10b has a lower loss) be a misleading measure of goodness of fit. DTW helps to match points in the two time series, while TDI helps minimise the offset of the predictions. (Graphically, in the example this corresponds to shortening the dotted grey lines.) For a full analysis, we refer you to the original DILATE paper of Vincent and Thome (2019), noting that all smoothing terms have been removed in the modification here.

E.3 Global constraints

We can see 2011–2015 and 2016–2019 fit the trend very closely, although despite having a low loss, the 2006–2010 simulation path does not follow the dip well, as no distinction is made about being above or below the trend in the loss function. Looking



(a) Without Constraints **(b)** Red shows the paths which will be removed. **(c)** Paths matching constraint

Fig. 11 Global constraint process

at the individual paths from every run, we can see that a peak and dip is predicted in many of the cases, although the distance is greater than the path with the lowest loss which was perfectly matching across a large portion of the training data but missing the dip. We apply a post-optimisation global constraint to 2006–2010, again only using this training period, that the midpoint of the simulation must be higher than the start and ending points (i.e. a peak must occur), and take the parameters with the lowest loss matching this criterion. The process is shown in Fig. 11 and the result is shown in Fig. 5p. We can see for 2006–2010, the ℓ is higher than before the constraint, however, clearly, the constraint allows for a closer overall trend following in the training period. The visualisation in Fig. 11 can also begin to show the wide range of possible market outcomes, for various combinations of the parameters. If a certain section occurs from many parameter outcomes (i.e. with the peak), we can deduce that such dynamics were likely to occur just due to the agent characteristics, regardless of the parameters used. This shows many combinations lead to a peak and dip, perhaps due to mortgage rates and worrying mortgage vs income ratios. This is more in line with suggestions in Edmonds and ní Aodha (2018), which suggest ABMs be used to determine a range of potential future outcomes, which in this case shows a variety of paths leading to a peak and dip.

E.4 Parameter space

The parameter space is defined in Table 3.

Even though there are only three parameters to tune, the number of potential combinations exceeds 4 million (this is assuming values are discretised values, so the true number is far greater), making a grid search impractical.

The three parameters are h, α, and β.

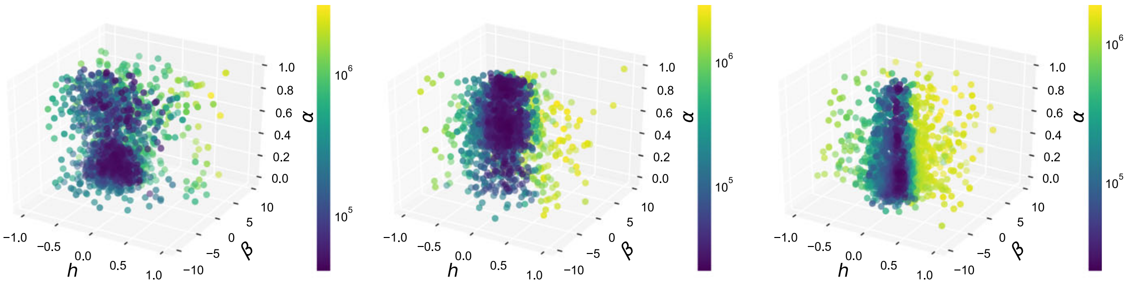
F: Sensitivity analysis

While in Sect. 5.1 we analysed the contribution of each new component by comparing the resulting optimised time series after introducing the components one at a time, here we verify and rank the importance of each of the contributions explicitly using global sensitivity analysis (GSA).

Table 3 Three tunable hyperparameters

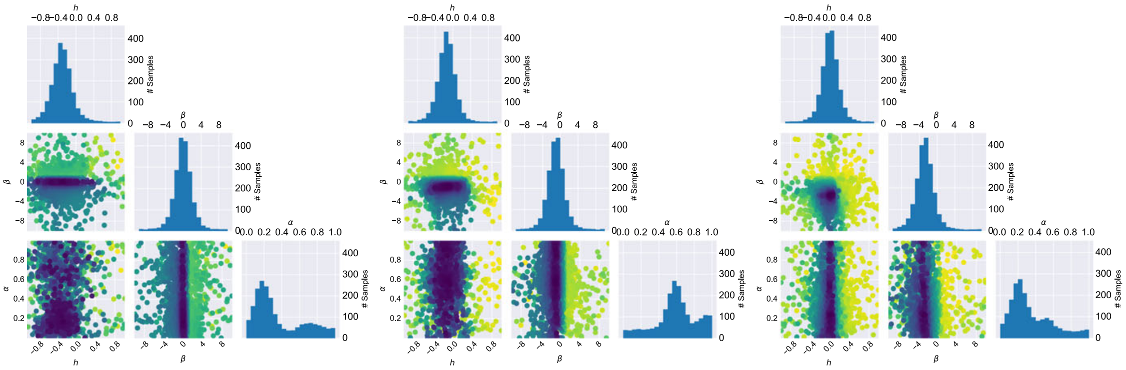
	Search space	Explanation
α	[0, 1]	Probability of viewing a listing, scaled by the outreach
β	[−10, +10]	The contribution of surrounding listing sales when considering selling a dwelling
h	[−1, +1]	Trend-following aptitude

All parameters are sampled uniformly within these ranges



(a) 2006–2010 **(b)** 2011–2015 **(c)** 2016–2019

Fig. 12 Search space exploration. Colour indicates the loss



(a) **(b)** **(c)**

Fig. 13 Parameter interactions and parameter sampling

Table 4 Morris method for sensitivity analysis

	2006–2010		2011–2015		2016–2019	
	μ^*	σ	μ^*	σ	μ^*	σ
h	692,019	708,678	1,390,890	1,693,784	871,388	882,436
β	434,626	843,830	714,507	961,618	623,839	650,213
α	317,806	573,236	467,376	729,330	325,056	475,900

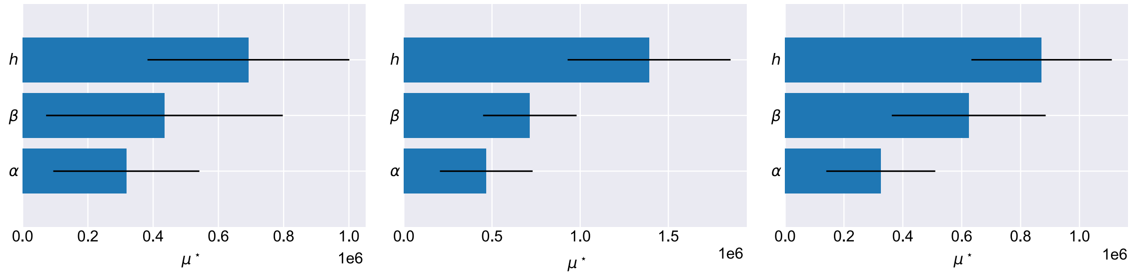
Specifically, we analyse the importance of the trend-following aptitude (h), the social contribution (β), and the role of α in minimising the loss function.

We use the Morris method (Morris 1991) for a GSA and present the revised μ^* as suggested in Saltelli et al. (2004) and σ. μ^* represents the mean absolute elementary effect and can be used to rank the contribution of each parameter, and this solves the problem of μ where elementary effects can cancel out. We also analyse σ, i.e. the standard deviation of the elementary effects, as a measure of the interactions.

For parameters for the Morris Method, we use $r = 20$ trajectories, $p = 10$ levels, and step size $\Delta = p/[2(p-1)]$, i.e. $\Delta \approx 0.52$ with $p = 10$. These are within the range of commonly used parameters, e.g. in Campolongo et al. (2007).

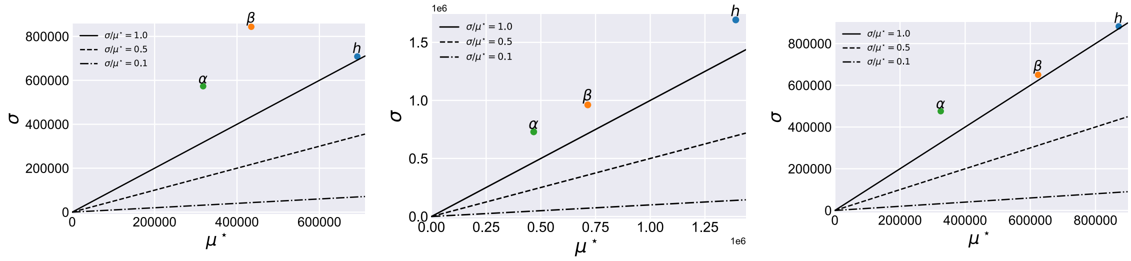
The results are presented in Table 4, and visualised in Figs. 15 and 14.

Checking the importance of each parameter, or $\mu^\star$, we can see h consistently ranks the most important, showing its changes have the largest effect on ℓ. This is followed



(a) 2006–2010 **(b)** 2011–2015 **(c)** 2016–2019

Fig. 14 Importance plot showing $\mu^\star$. Error bars are displayed at the 95% confidence level



(a) 2006–2010 **(b)** 2011–2015 **(c)** 2016–2019

Fig. 15 Global sensitivity analysis with Morris plots. Diagonal lines represent the ranges for $\sigma/\mu^\star$. One classification strategy proposed by Sanchez et al. (2014) says factors which are almost linear should be below the 0.1 line, factors which are monotonic between 0.1 and 0.5 lines, or almost monotonic between the 0.5 and 1 line, and factors with non-monotonic nonlinearities or interactions with other factors above the 1 line

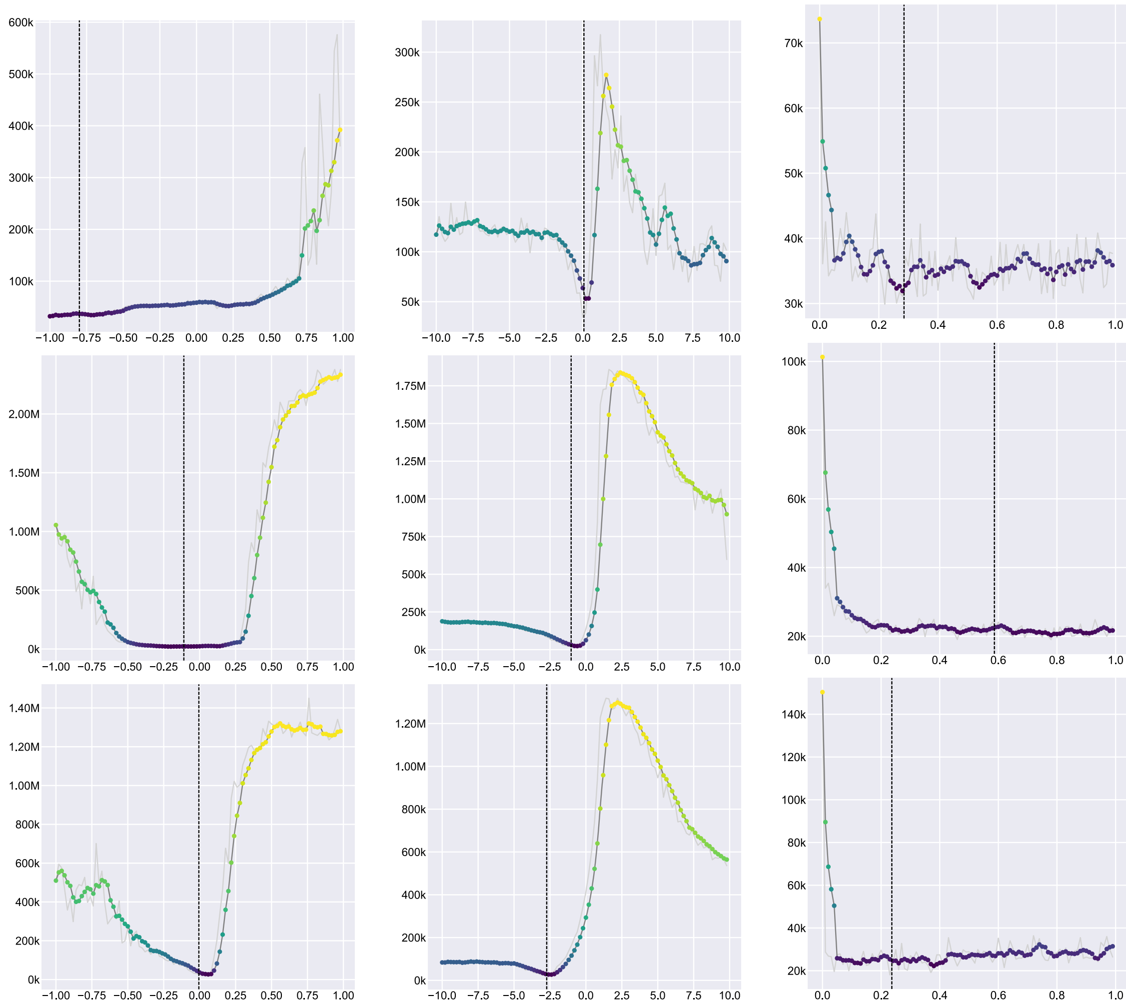
in importance by β, and then α each year. However, we see that confidence bars do overlap in Fig. 14.

Viewing the Morris plots in Fig. 15, we can see all parameters are deemed important, where unimportant parameters would show up in the bottom leftmost portion of the plot. Using the classification strategy of Sanchez et al. (2014), all parameters are all considered to be non-monotonic and/or with high levels of interaction, since $\frac{\sigma}{\mu^\star} > 1$ in all cases.

This analysis agrees with the preliminary parameter analysis in Sect. 5.2.

While the Morris method gives us the overall sensitivity across the parameter ranges (in a global way) and allows us to rank the factors in terms of importance, we also provide a fine-grained sensitivity analysis around the default values, i.e. a local sensitivity analysis (LSA). For this, we use $p = 100$ levels, but vary only one parameter at a time while keeping the others fixed at their default values. This is shown in Fig. 16. This analysis shows how robust the resulting default values are to small perturbations, but as this is a local method, the results should be interpreted with caution (and only in conjunction with the GSA method above), since this does not account for any parameter interactions as warned in Saltelli et al. (2019).

Viewing h (the left column), we can see all values surrounding the default have a similar loss, showing the model is robust to small changes in the aptitude. Looking across the entire search space, we can see choosing from within an appropriate range for the aptitude is important though, but the surrounding parameters are always relatively smooth to the resulting loss. Viewing β (the middle column), we can see the sharp

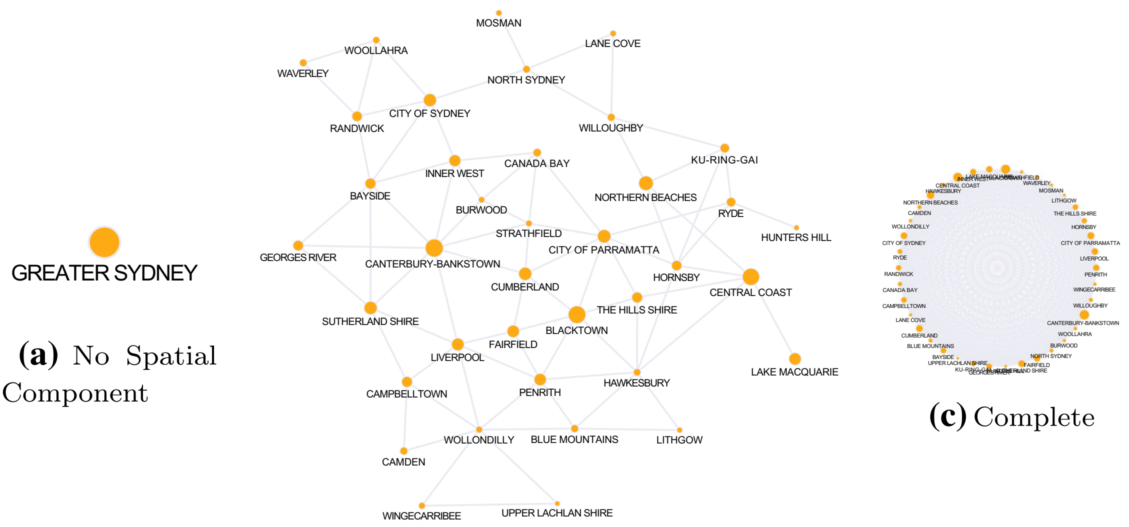


(a) h **(b)** β **(c)** α

Fig. 16 Univariate LSA of default parameters, varying one factor at a time with others at their **optimised** values. The plots give the change in parameter value (x-axis) versus ℓ (y-axis). The dotted vertical black line shows the optimised value

transition above zero. There is a clear optimal range for β, where the default lies. However, again, the area surrounding the default values is smooth showing robustness to the default parameters (assuming we do not vary past the sharp transition). Looking at α (the final column), the plots initially seem somewhat jagged, although when looking at the scale of the y-axis it becomes clear these are very small shifts in loss (as verified by the plotted time series with varying α levels). α was deemed the least important of the three parameters by the Morris method screening, but was still important based on the positioning on the Morris plot. We can verify this here, where changes in α do not have a huge impact on ℓ.

GSA was performed using SALib from Herman and Usher (2017).



(a) No Spatial Component **(b)** LGAs **(c)** Complete

Fig. 17 Various potential network architectures. The left represents a single node (i.e. no spatial component). The middle is the proposed graph-based approach constructed from the topological layout of the region. The right is a complete graph, with individual areas, but no concept of spatial neighbours (due to the fully connected nature)

G: Network topology

In Sect. 4, we introduced a novel graph-based structure for representing the region, at the same time introducing spatial submarkets into the simulation (based on nodes in the graph). In Sect. 5.1 and "Appendix F", we have validated the usefulness of the newly introduced parameters and performed a sensitivity analysis of the parameters, whereas in this section, we look to validate the usefulness of the network structure itself.

To do this, we compare the newly proposed model (with all parameters included), against an identical model with only a single node. We also compare to a fully connected network, i.e. where the spatial element (in terms of neighbourhoods) is not considered directly, but specific areas still exist. These structures are visualised in Fig. 17.

G.1 Topologies

G.1.1 No spatial component

To remove submarkets and all spatial components, we use a graph composed of a single node representing the overall Greater Sydney region. That is, agent characteristics and dwelling prices are assigned based on the overall Greater Sydney distributions, rather than specific area distributions. To implement this, rather than G being defined as in Sect. 4.1.1, instead, G contains a single node (i.e. a singleton graph) where the node represents the overall Greater Sydney region, i.e. it is the graph K_1. With this single-node configuration, the spatial outreach from Eq. (4.3) is removed, as there is no concept of space. Likewise, β is no longer defined, as this is expressed in spatial

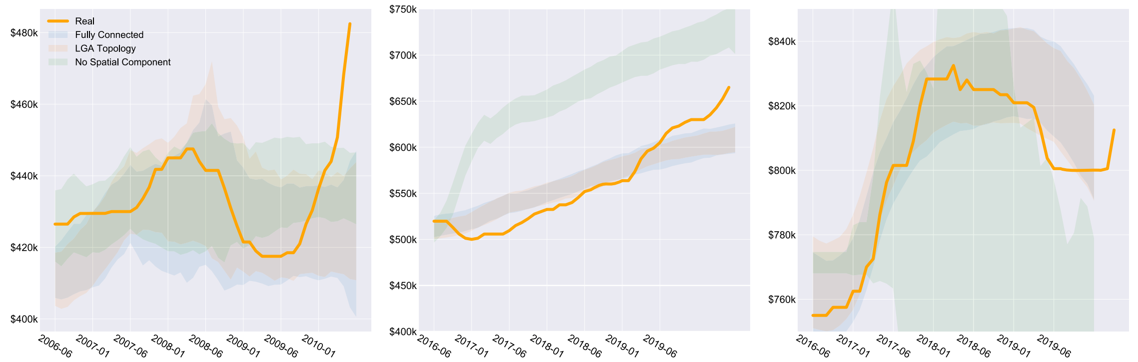


Fig. 18 Various network architectures (visualised as the mean ± standard deviation range), the actual trend is given in orange. We see the individual submarket methods perform well, whereas in this case the single area (meaning no submarkets) method fails to adequately capture the overall trends

terms. However, α remains, which controls the boundedness of the agent as discussed in "Appendix L", and h keeps the same interpretation.

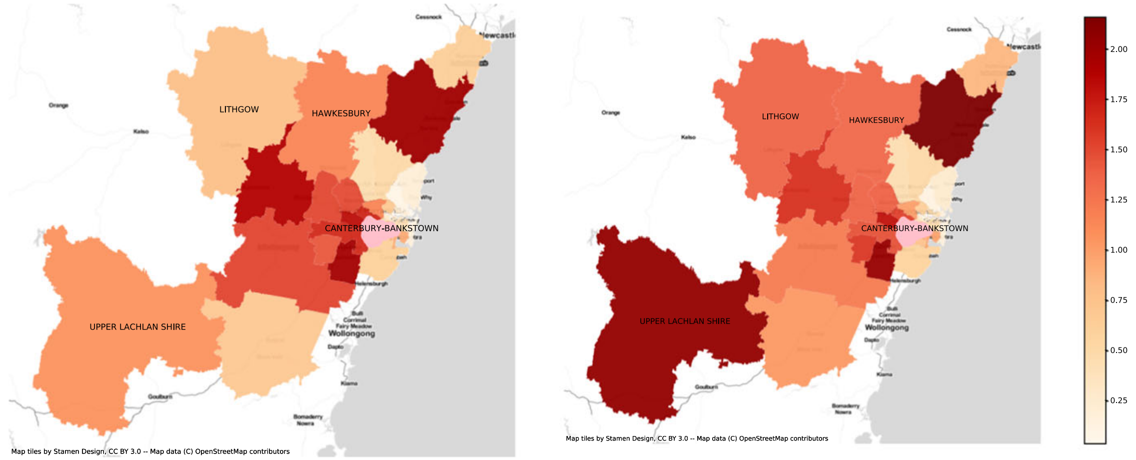
G.1.2 Fully connected graph

To represent the fully connected areas, we use a complete graph where every LGA is connected to every other LGA, i.e. $G = K_{38}$. Again, outreach [from Eq. (4.3)] need not be considered, since now every area is directly connected to one another. With this representation, the introduced parameters still remain (i.e α, β, h). Note that α again directly corresponds to the boundedness (discussed in "Appendix L") and does not encompass outreach. β and h have no change to their original meanings introduced in Sect. 4. This introduces individual submarkets (based on LGAs) into the simulation, but does not enforce any spatial-based search costs within the market. As in the proposed approach, agent calibration is also based on the area in which they reside, so agent characteristics match that of their area.

G.1.3 Analysis

The resulting comparisons are visualised in Fig. 18 which shows the models which include individual areas significantly outperforming the overall Greater Sydney model, indicating the usefulness of area-specific submarkets. Between the two area models, there was little difference in aggregate performance (as shown in Fig. 18) which shows the performance improvements come mainly from the introduction of submarkets, not necessarily on the overall spatial structure. This shows that β, which is based on the individual areas, initialising agent characteristics based on location, and the modification of price setting in terms of $\overline{Q}_{\mathrm{h}}$ based on the area are key for resulting prices (more so than the spatial outreach costs).

However, when considering the resulting agent preferences (in terms of suburbs to purchase in), the fully connected map had significantly more people moving to more remote regions, whereas the LGA connected spatial topology prevented as drastic movements (with the spatial outreach term), capturing the fact people are often tied to specific areas (i.e. those who work in the CBD and currently reside near there, are



(a) Spatial Structure Enforced **(b)** Fully Connected Structure

Fig. 19 Variations in agent preferences with the proposed spatial structure (left) and without a spatial structure enforced (right) from first-time home buyers situated in the Canterbury–Bankstown region (visualised in pink and labelled). The colour indicates the percentage of purchases in those areas, with dark indicating a high percentage of relative purchases (controlled for population sizes). Without a preserved spatial structure we see much higher rate of purchases on the outskirts of the Greater Sydney region, such as in Upper Lachlan Shire and Lithgow (labelled). We verify the proposed movements are logical in Sect. 6.3 (colour figure online)

unlikely to the outskirts of the Greater Sydney region). This is visualised in Fig. 19, where we show the proposed movements follow a diffusive-like pattern (Slavko et al. 220b) compared to the fully connected topology which had movements to more remote regions of Greater Sydney. We further discuss the resulting movements Sect. 6.3, where we show with the agent movement patterns with the proposed spatial structure are logical and consistent with the actual reported movements in the Greater Sydney region based on the observed trends reported in the recent literature.

In this section, we have shown the area-specific submarket extensions significantly outperform an equivalent model which does not include individual areas, highlighting the importance of capturing submarkets. Furthermore, the incorporation of area-specific submarkets allows for additional insights (such as those in Sect. 6.3.1 which would not otherwise be possible). We then further validated the choice of the network topology by comparing resulting agent preferences, and showing the proposed architecture (with spatial-based search costs) prevents drastic movements by the agent (in terms of relative distance from an agent's current location), allowing the agents to act based on their location (preferring closer areas) in a manner consistent with the actual observed trends as discussed in Sect. 6.3.

H: Experiment settings

Due to the stochastic and non-deterministic nature of ABMs, we run 100 Monte Carlo simulations per run (unless otherwise stated) and report the aggregate results over all runs to get a robust estimate.

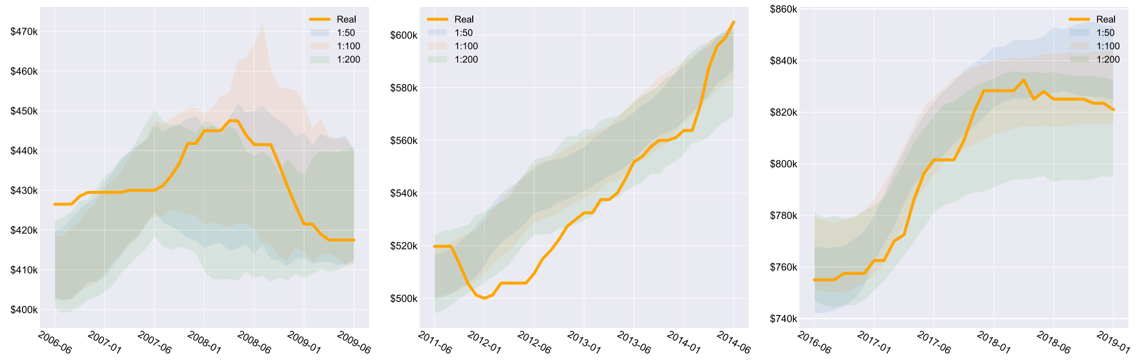


Fig. 20 Robustness of scale over the training phases (visualised as the mean $\pm$ standard deviation range for varying scales). We see with varying scales similar results are recovered (with large areas of overlap), indicating in-variance to the scales used (within acceptable bounds)

H.1 Scale

Experiments are run at a 1:100 scale of the true housing market, i.e. every one hundred households in the Greater Sydney region are represented by one household in the model. The 1:100 scale was chosen for efficiency, but results for 1:50, 1:100, and 1:200 are also presented in Fig. 20 to show robustness to scale. There is an upper limit on the scale where the performance will begin to degrade, for example, the number of overseas investments is given in Table 5, and by using a scale close to 1:1000, we would lose the contribution of foreign investments (since the values would be < 1). For lower scales (i.e. 1:1), the results may be more accurate but this comes at the expense of increased computational power, so the 1:100 provided a good trade-off between accuracy and efficiency.

I: Initialisation data

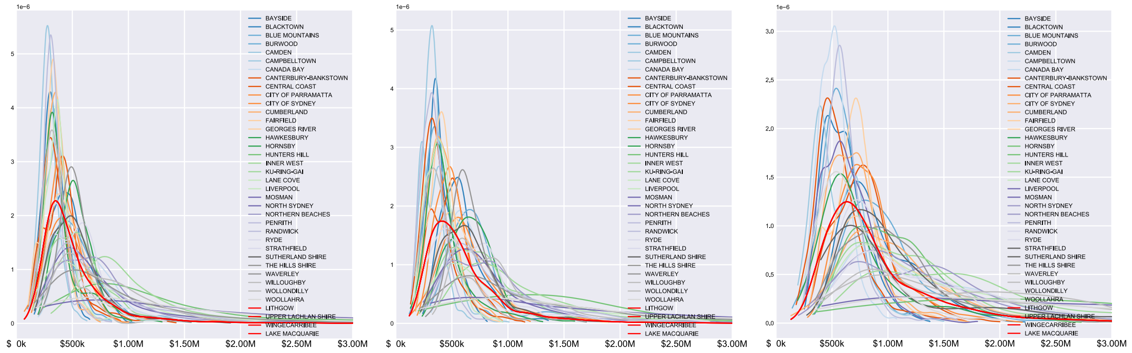
I.1 Data

All real estate listings and sales from 2006 to present (2020) were used from SIRCA–CoreLogic, including the sale price, LGA, and sale date. These data are used as the actual price, and to calibrate the ABM.

I.2 Spatial initialisation

I.2.1 Pricing distributions

Between LGAs, there is a wide range of dwelling sale prices, and different distributions of prices amongst the LGAs as well. To sample from this effectively, we use kernel density estimation (KDE) to create a probability density function for each LGA for each time period. The previous 3 months of sales from the beginning of the time period are used to generate the density function. Scott's Rule (Scott 2015) is used to assign the bandwidth, which sets the bandwidth to $n^{\frac{-1}{d+4}}$, where n is the number of data points



(a) 2006 **(b)** 2011 **(c)** 2015

Fig. 21 KDE plots for each LGA based on SIRCA–CoreLogic data. Dark red indicates the Greater Sydney average, and this is assigned to LGAs without enough data to generate their own reliable KDE

(in this case dwelling sales in the LGA at the beginning of the time period), and d is the number of dimensions (in this case $d = 1$). When new houses are created for an area, they are set with an initial quality based on this distribution. The resulting KDEs are shown in Fig. 21.

I.2.2 Positioning

Households are not assigned to an LGA directly, as households can freely move areas. Instead, the households area is based on the residential dwelling of the household (and thus can change over time). When we reference a households area, we are referring to the LGA of the dwelling where the household currently resides.

At the beginning of the simulation, households which are homeowners are assigned to dwellings to match the population distribution amongst LGAs. The income and liquid wealth for the household are then assigned based on the brackets from the dwellings LGA. Renters are assigned a random LGA to begin with (again weighted by the population of each LGA) and income and wealth based on the distribution of that LGA. Households then try and find a rental they can afford (on with a rental price approximately 10–30% of the household's income) which may mean some have to move LGAs.

I.3 Time periods

In line with the previous work of Glavatskiy et al. (2020), and following the Australian census timelines (which are performed every 5 years), we choose the three most recent census periods for analysis. These are 2006–2010, 2011–2015, and 2016–2019. The length was chosen such that upon new census information becoming available, a new simulation is run. Meaning, a separate model (and optimisation process) is run for each of the time periods to ensure the model is calibrated to the most recent data available. In doing so, we ensure the agent characteristics of the model most closely match those in the true Greater Sydney market. As each period corresponds to the census years, there is a large array of available data for calibration to ensure the models begin in a state as close to possible as the true populations state. Alternate (non-census) dates could be used; however, the model may not begin with as accurate of a reflection on the true underlying agent characteristics (depending on the data availability). While the models are calibrated for the time periods outlined here, such calibrations would also work well for surrounding dates (or alternate run duration's), or the model could be re-calibrated for alternative dates to provide additional forecasting—for example, after the 2021 census, the agent characteristics could be reassigned and a new optimisation process run to reflect updated agent behaviours, likewise for past market behaviour such as with the 2001 census.

I.4 Household characteristics

I.4.1 Area

Agents are initialised into an area based on the Australian census data, meaning the population of each area at the beginning of the simulations corresponds to the proportions from the census data for that time period. For example, if there are three areas "A", "B", "C", and the true proportions in each are 60:25:15, the model will also populate agents into the three areas according to this proportion. Throughout the simulation, agents may move areas. They may be forced to move to a cheaper area if they cannot afford their current area, or they may move to a more affluent area if they can afford a dwelling there. So once the simulation begins, the movement dynamics are controlled by the agents' cash flow position (again from census data, outlined below). Initialising agents into areas based on census data allows for correct agent characteristics (such as income and net worth) that directly line up with those observed throughout the Greater Sydney region.

I.4.2 Income

Income is assigned from the distribution based on the households area. This distribution comes from the census data. Income grows throughout the simulation. The income brackets follow those specified in the census data.

I.4.3 Liquid wealth

Again, the liquid wealth (liquidity) of a household is based on the true distributions from census data. However, in this case, liquidity is not available per LGA, only for Greater Sydney as a whole. So to map a household to an appropriate liquidity bracket, the households liquid is based on the income of the household. That is, if a household is in the top X% of earners in an LGA, the liquidity will be in the top X% as well (approximately, since liquidity is from brackets).

I.5 Population distribution

In this case, there are three measures of interest. The total number of dwellings, the total number of households, and the distribution of these households amongst LGAs.

The dwellings and households estimates from the census data are used for each year, and simple linear projections used for forecasting the growth of these. The distribution amongst LGAs is that recorded at the start of the simulation and is assumed to grow linearly with the overall population size. Individual LGA future population projections are available from 2016 onward, but as no projections existed before this date, we used this simplified measure instead of all LGAs growing by a fixed percentage within a given simulation period. As such, higher movements towards one particular LGA throughout simulation could indicate the requirement of additional dwellings being built here to cater for the growth, which is another contribution we consider in later sections of this work.

J: Movement pattern visualisations

Over 10 million total movements were tracked across the simulations (approximately 3.3 million per time period). All plots in this section represent the normalised heatmaps of these movements. The total number of movements to a particular LGA is scaled by the population size of this LGA, meaning the results can be interpreted as a preference for certain areas rather than visualising the population size of the LGAs. Therefore, movements are not just reflecting larger populations, instead, reflecting a larger portion of people moving there relative to the size. All movements are then normalised such that the summation of all cells in the plot is 1, meaning if a particular cell has a value of 0.05, this means 5% of all matched movements moved to this LGA.

The rows and columns of the plots are always sorted in ascending order based on median price, i.e. the most affordable LGAs first, and the most expensive LGA as the final row or column.

K: Exogenous variables

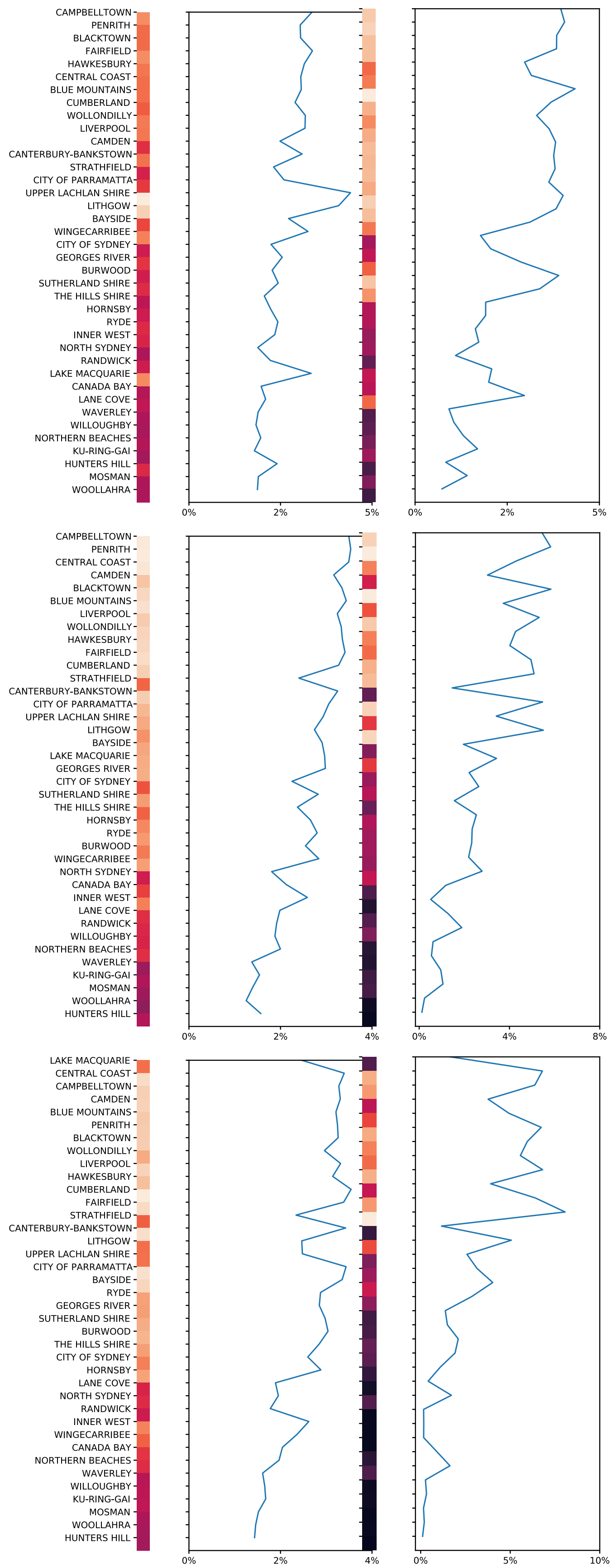
There are two main external influences on the model, which are governed by government approvals (in the case of overseas investments) and the central bank (in the case of mortgage rates).

K.1 Overseas investors

Overseas investments are often cited as a key driver of price growth in the Australian market (Rogers et al. 2017), and figures show the foreign investment has more than tripled since the mid-1990s (Haylen 2014). However, actual data on foreign investments are difficult to find. ABS has described their own data on overseas investments to parliament as “hit or miss” Iggulden (2014).

The purpose of this work is not a full investigation into overseas investments [overviews are given in Gauder et al. (2014), House of Representatives Standing Committee on Economics (2014)], but rather the contribution overseas might have in relation to many other factors with the readily available data (be this complete or not).

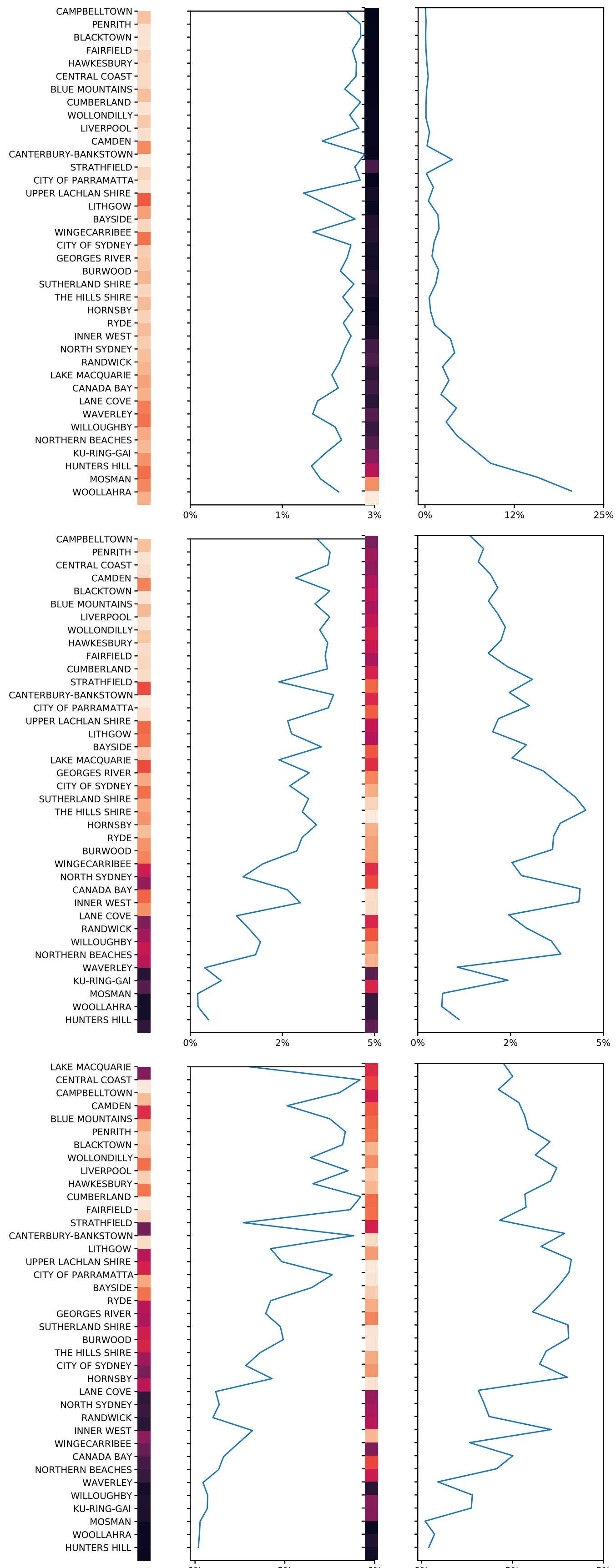
Fig. 22 Migrations. These plots capture new households in Greater Sydney throughout the simulation period, due to either migration or splitting of existing households. The first row is the 2006–2010 period, the middle row the 2011–2015 period, and the final row the 2016–2019 period



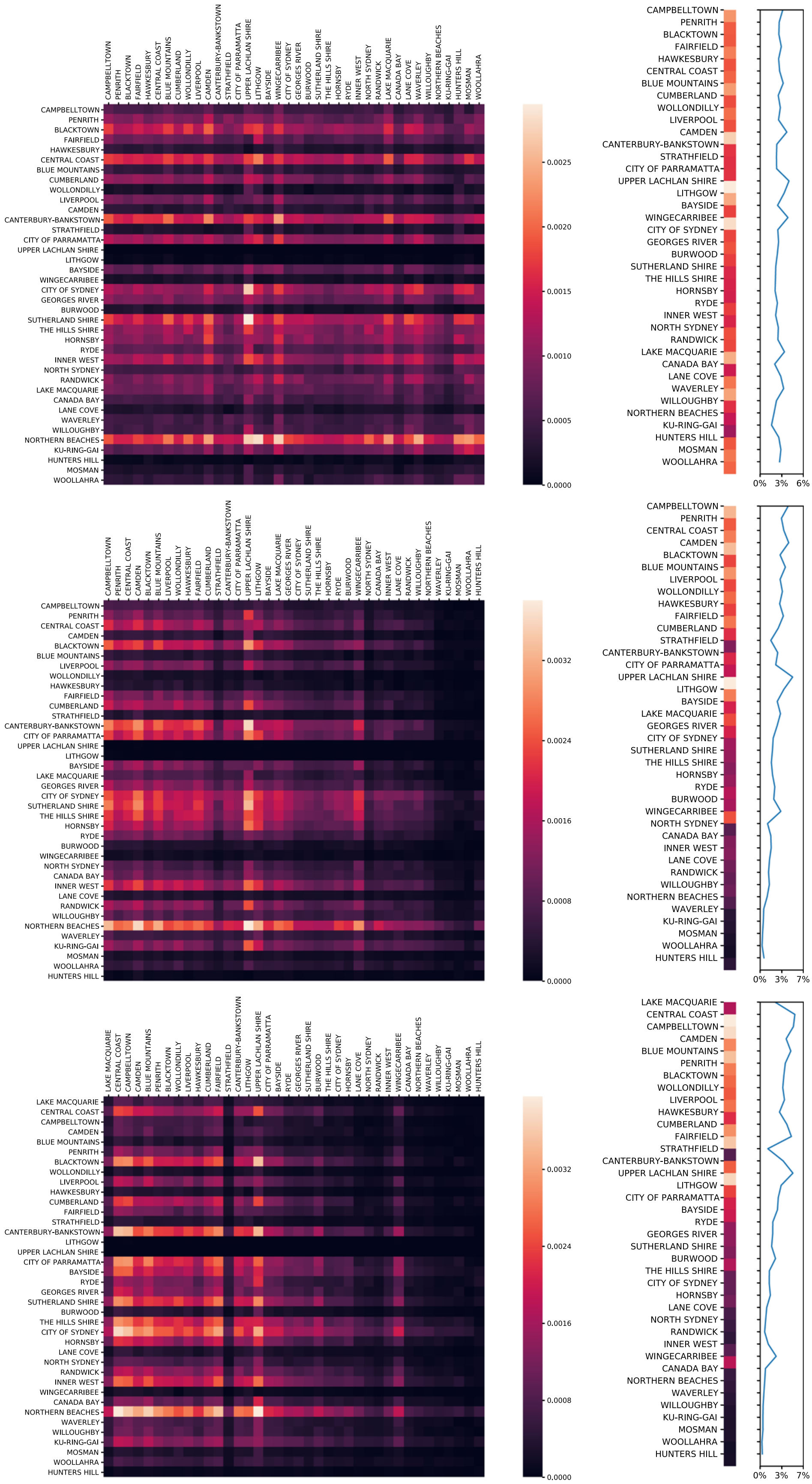
(a) New Renters **(b)** New Owners

68

Fig. 23 Investors. These plots show the simulation difference between local and overseas investment patterns. The first row is the 2006–2010 period, the middle row the 2011–2015 period, and the final row the 2016–2019 period



(a) Local **(b)** Overseas



(a) Renting LGA (rows), to Purchase LGA (columns) **(b)** Purchase LGAs

Fig. 24 First-time home buyers. The first row is the 2006–2010 period, the middle row the 2011–2015 period, and the final row the 2016–2019 period 70

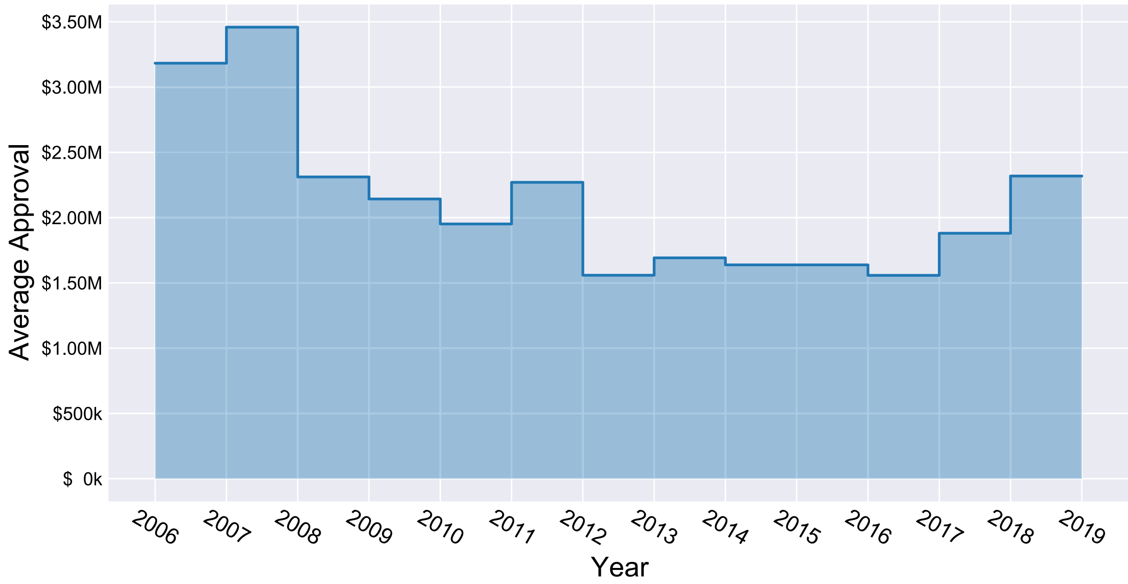


Fig. 25 Average overseas approval amount

Table 5 Overseas investment approval

Period	Number approved	Value approved	Average per approval
2018–2019	1337	$3,100,000,000	$2,318,624
2017–2019	2340	$4,400,000,000	$1,880,342
2016–2017	4224	$6,580,000,000	$1,557,765
2014–2015	12349	$20,230,000,000	$1,638,189
2013–2014	7814	$13,220,000,000	$1,691,835
2012–2013	3580	$5,580,000,000	$1,558,659
2011–2012	3048	$6,920,000,000	$2,270,341
2010–2011	2598	$5,070,000,000	$1,951,501
2009–2010	910	$1,950,000,000	$2,142,857
2008–2009	956	$2,210,000,000	$2,311,715
2007–2008	1223	$4,230,000,000	$3,458,708
2006–2007	908	$2,890,000,000	$3,182,819

For this, we use the annual reports from the Foreign Investment Review Board (FIRB) from June 2006 to June 2019. The June 2019–June 2020 report was not available at the time of this writing (in 2020), as reports are not made available until the following year. Data are provided yearly at a NSW level, which is converted to monthly (simply dividing by 12). Again, data in this area are sparse, so this is the closest estimate we could derive. These data are provided in Table 5, and the average approval per year given in Fig. 25.

While the data are provided for the entirety of NSW, it has been shown that foreign investors prefer the inner city over rural areas, and thus, the NSW levels have been used for Greater Sydney. This is a fair assumption since the numbers are relatively

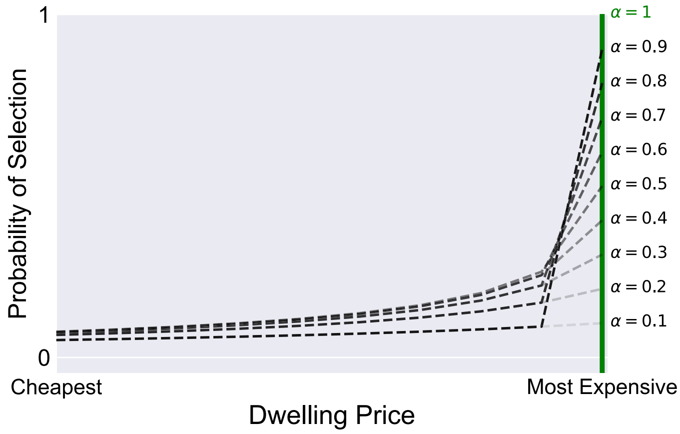


Fig. 26 α's effect on utility maximising behaviour

conservative anyway. For the testing period, the most recent overseas approval value from the training period is used.

K.2 Mortgage rates

Mortgage rates are those set by the RBA. The final training months mortgage rate is used throughout the testing period since no real value can be read.

L: Utility function

Following Axtell et al. (2014), agents are assumed to choose the most expensive house they can afford, that is, the house price directly corresponds to the utility for the agent.

However, the introduction of α alters this, such that there is some uncertainty or error in the agents choice. When $\alpha = 1$, the perfect utility maximisation behaviour is recovered where the agent attempts to purchase the most expensive dwelling they can afford. For $\alpha < 1$, the agent buys the most expensive dwelling they can afford with probability α, which then decreases for each subsequent listing in turn. This is visualised in Fig. 26. For high α, we can see the probability mass is contained only in the highest priced dwellings. For lower α, this probability mass becomes more distributed, meaning less focus on utility, and potential for cheaper houses to be purchased. For $\alpha = 0$, the utility is not considered at all and a random house within the agents budget is chosen (i.e. the probability mass is uniform across options). α, therefore, corresponds to the boundedness of the agent.

The above description considers the case of uniform knowledge, i.e. for investors where they are assumed to be invariant to the areas available. However, for first-time home buyers, we propose a space-based knowledge where buyers are more likely to consider listings close to where they are renting. The probability associated with the distance to the agents' location is visualised in Fig. 27. The uniform knowledge of investors is given in green, and the spatial knowledge of first-time home buyers is given as the dotted black line.

For first-time home buyers, the probability of viewing a listing is therefore controlled by both the proximity of the listing to the agents current (rental) location, and

Fig. 27 Probability of viewing based on distance to agents location

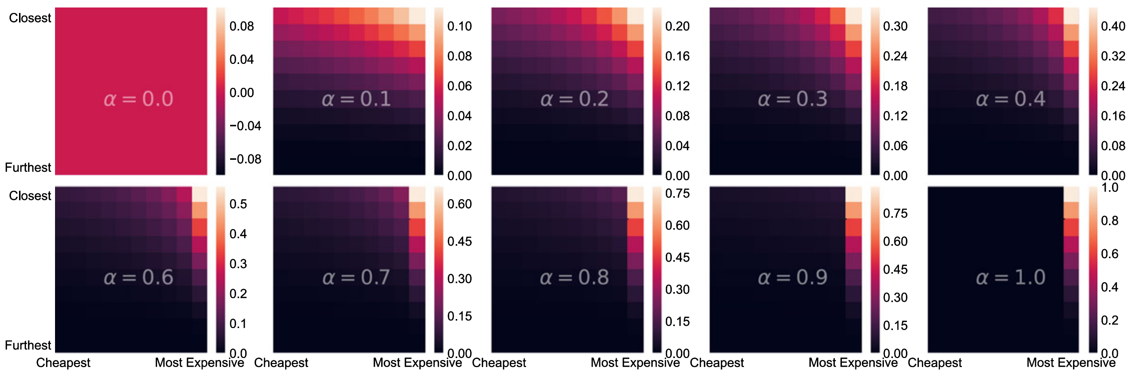


Fig. 28 First-time home buyers probability of viewing a listing for various α's. Showing the relationship between dwelling price (x-axis) and distance to dwelling (y-axis), and how α adjusts this distribution. Low α's correspond to higher dispersion, and less focus on utility maximising behaviour. High α's focus the agent on dwellings which maximise utility

the price of the listing. This is visualised in Fig. 28. For $\alpha = 0$, the agent preference is uniform across all choices, placing no emphasis on utility (from either price or difference). As α increases, the focus shifts to the more expensive dwellings, and does so based on the distance to the listing. This is shown in Fig. 28, where with increasing α the emphasis focuses on the top right corner, which is the optimal value for both distance (closest) and price (most expensive in the agents budget). We can see that price remains the most important term in the agents' utility though, with close listings with low prices having a low resulting probability, indicating the agent likely wants to move to a more affluent area if they can afford to do so. However, given an equal price, agents will prefer the closer listing.

References

Alhashimi H, Dwyer W (2004) Is there such an entity as a housing market. In: 10th annual pacific rim real estate conference (press), Bangkok

Arcaute E, Molinero C, Hatna E, Murcio R, Vargas-Ruiz C, Masucci AP, Batty M (2016) Cities and regions in Britain through hierarchical percolation. R Soc Open Sci 3(4):150691

Axtell R, Farmer D, Geanakoplos J, Howitt P, Carrella E, Conlee B, Goldstein J, Hendrey M, Kalikman P, Masad D, et al. (2014) An agent-based model of the housing market bubble in metropolitan Washington, DC. In: Whitepaper for Deutsche Bundesbank's Spring conference on "Housing markets and the macroeconomy: challenges for monetary policy and financial stability"

73

Bahadir B, Mykhaylova O (2014) Housing market dynamics with delays in the construction sector. J Hous Econ 26:94–108
Bangura M, Lee CL (2019) The differential geography of housing affordability in Sydney: a disaggregated approach. Aust Geogr 50(3):295–313
Bangura M, Lee CL (2020) Housing price bubbles in Greater Sydney: evidence from a submarket analysis. Hous Stud 1–36. https://doi.org/10.1080/02673037.2020.1803802
Baptista R, Farmer JD, Hinterschweiger M, Low K, Tang D, Uluc A (2016) Macroprudential policy in an agent-based model of the UK housing market. Bank of England working papers 619, Bank of England. https://doi.org/10.2139/ssrn.2850414
Barbosa H, Barthelemy M, Ghoshal G, James CR, Lenormand M, Louail T, Menezes R, Ramasco JJ, Simini F, Tomasini M (2018) Human mobility: models and applications. Phys Rep 734:1–74
Barthelemy M (2016) The structure and dynamics of cities. Cambridge University Press, Cambridge
Barthelemy M (2019) The statistical physics of cities. Nat Rev Phys 1(6)
Barthelemy M, Bordin P, Berestycki H, Gribaudi M (2013) Self-organization versus top-down planning in the evolution of a city. Sci Rep 3:2153
Bergstra J, Yamins D, Cox D (2013) Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. In: International conference on machine learning, pp 115–123
Berndt DJ, Clifford J (1994) Using dynamic time warping to find patterns in time series. In: Proceedings of the 3rd international conference on knowledge discovery and data mining. AAAI Press, AAAIWS'94, pp 359–370
Bessant JC, Johnson G (2013) Dream on declining homeownership among young people in Australia? Hous Theory Soc 30(2):177–192
Burnside C, Eichenbaum M, Rebelo S (2016) Understanding booms and busts in housing markets. J Polit Econ 124(4):1088–1147
Campolongo F, Cariboni J, Saltelli A (2007) An effective screening design for sensitivity analysis of large models. Environ Model Softw 22(10):1509–1518
Carstensen CL (2015) An agent-based model of the housing market: steps toward a computational tool for policy analysis. University of Copenhagen, MSc-szakdolgozat
Chang SL, Harding N, Zachreson C, Cliff OM, Prokopenko M (2020) Modelling transmission and control of the COVID-19 pandemic in Australia. arXiv preprint arXiv:2003.10218
Cheng IH, Raina S, Xiong W (2014) Wall street and the housing bubble. Am Econ Rev 104(9):2797–2829
Cliff OM, Harding N, Piraveenan M, Erten EY, Gambhir M, Prokopenko M (2018) Investigating spatiotemporal dynamics and synchrony of influenza epidemics in Australia: an agent-based modelling approach. Simul Model Pract Theory 87:412–431
Conlisk J (1996) Why bounded rationality? J Econ Lit 34(2):669–700
Crosato E, Nigmatullin R, Prokopenko M (2018) On critical dynamics and thermodynamic efficiency of urban transformations. R Soc Open Sci 5(10):180863
Crosato E, Prokopenko M, Harré MS (2021) The polycentric dynamics of Melbourne and Sydney: suburb attractiveness divides a city at the home ownership level. Proc R Soc A 477(2245):20200514
Edmonds B, ní Aodha L (2018) Using agent-based modelling to inform policy—what could possibly go wrong? In: International workshop on multi-agent systems and agent-based simulation. Springer, pp 1–16
Fernald M (2020) Americas rental housing 2020. Joint Center for Housing Studies of Harvard University, Cambridge
Frías-Paredes L, Mallor F, León T, Gastón-Romeo M (2016) Introducing the temporal distortion index to perform a bidimensional analysis of renewable energy forecast. Energy 94:180–194
Frías-Paredes L, Mallor F, Gastón-Romeo M, León T (2017) Assessing energy forecasting inaccuracy by simultaneously considering temporal and absolute errors. Energy Convers Manag 142:533–546
Gallegati M, Kirman A (1999) Beyond the representative agent. Edward Elgar Publishing, Cheltenham
Gauder M, Houssard C, Orsmond D, et al. (2014) Foreign investment in residential real estate. RBA Bulletin, June pp 11–18
Ge J (2013) Who creates housing bubbles? An agent-based study. In: International workshop on multi-agent systems and agent-based simulation. Springer, pp 143–150
Ge J (2017) Endogenous rise and collapse of housing price: an agent-based model of the housing market. Comput Environ Urban Syst 62:182–198

Geanakoplos J, Axtell R, Farmer JD, Howitt P, Conlee B, Goldstein J, Hendrey M, Palmer NM, Yang CY (2012) Getting at systemic risk via an agent-based model of the housing market. Am Econ Rev 102(3):53–58

Gilbert N, Hawksworth JC, Swinney PA (2009) An agent-based model of the English housing market. In: AAAI spring symposium: technosocial predictive analytics, pp 30–35

Glavatskiy KS, Prokopenko M, Carro A, Ormerod P, Harre M (2020) Explaining herding and volatility in the cyclical price dynamics of urban housing markets using a large scale agent-based model. arXiv preprint arXiv:2004.07571

Goldstein J (2017) Rethinking housing with agent-based models: Models of the housing bubble and crash in the Washington DC area 1997–2009. PhD thesis, George Mason University

Greater Sydney Commission (2018) Greater Sydney region plan: a metropolis of three cities. NSW Department of Planning and Environment. https://www.greater.sydney/metropolis-of-three-cities. Accessed 22 Aug 2020

Guest R, Rohde N (2017) The contribution of foreign real estate investment to housing price growth in Australian capital cities. Abacus 53(3):304–318

Haylen A (2014) House prices, ownership and affordability: trends in New South Wales. NSW Parliamentary Library

Herman J, Usher W (2017) SALib: an open-source Python library for sensitivity analysis. J Open Source Softw. https://doi.org/10.21105/joss.00097

House of Representatives Standing Committee on Economics (2014) Report on Foreign Investment in Residential Real Estate. The Parliament of the Commonwealth of Australia

Huang Y, Ge J (2009) House prices and the collapse of stock market in mainland China?-an empirical study on house price index. In: Pacific Rim real estate conference

Iggulden T (2014) ABS admits data on foreign real estate buyers is 'hit and miss'. https://www.abc.net.au/news/2014-06-25/abs-admits-foreign-real-estate-purchase-data-unreliable/5549926 Accessed 10 May 2020

Kim JH, Pagliara F, Preston J (2005) The intention to move and residential location choice behaviour. Urban Stud 42(9):1621–1636

Kouwenberg R, Zwinkels R (2014) Forecasting the US housing market. Int J Forecast 30(3):415–425

Kouwenberg R, Zwinkels RC (2015) Endogenous price bubbles in a multi-agent system of the housing market. PLoS ONE 10(6):e0129070

Kupke V, Rossini P (2011) Housing affordability in Australia for first home buyers on moderate incomes. Property Management

La Cava G, Leal H, Zurawski A et al (2017) Housing accessibility for first home buyers. Reserve Bank of Australia Bulletin, pp 19–28

LeBaron B, Tesfatsion L (2008) Modeling macroeconomies as open-ended dynamic systems of interacting agents. Am Econ Rev 98(2):246–50

Louf R, Barthelemy M (2013) Modeling the polycentric transition of cities. Phys Rev Lett 111(19):198702

Mc Breen J, Goffette-Nagot F, Jensen P (2010) Information and search on the housing market: an agent-based model. In: Li Calzi M, Milone L, Pellizzari P (eds) Progress in artificial economics. Springer, Berlin, pp 153–164

McMaster R, Watkins C (1999) The economics of housing: the need for a new approach. In: PRRES/AsRES/IRES conference. Kuala Lumpur

Miles W (2008) Boom-bust cycles and the forecasting performance of linear and non-linear models of house prices. J Real Estate Finance Econ 36(3):249–264

Morris MD (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33(2):161–174

Myers C, Rabiner L, Rosenberg A (1980) Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. IEEE Trans Acoust Speech Signal Process 28(6):623–635

Pangallo M, Nadal JP, Vignes A (2019) Residential income segregation: a behavioral model of the housing market. J Econ Behav Organ 159:15–35

Pawson H, Martin C (2020) Rental property investment in disadvantaged areas: the means and motivations of Western Sydney's new landlords. Hous Stud 1–23 https://doi.org/10.1080/02673037.2019.1709806

Piazzesi M, Schneider M, Tuzel S (2007) Housing, consumption and asset pricing. J Financ Econ 83(3):531–569

Piovani D, Arcaute E, Uchoa G, Wilson A, Batty M (2018) Measuring accessibility using gravity and radiation models. R Soc Open Sci 5(9):171668

Poledna S, Miess MG, Hommes CH (2019) Economic forecasting with an agent-based model. Available at SSRN 3484768
Polhill G (2018) Why the social simulation community should tackle prediction. Rev Artif Soc Soc Simul. https://rofasss.org/2018/08/06/gp/. Accessed 1 May 2020
Power C (2009) A spatial agent-based model of n-person prisoner's dilemma cooperation in a socio-geographic community. J Artif Soc Soc Simul 12(1):8
Raimbault J, Broere J, Somveille M, Serna JM, Strombom E, Moore C, Zhu B, Sugar L (2020) A spatial agent based model for simulating and optimizing networked eco-industrial systems. Resour Conserv Recycl 155:104538
Randolph B, Pinnegar S, Tice A (2013) The first home owner boost in Australia: a case study of outcomes in the Sydney housing market. Urban Policy Res 31(1):55–73
Rogers D, Lee CL, Yan D (2015) The politics of foreign investment in Australian housing: Chinese investors, translocal sales agents and local resistance. Hous Stud 30(5):730–748
Rogers D, Wong A, Nelson J (2017) Public perceptions of foreign and Chinese real estate investment: intercultural relations in Global Sydney. Aust Geogr 48(4):437–455
Sakoe H, Chiba S (1978) Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans Acoust Speech Signal Process 26(1):43–49
Saltelli A, Tarantola S, Campolongo F, Ratto M (2004) Sensitivity analysis in practice: a guide to assessing scientific models, vol 1. Wiley Online Library, New York
Saltelli A, Aleksankina K, Becker W, Fennell P, Ferretti F, Holst N, Li S, Wu Q (2019) Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. Environ Model Softw 114:29–39
Sanchez DG, Lacarrière B, Musy M, Bourges B (2014) Application of sensitivity analysis in building energy simulations: combining first-and second-order elementary effects methods. Energy Build 68:741–750
Scott DW (2015) Multivariate density estimation: theory, practice, and visualization. Wiley, New York
Shahriari B, Swersky K, Wang Z, Adams RP, De Freitas N (2015) Taking the human out of the loop: a review of Bayesian optimization. Proc IEEE 104(1):148–175
Shimizu C, Nishimura KG, Watanabe T (2010) Housing prices in Tokyo: a comparison of hedonic and repeat sales measures. Jahrbücher für Nationalökonomie und Statistik 230(6):792–813
Simini F, González MC, Maritan A, Barabási AL (2012) A universal model for mobility and migration patterns. Nature 484(7392):96
Simon HA (1955) A behavioral model of rational choice. Q J Econ 69(1):99–118
Simon HA (1957) Models of man; social and rational. Wiley, New York
Sinai TM (2012) House price moments in boom-bust cycles. Technical report, National Bureau of Economic Research
Slavko B, Glavatskiy K, Prokopenko M (2019) Dynamic resettlement as a mechanism of phase transitions in urban configurations. Phys Rev E 99(4):042143
Slavko B, Glavatskiy K, Prokopenko M (2020a) City structure shapes directional resettlement flows in Australia. Sci Rep 10(1):1–11
Slavko B, Prokopenko M, Glavatskiy KS (2020b) Diffusive resettlement: irreversible urban transitions in closed systems. arXiv preprint arXiv:2009.04094
Snoek J, Larochelle H, Adams RP (2012) Practical Bayesian optimization of machine learning algorithms. In: Advances in neural information processing systems, pp 2951–2959
Tesfatsion L (2002) Agent-based computational economics: growing economies from the bottom up. Artif Life 8(1):55–82
Thomas M, Hall A (2016) Housing affordability in Australia. Briefing Book: Key Issues for the 45th Parliament pp 86–90
Ustvedt S (2016) An agent-based model of a metropolitan housing market-linking micro-level behavior to macro-level analysis. Master's thesis, NTNU
Vallance L, Charbonnier B, Paul N, Dubost S, Blanc P (2017) Towards a standardized procedure to assess solar forecast accuracy: a new ramp and time alignment metric. Sol Energy 150:408–422
Vincent L, Thome N (2019) Shape and time distortion loss for training deep time series forecasting models. In: Advances in neural information processing systems, pp 4189–4201
Wang W, Yang S, Hu F, Han Z, Jaeger C (2018) An agent-based modeling for housing prices with bounded rationality. In: Journal of physics: conference series, vol 1113. IOP Publishing, p 012014
Watkins CA (2001) The definition and identification of housing submarkets. Environ Plan A 33(12):2235–2253

Wei SJ, Zhang X, Liu Y (2012) Status competition and housing prices. Technical report, National Bureau of Economic Research
Wilkins R, Lass I (2015) The household, income and labour dynamics in Australia survey: Selected findings from waves 1 to 12. Melbourne Institute of Applied Economic and Social Research, University of
Wong PY (2017) Foreign real estate investment and the Australian residential property market: a study on Chinese investors. Int J Soc Behav Educ Econ Bus Ind Eng 11:1529–1538
Yetsenga R, Emmett F (2020) The ANZ CoreLogic housing affordability report 2020. ANZ Media Centre

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

CHAPTER 4

A maximum entropy model for inference of agent decisions in economic markets

---

In interacting social systems, such as markets, the negative feedback generated from actions can be enough to (temporarily) stabilise resulting outcomes to a *statistical* equilibrium (Scharfenaker and Foley, 2017). In contrast to typical notions of equilibrium, statistical equilibrium represents a probability distribution of all possible states of the system. In this section, we develop a method of inferring agent decisions in statistical equilibrium based on prior beliefs and market feedback loops, and consider the temporal evolution of these beliefs.

The approach proposed here is complementary to the agent-based model developed in the previous chapter (Yang and Carro, 2020). While the ABM provided a rich overall picture of market dynamics, the behavioural rules, including when to buy and sell, and the feedback effects, were hand-crafted. In contrast, the approach proposed in this section provides a more generalised view, based on a least-biased method of inference for decision-making and feedback loops in competitive markets, helping to provide support for this hand-crafted element.

entropy

MDPI

Article

A Maximum Entropy Model of Bounded Rational Decision-Making with Prior Beliefs and Market Feedback

Benjamin Patrick Evans * and Mikhail Prokopenko

Centre for Complex Systems, The University of Sydney, Sydney, NSW 2006, Australia; mikhail.prokopenko@sydney.edu.au

* Correspondence: benjamin.evans@sydney.edu.au

Abstract: Bounded rationality is an important consideration stemming from the fact that agents often have limits on their processing abilities, making the assumption of perfect rationality inapplicable to many real tasks. We propose an information-theoretic approach to the inference of agent decisions under Smithian competition. The model explicitly captures the boundedness of agents (limited in their information-processing capacity) as the cost of information acquisition for expanding their prior beliefs. The expansion is measured as the Kullblack–Leibler divergence between posterior decisions and prior beliefs. When information acquisition is free, the homo economicus agent is recovered, while in cases when information acquisition becomes costly, agents instead revert to their prior beliefs. The maximum entropy principle is used to infer least biased decisions based upon the notion of Smithian competition formalised within the Quantal Response Statistical Equilibrium framework. The incorporation of prior beliefs into such a framework allowed us to systematically explore the effects of prior beliefs on decision-making in the presence of market feedback, as well as importantly adding a temporal interpretation to the framework. We verified the proposed model using Australian housing market data, showing how the incorporation of prior knowledge alters the resulting agent decisions. Specifically, it allowed for the separation of past beliefs and utility maximisation behaviour of the agent as well as the analysis into the evolution of agent beliefs.

Keywords: decision-making; bounded rationality; complexity economics; information-theory; maximum entropy principle; quantal response statistical equilibrium

JEL Classification: D91; G41; D83; C61; C60; C50

check for updates

Citation: Evans, B.P.; Prokopenko, M. A Maximum Entropy Model of Bounded Rational Decision-Making with Prior Beliefs and Market Feedback. *Entropy* **2021**, *23*, 669. https://doi.org/10.3390/e23060669

Academic Editors: Ryszard Kutner, Christophe Schinckus and Eugene Stanley

Received: 21 April 2021
Accepted: 21 May 2021
Published: 26 May 2021

1. Introduction

Economic agents are often faced with partial information and make decisions under pressure, yet many canonical economic models assume perfect information and perfect rationality. To address these challenges, Simon [1] introduced bounded rationality as an alternate attribute of decision-making. Bounded rationality aims to represent partial access to information, with possible acquisition costs, and limited computational cognitive processing abilities of the decision-making agents.

Information theory offers several natural advantages in capturing bounded rationality, interpreting the economic information as the source data to be delivered to the agent (receiver) through a noisy communication channel (where the level of noise is related to the "boundedness" of the agent). This representation has spurred the creation of information-theoretic approaches to economics, such as Rational Inattention (R.I.) [2], and more recently, the application of R.I. to discrete choice [3]. Another approach represents decision-making as a thermodynamic process over state changes and employs the energy-minimisation principle to derive suitable decisions [4].

These approaches have shown how one can incorporate a priori knowledge into decision-making, but place no consideration to inferring these decisions based on observed macroeconomic outcomes (e.g., a distribution of profit rates within a financial market) and market feedback loops. Independently, another recent information-theoretic framework, Quantal Response Statistical Equilibrium (QRSE) [5], was developed aiming to infer least biased (i.e., "maximally noncommittal with regard to missing information" [6]) decisions through the maximum entropy principle, given only the macroeconomic outcomes (e.g., when the choice data is unobserved). However, the ways to incorporate prior knowledge into such a system remain mostly unexplored.

In this work, we provide a unification of these approaches, showing how to incorporate prior beliefs into QRSE in a generic way. In doing so, we provide a least biased inference of decision-making, given an agent's prior belief. Specifically, we show how the incorporation of prior beliefs affects the agent's resulting decisions when their individual choices are unobserved (as is common in many real-world economic settings). The proposed information-theoretic approach achieves this by considering a cost of information acquisition (measured as the Kullback-Leibler divergence), where this cost controls deviations from an agent's prior knowledge on a discrete choice set. When the cost of information acquisition is prohibitively high (i.e., when an agent is faced with limitations through time, cognition, cost, or other constraints), the agent falls back to their prior beliefs. When information acquisition is free, the agent becomes a perfect utility maximiser. The cost of information acquisition therefore measures the boundedness of the agent's decision-making.

The proposed approach is general, allowing the incorporation of any form of prior belief, while separating the agents' current expectations from their built-up beliefs. In particular, we show how incorporating prior beliefs into the QRSE framework allows for modelling decisions in a rolling way, when previous decisions "roll" into becoming the latest beliefs. Furthermore, we place the original QRSE in the context of related formalisms, and show that it is a special case of the general model proposed in our study, when the prior preferences (beliefs) are assumed to be uniform across the agent choices. Finally, we verify and demonstrate our approach using actual Australian housing market data, in terms of agent buying and selling decisions.

The remainder of the paper is organised as follows. Section 2 provides a background of information-theoretic approaches to economic decision-making, Section 3 describes QRSE and relevant decision-making literature. Section 4 outlines the proposed model, and Section 5 applies the developed model to the Australian housing market. Section 6 presents conclusions.

2. Background and Motivation

The use of statistical equilibrium (and more generally, information-theoretic) models remains a relatively new concept in economics [7]. For example, Yakovenko [8] outlines the use of statistical mechanics in economics. Scharfenaker and Semieniuk [9] detail the applicability of maximum entropy for economic inference, Scharfenaker and Yang [10] give an overview of maximum entropy and statistical mechanics in economics outlining the benefits of utilising the maximum entropy principle for rational inference, and Wolpert et al. [11] outline the use of maximum entropy for deriving equilibria with bounded rational players in game theory. Earlier, Dragulescu and Yakovenko [12] showed how in a closed economic system, the probability distribution of money should follow the Boltzmann-Gibbs law [13]. Foley [14] discusses Rational expectations and boundedly rational behaviour in economics. Harré [15] gives an overview of information-theoretic decision-theory and applications in economics, and Foley [16] analyses information-theory and results on economic behaviour.

Ömer [17] provides a comparison of "conventional" economic models and newly proposed ideas from complex systems such as maximum entropy methods and Agent-based models (ABM), which deviate from the assumption of homo economicus—a perfectly rational representative agent. Yang and Carro [18] discuss how a combination of agent-based modelling and maximum entropy models can be complementary, leveraging the analytical rigour of maximum entropy methods and the relative richness of agent-based modelling.

One of the key developments in this area is Quantal Response Statistical Equilibrium (QRSE) proposed by Scharfenaker and Foley [5]. This approach enabled applications of the maximum entropy method [6,19,20] to a broad class of economic decision-making. The QRSE model was further explored in [21], arguing that "any system constrained by negative feedbacks and boundedly rational individuals will tend to generate outcomes of the QRSE form". The QRSE approach is detailed in Section 3.1.

Ömer [22–24] applies QRSE to housing markets (which we also use as a validating example), modelling the change in the U.S. house price indices over several distinct periods, and explaining dynamics of growth and dips. Yang [25] applies QRSE to a technological change, modelling the adoption of new technology for various countries over multiple years and successfully recovering the macroeconomic distribution of rates of cost reduction. Wiener [26–28] applies QRSE to labour markets, modelling the competition between groups of workers (such as native and foreign-born workers in the U.S.), and capturing the distribution of weekly wages. Blackwell [29] provides a simplified QRSE for understanding the behavioural foundations. Blackwell further extends this in [30], introducing an alternate explanation for skew, which arises due to the agents having different buy (enter) and sell (exit) preferences. Scharfenaker [31] introduces Log-QRSE for income distribution, and importantly, (briefly) mentions informational costs as a possible cause for asymmetries in QRSE. This is captured by measuring utility U as a sum $U[a,x] + C(a|x)$, allowing for higher costs (C) of entrance or exit into a market, where a is an action and x is a rate. Such a separation allows for an "alternative interpretation of unfulfilled expectations".

These developments show the usefulness of maximum entropy methods, where we have placed particular focus on QRSE, for inferring decisions from only macro-level economic data. However, these approaches do not consider the contribution of a priori knowledge to the resulting decision-making process. The key objective of our study is to generalise the QRSE framework by the introduction of the prior beliefs, as well as the information acquisition costs as a measure of deviation from such priors.

3. Underlying Concepts

Two main concepts form the basis for the proposed model. The first is the QRSE approach developed by [5], and the second is a thermodynamics-based concept of decision-making derived from minimising negative free energy, proposed by [4].

3.1. QRSE

The QRSE framework aims to explain macroeconomic regularities as arising from social interactions between agents. There are two key assumptions stemming from the idea of Smithian competition: Agents observe and respond to macroeconomic outcomes, and agent actions affect the macroeconomic outcome, i.e., a feedback loop is assumed. It is this feedback that is deemed to cause the macroeconomic outcome to have a distribution that stabilises around an average value. Given only the macroeconomic outcome, QRSE infers the least biased distribution of decisions, which result in the observed macroeconomic distribution using the principle of maximum entropy. This makes QRSE particularly useful for inferring decisions when the individual decision level data is unobserved. In the following section, we outline the key notions behind QRSE [5].

3.1.1. Deriving Decisions

Agents are assumed to respond (i.e., make decisions) based on the macroeconomic outcome, for example, based on profit rates x. This is captured by the agents' utility U. However, agents are assumed to act in a boundedly rational way, such that they may not always choose the option with the highest U, for example, if it becomes impractical to consider all outcomes. That is, agents are attempting to maximise their expected utility, subject to an entropy constraint capturing the uncertainty:

$$\max \sum_{a \in A} f[a|x] U[a,x] \tag{1}$$

$$\text{subject to} \sum_{a \in A} f[a|x] = 1$$
$$- \sum_{a \in A} f[a|x] \log f[a|x] \geq H_{min} \tag{2}$$

where $f[a|x]$ represents the probability of an agent choosing action a if rate x is observed. The first constraint ensures the probabilities sum to 1, while the second is a constraint on the minimum entropy. The minimum entropy constraint implies a level of boundedness such that there is some limit to the agents' processing abilities, which allows QRSE to deviate from perfect rationality.

Lagrange multipliers can be used to turn the constrained optimization problem of Equation (2) into an unconstrained one, which forms the following Lagrangian function:

$$\mathcal{L} = -\sum_{a \in A} f[a|x] U[a,x] - \lambda \left(\sum_{a \in A} f[a|x] - 1 \right) + T \left(- \sum_{a \in A} f[a|x] \log f[a|x] - H_{min} \right) \tag{3}$$

taking the first order conditions of Equation (3), and solving for $f[a|x]$ yields:

$$f[a|x] = \frac{1}{Z} e^{\frac{U[a,x]}{T}} \tag{4}$$

representing a choice of a mixed strategy by maximising the expected utility subject to an entropy constraint. This problem is dual to maximising entropy of the mixed strategy, subject to a constraint on the expected utility as detailed in Appendix A.1.

3.1.2. Deriving Statistical Equilibrium

From Section 3.1.1 we have a derivation for a decision function, where agents maximise expected utility subject to an entropy constraint introducing bounds in the agents processing abilities. In order to infer the statistical equilibrium based on observed macroeconomic outcomes, the joint probability $f[a,x]$ must be computed.

The joint distribution captures the resulting statistical equilibrium which arises from the individual agent decisions. While there are many potential joint distributions, using the principle of maximum entropy allows for inference of the least biased distribution. From an observer perspective, maximising the entropy of the model accounts for model uncertainty, by providing the maximally noncommittal joint distribution. To compute this, Scharfenaker and Foley [5] maximise the joint entropy with respect to the marginal probabilities (since individual action data is not available), by decomposing the joint entropy into a sum of the marginal entropy and the (average) conditional entropy.

The solution for $f[a|x]$, given by Equation (4), can be used to compute the joint probability $f[a,x]$, as long as marginal $f[x]$ is determined (since $f[a,x] = f[a|x]f[x]$). In order to derive $f[x]$, the approach considers the state dependant conditional entropy, represented as

$$H[A|x] = -\sum_{a \in A} f[a|x] \log f[a|x] \tag{5}$$

Scharfenaker and Foley [5] then use the principle of maximum entropy to find the distribution of $f[x]$ which maximises

$$\max_{f[x] \geq 0} H = -\int_x f[x] \log f[x] dx + \int_x f[x] H[A|x] dx \tag{6}$$

$$\text{subject to} \int_x f[x] dx = 1$$
$$\int_x f[x] x dx = \xi \tag{7}$$

The first constraint ensures the probabilities sum to 1, and the second constraint applies to the mean outcome (with ξ being the mean from the actual observed data $\bar{f}[x]$). Importantly, there is also an additional constraint which models Smithian competition [32] in the market. Smithian competition models the feedback structure for competitive markets, for example, entrance into a market tends to lower the profit rates, and exit tends to raise the profit rates. This is captured as the difference between the expected returns conditioned on entrance, and the expected returns conditioned on exiting. This competition constraint can be represented as

$$\text{subject to } \int_x f[x](f[a|x] - f[\bar{a}|x])x dx = \delta \tag{8}$$

The combination of the conditional probabilities of Equation (4), which stipulate that the agents enter and exit based on profit rates, and the competition constraint of Equation (8) models a negative feedback loop that results in a distribution of the profit rates around an average (ξ).

Again, using the method of Lagrange multipliers, the associated Lagrangian becomes

$$\begin{aligned} \mathcal{L} = -\int_x f[x] \log f[x] dx + \int_x f[x] H[A|x] dx - \\ \lambda\left(\int_x f[x] dx - 1\right) - \gamma\left(\int_x f[x] x dx - \xi\right) - \rho\left(\int_x f[x](f[a|x] - f[\bar{a}|x]) x dx - \delta\right) \end{aligned} \tag{9}$$

where taking the first order conditions of Equation (9), and solving for $f[x]$ yields

$$f[x] = \frac{1}{Z_A} e^{H[A|x] - \gamma x - \rho x (f[a|x] - f[\bar{a}|x])} \tag{10}$$

where Z_A is the partition function $Z_A = \int_x e^{H[A|x] - \gamma x - \rho x (f[a|x] - f[\bar{a}|x])} dx$. Note that in Equation (9) we use ρ as the Lagrangian multiplier for the competition constraint. Parameter ρ is referred to as β in [5], we have avoided this notation to avoid confusion with the thermodynamic β (inverse temperature) discussed in later sections.

Equations (4) and (10) comprise a fully defined joint probability. Crucially, QRSE allows for modelling the resultant statistical equilibrium even when the individual actions are unobserved—by inferring these decisions based on the principle of maximum entropy.

3.1.3. Limitations of Logit Response

In Section 3.1.1 we have seen how the logit response function used for decision-making in QRSE is derived from entropy maximisation. Following the Boltzmann distribution well known in thermodynamics, this logit response has seen extensive use throughout the literature arising in a variety of domains. For example, the logit function is used as sigmoid or softmax in neural networks, logistic regression, and in many applications in economics and game theory [33,34]. However, one important development not yet discussed is the incorporation of prior knowledge into the formation of beliefs. Up until now, we have considered a choice to be the result of expected utility maximisation based on entropy constraints from which the logit models have arisen. However, from psychology [35], behavioural economics [36,37], and Bayesian methods [38,39] we know that the incorporation of a priori information is often an important factor in decision-making. Thus, we explore the incorporation of prior beliefs into agent decisions in more detail in the following section (and the remainder of the paper).

Furthermore, one criticism of the logit response arises from the independence of irrelevant alternatives (IIA) property of multinomial logit models (which would extend to the conditional function used in QRSE in a multi-action case), which states that the ratio between two choice probabilities should not change based on a third irrelevant alternative. Initially, this may seem desirable, however, this can become problematic for correlated outcomes (of which many real examples possess). This criticism has been proved correct in several thought experiment studies, showing violations of the IIA assumption [40]. The classical example is the Red Bus/Blue Bus problem [41,42].

Consider a decision-maker who must choose between a car and a (blue) bus, $A = \{\text{car}, \text{blue bus}\}$. The agent is indifferent to taking the car or bus, i.e., $p(\text{car}) = p(\text{blue bus}) = 0.5$. However, suppose a third option is added, a red bus which is equivalent to the blue bus (in all but colour). The agent is indifferent to the colour of the bus, so when faced with $A_1 = \{\text{blue bus}, \text{red bus}\}$ the agent would choose $p(\text{red bus}) = p(\text{blue bus}) = 0.5$. Now suppose the agent is faced with a choice between $A_2 = \{\text{car}, \text{blue bus}, \text{red bus}\}$. As per the IIA property, the ratio $\frac{p(\text{blue bus})}{p(\text{car})}$ (from A, $\frac{0.5}{0.5}$) must remain constant. So adding in a third option, the probability of taking any a becomes $p(a) = \frac{1}{3}$ (for all a), maintaining $\frac{p(\text{blue bus})}{p(\text{car})} = 1$. However, this has reduced the odds of taking the car from 0.5 to 0.33 based on the addition of an irrelevant alternative (i.e., the red bus in which the agent does not care about colour of the bus). In reality, the probability for taking the car should have stayed fixed at $p(\text{car}) = 0.5$, and the probability of taking a bus reduced to 0.25 each. This reduction in the probability of $p(\text{car})$ does not make sense for a decision-maker who is indifferent to the colour of the bus and is the basis for the criticism. This may not be immediately relevant for current QRSE models (especially binary ones), but with potential future applications, for example, in portfolio allocation, this could become an important consideration. For example, if adding an additional stock to a portfolio which is similar to an existing stock, it may not be desriable to reduce the likelihood of selecting other (unrelated) stocks.

3.2. Thermodynamics of Decision-Making

A thermodynamically inspired model of decision-making which explicitly considers information costs, as well as the incorporation of prior knowledge, is proposed by [4]. The proposed approach can be seen as a generalisation of the logit function, where the typical logit function can be recovered as a special case, but in the more general case manages to avoid the IIA property.

Ortega and Braun [4] represent changing probabilistic states as isothermal transformations. Given some initial state $x \in X$ with initial energy potential $\phi_0[x]$, the probability of being in state x is $p[x] = \frac{e^{-\beta\phi_0[x]}}{\sum_{x' \in X} e^{-\beta\phi_0[x']}}$ (from the Boltzmann distribution). Updating state to $f[x]$ corresponds to adding new potential $\Delta\phi_0[x]$. The transformation requires physical work, given by the free-energy difference $\Delta F[f]$. The free energy difference between the initial and resulting state is then

$$\begin{aligned} \Delta F[f] &= F[f] - F[p] \\ &= \sum_{x \in X} f[x]\Delta\phi(x) + \frac{1}{\beta} \sum_{x \in X} f[x] \log\left(\frac{f[x]}{p[x]}\right) \end{aligned} \tag{11}$$

which allows the separation of the prior $p[x]$ and the new potential $\Delta\phi_0[x]$. In economic sense, representing the negative of the new potential as the utility gain, i.e., $U(x) = -\Delta\phi_0[x]$, allows for reasoning about utility maximisation subject to an informational constraint, given here as the Kullback-Leibler (KL) divergence from the prior distribution [4]. Golan [43] shows how the KL-divergence naturally arises as a generalisation of Shannon entropy (of Equation (2)) when considering prior information, and Hafner et al. [44] show how various objective functions can be seen as functionally equivalent to minimising a (joint) KL-divergence, even those not directly motivated by the free energy principle. Such analysis makes the KL-divergence a logical and fundamentally grounded measure of information acquisition costs, captured as the divergence from a prior distribution.

Ortega and Stocker [45] then apply this formulation to discrete choice by introducing a choice set A (space of actions), which leads to the following negative free energy difference, for a given observation x:

$$-\Delta F[f[a|x]] = \sum_{a \in A} f[a|x] U[a,x] - \frac{1}{\beta} \sum_{a \in A} f[a|x] \log\left(\frac{f[a|x]}{p[a]}\right) \quad (12)$$

where again a represents a choice (or action), and U the utility for the agent. The first term of Equation (12) is maximising the expected utility, and the second term is a regularisation on the cost of information acquisition. Again, in this representation, information cost is measured as the KL-divergence from the prior distribution.

Taking the first order conditions of Equation (12) and solving for $f[a|x]$ yields

$$f[a|x] = \frac{p[a] e^{\frac{U[x,a]}{T}}}{\sum_{a' \in A} p[a'] e^{\frac{U[a',x]}{T}}} \quad (13)$$

where we have moved from inverse temperature β to temperature T for notational convenience, i.e., $T = \frac{1}{\beta}$. The key formulation here is the separation of the prior probability p from the utility gain (or the new potential from the initial potential). T then arises as the Lagrange multiplier for the cost of information acquisition (as opposed to the entropy constraint of QRSE, described in Section 3.1). We emphasise this aspect in later sections.

Revisiting the IIA property, the incorporation of the prior probabilities in Equation (A7) can adjust the choices away from the logit equation, and thus managing to avoid IIA. However, if desired, the free energy model reverts to the typical logit function in the case of uniform priors, and so this property can be recovered. In economic literature, a similar model is given by Rational Inattention (R.I.) by [2]. The relationship between R.I. and the free energy approach of [4,45] is detailed in Appendix C.

4. Model

In this section, we propose an information-theoretic model of decision-making with prior beliefs in the presence of Smithian competition and market feedback. Given an agent's prior beliefs and an observed macroeconomic outcome (such as the distribution of returns), the model can infer the least biased decisions that would result in such returns. Importantly, the incorporation of prior beliefs allows for reasoning about the decision-making of the agent based upon both their prior beliefs and their utility maximisation behaviour.

We develop upon the maximum-entropy model of inference from [5], and the thermodynamic treatment of prior beliefs formalised by [4], as outlined in Section 3.

4.1. Maximum Entropy Component

The proposed approach can be seen as a generalisation of QRSE, allowing for the incorporation of heterogeneous prior beliefs based on the free-energy principle. The key element is the information acquisition cost, measured as the KL-divergence which arises from the free-energy principle and has been shown to provide a fundamentally grounded application of Bayesian inference [46]. In order to derive decisions $f[a|x]$ for an action or choice a (e.g., buy, hold or sell) given an observed return x (e.g., a return on investment), we maximise the expected utility U subject to a constraint on the acquisition of information measured as the maximal divergence d between the posterior decisions and prior beliefs

$p[a]$. As mentioned, d is measured as the KL-divergence, which is the generalised extension of the original (Shannon) entropy constraint [43] introduced in Equation (2)):

$$\begin{aligned} &\max \sum_{a \in A} f[a|x] U[a,x] \\ \text{subject to } &\sum_{a \in A} f[a|x] \log\left(\frac{f[a|x]}{p[a]}\right) \leq d \\ &\sum_{a \in A} f[a|x] = 1 \end{aligned} \tag{14}$$

The Lagrangian for Equation (14) then becomes

$$\mathcal{L} = \sum_{a \in A} f[a|x] U[a,x] - \lambda \left(\sum_{a \in A} f[a|x] - 1 \right) - T \left(\sum_{a \in A} f[a|x] \log\left(\frac{f[a|x]}{p[a]}\right) - d \right) \tag{15}$$

There are two distinct modelling views on such a formulation [47–50]. The first assumes that specific constraints are known from the data, for example, a maximal divergence d may be specified based on actual observations of agent behaviour. The second view, instead, would consider the Lagrange multiplier T to be a free parameter of the model, with the constraint d representing an arbitrary maximum value: Thus, this approach would optimise T in finding the best fit. In this work, we take the second perspective since underlying decision data is unavailable, and a specific restriction on divergent information costs should not be enforced. In other words, T is considered to be a free model parameter corresponding to different information acquisition costs, mapping to different (unknown) cognitive and information-processing limits d.

Looking at the final term in Equation (15), in the case of homogeneous priors, $\log p[a]$ is a constant which drops out of the solution, which is equivalent to the optimisation problem of Equation (3), and thus, recovers the original QRSE model. In the general case, the dependence on $\log(p[a])$ means that T instead serves as the Lagrange multiplier for the cost of information acquisition. Taking the first order conditions of Equation (15) and solving for $f[a|x]$ (as shown in Appendix A.2) yields

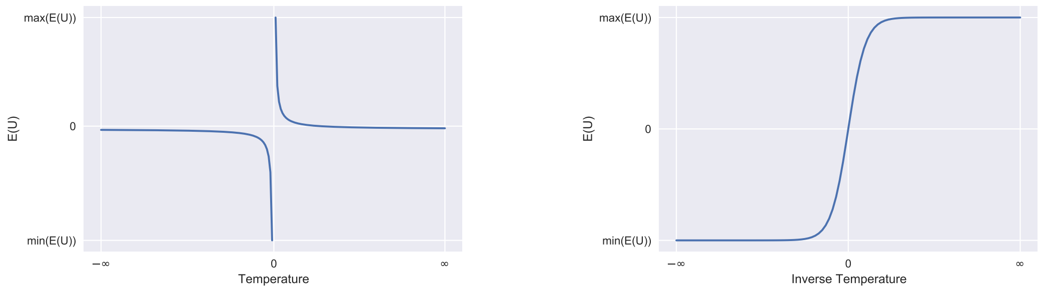
$$f[a|x] = \frac{1}{Z_{A|x}} p[a] e^{\frac{U[a,x]}{T}} \tag{16}$$

we see this as a generalisation of the logit function, which allows for the separation of the prior beliefs and the agent's utility function.

In the more general case, $p[a]$ can be heterogeneous for all a. Parameter T therefore controls the deviations from the prior (rather than from the base case of uniformity), that is, it controls the cost of information acquisition. Following [4], we observe the following limits

$$\begin{aligned} \lim_{T \to \infty} f[a|x] &= p[a] \\ \lim_{T \to 0, T \geq 0} f[a|x] &= e^{\frac{U[x,a]}{T}} = \max U[x,a] \\ \lim_{T \to 0, T < 0} f[a|x] &= e^{\frac{U[x,a]}{T}} = \min U[x,a] \end{aligned} \tag{17}$$

In the limit $T \to \infty$ (i.e., infinite information acquisition costs), the agent just falls back to their prior beliefs as it becomes impossible to obtain new information. In the limit $T \to 0$, the agent becomes a perfect utility maximiser (i.e., if information is free to obtain, the agent could obtain it all and choose the option that best maximises payoff with probability 1). In the $T < 0$ case, we see this corresponds to anti-rationality. For economic decision-making, we can limit temperatures to be non-negative, $T \geq 0$, although there are specific cases where such anti-rationality may be useful (e.g., modelling a pessimistic observer or adversarial environments [4]). The relationship between temperature and utility is visualised in Figure 1.



(**a**) Temperate vs. Utility (**b**) Inverse Temperate ($\frac{1}{T}$) vs. Utility

Figure 1. The effect of decision temperature T on the resulting expected payoffs (**a**). for the limits given by Equation (17). The inverse temperature $\frac{1}{T}$ (**b**) conveys the same information but may offer a more useful visualisation due to the continuity.

Crucially, large temperatures (costly acquisition) do not revert to the uniform distribution (as in the typical QRSE case, unless the prior is uniform), instead reverting to prior beliefs. This is visualised in Figure 2, and discussed in more detail in Section 4.3.

4.2. Feedback Between Observed Outcomes and Actions

Following [5], we use a joint distribution to model the interaction between the economic outcome x, and the action of agents a.

To recover a joint probability, we need to determine $f[x]$ (since $f[a,x] = f[a|x]f[x]$) which we do with the maximum entropy principle, as shown in Section 3.1. To do this, we maximise the joint entropy with respect to the marginal probabilities. That is,

$$\begin{aligned}\mathcal{L} = &-\int_x f[x]\log f[x]dx + \int_x f[x]H[A|x]dx - \lambda\left(\int_x f[x]dx - 1\right)\\ &-\gamma\left(\int_x f[x]xdx - \xi\right) - \rho\left(\int_x f[x]\,\frac{p[a]e^{\frac{U[a,x]}{T}} - p[\bar{a}]e^{\frac{U[\bar{a},x]}{T}}}{Z_{A|x}}xdx - \delta\right)\end{aligned} \tag{18}$$

with

$$\begin{aligned}H[A|x] &= -\sum_{a\in A} f[a|x]\log f[a|x]\\ &= -\frac{1}{Z_{A|x}}\sum_{a\in A} p[a]e^{\frac{U[a,x]}{T}}\left(\log p[a] + \frac{U[a,x]}{T} - \log Z_{A|x}\right)\end{aligned} \tag{19}$$

An important point to be made here is that $H[A|x]$ still measures (Shannon) entropy. We have seen above how the new definition for $f[a|x]$ uses the KL-divergence as a generalised extension of entropy when incorporating prior information. In Equation (19), we do not use this divergence for an important reason. In Equation (14) we are measuring divergence from known prior beliefs, however, now when optimising Equation (18) we wish to infer decisions from unobserved decision data. This is where the principle of maximum entropy comes into play, i.e., we wish to maximise the entropy of our new choice data (which was derived from KL-divergence of prior beliefs), but we do not wish to perform cross-entropy minimisation as we do not have the true decisions $\bar{f}[a|x]$. With this in mind, we still utilise the principle of maximum entropy as is done in QRSE for inference to obtain the least biased resulting decisions. This keeps the proposed extensions in the realm of QRSE, but comparisons to the principle of minimum cross-entropy [51,52] could be considered in future work particularly when some target distributions are known directly.

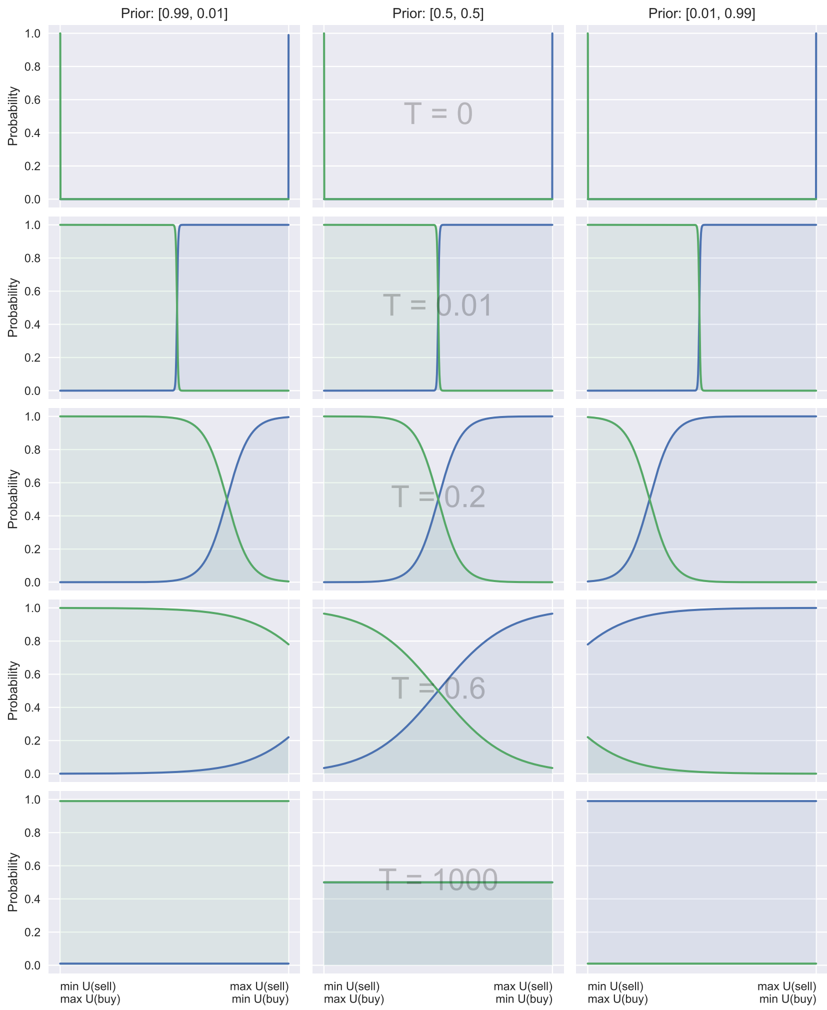


Figure 2. Decision Functions. All cases have equivalent utility functions. Each row has equivalent temperatures, showing how with matched parameters and utility, having an alternate prior can shift the decision-makers preference. Each column has different priors, given along the top of the first row to show how decision-makers decisions change based on their prior beliefs. On the left-hand side, preference is shifted towards the buying case. Likewise, on the right-hand side, preference is given to the selling case. The uniform case with equal preference is shown in the middle.

In Equation (18), ξ is known from the mean of the observed macroeconomic outcome, and so this constraint is used explicitly. This is in contrast to d (and δ) which are unknown as outlined in Section 4.1. The important distinction with Equation (18) is that the $f[a|x]$ functions (and $H[A|x]$) now use the updated expressions for $f[a|x]$, which incorporate the prior beliefs. Taking the partial derivative of $\mathcal{L}$ with respect to $f[x]$, and solving for $f[x]$ gives

$$f[x] = \frac{1}{Z_A} e^{H[A|x] - \gamma x - \rho x \left(\frac{p[a] e^{\frac{U[a,x]}{T}} - p[\bar{a}] e^{\frac{U[\bar{a},x]}{T}}}{Z_{A|x}} \right)} \tag{20}$$

Equation (20) expresses the information acquisition cost in the form of the Lagrange multiplier T (from Equation (15)), and a competition cost in the form of the multiplier ρ.

As we have a solution for $f[a|x]$ (Equation (16)) and $f[x]$ (Equation (20)) in terms of prior beliefs and information acquisition costs, we can then derive all other probability functions using the Bayes rule. That is, we can obtain $f[a,x]$, $f[x|a]$ and $f[a]$ which in turn incorporate these prior beliefs/acquisition costs:

$$\begin{aligned} f[a,x] &= f[a|x]f[x] \\ &= \frac{p[a]e^{\frac{U[a,x]}{T}+H[A|x]-\gamma x-\rho x\left(\frac{p[a]e^{\frac{U[a,x]}{T}}-p[\bar{a}]e^{\frac{U[\bar{a},x]}{T}}}{Z_{A|x}}\right)}}{Z_{A|x}Z_A} \end{aligned} \tag{21}$$

We can obtain $f[a]$ by marginalising out x from the joint distribution:

$$\begin{aligned} f[a] &= \int_x f[a,x] \\ &= \frac{1}{Z_A}\int_x \frac{1}{Z_{A|x}}p[a]e^{\frac{U[a,x]}{T}+H[A|x]-\gamma x-\rho x\left(\frac{p[a]e^{\frac{U[a,x]}{T}}-p[\bar{a}]e^{\frac{U[\bar{a},x]}{T}}}{Z_{A|x}}\right)} \end{aligned} \tag{22}$$

Finally, $f[x|a]$ can then be computed by a direct application of the Bayes rule: $f[x|a] = f[a,x]/f[a]$.

Given only an expected average value ξ (and the usual normalisation constraints), we have derived a joint probability distribution, which maximises the entropy subject to some information acquisition cost d, along with a competition cost δ. The resulting distribution free parameters (the Lagrange multipliers) are those which fit most closely to the true underlying distribution of returns. Thus, we have provided a generalisation of QRSE, which is fully compatible with the incorporation of prior beliefs.

4.3. Priors and Decisions

The introduced priors affect the conditional probabilities of agent decisions by shifting focus towards these preferred choices. The introduced priors allow the decision-maker to place more focus on particular actions if they have been deemed important a priori.

In Section 3.2 we showed how to separate the initial energy potential and new energy potential for distinguishing prior beliefs and utility functions. It is instructive to interpret these again as potentials, by setting $\alpha_a = T \log p[a]$, which allows us to represent the choice probability as

$$f[a|x] = \frac{1}{Z_{A|x}}e^{\frac{U[x,a]+\alpha_a}{T}}\,. \tag{23}$$

Equation (23) shows how α shifts the likelihood based on the prior preferences. An example of these shifts is visualised in Figure 2. This can be interpreted as placing more emphasis on actions deemed useful a priori as T increases. The information acquisition cost component T then controls the sensitivity between the utility and a priori knowledge, with a high T meaning higher dependence on prior information, and low T indicating a stronger focus on the utility alone.

The majority of binary QRSE models use a simple linear payoff definition for utility:

$$U[x,a] = x - \mu, \;\; U[x,\bar{a}] = -(x-\mu)\,.$$

With this definition, a tunable shift parameter μ serves as the expected fundamental rate of return. The relationship between μ and the real markets returns ξ (which was used as a constraint in Equation (7)), serves then as a measure of fulfilled expectations (i.e., if $\mu = \xi$) or unfulfilled expectations ($\mu \neq \xi$). This implies a symmetric shift parameter μ. As a specific example, if $a =$ sell and $\bar{a} =$ buy, $\mu = 0.25$ means that at $x = 0.25$, buyers and sellers will be equally likely to participate in the market, i.e., $f[\text{sell}|\mu] = f[\text{buy}|\mu] = 0.5$. In this sense, μ can be seen as the indifference point. The symmetry arises from the fact that $f[\text{buy}|x] + f[\text{sell}|x] = 1$. Therefore, in the binary action case, it is possible to find a μ^* with the uniform priors $p = [0.5, 0.5]$ such that the decision functions will be equivalent to μ with any arbitrary priors $p = [c, 1-c]$, with $c \in [0,1]$. In this sense, μ can be seen as encapsulating a prior belief.

However, explicit incorporation of prior beliefs on actions is useful here as it helps to separate the agents' expectations in relation to their prior belief (e.g., a higher μ resulted from needing to change from their past behaviour) and choose the actions for which an agent should emphasise acquiring more information. The introduced prior beliefs are strictly known before any inference is performed, whereas μ is the result of the inference process. The separation of prior beliefs and current expectations is important, as with μ alone this can not capture an agent's predisposition prior to performing any information processing. In addition, this applies more generally to any arbitrary utility functions (as QRSE is, of course, not limited to the linear shift utility function with μ outlined above), or when any preference is known about decisions a priori.

Consider also the three action case, $A = \{\text{buy, hold, sell}\}$, with the same utility functions as above but with the extra utility for holding being $U[x, \text{hold}] = 0$. We can see that it would be desirable if buying and selling no longer required this symmetry. The use of priors can introduce this asymmetry, by providing separate indifference points for buy/hold and sell/hold. Such asymmetry alters the resulting frequency distribution of transactions, and may help to explain various trading patterns [16]. The difference of symmetric and asymmetric buy and sell curves is shown in Figure 3. Figure 3 shows that such functions could be recovered by introducing a secondary shift parameter μ_2. Parameter μ_1 (the original μ) then becomes the indifference point for buy and hold, and μ_2 for sell and hold. This is the method proposed in [30]. Introducing priors into this case again allows for separation of expectation μ, from prior belief and follows the same methodology as outlined above for the binary case. Furthermore, if we set $p[\text{hold}] = 0$, we recover the binary case. This highlights that the standard QRSE with binary actions and uniform priors is a special case of the ternary action case with heterogeneous priors.

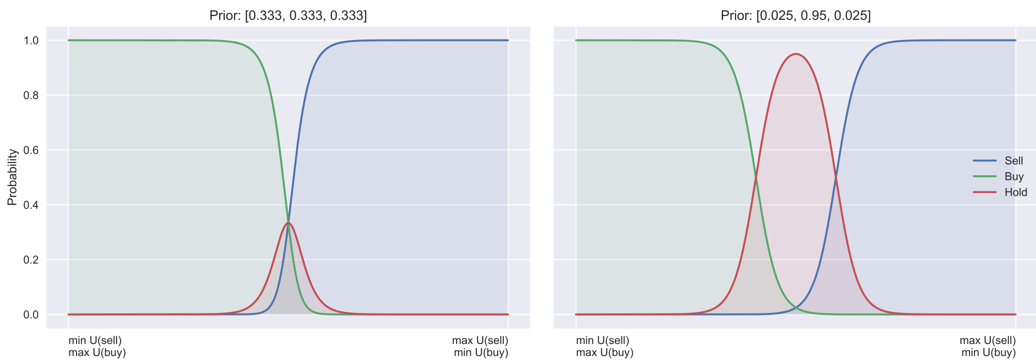


Figure 3. In the three-action case, the priors can introduce asymmetries by biasing the decision functions. This allows for separate indifferent points (**right**) vs. the uniform priors implying a single intersect (**left**).

From this, we can see how introducing priors alters the decision functions by allowing agents to focus on suitable a priori candidate actions. We have also shown how the binary case of a utility function with a shift parameter can be formalised to achieve equivalent results with a uniform prior and altered shift parameter. However, in the multi-action case, the priors allow for asymmetry, and in general, the priors may help with the optimisation process (by providing an alternate initial configuration). This approach also allows for the explicit separation of the two factors affecting an agent's choice, by distinguishing the contributions of prior beliefs and the utility maximisation.

4.4. Rolling Prior Beliefs

The proposed extension is general and allows for the incorporation of any form of prior beliefs, and in this section, we illustrate an example where the priors at time t are set as the resulting marginal probabilities from the previous time $t-1$:

$$p_t[a] = f_{t-1}[a]$$

i.e., the prior belief $p_t[a]$ is set as the previous marginal probability $f_{t-1}[a]$ for taking action a (at $t = 0$, we use a uniform prior). Using the previous marginal probability as a prior introduces an "information-switching" cost, where T relates to the divergence from the previous actions, resulting in the following decision function:

$$f_t[a|x] = \frac{1}{Z_{A|x}} f_{t-1}[a] e^{\frac{U[x,a]}{T}}$$

That is, acquiring information on top of the previous knowledge comes at a cost (controlled by T). When the cost of information acquisition is high (large T), the agent reverts to the previously learnt knowledge (i.e., the marginal probabilities from $t-1$). In contrast, when T is extremely small, the agent is able to acquire new information allowing deviation from their prior knowledge at $t-1$. In the special case of $T = 0$, information is free, and the agent can become a perfect utility maximiser.

Given the expression for $f_t[a|x]$, we obtain the following solution for $f_t[x]$:

$$f_t[x] = \frac{1}{Z_A} e^{H[A|x] - \gamma x - \rho x \left(\frac{f_{t-1}[a] e^{\frac{U[a,x]}{T}} - f_{t-1}[\bar{a}] e^{\frac{U[\bar{a},x]}{T}}}{Z_{A|x}} \right)}$$

from which we can derive the joint and other probabilities, as shown in Section 4.1. This is exemplified in Section 5, in which we examine various priors for time-dependent applications.

5. Australian Housing Market

To exemplify the model, we use the Greater Sydney house price dataset provided by SIRCA-CoreLogic and utilised in [53,54]. This dataset is outlined in Appendix B. In [54], an agent-based model is used to explain and forecast house price trends and movement patterns as arising from the individual agent's buy and sell decisions. Furthermore, the ABM implemented bounded rational agents driven by social influences (e.g., fear of missing out) and partial information about submarkets. While the resulting dynamics produced by the ABM accurately match the actual price trends, the decision-making mechanism and the bounded rationality of the agents were not theoretically grounded. In the following section, we aim to explain how the bounded rational behaviour of the agents operating in the housing market can be aligned with the model proposed in this study based on prior beliefs of agents and Smithian competition within the market. With this example, Smithian competition can be seen as agent decisions (buying or selling) affecting returns for an area, and agents decisions also being made based on returns for particular areas, i.e., a feedback loop is assumed in the market.

In particular, we want to explore what role an agent's prior beliefs play in their resulting decisions. For example, given equivalent configurations (e.g., utility and returns) and different prior knowledge, how would the agent's behaviour differ? Furthermore, we would like to explore the rationality of the agents, measured in terms of the cost of information acquisition, in order to see how the agents behave. For example, are agents predominantly reliant on past knowledge in times of market growth, resulting in unexpected downturns from mismanaged agent expectations? Alternatively, in deciding if it is a good time to buy or sell, the agents may balance their past knowledge with utility and current returns (i.e., the past knowledge would not be a predominant factor). The proposed model is particularly suited for answering such questions due to the low number of free (and microeconomically) interpretable parameters, as well as the explicit separation of prior beliefs (as opposed to previous QRSE approaches). Our goal is not to infer the "best" prior, but rather to explore and compare dynamics resulting from various priors. In addition, we aim to verify the conjecture that during crises, and periods exhibiting non-linear market dynamics, macroeconomic conditions may become more heterogeneous, and thus, non-uniform priors may outperform uniform ones in such times.

5.1. Model

We use our model of binary actions with prior beliefs introduced in Section 4.1, with actions $A = \{\text{buy}, \text{sell}\}$. The decision functions are then given by

$$\begin{aligned} f_t[\text{buy}|x] &= \frac{1}{Z_{t,x}} p_t[\text{buy}] e^{\frac{U[x,\text{buy}]}{T}} \\ f_t[\text{sell}|x] &= \frac{1}{Z_{t,x}} p_t[\text{sell}] e^{\frac{U[x,\text{sell}]}{T}} \\ Z_{t,x} &= f_t[\text{buy}|x] + f_t[\text{sell}|x] \end{aligned} \tag{24}$$

where we explore a range of p_t (prior at time t) functions, discussing their effects on decision-making and resulting probability distributions.

5.1.1. Priors

While the proposed approach is capable of incorporating any form of prior belief on the choice set A, below we outline several example priors which we explore. In exploring these priors, we highlight differences in resulting agent posterior decisions based on various prior beliefs.

Uniform

We begin with a uniform prior. The uniform probability represents the default case of QRSE, where each action has an equally weighted prior. In the binary case, this corresponds to $p_t[a] = 0.5$ for all t and a. This corresponds to an agent who is agnostic to the available actions before observing U.

Previous

Next we look at a "previous" prior. The previous prior uses the marginal action probabilities from the previous time step as priors to the current timestep. This means at time t, $f_t[a]$ plays the role of a posterior probability of making a decision, however, at time $t+1$ $f_t[a]$ now serves as the empirical prior. This is the example introduced in Section 4.4. This corresponds to $p_t[a] = f_{t-1}[a]$ for $t > 0$, and $p_t[a] = 0.5$ for $t = 0$. The previous prior represents an empirical prior where the decision is conditioned on previous market information, where T controls the level of influence from the previous market stage (in our case, each year). A high T means high influence from the past market state, whereas low T means focusing on current market conditions alone (as measured by U). In the extreme case of $T = \infty$, a backward looking expectations [55] approach is recovered where decisions are assumed to be a function purely of past decisions, however, in the more general case with $T < \infty$, U adjusts the decisions based on the current market state.

Mean

We also consider a mean prior. The mean prior uses the average marginal action probability from all previous timesteps. This corresponds to $p_t[a] = \frac{\sum_{t'=0}^{t-1} f_{t'}[a]}{t}$, for $t > 0$, and $p_t[a] = 0.5$ for $t = 0$. This can be seen as belief evolution, where over time, the previous decisions help build the current prior (modulated by T) at each stage.

Extreme Priors

As two further examples, we introduce extreme priors (more for visualisation/discussion sake as opposed to being particularly useful). The extreme buy prior corresponds to a strong prior preference for the buy action, $p_t[\text{buy}] = 0.99, p_t[\text{sell}] = 0.01$, for all t. Likewise, the extreme sell case is simply the inverse of the buy case, a strong prior preference for selling, i.e., $p_t[\text{sell}] = 0.99, p_t[\text{buy}] = 0.01$, for all t.

However, the formulations provided above by no means represent an exhaustive set of possible priors. For example, Genewein et al. [56] discuss "optimal" priors, which draws parallels with rate-distortion theory and can be seen as building abstractions of decisions (see Appendix C). Adaptive expectations [57] are discussed in [58–60], where priors could be partially adjusted based on some strength term (λ_E), where the strength term adjusts the contribution from some error. For example, an adaptive prior could be represented as $p_t = p_{t-1} + \lambda(p_{t-1} - \hat{p}_{t-1})$, where $\hat{p}_{t-1}$ is the actual known likelihood of actions from the previous time period. With our specific housing market data, we do not have $\hat{p}$, i.e., we do not have the true buying and selling likelihoods, but if known, such information could be used to adjust future beliefs, i.e., over time the adaptive priors would adjust decisions based on the previously observed likelihoods (controlled by λ). The proposed approach makes no assumption about the forms of prior beliefs, so the ideas outlined above can be incorporated into the method outlined here by adjusting the definition of p_t.

5.2. Results

We fit the distributions with the various priors outlined in Section 5.1.1 to the actual underlying return data, to estimate how well we are able to capture this distribution and explore the effects that these priors have on the resulting distribution. The results are presented in Table 1, which summarises the likelihood and the percentage of the explained variability (measured as Information Distinguishability (I.D.) [61]) compared to the underlying distribution. We see that there are no large differences in general between the priors in terms of the explained variability. However, the goal here is not to argue for the "best" prior fitting the dataset in terms of the explained variability, but rather to explore differences in the agent behaviour based on the prior knowledge (using the housing dataset as an example). Thus, the resulting fitted distributions $f[x]$, which are visualised in Figure A5, are more interesting. We observe how altering prior beliefs result in different resulting distributions and discuss how the incorporation of prior beliefs allows for a separation of the agents' utility maximisation behaviour from their previous knowledge. From Figure A5 we can also see how the priors can alter the optimisation process, for example, a good (bad) prior may help (harm) the optimisation by providing alternate initial configurations. The extreme priors can be seen as harmful, for example, in 2012 where the resulting distributions are unable to capture the true underlying distribution. The reason for this is being unable to find suitable T to enable appropriate divergence from the extreme prior beliefs. In contrast, well selected priors can help the optimisation process and result in better fitting distributions, such as in 2016 where the decisions resulting from the mean and previous prior fit the true data significantly better than the uniform prior.

The agents' decision functions $f[a|x]$ are visualised in Figure A7 which makes it clear how each prior adjusts the resulting probability of taking an action (and thus, alters the decisions). From this, we can see different probabilistic behaviours despite having equivalent utility functions and optimisation processes due to varying prior beliefs. For example, with the extreme priors, we observe a clear shift towards the strongly preferred action.

Figure A6 shows the resulting joint distributions $f[a,x]$, combining the results of Figures A5 and A7, since $f[a,x] = f[a|x]f[x]$. Looking at the second row of each plot in Figure A6, we can see a visual representation of how the joint probabilities adjust over time when using the previous year as the prior belief.

Table 1. Resuling likelihood and percentage of variability explained for each year, when compared to the actual underlying distribution (i.e., those given in Figure A2). Optimisation is done by minimising the negative log-likelihood between the resulting distributions and the actual distribution of returns.

	Uniform	Previous	Mean	Extreme Buy	Extreme Sell
2006	1082 (93%)	1082 (93%)	1082 (93%)	885 (59%)	1005 (74%)
2007	1089 (92%)	1089 (92%)	1090 (90%)	939 (68%)	1042 (83%)
2008	998 (95%)	905 (78%)	998 (95%)	998 (95%)	998 (95%)
2009	918 (96%)	918 (96%)	866 (88%)	880 (85%)	875 (85%)
2010	857 (95%)	857 (95%)	857 (95%)	740 (62%)	857 (95%)
2011	1045 (92%)	1044 (91%)	1047 (92%)	1045 (91%)	873 (62%)
2012	1067 (96%)	1067 (96%)	1067 (96%)	162 (6%)	142 (8%)
2013	1080 (90%)	1076 (90%)	1083 (90%)	983 (77%)	1075 (91%)
2014	938 (98%)	851 (74%)	938 (98%)	875 (71%)	938 (98%)
2015	860 (96%)	860 (96%)	860 (96%)	33 (10%)	808 (71%)
2016	873 (84%)	932 (95%)	908 (86%)	817 (70%)	932 (95%)
2017	916 (97%)	916 (97%)	916 (97%)	812 (76%)	916 (97%)
2018	989 (88%)	932 (85%)	933 (85%)	955 (82%)	998 (91%)
2019	1101 (92%)	1103 (92%)	1067 (94%)	1101 (92%)	952 (76%)

The resulting marginal action probabilities are visualised in Figure 4, where we observe clear market peaks and dips which match the actual returns of Figure 5, aligning with the general trends observed in Figure A1. The priors work on either increasing or decreasing the resulting marginal probabilities. For example, in the extreme sell case we see much higher resulting probabilities for $f[\text{sell}]$, likewise in the extreme buying case, we see much higher probabilities for $f[\text{buy}]$. The general peaks/dips remain in both cases. Overall, this shows how the prior belief can influence the resulting marginal probabilities.

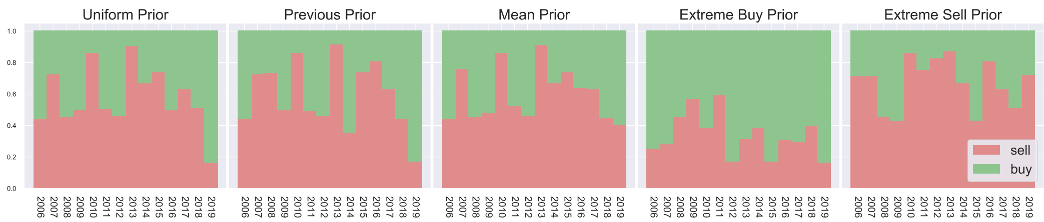


Figure 4. Resulting marginal probabilities $f[a]$ for varying priors. Green represents $f[\text{buy}]$, and red represents $f[\text{sell}]$.

Figure 5. Real Average Returns.

Using the previous year's marginal probability as a prior for the current year has a smoothing effect on the resulting year-to-year marginal probabilities. Comparing the previous prior with the uniform prior in Figure 4, we observe, particularly during 2015–2018, a more defined/well-behaved step-off in $f[\text{sell}]$. This indicates the slowing of returns during these years. At the same time, the uniform priors are more affected by local noise, potentially overfitting to only the current time period, since no consideration can be given to the past behaviour of the market. This results in larger fluctuations in the agent behaviour as they have no concept of market history.

5.3. Role of Parameters

One of the benefits of QRSE is the low number of free parameters which results in a relatively interpretable model. There are four free parameters in the typical QRSE distribution: T, μ, ρ and γ, each with a corresponding microeconomic foundation. In this section, we discuss the two main parameters of interest in this work: The decision temperature T and agent expectations μ, and the effect that prior beliefs have on the resulting values (and interpretation) of these parameters. We also include discussion on the impact of decisions on resulting outcomes ρ and skewness of the resulting distributions γ in Appendix D, since ρ and γ were less affected by the introduced extensions. There is an additional parameter ξ (shown in Figure 5), which is not a free parameter, representing the mean of the actual returns and serving as a constraint on the mean outcome in Equation (7).

5.3.1. Decision Temperature

The decision temperature T controls the level of rationality and deviations from an agent's prior beliefs. An extremely high temperature corresponds to high information acquisition cost and results in choosing actions simply based on the prior belief. In contrast, an extremely low temperature corresponds to utility maximisation, and in the case of free information ($T = 0$) a perfect utility maximiser is recovered (i.e., homo economicus). In the housing example used here, T relates to the ability of an agent to learn all the required knowledge of the market, i.e. the actual profit rates for various areas. With $T = 0$, the agent has perfect knowledge of the current market profitability. With $T > 0$, this represents some friction with acquiring such information, e.g., it can be difficult to gather all the required information to make an informed choice due to, for example, search costs. From a psychological perspective, T can be a measure of the "just-noticeable difference" [62], meaning microeconomically, T is related to the ability of an agent to observe quantitative differences in resulting choices. High T means the agent is unable to distinguish choices based on U, due to high information-processing costs, so instead acts according to their previously learnt knowledge.

Since T is related to the prior, we see differences in the resulting values visualised in Figure 6. What can be observed from looking at the general trends of T is that it peaks in the years with high average growth (large ξ), such as 2015, as these years correspond to a growing market, and agents require less attention to market conditions, although this depends on the prior used.

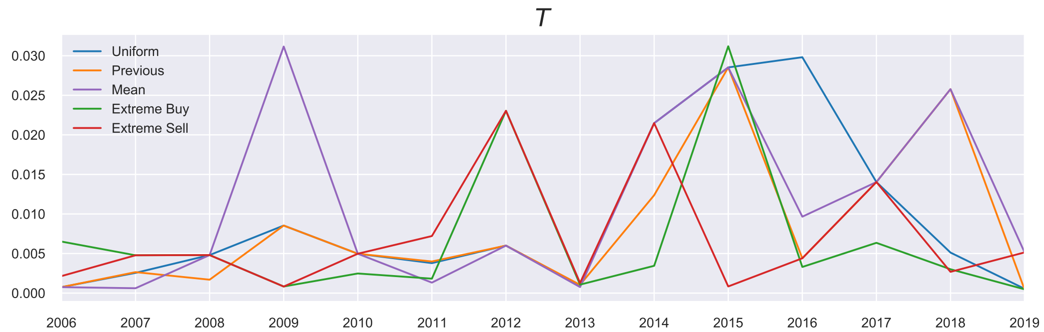


Figure 6. Decision Temperature.

Looking at the previous marginal probability as the prior (the orange profile), we observe in the build-up phase to 2015 increasing decision temperatures corresponding to agents acting on these previous beliefs. As these beliefs were also positive (i.e., agents expected favourable returns), these large returns can be explained by the agents continuously expecting this growth. This pattern changed in 2016, when the market "reverses": Now the agents must focus instead on their current utility since their prior beliefs no longer reflect the current market state. Such market reversals are categorised by low decision temperatures, since using the previous action probabilities now becomes misinformative (in contrast to the "building"/trend-following stages). This indicates an increased focus on agent rationality in times of market reversals. The incorporation of prior beliefs (particularly using the previous priors) is useful as it allows for the discussion to be extended in the temporal sense (as is done here). In other words, we can consider "building" the agent's beliefs as possible underlying causes for market collapses and relating the rationality of agents to the relative state of the market.

5.3.2. Agent Expectations

In microeconomic terms, parameter μ captures the agent's expectations. A large μ corresponds to an optimistic agent, who is expecting high returns from the market. In contrast, a low μ corresponds to a pessimistic agent, who is expecting poor returns from the market. As this works to shift the decision functions, there is a relation between the prior and parameter μ, since the prior also works as shifting preferences towards a priori preferred actions as shown in Section 4.3. There is also a relationship between μ and γ (outlined in Appendix D.2), since γ can help to account for unfulfilled agent expectations by adjusting the skew of the resulting distributions.

Generally, the agent's beliefs are within the $\pm 2.5\%$ range (expecting between a 2.5% quarterly growth or 2.5% dip), which corresponds to the bulk of the area under the curve in Figure A2. This means that the agent's expectations develop in accordance with actual market conditions, as can be seen in Figure 7.

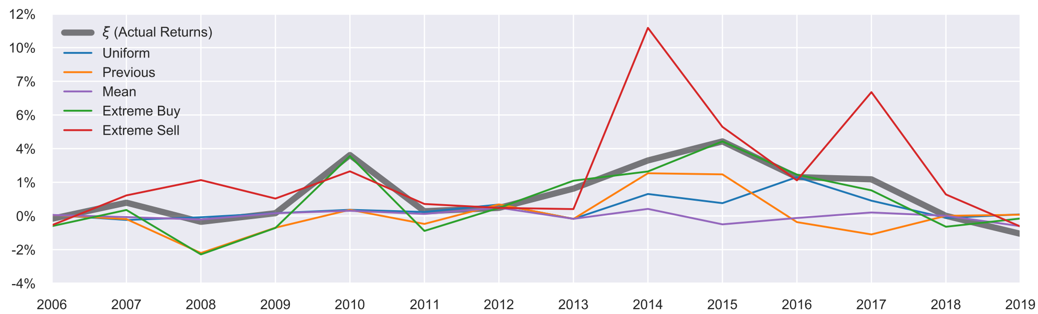


Figure 7. Agent Expectations vs. Actual Returns (in black).

The extreme priors result in larger absolute values of μ since larger shifts are needed to offset the (perhaps) poor prior beliefs. This can be seen in 2014 particularly, where the extreme sell prior has $\mu = 10\%$.

The values of previous prior μ tend to have a larger magnitude than the uniform priors, since as mentioned, these priors can capture build-up of beliefs (and as such some "trend-following" can be captured). For example, the year 2008 saw the lowest average returns ξ, as shown in Figure 5. Using the previous prior, the agents' expectations correctly match the sign of the actual returns in 2008 (i.e., agents correctly expected a decline in house prices). This results in more pessimistic agents than those using the uniform prior since they can reflect on the market performance from 2007. Likewise, during 2013–2015, the values of previous prior μ become larger than those for the uniform prior, since they are building on the previous years expectations which were all positive. In contrast, the period 2015–2017 saw a steady decline in agents expectations of returns with previous priors, reflecting the overall market state which appeared to be in a downward trend. The previous priors were able to capture this trend. Using the uniform priors, the year 2016 had a higher μ than the market peak of 2015. The reason is that uniform priors are unable to capture the fact that the previous timestep had higher (or lower) returns than the current timestep. In this case, the discussion can not be extended in the temporal sense of "building" on beliefs, and agents may miss such crucial temporal information without the incorporation of prior beliefs. This is evidenced by the significantly lower performance of the uniform prior in 2016 in comparison to the previous prior, as shown in Table 1, highlighting the usefulness of non-uniform (and temporal-based) priors in times of market crises and reversals.

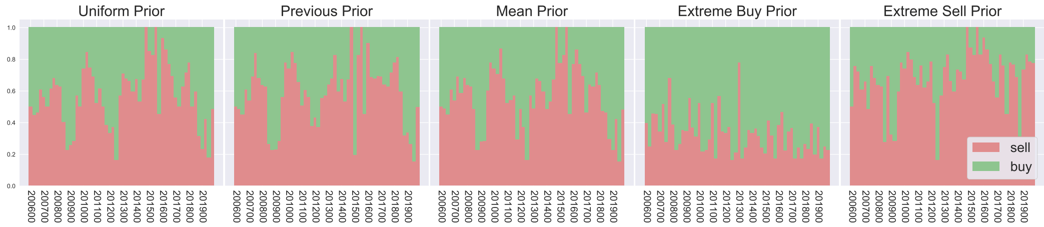
5.4. Temporal Effects of Data Granularity on Decisions

In Section 5.2, we have analysed agent decisions over the previous 15 years, where decisions were grouped annually. This level of granularity was chosen to examine different agent behaviour from year to year. However, other levels of grouping can also be explored to give an insight into the impact of noise on the inference process. For example, an extremely granular grouping will likely result in additional noise in the decision-making process, which may or may not be impacted by the incorporation of prior beliefs. Likewise, a low granular grouping can be seen as "pre-smoothed", which may work in a similar fashion to the incorporation of prior temporal-based beliefs at a higher granularity, which we have seen can smooth the resulting decisions. In this section, we examine the usefulness of prior beliefs in such situations, providing comparisons with alternate data representations.

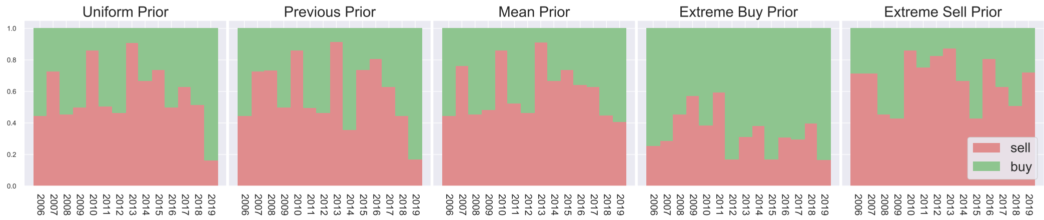
Two additional levels of granularity are considered, one more granular and one less granular than the annual groupings introduced in Section 5.2. We look at quarterly data, as well as aggregate groupings based on market state. In doing so, we have three levels for categorising agent behaviour: Quarterly, annually, and aggregated market state. This allows us to compare resulting agent decisions across different temporal scales, comparing the differences generated by the incorporation of prior beliefs and various data-level modifications.

The aggregate market state data groups years into "terms", which correspond with various "stages" of the market. These are growth and crash phases, highlighted as "Pre Crash" (Mid 2006–2007), "Crash" (2008), "Recovery 1" (2009–Mid 2011), "Small Crash" (Mid 2011–Mid 2012), "Recovery 2" (Mid 2012–Mid 2018) and "Recent Crash" (Mid 2018 to 2020). The overall market trends can be visualised in Figure A1 to see market returns for each corresponding "term".

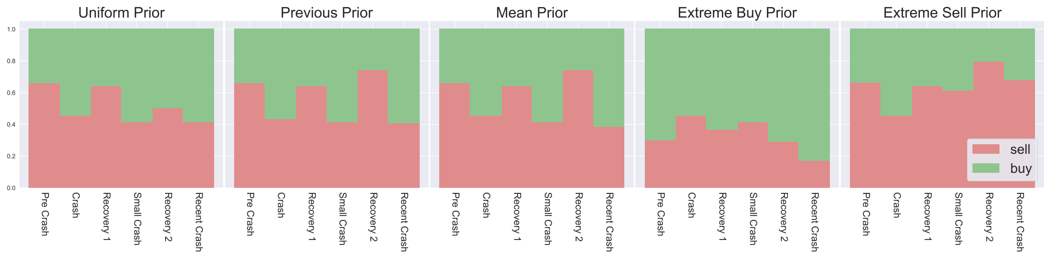
The resulting decision likelihoods $f[A]$ are presented in Figure 8. In analysing the differences in resulting marginal probabilities between the various granularities, we can observe the impact from data-level modifications, i.e., performing inference on a larger time scale for macroeconomic observations, and how the incorporation of prior information affects such results. In Section 5.2 we have mentioned the previous and mean priors can have a smoothing effect on resulting decisions, in this sense, the lower granularity groupings (the market state based grouping) can also be seen as a smoothed version of the macroeconomic outcomes, i.e. pre-smoothing the data by considering a much larger interval composed of several years for groupings. We see that the incorporation of prior information helps preserve some important information in such settings. Looking at the left-most column of Figure 8 (the uniform priors), we can see the overall "shape" of the peaks and dips in preferences $f[a]$ is lost with aggregate groupings. For example, in the quarterly breakdown, there is a clear preference for selling in the later region in the range 2014–2017, corresponding to the highest growing market, which is labelled as "Recovery 2" in the aggregated version. When considering the "Recovery 2" with uniform priors, such a clear preference is lost, and the "Pre Crash" and "Initial Recovery" have a higher corresponding preference. This is because the agents can not separate past market information from the current market state and act purely based on the current utility. In contrast, with both the mean and the previous prior, such overall trends are preserved across the various granularities since agents can distinguish favourable environments when compared with previous market states (as captured by their prior beliefs). This additional temporal insight provides an important consideration and shows that even with various data-level smoothing or preprocessing (i.e., considering alternate data groupings) the prior information remains useful and highlights various market states and corresponding agent preferences.



(**a**) Quarter



(**b**) Annual



(**c**) Terms

Figure 8. $f[a]$ for varying granularities.

A key takeaway from this exploration is that the potential for temporal analysis introduced by the prior beliefs provides additional insights into decision-making. These insights can not be generated by simple data-level modifications. Furthermore, the decision temperature T provides a way to modulate market state changes when considering agent decision-making.

6. Discussion and Conclusions

Despite many well-founded doubts of perfect rationality in decision-making, agents are often still modelled as perfect utility maximisers. In this paper, we proposed an approach for inference of agent choice based on prior beliefs and market feedback, in which agents may deviate from the assumption of perfect rationality.

The main contribution of this work is a theoretically grounded method for the incorporation of an agent's prior knowledge in the inference of agent decisions. This is achieved by extending a maximum entropy model of statistical equilibrium (specifically, Quantal Response Statistical Equilibrium, QRSE), and introducing bounds on the agent processing abilities, measured as the KL-divergence from their prior beliefs. The proposed model can be seen as a generalization of QRSE, where prior preferences across an action set do not necessarily have to be uniform. However, when uniform prior preferences are assumed, the typical QRSE model is recovered. The result is an approach that can successfully infer least biased agent choices, and produce a distribution of outcomes matching that of the actual observed macroeconomic outcomes when individual choice level data is unobserved.

In the proposed approach, the agent rationality can vary from acting purely on prior beliefs, to perfect utility maximisation behaviour, by altering the decision temperature. Low decision temperatures correspond to rational actors, while high decision temperatures represent a high cost of information acquisition and, thus, revert to prior beliefs. We showed how varying an agent's prior belief altered the resulting decisions and behaviour of agents, even those with equivalent utility functions. Importantly, the incorporation of prior beliefs into the decision-making framework allowed the separation of two key elements: The agent's utility maximisation, and the contribution of the agent's past beliefs. This separation allowed for a discussion on the decision-making process in a temporal sense, being able to refer to the previous decisions. This allows for investigation into the building of beliefs over time, elucidating resulting microeconomic foundations in terms of the underlying parameters.

It is worth pointing out some parallels with, and differences from, the frameworks of embodied intelligence and information-driven (guided) self-organisation, in which embodiment is seen as a fundamental principle for the organisation of biological and cognitive systems [63–66]. Similar to these approaches, we consider information-processing as a dynamic phenomenon and treat information as a quantity that flows between the agent and its environment. As a result, an adaptive decision-making behaviour emerges from these interactions under some constraints. Maximisation of potential information flows is often proposed as a universal utility for such emergent agent behaviour, guiding and shaping relevant decisions and actions within the perception-action loops [67–70]. Importantly, these studies incorporate a trade-off between minimising generic and task-independent information-processing costs and maximising expected utility, following the tradition of information bottleneck [71].

In our approach, we instead consider specific information acquisition costs incurred when the agents need to update their relevant beliefs in the presence of (Smithian) competition and market feedback. The adopted thermodynamic treatment of decision-making allows us to interpret relevant economic parameters in physical terms, e.g., agent's decision temperature T, the strength of negative feedback ρ, and skewness of the resulting energy distribution γ. Interestingly, the decision temperature appears in our formalism as the Lagrange multiplier of the information cost incurred when switching posterior and prior beliefs (KL-divergence). The KL-divergence can be interpreted as the expected excess code-length that is needed if a non-optimal code that was optimal for the prior (outdated) belief is used instead of an optimal code based on the posterior (correct) belief. Thus, the decision temperature modulates the inference problem of determining the true distribution given new evidence, in a forward time direction [72]. Moreover, the thermodynamic time arrow (asymmetry) is maintained only when decision temperatures are non-zero.

We demonstrated the applicability of the method using actual Australian housing data, showing how the incorporation of prior knowledge can result in agents building on past beliefs. In particular, the agent focus can be shown to shift from utility maximisation to acting on previous knowledge. In other words, during the periods when the market has been performing well, the agents were shown to become overly optimistic based on the past performance.

The generality of the proposed approach makes it useful for incorporating any form of prior information on the agent's choice set. Moreover, we have shown that the default QRSE is a special case of the proposed extension with uniform (i.e., uninformative) priors. Therefore, the proposed approach can be seen as an extension of QRSE, which accounts for prior agent beliefs based on information acquisition costs. As the QRSE framework continues to be expanded, the generalised model proposed here could become an important approach. Particularly, this would be useful whenever prior knowledge on agent decisions is known, as well as in multi-action cases when the IIA property of the general logit function is undesirable. Other relevant applications include scenarios with multiple time periods, allowing for a detailed temporal analysis and exploration of the cost of switching between equilibria (measured as an information acquisition cost from prior beliefs).

Author Contributions: B.P.E. and M.P.; Funding acquisition, M.P.; Software, B.P.E.; Supervision, M.P.; Writing—original draft, B.P.E. and M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Australian Research Council Discovery Project DP170102927.

Data Availability Statement: The real-estate pricing data used in this work were made available under license for this study by SIRCA-CoreLogic (https://www.corelogic.com.au/industries/residential-real-estate).

Acknowledgments: The authors would like to thank Kirill Glavatskiy and Michael S. Harré for many helpful discussions regarding the Australian housing market, as well as Adrián Carro, Jangho Yang and anonymous reviewers for various comments. The authors would also like to acknowledge the Securities Industry Research Centre of Asia-Pacific (SIRCA) and CoreLogic, Inc. (Sydney, Australia) for their data on Greater Sydney housing prices.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Derivations

Appendix A.1. Decision Duality

There are two main perspectives, the first is of the agent performing actions within the system, and the second is of the system observer [29].

Each of the two perspectives allows to capture the uncertainty faced by either the actor or the observer, by imposing a constraint on entropy. In this section, we outline the duality that arises from these perspectives, showing that a duality exists between maximum entropy models, and entropy constrained models [7]. Additional discussion on such perspectives is given in [21].

Modelling the actor corresponds to maximising the expected utility subject to a fixed entropy constraint. This is the method outlined in Section 3.1.1. In this case, the agent can be seen as a boundedly rational decision-maker, in that they might not have all of the information required to make a perfectly rational choice.

The alternate perspective, modelling an observer, corresponds to maximising the entropy of the decisions subject to a fixed expected utility. With this perspective, we capture modelling uncertainty from the observer. The observers problem is formulated as follows

$$\begin{aligned} \max -\sum_{a\in A} f[a|x]\log f[a|x] \\ \text{subject to} \sum_{a\in A} f[a|x] = 1 \\ \sum_{a\in A} f[a|x]U[a,x] \geq U_{min} \end{aligned} \tag{A1}$$

where U_{min} represents the minimum expected utility. In order to see the duality of Equations (A1) and (1), we formulate the following Lagrangian for converting Equation (A1) into an unconstrained optimization problem.

$$\mathcal{L} = -\sum_{a\in A} f[a|x]\log f[a|x] - \lambda\left(\sum_{a\in A} f[a|x] - 1\right) + \beta\left(\sum_{a\in A} f[a|x]U[a,x] - U_{min}\right) \tag{A2}$$

where again, taking the first order conditions and solving for $f[a|x]$ yields

$$f[a|x] = \frac{1}{Z}e^{\beta U[a,x]} \tag{A3}$$

We can see Equation (A3) is equivalent with Equation (4) with $\beta = \frac{1}{T}$, which highlights an important dualism between the two perspectives.

Appendix A.2. Decision Function

By setting the partial derivative of the unconstrained optimisation problem given in Equation (15) with respect to $f[a|x]$ to 0, we can obtain the following definition for $f[a|x]$:

$$\begin{aligned} \frac{d\mathcal{L}}{f[a|x]} &= U[a,x] - \lambda - T\log\left(\frac{f[a|x]}{p[a]}\right) = 0 \\ f[a|x] &= e^{\frac{U[a,x]}{T} - \lambda + \log p[a]} \end{aligned} \tag{A4}$$

and, using the normalisation constraint $\sum_{a \in A} f[a|x] = 1$, we obtain the following decision function

$$\begin{aligned} f[a|x] &= \frac{1}{Z_{A|x}} e^{\frac{U[a,x]}{T} + \log p[a]} \\ &= \frac{1}{Z_{A|x}} p[a] e^{\frac{U[a,x]}{T}} \end{aligned} \tag{A5}$$

with the partition function $Z_{A|x} = \sum_{a' \in A} p[a'] e^{\frac{U[a',x]}{T}}$.

Appendix B. Australian Housing Market Data

Data from 2006–2020 is used. Data is split into individual years. We use the rolling median price for each area and then measure the quarterly percentage growth rate for the areas. The month-to-month percentage changes are visualised in Figure A1. The distributions of the returns are visualised in Figure A2.

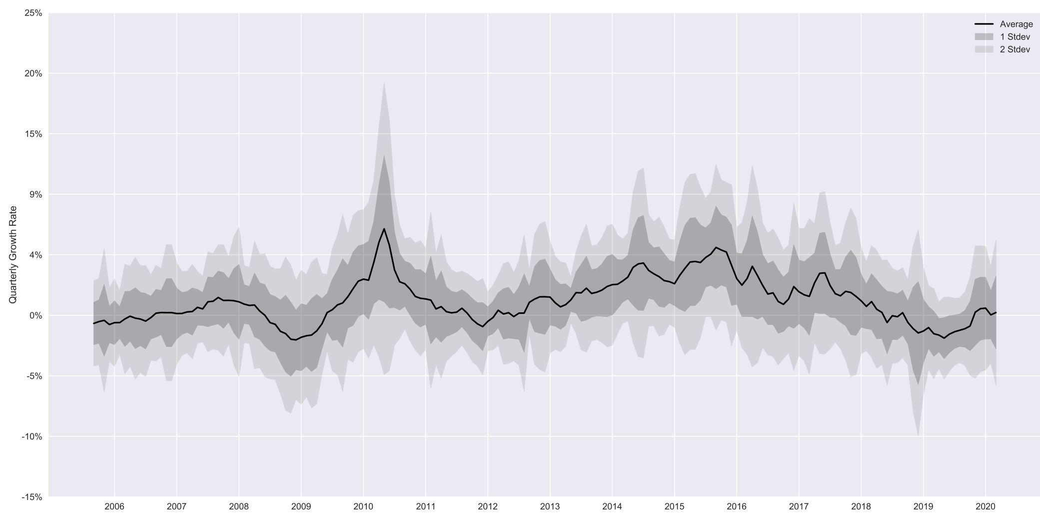


Figure A1. Quarterly returns in the Sydney housing market.

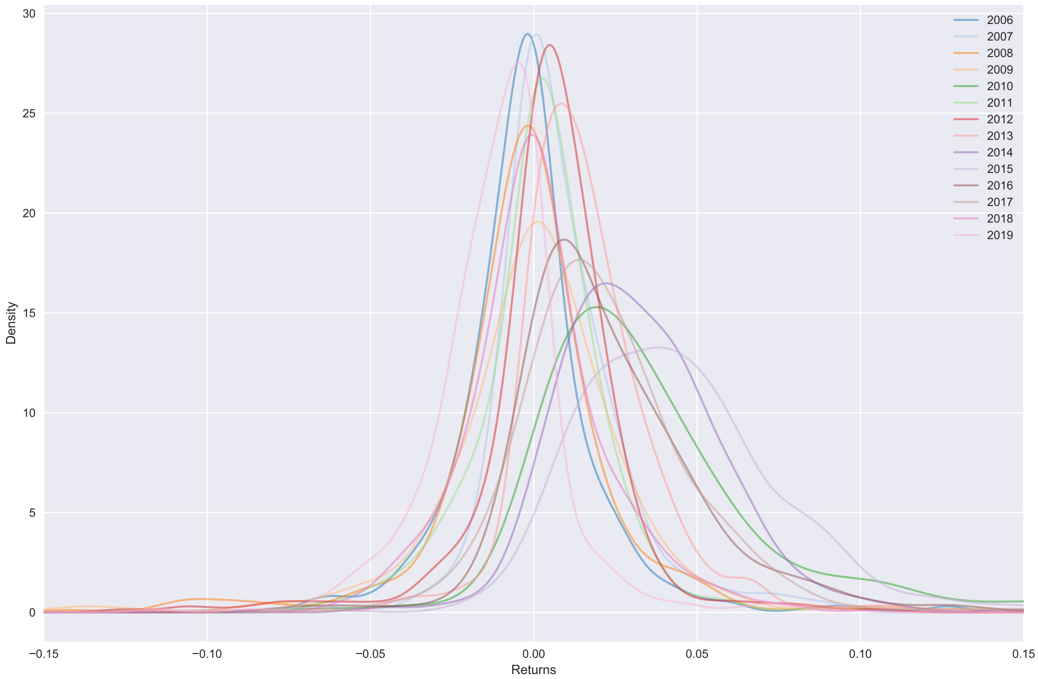


Figure A2. Density plots of returns grouped by year. We can see each year follows a different shape, but shows some striking regularities representing a statistical equilibrium.

Appendix C. Relation to Rational Inattention

In his seminal work, [2] outlined rational inattention "based on the idea that individual people have limited capacity for processing information". This work introduced information-processing constraints into the macroeconomic literature, using mutual information as a measure of such information costs.

Of particular interest are the developments of [3] who showed how to apply rational inattention (RI) to discrete decision-making. The key contribution was the modification to the logit function that arises from considering a cost to decision-makers from deviating from prior knowledge. In this section, we highlight the similarities of R.I. with the thermodynamic approach of [4] and the work proposed here.

The problem to be solved is formulated as follows. A utility-maximising agent must make a discrete choice, while it is costly to acquire information about the options A available:

$$\max f[a,x] \sum_{a \in A} \int_x f[a,x] U[a,x] dx - T\left(- \sum_{a \in A} f[a,x] \log\left(\frac{f[a,x]}{p[x] f[a]}\right)\right) \\ \text{subject to} \sum_{a \in A} f[a|x] = 1 \tag{A6}$$

where the first term is the expected utility, and the second a cost of information (following Sims [2], the mutual information). We see this as a similar setup to that of [4], which also corresponds to maximising the expected utility subject to an information cost, however, the information cost in [4] is instead measured as the KL-divergence. A key difference between the two is that Equation (A6) adds a dependence on $f[a]$ into the denominator of the information cost term. We can take the first order conditions of the resulting Lagrangian for (A6) and solve for $f[a|x]$, yielding:

$$f[a|x] = \frac{e^{\frac{U(a,x)}{T} + \log(f[a])}}{\sum_{a' \in A} e^{\frac{U(a',x)}{T} + \log(f[a'])}} = \frac{f[a] e^{\frac{U(a,x)}{T}}}{\sum_{a' \in A} f[a'] e^{\frac{U(a',x)}{T}}} \tag{A7}$$

which is not yet fully solved, as there is a dependence on the unconditional probability $f[a]$. Since $f[a] = \int_x f[a|x]p[x]dx$, $f[a]$ depends on $f[a|x]$, and $f[a|x]$ depends on $f[a]$, this must (generally) be solved numerically, for example, with the Blahut–Arimoto algorithm by first making a guess for $f[a]$ and then iterating from there (see Caplin et al. [73] or Matějka and McKay [3] for solutions). It is for this reason, we utilise the configuration of [4] for the decision-making component, which depends only on the prior probabilities, and not the unconditional action probabilities $f[a]$ meaning an analytical solution can be obtained. However, the R.I. framework can be seen as equivalent to choosing an "optimal" prior in the free energy framework of [4], as both can be seen as applications of rate-distortion theory [56].

Further discussion on the relationship between R.I. and QRSE is given in [30].

Appendix D. Additional Parameters

While μ and T are the main parameters of interest in this work, since they have a direct contribution to the modified decision function introduced, ρ and γ are still important, although to a lesser extent as they are indirectly impacted. ρ is the Lagrange multiplier for the competition constraint, and γ controls the skewness of the resulting distribution.

Appendix D.1. Impact of Decisions on Outcomes

Parameter ρ measures the impact of individual decisions on housing prices. A large ρ corresponds to a highly effective market (high impact of actions on the response). In contrast, a low ρ corresponds to a weaker market response, and thus, lower market effectiveness. Parameter ρ, therefore, corresponds to the strength of the negative feedback mechanism, with the case of $\rho = 0$ implying no market feedback (i.e., no impact on the outcome based on the actions). In all cases, we see relatively large ρ's, peaking in 2013 and 2019, indicating the presence of a well-functioning feedback loop across the years. We see little variation between the uniform, previous, and mean prior in Figure A3, perhaps drawn from the fact the priors work as linear weightings in the difference between the conditional action probabilities, as shown in Equation (20).

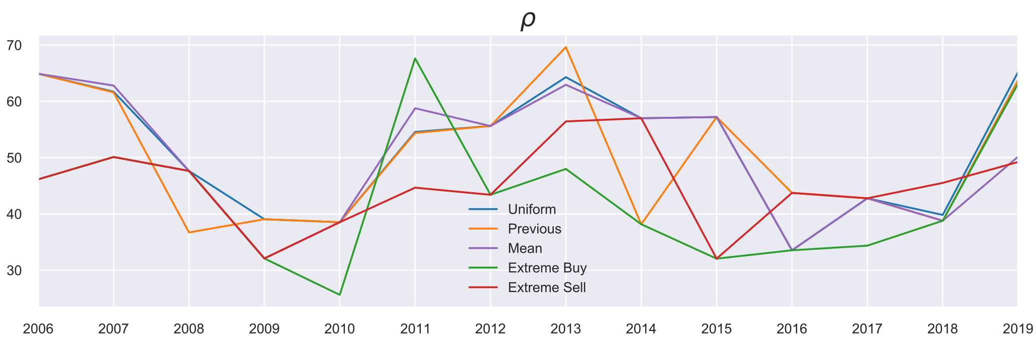


Figure A3. Competition.

Appendix D.2. Skewness

The parameter γ affects the skew of the resulting exponential distribution. This skew arises from (potentially) unfulfilled agent expectations, i.e., where $\mu \neq \xi$ [21]. Parameter γ, therefore, is a measure of skewness in the binary action case. In the asymmetric multi-action QRSE case, γ is replaced by alternate μ's explaining such skew. As mentioned, the priors can also introduce such a skew (without the need for a γ). This is shown in the extreme buy γ in Figure A4 which was almost always near zero, as the buying preference already creates the skew needed to describe the underlying distribution (i.e., the skewness was already explained by p). In contrast, extreme sell needs small γ's to switch their (incorrect) skew.

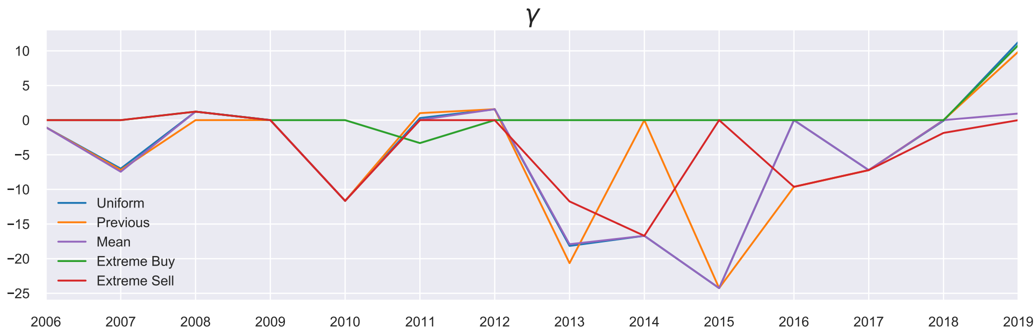
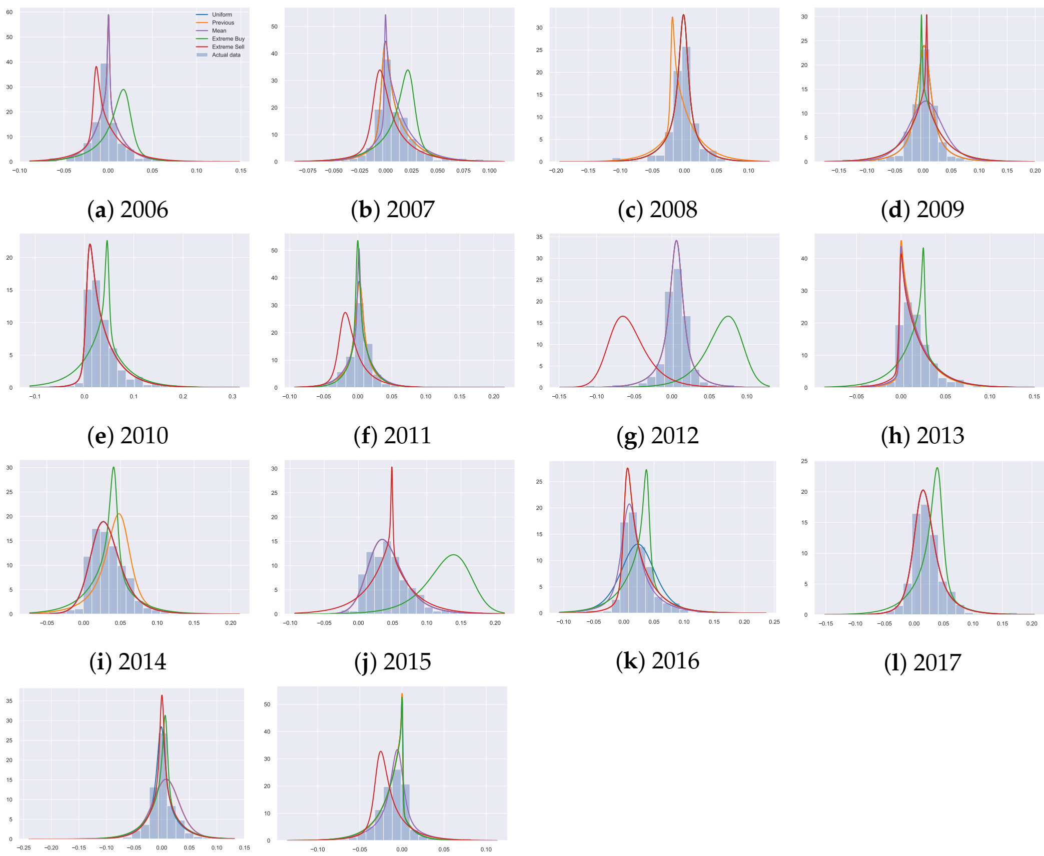


Figure A4. Skewness.

Negative γ corresponds to positive skewness, and positive γ corresponds to negative skewness. In most cases here, we see (at least slightly) positively skewed distributions (resulting in negative γ's), with the exception of 2019, which is negatively skewed, as can be verified in Figure A5.

Generally, γ's for the mean, previous, and uniform priors follow similar paths, except for the 2013–2016 years. In 2014 and 2016, γ's for the previous priors differs from the other priors. This can be explained by the fact that in both cases, the prior had a strong sell preference (shown in Figure 4), meaning an adjusted γ was needed to capture the current distributions shift correctly (and offset the influence of the prior).

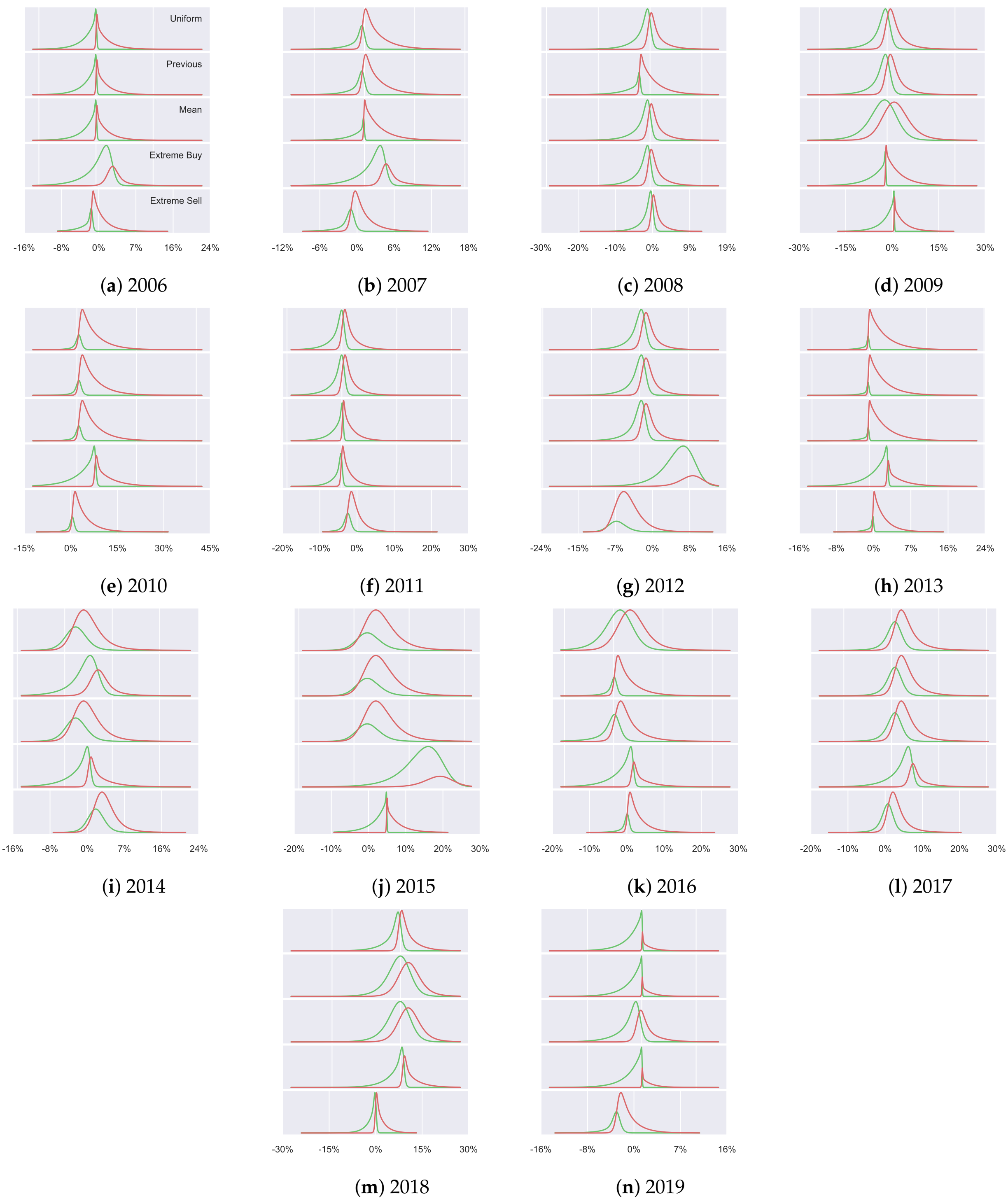


(**a**) 2006 (**b**) 2007 (**c**) 2008 (**d**) 2009
(**e**) 2010 (**f**) 2011 (**g**) 2012 (**h**) 2013
(**i**) 2014 (**j**) 2015 (**k**) 2016 (**l**) 2017
(**m**) 2018 (**n**) 2019

Figure A5. Resulting fitted marginals distributions $f[x]$ for each year. Each coloured line represents a different prior (with the legend given in the top left). The blue bars show the (discretized) actual return distribution.

Appendix E. Probability Plots

In this section, we provide the resulting probability plots for $f[x]$ (Figure A5), $f[a,x]$ (Figure A6), and $f[a|x]$ (Figure A7) across all years analysed.



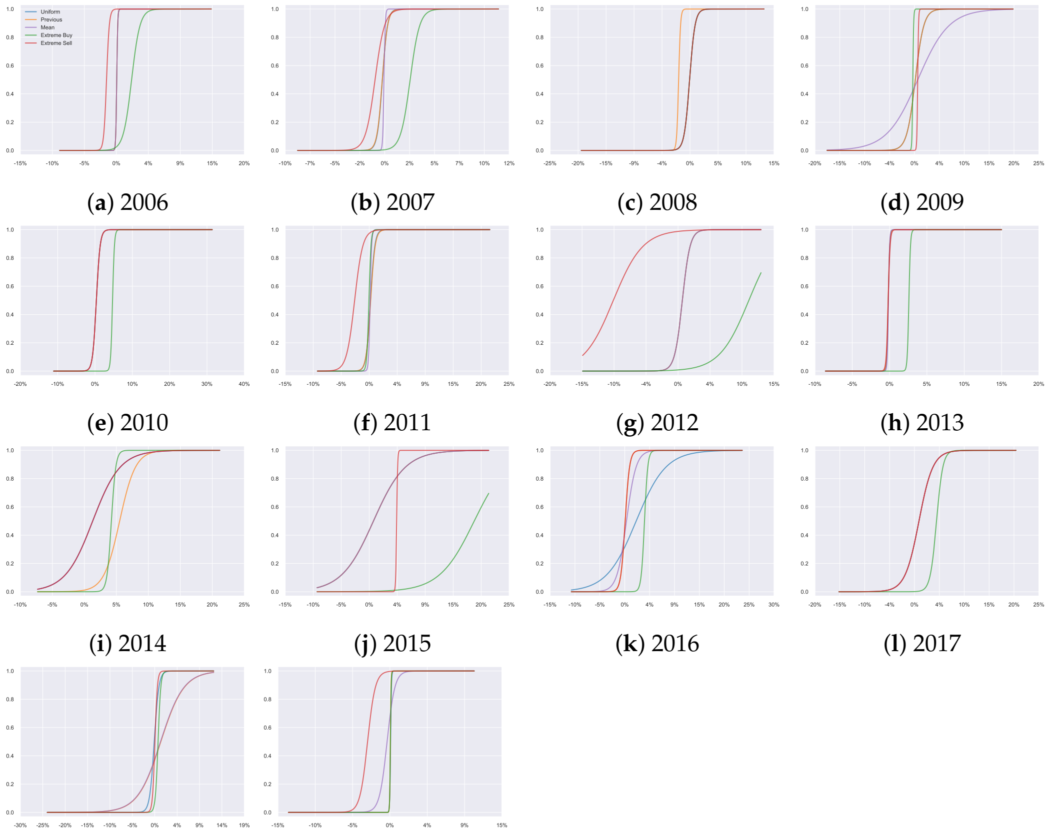
(**a**) 2006 (**b**) 2007 (**c**) 2008 (**d**) 2009

(**e**) 2010 (**f**) 2011 (**g**) 2012 (**h**) 2013

(**i**) 2014 (**j**) 2015 (**k**) 2016 (**l**) 2017

(**m**) 2018 (**n**) 2019

Figure A6. Resulting Joint Distributions. Red lines represent $f[\text{sell}, x]$, and green lines represent $f[\text{buy}, x]$. Each plot from top to bottom shows: Uniform, previous, mean and extreme buy and extreme sell priors (in that order).



(**a**) 2006 (**b**) 2007 (**c**) 2008 (**d**) 2009

(**e**) 2010 (**f**) 2011 (**g**) 2012 (**h**) 2013

(**i**) 2014 (**j**) 2015 (**k**) 2016 (**l**) 2017

(**m**) 2018 (**n**) 2019

Figure A7. Decision functions for selling. Buying curves are excluded as they are simply the complement (1 − sell). The green lines represent the extreme buy a priori preference, which means the resulting probabilities of selling are shifted far to the right, i.e., the majority of the area comprises buying actions, and only the extreme positive growth rates for sell. In contrast, the red lines represent the sell preference, which "pulls" the area to the left, resulting in a strong resulting conditional preference for selling.

References

1. Simon, H.A. *Models of Man; Social And Rational.*; Wiley: Hoboken, NJ, USA, 1957.
2. Sims, C.A. Implications of rational inattention. *J. Monet. Econ.* **2003**, *50*, 665–690. [CrossRef]
3. Matějka, F.; McKay, A. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *Am. Econ. Rev.* **2015**, *105*, 272–298. [CrossRef]
4. Ortega, P.A.; Braun, D.A. Thermodynamics as a theory of decision-making with information-processing costs. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2013**, *469*, 20120683. [CrossRef]
5. Scharfenaker, E.; Foley, D.K. Quantal response statistical equilibrium in economic interactions: Theory and estimation. *Entropy* **2017**, *19*, 444. [CrossRef]
6. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620. [CrossRef]
7. Yang, J. Information theoretic approaches in economics. *J. Econ. Surv.* **2018**, *32*, 940–960. [CrossRef]
8. Yakovenko, V.M. Econophysics, Statistical Mechanics Approach to. In *Encyclopedia of Complexity and Systems Science*; Meyers, R.A., Ed.; Springer: New York, NY, USA, 2009; pp. 2800–2826.
9. Scharfenaker, E.; Semieniuk, G. A statistical equilibrium approach to the distribution of profit rates. *Metroeconomica* **2017**, *68*, 465–499. [CrossRef]
10. Scharfenaker, E.; Yang, J. Maximum entropy economics. *Eur. Phys. J. Spec. Top.* **2020**, *229*, 1577–1590. [CrossRef]
11. Wolpert, D.H.; Harré, M.; Olbrich, E.; Bertschinger, N.; Jost, J. Hysteresis effects of changing the parameters of noncooperative games. *Phys. Rev. E* **2012**, *85*, 036102. [CrossRef]
12. Dragulescu, A.; Yakovenko, V.M. Statistical mechanics of money. *Eur. Phys. J. Condens. Matter Complex Syst.* **2000**, *17*, 723–729. [CrossRef]

13. Yakovenko, V.M.; Rosser Jr, J.B. Colloquium: Statistical mechanics of money, wealth, and income. *Rev. Mod. Phys.* **2009**, *81*, 1703. [CrossRef]
14. Foley, D.K. Unfulfilled Expectations: One Economist's History. In *Expectations*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 3–17.
15. Harré, M.S. Information Theory for Agents in Artificial Intelligence, Psychology, and Economics. *Entropy* **2021**, *23*, 310. [CrossRef] [PubMed]
16. Foley, D.K. Information theory and behavior. *Eur. Phys. J. Spec. Top.* **2020**, *229*, 1591–1602. [CrossRef]
17. Ömer, Ö. Maximum entropy approach to market fluctuations as a promising alternative. *Eur. Phys. J. Spec. Top.* **2020**, *229*, 1715–1733. [CrossRef]
18. Yang, J.; Carro, A. Two tales of complex system analysis: MaxEnt and agent-based modeling. *Eur. Phys. J. Spec. Top.* **2020**, *229*, 1623–1643. [CrossRef]
19. Jaynes, E.T. Information theory and statistical mechanics. II. *Phys. Rev.* **1957**, *108*, 171. [CrossRef]
20. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: Cambridge, UK, 2003.
21. Scharfenaker, E. Implications of quantal response statistical equilibrium. *J. Econ. Dyn. Control.* **2020**, *119*, 103990. [CrossRef]
22. Ömer, Ö. Dynamics of the US Housing Market: A Quantal Response Statistical Equilibrium Approach. *Entropy* **2018**, *20*, 831. [CrossRef]
23. Ömer, Ö. Essays on Modeling Housing Markets, Income Distribution, and Wealth Concentration. Ph.D. Thesis, The New School, New York, NY, USA, 2018.
24. Ömer, Ö. Equilibrium-Disequilibrium Dynamics of the US Housing Market, 2000–2015: A Quantal Response Statistical Equilibrium Approach. Working Papers 1809, New School for Social Research, Department of Economics, 2018. Available online: https://econpapers.repec.org/paper/newwpaper/1809.htm (accessed on 30 September 2020)
25. Yang, J. A quantal response statistical equilibrium model of induced technical change in an interactive factor market: Firm-level evidence in the EU economies. *Entropy* **2018**, *20*, 156. [CrossRef]
26. Wiener, N. Measuring Labor Market Segmentation from Incomplete Data. Working Paper 2018-01, Amherst, MA, 2018. Available online: https://scholarworks.umass.edu/econ_workingpaper/238/ (accessed on 3 October 2020)
27. Wiener, N. Essays on Labor Mobility and Segmentation. Ph.D. Thesis, The New School, New York, NY, USA, 2019.
28. Wiener, N.M. Labor market segmentation and immigrant competition: A quantal response statistical equilibrium analysis. *Entropy* **2020**, *22*, 742. [CrossRef]
29. Blackwell, K. A Behavioral Foundation for Commonly Observed Distributions of Financial and Economic Data. Working Papers 1912, New School for Social Research, Department of Economics, 2019. Available online: https://ideas.repec.org/p/new/wpaper/1912.html (accessed on 8 October 2020)
30. Blackwell, K. Entropy Constrained Behavior in Financial Markets A Quantal Response Statistical Equilibrium Approach to Financial Modeling. Ph.D. Thesis, The New School, New York, NY, USA, 2018.
31. Scharfenaker, E. Statistical Equilibrium Methods in Analytical Political Economy. *J. Econ. Surv.* **2020**. [CrossRef]
32. Smith, A. *The Wealth of Nations: An inquiry into the nature and causes of the Wealth of Nations*; Harriman House Limited: Petersfield, UK, 2010.
33. McKelvey, R.D.; Palfrey, T.R. Quantal response equilibria for normal form games. *Games Econ. Behav.* **1995**, *10*, 6–38. [CrossRef]
34. McKelvey, R.D.; Palfrey, T.R. Quantal response equilibria for extensive form games. *Exp. Econ.* **1998**, *1*, 9–41. [CrossRef]
35. Lord, C.G.; Ross, L.; Lepper, M.R. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *J. Personal. Soc. Psychol.* **1979**, *37*, 2098. [CrossRef]
36. K Levine, D. *Is Behavioral Economics Doomed?: The Ordinary Versus the Extraordinary*; Open Book Publishers: Cambridge, UK, 2012.
37. DellaVigna, S. Psychology and economics: Evidence from the field. *J. Econ. Lit.* **2009**, *47*, 315–372. [CrossRef]
38. Daunizeau, J.; Den Ouden, H.E.; Pessiglione, M.; Kiebel, S.J.; Stephan, K.E.; Friston, K.J. Observing the observer (I): Meta-bayesian models of learning and decision-making. *PLoS ONE* **2010**, *5*, e15554. [CrossRef] [PubMed]
39. Khalvati, K.; Park, S.A.; Mirbagheri, S.; Philippe, R.; Sestito, M.; Dreher, J.C.; Rao, R.P. Modeling other minds: Bayesian inference explains human choices in group decision-making. *Sci. Adv.* **2019**, *5*, eaax8783. [CrossRef] [PubMed]
40. Kruis, J.; Maris, G.; Marsman, M.; Bolsinova, M.; van der Maas, H.L. Deviations of rational choice: An integrative explanation of the endowment and several context effects. *Sci. Rep.* **2020**, *10*, 1–16. [CrossRef]
41. Debreu, G. Review of individual choice behavior by RD Luce. *Am. Econ. Rev.* **1960**, *50*, 186–188.
42. McFadden, D. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, Zarembka, P., Ed.; Academic Press: Cambridge, MA, USA, 1973; pp. 105–142.
43. Golan, A. *Foundations of Info-Metrics: Modeling, Inference, and Imperfect Information*; Oxford University Press: Oxford, UK, 2018.
44. Hafner, D.; Ortega, P.A.; Ba, J.; Parr, T.; Friston, K.; Heess, N. Action and perception as divergence minimization. *arXiv* **2020**, arXiv:2009.01791.
45. Ortega, P.A.; Stocker, A.A. Human decision-making under limited time. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 100–108.
46. Gottwald, S.; Braun, D.A. The two kinds of free energy and the Bayesian revolution. *PLoS Comput. Biol.* **2020**, *16*, e1008420. [CrossRef] [PubMed]

47. Wilson, A. Boltzmann, Lotka and Volterra and spatial structural evolution: An integrated methodology for some dynamical systems. *J. R. Soc. Interface* **2008**, *5*, 865–871. [CrossRef]
48. Crosato, E.; Nigmatullin, R.; Prokopenko, M. On critical dynamics and thermodynamic efficiency of urban transformations. *R. Soc. Open Sci.* **2018**, *5*, 180863. [CrossRef]
49. Slavko, B.; Glavatskiy, K.; Prokopenko, M. Dynamic resettlement as a mechanism of phase transitions in urban configurations. *Phys. Rev. E* **2019**, *99*, 042143. [CrossRef]
50. Harding, N.; Spinney, R.E.; Prokopenko, M. Population mobility induced phase separation in SIS epidemic and social dynamics. *Sci. Rep.* **2020**, *10*, 1–11. [CrossRef]
51. Shore, J.; Johnson, R. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory* **1980**, *26*, 26–37. [CrossRef]
52. Kesavan, H.; Kapur, J. Maximum Entropy and Minimum Cross-Entropy Principles: Need for a Broader Perspective. In *Maximum Entropy and Bayesian Methods*; Springer: Berlin/Heidelberg, Germany, 1990; pp. 419–432.
53. Glavatskiy, K.S.; Prokopenko, M.; Carro, A.; Ormerod, P.; Harre, M. Explaining herding and volatility in the cyclical price dynamics of urban housing markets using a large-scale agent-based model. *SN Bus. Econ.* **2021**, *1*, 1–21. [CrossRef]
54. Evans, B.P.; Glavatskiy, K.; Harré, M.S.; Prokopenko, M. The impact of social influence in Australian real estate: Market forecasting with a spatial agent-based model. *J. Econ. Interact. Coord.* **2021**, 1–53.
55. Hommes, C.H. On the consistency of backward-looking expectations: The case of the cobweb. *J. Econ. Behav. Organ.* **1998**, *33*, 333–362. [CrossRef]
56. Genewein, T.; Leibfried, F.; Grau-Moya, J.; Braun, D.A. Bounded rationality, abstraction, and hierarchical decision-making: An information-theoretic optimality principle. *Front. Robot. AI* **2015**, *2*, 27. [CrossRef]
57. Friedman, M. *Theory of the Consumption Function*; Princeton University Press: Princeton, NJ, USA, 2018.
58. Hommes, C.; Wagener, F. Complex evolutionary systems in behavioral finance. In *Handbook of Financial Markets: Dynamics and Evolution*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 217–276.
59. Evans, G.W.; Honkapohja, S. *Learning and Expectations in Macroeconomics*; Princeton University Press: Princeton, NJ, USA, 2012.
60. Chow, G.C. *Usefulness of Adaptive and Rational Expectations in Economics*; Center for Economic Policy Studies, Princeton University: Princeton, NJ, USA, 2011.
61. Soofi, E.S.; Retzer, J.J. Information indices: Unification and applications. *J. Econom.* **2002**, *107*, 17–40. [CrossRef]
62. Dziewulski, P. Just-noticeable difference as a behavioural foundation of the critical cost-efficiency index. *J. Econ. Theory* **2020**, *188*, 105071. [CrossRef]
63. Pfeifer, R.; Bongard, J. *How the Body Shapes the Way We Think: A New View of Intelligence*; MIT Press: Cambridge, MA, USA, 2006.
64. Polani, D.; Sporns, O.; Lungarella, M. How information and embodiment shape intelligent information processing. In *50 Years of Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 99–111.
65. Ay, N.; Bernigau, H.; Der, R.; Prokopenko, M. Information-driven self-organization: The dynamical system approach to autonomous robot behavior. *Theory Biosci.* **2012**, *131*, 161–179. [CrossRef]
66. Montúfar, G.; Ghazi-Zahedi, K.; Ay, N. A theory of cheap control in embodied systems. *PLoS Comput. Biol.* **2015**, *11*, e1004427. [CrossRef]
67. Polani, D.; Nehaniv, C.L.; Martinetz, T.; Kim, J.T. Relevant information in optimized persistence vs. progeny strategies. In Proceedings of the Artificial Life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems, Bloomington, IN, USA, 3–6 June 2006; MIT Press: Cambridge, MA, USA, 2006.
68. Prokopenko, M.; Gerasimov, V.; Tanev, I. Measuring spatiotemporal coordination in a modular robotic system. In Proceedings of the Artificial Life X: Proceedings of the 10th International Conference on the Simulation and Synthesis of Living Systems, Bloomington, IN, USA, 3–6 June 2006; pp. 185–191.
69. Capdepuy, P.; Polani, D.; Nehaniv, C.L. Maximization of potential information flow as a universal utility for collective behaviour. In Proceedings of the 2007 IEEE Symposium on Artificial Life, Honolulu, HI, USA, 1–5 April 2007; pp. 207–213.
70. Tishby, N.; Polani, D. Information theory of decisions and actions. In *Perception-Action Cycle*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 601–636.
71. Tishby, N.; Pereira, F.C.; Bialek, W. The information bottleneck method. *arXiv* **2000**, arXiv:physics/0004057.
72. Spinney, R.E.; Lizier, J.T.; Prokopenko, M. Transfer entropy in physical systems and the arrow of time. *Phys. Rev. E* **2016**, *94*, 022135. [CrossRef] [PubMed]
73. Caplin, A.; Dean, M.; Leahy, J. Rational inattention, optimal consideration sets, and stochastic choice. *Rev. Econ. Stud.* **2019**, *86*, 1061–1094. [CrossRef]

CHAPTER 5

An information-theoretic model for bounded rational strategic reasoning

Until this point, we have modelled inductive decision-making in markets using behavioural heuristics. We now consider the strategic decision-making between interacting (and competing) agents. Strategic decision-making is an important consideration, as often, an agent is not making a decision in isolation. Instead, an agent must consider the behaviour of other agents when deciding what action to take. For example, when considering potential profit rates in markets, one must consider the reasoning of other market participants, as the resulting decisions will impact potential profit.

In the previous chapters, this type of strategic reasoning had been approximated using various heuristics; however, here, we propose a more generalised way of modelling generic strategic reasoning interactions. The resulting outcomes do not require any equilibrium (statistical or otherwise) to arise.

Theory and Decision
https://doi.org/10.1007/s11238-023-09941-z

Bounded rationality for relaxing best response and mutual consistency: the quantal hierarchy model of decision making

Benjamin Patrick Evans[1] · **Mikhail Prokopenko**[1]

Accepted: 24 April 2023

Abstract

While game theory has been transformative for decision making, the assumptions made can be overly restrictive in certain instances. In this work, we investigate some of the underlying assumptions of rationality, such as mutual consistency and best response, and consider ways to relax these assumptions using concepts from level-k reasoning and quantal response equilibrium (QRE) respectively. Specifically, we propose an information-theoretic two-parameter model called the quantal hierarchy model, which can relax both mutual consistency and best response while still approximating level-k, QRE, or typical Nash equilibrium behavior in the limiting cases. The model is based on a recursive form of the variational free energy principle, representing higher-order reasoning as (pseudo) sequential decision-making in extensive-form game tree. This representation enables us to treat simultaneous games in a similar manner to sequential games, where reasoning resources deplete throughout the game-tree. Bounds in player processing abilities are captured as information costs, where future branches of reasoning are discounted, implying a hierarchy of players where lower-level players have fewer processing resources. We demonstrate the effectiveness of the quantal hierarchy model in several canonical economic games, both simultaneous and sequential, using out-of-sample modelling.

Keywords Quantal hierarchy · Bounded rationality · Economic decision · Level-k · Cognitive hierarchy · Quantal response equilibrium

1 Introduction

A crucial assumption made in Game Theory is that all players behave perfectly rationally. Nash Equilibrium (Nash, 1951) is a key concept that arises based on all players being rational and assuming other players will also be rational, requiring correct and consistent beliefs amongst players. However, despite being the traditional

Extended author information available on the last page of the article

Published online: 17 May 2023

economic assumption, perfect rationality is incompatible with many human processing tasks, with models of limited rationality better matching human behaviour than fully rational models (Camerer, 2010; Gächter, 2004; Gabaix et al., 2006). Further, if an opponent is irrational, then it would be rational for the subject to play "irrationally" (Raiffa & Luce, 1957). That is, "[Nash equilibrium] is consistent with perfect foresight and perfect rationality only if both players play along with it. But there are no rational grounds for supposing they will" (Koppl et al., 2002). These assumptions of perfect foresight and rationality often lead to contradictions and paradoxes (Cournot et al., 1838; Glasner, 2022; Hoppe, 1997; Morgenstern, 1928).

Alternate formulations have been proposed that relax some of these assumptions and model boundedly rational players, better approximating actual human behaviour and avoiding some of these paradoxes (Simon, 1976). For example, relaxing mutual consistency allows players to form different beliefs of other players (Camerer et al., 2004; Stahl & Wilson, 1995) avoiding the infinite self-referential higher-order reasoning which emerges as the result of interaction between rational players (Knudsen, 1993) [I have a model of my opponent who has a model of me ... *ad infinitum* (Morgenstern, 1935)] and non-computability of best response functions (Koppl et al., 2002; Rabin, 1957). The ability to break at various points in the higher-order reasoning chain can be considered as "partial self-reference" (Löfgren, 1990; Mackie, 1971). Importantly, rather than implicating negation (Prokopenko et al., 2019), this type of self-reference represents higher-order reasoning as a logically non-contradictory chain of recursion (reasoning about reasoning...). Hence, bounded rationality arises from the ability to break at various points in the chain, discarding further branches. "Breaking" the chain on an otherwise potentially infinite regress of reasoning about reasoning can be seen as the players limitations in information processing, determining when to end the recursion.

Another example of bounded rationality based on information processing constraints is the relaxation of the best response assumption of players, which allows for erroneous play, with deviations from the best response governed by a resource parameter (Goeree et al., 2005; Haile et al., 2008).

In this work, we adopt an information-theoretic perspective on reasoning (decision-making). By enforcing potential constraints on information processing, we are able to relax both mutual consistency and best response, and hence, players do not necessarily act perfectly rational. The proposed approach provides an information-theoretic foundation for level-k reasoning and a generalised extension where players can make errors at each of the k levels.

Specifically, in this paper, we focus on three main aspects:

- Players reasoning abilities decrease throughout recursion in a wide variety of games, motivating an increasing error rate at deeper levels of recursion. That is, it becomes more and more difficult to reason about reasoning about reasoning ... (necessitating a relaxation in best response decisions).
- Finite higher-order reasoning can be captured by discounting future chains of recursion and ultimately discarding branches once resources run out. This representation introduces an implicit hierarchy of players, where a player assumes

they have a higher level of processing abilities than other players, motivating relaxation of mutual consistency.

- Existing game-theoretic models can be explained and recovered in limiting cases of the proposed approach. This fills an important gap between methods relaxing best response and methods relaxing mutual consistency.

The proposed approach features only two parameters, β and γ, where β quantifies relaxation of players' best response, and γ governs relaxation of mutual consistency between players. In the limit, $\beta \to \infty$ best response can be recovered, and in the limit $\gamma = 1$ mutual consistency can be recovered. Equilibrium behaviour is recovered in the limit of both $\beta \to \infty, \gamma = 1$. For other values of β and γ, interesting out-of-equilibrium behaviour can be modelled which concurs with experimental data, and furthermore, in repeated games that converge to Nash equilibrium, player learning can be captured in the model with increases in β and γ. Importantly, we also show how fitted values of β and γ also generalise well to out-of-sample data.

The remainder of the paper is organised as follows. Section 2 analyses bounded rationality in the context of decision-making and game theory, Sect. 3 introduces information-theoretic bounded rational decision-making, Sect. 4 extends the idea to capture higher-order reasoning in game theory. We then use canonical examples highlighting the use of information-constrained players in addressing bounded rational behaviour in games in Sect. 5. We draw conclusions and outline future work in Sect. 6.

2 Background and motivation

While a widespread assumption in economics, perfect rationality is incompatible with the observed behaviour in many experimental settings, motivating the use of bounded rationality (Camerer, 2011). Bounded rationality offers an alternative perspective, by acknowledging that players may not have a perfect model of each other or may not play perfectly rationally. In this section, we explore some common approaches to modelling bounded rational decision-making.

2.1 Mutual consistency

Equilibrium models assume mutual consistency of beliefs and choices (Camerer et al., 2003; Camerer, 2003), however, this is often violated in experimental settings (Polonio & Coricelli, 2019) where "differences in belief arise from the different iterations of strategic thinking that players perform" (Chong et al., 2005).

Level-*k* reasoning (Stahl & Wilson, 1995) is one attempt at incorporating bounded rationality by relaxing mutual consistency, where players are bound to some level k of reasoning. A player assumes that other players are reasoning at a lower level than themselves, for example, due to over-confidence. This relaxes the mutual consistency assumption, as it implicitly assumes other players are not as advanced as themselves. Players at level 0 are not assumed to perform any information processing, and simply choose uniformly over actions (i.e., a Laplacian assumption due to the principle of insufficient reason), although alternate level-0 configurations can be considered (Wright & Leyton-Brown, 2019). Level-1 players then exploit these level-0 players and act based on this. Likewise, level-2 players act based on other players being level-1, and so on and so forth for level-k players acting as if the other players are at level-$(k-1)$. Various extensions have also been proposed (Levin & Zhang, 2022).

A similar approach to level-k is that of cognitive hierarchies (CH) (Camerer et al., 2004), where again it is assumed other players have lower reasoning abilities. However, rather than assuming that the other players are at $k-1$, players can be distributed across the k levels of cognition according to Poisson distribution with mean and variance τ. The validation of the Poisson distribution has been provided in Chong et al. (2005), where an unconstrained general distribution offered only marginal improvement. Again, various extensions have been proposed (Chong et al., 2016; Koriyama & Ozkes, 2021) and there are many examples of successful applications of such depth-limited reasoning in literature, for example, Goldfarb and Xiao (2011).

Endogenous depth of reasoning (EDR) is a similar approach to level-k and CH, but it separates the player's cognitive bounds from their beliefs of their opponent's reasoning (Alaoui & Penta, 2016). EDR captures player reasoning as if they are following a cost-benefit analysis (Alaoui & Penta, 2022), with cognitive abilities (costs) and payoffs (benefits).

One fundamental similarity across these methods is that they all maintain best-response. That is, they best respond based on the lower-level play assumptions. The following section introduces methods that instead maintain mutual consistency, but relax the best-response assumption.

2.2 Best response

Alternate approaches assume that a player may make errors when deciding which strategy to play, rather than playing perfectly rationally. That is, they relax the best response assumption. Quantal response equilibrium (McKelvey & Palfrey, 1995, 1998) (QRE) is a well-known example, where rather than choosing the best response with certainty, players choose noisily based on the payoff of the strategies and a resource parameter controlling this sensitivity. Another method of capturing this erroneous play is Noisy Introspection (Goeree & Holt, 2004). Utility proportional beliefs (Bach & Perea, 2014) is another method that relaxes the best response assumption, where the authors note that "possibly, the requirement that only rational choices are considered and zero probability is assigned to any irrational choice is too strong and does not reflect how real world [players] reason", giving merit to the relaxation of best-response. By allowing for errors in decision-making, these methods offer a more realistic perspective on how individuals make choices.

2.3 Infinite-regress

When considering reasoning about reasoning, infinite regress can emerge Knudsen (1993). The problem of infinite regress can be formulated as a sequence

$$A, f(A), f(f(A)), f(f(f(A)))$$

where the player is first confronted with an initial choice of an action a from the set of actions A, or a computation $f(A)$. If the action is chosen, the decision process is complete. If, instead, the computation is chosen, the outcome of $f(A)$ must be calculated, and the player is then faced with another choice between the obtained result or yet another computation. This process is repeated until the player chooses the result, as opposed to performing an additional computation. Lipman (1991) investigates whether such sequence converges to a fixed point. A similar approach by Mertens and Zamir (1985) formally approximates Harsanyi's infinite hierarchies of beliefs (Harsanyi, 1967, 1968) with finite states.

2.4 Contributions of our work

In contrast to existing works, we propose an information-theoretic approach to higher-order reasoning, where each level or hierarchy corresponds to additional information processing for the player. While Alaoui and Penta (2016) capture the trade-off between additional reasoning and payoff, by measuring the first intersection of the players' payoff improvement (from k to $k+1$), and the cost of performing this additional reasoning $c(k+1)$. This cost function c has to be determined (for all $c(k)$), for example, with maximum likelihood estimation to estimate the average cost of performing this extra level of reasoning (Alaoui & Penta, 2022). In contrast, we propose capturing this trade-off with information processing costs by using the Kullback–Leibler divergence to constrain the overall change in action probabilities at each stage of reasoning.

Our approach constrains the overall amount of information processing available to the players, leading to potential errors at each stage of reasoning, which is not present in existing level-k type approaches. By doing so, we establish a foundation for "breaking“ the chain of higher-order reasoning based on the depletion of players' information processing resources.

This results in a principled information-theoretic explanation for decision-making in games involving higher-order reasoning. Best response is relaxed with $\beta < \infty$ (linking to quantal response equilibrium) and mutual consistency is relaxed with $\gamma < 1$ (linking to level-k type models). Best response and mutual consistency are recovered with $\beta \to \infty$ and $\gamma = 1$. This contributes to the existing literature on game theory and decision-making by adopting an information-theoretic perspective on bounded rationality, quantified by information processing abilities. A key benefit of the proposed approach is while level-k models relax mutual consistency, but retain best response, and QRE models relax best response but retain mutual consistency (Chong et al., 2005), the proposed approach is able to relax either assumption through the introduction of two tunable parameters. We apply this model to various games, demonstrating the usefulness of the proposed approach for capturing human behaviour when compared to these existing approaches.

3 Technical preliminaries: information-theoretic bounded rationality

Information theory provides a natural way to reason about limitations in player cognition, as it abstracts away specific types of costs (Wolpert, 2006; Harre, 2021). This means that we can assume the existence of cognitive limitations without speculating about the underlying behavioural foundations. As pointed out by Sims (2003), an information-theoretic treatment may not be desirable for a psychologist, as this does not give insights to where the costs arise from, however, for an economist, reasoning about optimisation rather than specific psychological details may be preferable, for example, in the Shannon model (Caplin et al., 2019). Such models have seen considerable success in a variety of areas for information processing, for example, embodied intelligence (Polani et al., 2007), self-organisation (Ay et al., 2012) and adaptive decision-making of agents (Tishby & Polani, 2011) based on the information bottleneck formalism (Tishby et al., 1999).

In this work, we adopt the information-theoretic representation of bounded rational decision-making proposed by Ortega and Braun (2013), which has been further developed in Ortega and Braun (2011), Braun and Ortega (2014), Ortega and Stocker (2016), Gottwald and Braun (2019). This approach has a solid theoretical foundation based on the (negative) free energy principle and has been successfully applied to several tasks (Evans & Prokopenko, 2021). We begin by providing an overview of single-step decisions and then sequential decisions, before discussing extensions for capturing the relationship between processing limitations and higher-order reasoning.

3.1 Single-step decisions

A boundedly-rational decision maker who is choosing an action $a \in A$ with payoff $U[a]$ is assumed to follow the following negative free energy difference when moving from a prior belief $p[a]$, e.g., a default action, to a (posterior) choice $f[a]$, given by:

$$-\Delta F[f[a]] = \sum_{a \in A} f[a] U[a] - \frac{1}{\beta} \sum_{a \in A} f[a] \log\left(\frac{f[a]}{p[a]}\right) \tag{1}$$

The first term represents the expected payoff, while the second term represents a cost of information acquisition that is regularised by the parameter β. Formally, the second term quantifies information acquisition as the Kullback–Leibler (KL) divergence from the prior belief $p[a]$. Parameter β, therefore, serves as the resource allowance for a decision-maker.

By taking the first order conditions of Eq. (1) and solving for the decision function $f[a]$, we obtain the equilibrium distribution:

$$f[a] = \frac{1}{Z} p[a] e^{\beta U[a]} \tag{2}$$

where $Z = \sum_{a' \in A} p[a'] e^{\beta U[a',x]}$ is the partition function. This representation is equivalent to the logit function (softmax) commonly used in QRE models, and relates to control costs derived in economic literature (Mattsson & Weibull, 2002; Stahl, 1990).

The parameter β serves as a resource allowance for the decision-maker, modulating the cost of information acquisition from the prior belief. Low values of β correspond to high costs of information acquisition (and high error play), while as $\beta \to \infty$, information becomes essentially free to acquire, and the perfectly rational *homo economicus* player is recovered.

3.2 Sequential decisions

This free energy definition can be extended to sequential decision-making by considering a recursive negative free energy difference, as described in Ortega and Braun (2013, 2014). This corresponds to a nested variational problem, and involves introducing new inverse temperature parameters β_k for K sequential decisions, which allows for different reasoning at different depths of recursion to choose a sequence of K actions $a_{\leq K}$.

Therefore, for sequential decision-making, Eq. (1) can be represented as:

$$-\Delta F[f] = \sum_{a_{\leq K}} f[a_{\leq K}] \sum_{k=1}^{K} \left(U[a_k \mid a_{<k}] - \frac{1}{\beta_k} \log \frac{f[a_k \mid a_{<k}]}{p[a_k \mid a_{<k}]} \right) \tag{3}$$

where $a_{<k}$ abbreviates the history $a_o, .., a_{k-1}$ of decisions. We can expand the sum:

$$\begin{aligned} &= \sum_{a_1} f[a_1] \Bigg(U[a_1] - \frac{1}{\beta_1} f[a_1] \log \frac{f[a_1]}{p[a_1]} \\ &+ \sum_{a_2} f[a_2 \mid a_1] \Bigg(U[a_2 \mid a_1] - \frac{1}{\beta_2} f[a_2 \mid a_1] \log \frac{f[a_2 \mid a_1]}{p[a_2 \mid a_1]} \\ &+ \dots \\ &+ \sum_{a_K} f[a_K \mid a_{<K}] \left(U[a_K \mid a_{<K}] - \frac{1}{\beta_K} f[a_K \mid a_{<K}] \log \frac{f[a_K \mid a_{<K}]}{p[a_K \mid a_{<K}]} \right) \Bigg) \Bigg) \end{aligned} \tag{4}$$

which we can see as first choosing an action a_1 at $k = 1$, while considering that to choose this action, we must consider the future stages by analysing the result at $k = 2$ given the choice a_1, and so forth. To compute this, we can solve the innermost sum first:

$$f[a_K \mid a_{<K}] = \frac{1}{Z_K} p[a_K \mid a_{<K}] e^{\beta_K U[a_K \mid a_{<K}]} \tag{5}$$

which recovers Eq. (2) with the introduction of conditioning on decision histories. This represents the base-case for recursion. For steps where $k<K$, we get the following equilibrium solution for sequential decisions:

$$
\begin{aligned}
f[a_k \mid a_{<k}] &= \frac{1}{Z_k} p[a_k \mid a_{<k}] \exp\left(\beta_k \left(U[a_k \mid a_{<k}] + \frac{1}{\beta_{k+1}} \log Z_{k+1}\right)\right) \\
&= \frac{1}{Z_k} \underbrace{p[a_k \mid a_{<k}]}_{\text{PriorBelief}} \times \underbrace{Z_{k+1}^{\beta_k/\beta_{k+1}}}_{\text{FutureContribution}} \times \underbrace{e^{\beta_k U[a_k|a_{<k}]}}_{\text{CurrentPayoff}}
\end{aligned}
\tag{6}
$$

where decisions are now dependent on the history of decisions as well as on a recursive component based on the future contribution of each decision. Here

$$
Z_k = \sum_{a_k} p[a_k \mid a_{<k}] Z_{k+1}^{\beta_k/\beta_{k+1}} e^{\beta_k U[a_k|a_{<k}]} \tag{7}
$$

where $Z_K = 1$, i.e., the base-case for recursion at the final level.

3.3 Extension

With this formal and generic treatment of information-processing costs for sequential decision-making, it is desirable to use this to capture bounded rational reasoning of an agent making sequential *or* simultaneous decisions in games. By representing reasoning as an extensive-form game tree, these two can be captured in a similar manner. Simultaneous decisions can be treated as if they are pseudo-sequential decisions, considering possible repercussions for various choices.

To model higher-order reasoning, we extend the information-theoretic formulation for sequential decisions discussed in Sect. 3.2. This involves representing reasoning as a (pseudo) sequence of decisions, where each decision corresponds to a ”level“ of reasoning. At each level, players may play incorrectly, producing a level-k play that is modulated by β_k. A high β_k corresponds to an exact level-k thinker, while a low β_k corresponds to an error-prone level-k thinker. We can formalise this chain of reasoning as an extensive-form game tree, where at the root node, a player is faced with a decision to choose from a set of available options A or perform additional processing $f(A)$ to acquire new information on the beliefs and repercussions of each choice. If the player chooses not to process additional information, each branch is terminated early, and the player makes a decision based solely on the immediately available information. However, if the player chooses to process additional information, each action branch is (potentially) extended to analyze the possible repercussions of taking an action, and a higher level thinker can examine these repercussions. This process continues until the player runs out of processing resources or, in a finite problem, converges to a solution. For example, in the p-beauty contest analyzed in later sections, convergence occurs once the guess hits 0.

This extensive-form game tree representation means reasoning about simultaneous decisions can be treated in the same manner as decisions in sequential games. However, the key issue here is the requirement of Eq. (6) to have K information processing parameters for each step or level of reasoning. To analyse the purpose of these parameters, we consider a simple case, setting $\beta_k = \beta$ for all k, giving discrete control (Braun et al., 2011):

$$f[a_k \mid x, a_{<k}] = \frac{1}{Z_k} p[a_k \mid a_{<k}] Z_{k+1} e^{\beta U[a_k, x|a_{<k}]} \tag{8}$$

which clarifies that the boundedness applies to the computation of payoff U, but the depth of reasoning (or recursion-depth) is dependent on the length of the sequence (or level) K, not on β. To represent higher-order reasoning more succinctly, it would be desirable to implicitly base the sequence length K on the resource parameter β rather than keeping them separate. This would allow us to treat recursive reasoning in information-theoretic terms. For instance, when $\beta \to \infty$, $K \to \infty$, which captures (potential) infinite-regress, while $\beta = 0$ would imply $K = 0$, leaving the player with no information processing abilities.

In order to achieve this, β must be reduced at each level k. This reduction in β at each level is related to the relaxation of mutual consistency, reflecting how a player perceives other (lower-level) players' reasoning about the problem. In the next section, we outline our proposed approach for modelling this. This extension allows us to reason directly about resource constraints instead of sequential level-k thinking.

4 Proposed approach

The proposed approach implicitly enforces the assumption that, in sequential games, it becomes more difficult to reason to later stages in the game (e.g. in chess, it is difficult to reason 5 steps-ahead), and likewise in simultaneous games, it is difficult to perform the level of higher-order reasoning required to arrive at the equilibrium solution. The further a player mentally tries to reason, the more likely an error is to occur, as the processing resources deplete. This assumption is captured under the proposed model in a generic information-theoretic sense.

The key concept is that it becomes more difficult to reason about reasoning, that is the further one tries to explore through the extensive-form tree. This overall process is visualised Fig. 1, where the players reasoning error increases throughout the steps of reasoning. This representation can be thought of as a hierarchy of players noisily

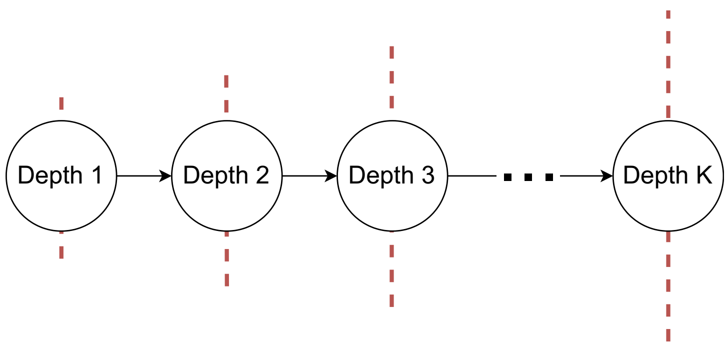


Fig. 1 The effect of discounting resources over time. In the beginning, the player has β resources and may make some error (red bars). Parameter γ controls how this error grows over time and, implicitly, how the player believes lower-level players will respond. A low γ means the errors increase drastically at each level, assuming opponents with much lower reasoning abilities. As $\beta\gamma^k \to 0$, the noise increases, and the recursion eventually stops once the utilities become indistinguishable (i.e., the player can not reason any deeper)

responding to lower level players, or as a sequential decision with increasing noise at each step. Once a player's resources deplete, the later depths simply echo the prior beliefs as the noise obstructs the payoffs.

4.1 Quantal hierarchy model

To model higher-order reasoning with processing costs, we propose a more flexible and succinct approach than simply setting a maximum depth K and corresponding parameters β_k in a pseudo-sequential decision-making task. Instead, we introduce an overall information processing parameter β, which captures the available information processing resources for a player, and a discount parameter $\gamma \in [0, 1]$, which controls the reduction in player rationality throughout the reasoning chain. This approach allows for heterogeneous bounds on player reasoning, relaxing the assumptions of best response and mutual consistency. We refer to this approach as the quantal hierarchy (QH) model, as it shares formal similarities with quantal response equilibrium and cognitive hierarchy models, as discussed in previous sections.

4.1.1 Formulation

We represent reasoning as information-constrained sequential decision-making. The proposed formulation features only two parameters, β and $\gamma \in [0, 1]$ (as opposed to the vector β_k and number of levels K). Reasoning resources are then set as

$$\beta_k = \beta\gamma^k$$

i.e., β_k is β discounted based on γ and the current depth of reasoning. This can be represented by the following recursive free energy difference:

$$-\Delta F[f] = \sum_{a_{\leq K}} f[a_{\leq K}] \sum_{k=0}^{\infty} \left(U[a_k \mid a_{<k}] - \frac{1}{\beta\gamma^k} \log \frac{f[a_k \mid a_{<k}]}{p[a_k \mid a_{<k}]} \right) \tag{9}$$

where we have represented the sequence as an infinite-sum. The sum converges due to the discount parameter as the later (inner-most sums) simply echo the prior beliefs once the player's computational resources are exhausted. This formalisation draws parallels with the reinforcement learning (RL) methods, where such representations are common for reasoning about future states for the player. In RL, γ is used to discount future timesteps. Here, γ is used to discount future chains of reasoning about reasoning (i.e., the depth of recursion is governed by β discounted by γ), and to represent the limited resources that we believe other players have. We have represented this as in infinite sum with $k \to \infty$, however, in various problems (such as sequential games) K can be assumed to be finite. The solution for the decision function $f[a_k \mid a_{<k}]$ with the discounted β becomes:

$$f[a_k \mid a_{<k}] = \begin{cases} \frac{1}{Z_k} \underbrace{p[a_k \mid a_{<k}]}_{\text{Prior Belief}}, & \text{if } \beta\gamma^k \approx 0 \\ \frac{1}{Z_k} \underbrace{p[a_k \mid a_{<k}]}_{\text{Prior Belief}} \times \underbrace{e^{\beta U[a_k|a_{<k}]}}_{\text{Current Payoff}}, & \text{if } \gamma = 0 \\ \frac{1}{Z_k} \underbrace{p[a_k \mid a_{<k}]}_{\text{Prior Belief}} \times \underbrace{Z_{k+1}^{1/\gamma}}_{\text{Future Contribution}} \times \underbrace{e^{\beta\gamma^k U[a_k|a_{<k}]}}_{\text{Current Payoff}}, & \text{otherwise} \end{cases} \tag{10}$$

which aims to capture decisions based on the beliefs of other players reasoning at later stages. With the assumption of a discount rate γ, we can see the recursion is now depth-bound based on the (discounted) resource parameter. Once $\beta\gamma^k \to 0$,[1] the recursion will stop since the result will simply echo the prior belief as no focus will be placed on the payoff, becoming the base case for recursion (the "naive" player). This means that in the new implicit limit K[2] (denoted by the case where $\beta_k \approx 0$), reasoning about the future provides no new information, which recovers the original form of Eq. (5) (when $\gamma = 1$) with the introduction of conditioning on histories.

With the proposed representation, what was previously thought of in the context of sequential decisions, can be extended and modified to think about hierarchies of beliefs. The informationally constrained players "break" the chain of reasoning due to depleting their cognitive or computational capabilities as bounded by β and γ. Formally, this is represented as a (potentially) infinite-sum that converges based on γ. By discounting future computation in a chain of recursion, we can approximate higher-order reasoning, where players become increasingly limited as the reasoning chain progresses, making it more challenging to reason about reasoning.

4.1.2 Parameters

Resource parameter The resource parameter β quantifies the amount of processing a player can perform. Perfect utility maximisation behaviour is recovered with $\beta \to \infty$. With $\beta < \infty$, players are assumed to be limited in computational resources and must now balance the trade-off between their computation cost and payoff. With $\beta \to 0$, players have no processing resources, and choose based on their prior beliefs (default actions). Anti-rational (or adversarial) play can be modelled with $\beta \to -\infty$.

Discount parameter The discount parameter γ quantifies other players mental processing abilities in terms of level-k thinking. A high γ assumes other players play at a relatively similar cognitive level, whereas a low γ assumes other players have less playing ability. With $\gamma < 1$, Eq. (10) is guaranteed to converge to a finite sequence of decisions, where the ability to process information decreases the further we get through the sequence. This captures the *belief* about play at later stages of reasoning, where other players are assumed to be less rational (and thus, more noisy) as governed by γ. The case $\gamma < 1$ implicitly relaxes mutual consistency as lower-level

[1] e.g. when $\beta\gamma^k < \epsilon$ where ϵ is some small enough term where the payoffs become indistinguishable, here, $\epsilon = 10^{-8}$.

[2] In contrast to level-k, here K indicates the level with the lowest resources.

thinkers are then governed by a lower resource parameter, and allows for players to believe that players at later nodes behave more noisily. In the case of otherwise infinite regress (where backward induction can not be used), the limited foresight approach proposed converges to a finite approximation of the sequence by relaxing the assumption of mutual consistency. In tractable problems, we recover an approximation of backward induction where the player performing such induction may make errors at each step due to limited computational processing abilities. With $\gamma = 1$, we recover mutual consistency, as we assume other players are just as rational as ourselves, and in the special case with uniform prior beliefs and $\gamma = 1$, we collapse to a logit form of agent QRE (McKelvey & Palfrey, 1998; Turocy, 2010). The special case for $\gamma = 0$ exists in Eq. (10) as in this case, no future processing will be performed.

In the limit, with $\beta \rightarrow \infty, \gamma = 1$, perfect backward induction is recovered (see Sect. B.1). Crucially, the proposed QH model allows for a general representation, relaxing the perfect rationality assumption (with $\beta < \infty, \gamma < 1$), which can model out-of-equilibrium behaviour compatible with observed experimental data. We explore the role of these parameter values in more detail in the following section.

Parameter interactions In Fig. 2 we visualise how β and γ interact in a general setting. For $\beta \rightarrow \infty$ and $\gamma = 1$, we approach payoff maximisation behaviour, i.e., the perfectly rational (Nash Equilibrium) player is recovered. For $\beta \rightarrow -\infty$ and $\gamma = 1$, payoff minimisation behaviour (an adversarial player) is recovered. In between, we can see how γ adjusts β. It is these values in between random play ($\beta = 0$) and perfect payoff maximisation behaviour which are particularly interesting, as they give rise to out-of-equilibrium behaviour not predicted by traditional methods.

4.2 Explanation

We work through a generic example of the QH model on an extensive-form game tree. A player is given decision-making resources, governed by β, to make a decision f. At each stage of reasoning k, the player's resources are discounted by γ. Ultimately, the player's resources become depleted (at K, i.e., once $\beta\gamma^k \approx 0$), and the game tree is considered terminated, and the naive player chooses based on their prior belief (which we assume to be uniform). This decision is then propagated backwards, and

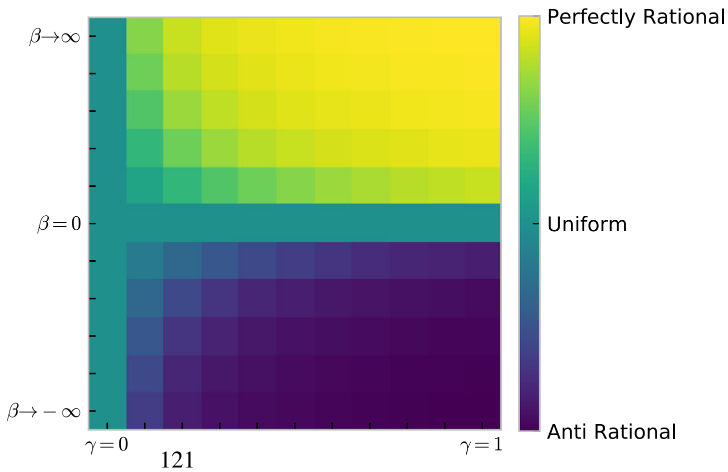


Fig. 2 A heatmap visualising the resulting players expected payoff based on the values of β and γ. The yellow colour represents the maximal expected payoff, and the purple colour represents the minimum expected payoff. When either parameter is 0, the result is a random choice amongst the actions

becomes the prior belief at the higher stage of reasoning $p[a_{K-1}]$, where the player has processing resources $\beta\gamma^{K-1}$, and hence, noisily responds to the lower-level play based on these resources. This process is continued, with noisy responses from the lower-level thinker captured by the resource constraint. Finally, once all results have been recursively propagated, the higher-level players' decision is made (which may still be noisy, as captured by β). That is, their decision is made recursively, starting from the most basic level of player reasoning (the naive player) and reasoning upwards.

4.2.1 Basic example

To help illustrate the proposed approach, we will use a simplified version of the Ultimatum game (see Sect. C.4) as a specific example. In this game, a player must decide what percentage of a pie to take. We assume that there are uniform priors among the available options at each stage.

At the first stage, Player 1 must request the percentage of the pie they want to take, denoted by $a_1 \in [0, 100]$, with 100 giving the highest payoff (i.e., they receive the entire pie). However, at the second stage, Player 1 encounters a fairness calculator (Player 2). Player 2 decides whether or not to approve Player 1's request, denoted by $A_2 = \{\text{accept}, \text{reject}\}$. The decisions are based on the following utilities:

$$\begin{aligned} U_1[a_1] &= a_1 \times f_2[\text{accept} \mid a_1] \\ U_2[\text{accept}_2 \mid a_1] &= 100 - a_1 \\ U_2[\text{reject}_2 \mid a_1] &= 50 \end{aligned} \tag{11}$$

where U_n corresponds to Player n's payoff, and f_n the probability with which Player n chooses the action. A player who has no look-ahead, i.e., one who assigns zero weight to future decisions (or assumes that their opponent has zero processing abilities), can be represented with $\beta \to \infty$ and $\gamma = 0$. Such a player simply looks at the first stage and sees that it is in their best interest to request 100% of the pie. However, this player fails to take into account the repercussions of their chosen action, as they did not consider the future decisions. They did not compute $f_2[a_2 \mid a_1]$, and thus assumed that $f_2[a_2 \mid a_1]$ is uniform and that their opponent would be indifferent to accepting or rejecting their request regardless of the value of a_1.

A perfectly rational player with unlimited computational resources, i.e., $\beta = \infty$ and $\gamma = 1$, would request 49 (assuming integer requests). They assign weight to the future of their actions, and can see that for any $a > 50$, the fairness calculator will deny their request, and they will be left with nothing (at $a = 50$, the calculator will be indifferent to accepting or rejecting their request). This corresponds to the subgame perfect equilibrium, where the player performed backward induction. That is, the player examined the future until they reached the end of the game and then reasoned backwards to request the optimal choice.

A player with limited computational resources, i.e., $\beta < \infty$, requests the best action they can subject to their resource constraint. For example, they may only request $a = 40$, as they are unable to complete the search for $a = 49$. A player with no information processing abilities, i.e., $\beta = 0$, cannot search for any optimal choices and therefore chooses based on the prior distribution, which we assumed to be uniform. Therefore, the player is equally likely to choose any $a \in A$.

An interesting question that arises is what if we assume that the ”fairness“ calculator may make errors, and that it is not necessarily defined by a step function that rejects all requests above 50 and accepts all below. For example, there may be a range where 100 will get rejected, but perhaps 75 would not. This can be captured with $\beta\gamma < \infty$, where the calculator is assumed to make errors for low values, and for $\beta\gamma \to \infty$, it is assumed to be perfectly rational. If the fairness calculator is broken and is indifferent to accepting or rejecting values, this can be represented with $\gamma = 0$, which gives 0 processing ability to the calculator. This means $f_2[\text{accept}_2 \mid a_1] = f_2[\text{reject}_2 \mid a_1] = 0.5$, and therefore, a rational Player 1 should request $a_1 = 100$.

This example shows the usefulness of the proposed approach, and how modifying β and γ can capture a variety of heterogeneous behaviours between the two players.

5 Results

In this section, we perform out-of-sample comparisons across various canonical economic games, including both sequential and simultaneous games. We compare the proposed quantal hierarchy model against well-known approaches to capturing bounded rational reasoning, including QRE, level-K and cognitive hierarchy, as well as the Nash equilibrium predicted solutions. To assess the performance of each method, we fit the corresponding parameter values to experimental data and then evaluate the performance on hold-out data. For the quantal hierarchy method, these parameter values are β and γ. For QRE, the parameter value is λ, which serves a similar purpose as β in our approach, i.e., relaxing best response. For level-k, the parameter value is the steps of reasoning k. For cognitive hierarchy, the parameter value is τ, corresponding to the Poisson distribution of level-k thinkers. Further information on model fitting is given in Sect. A.

We show how the proposed approach convincingly captures human behaviour and generalises beyond the training examples, outperforming existing approaches on a wide range of games.

5.1 Performance on canonical games

For this work, we use various experimental data from canonical economic sequential and simultaneous games. Specifically, for simultaneous games, we analyse market entrance and beauty contest games, and for sequential games, centipede, and bargaining games.

For market entrance games, we use the data of Camerer (2011), originally presented in Sundali et al. (1995). For the beauty contest game, we use p-beauty contest results from Bosch-Domenech et al. (2002). For the Centipede games, we use the four and six-level data from McKelvey and Palfrey (1992). For the sequential bargaining games, we use the Ultimatum game and two-stage game from Binmore et al. (2002) (Game 1 and 3 in their paper). Further discussion on game specifics, utilities, and experimental analysis is given in Sect. C.

For each game, we perform 5 $\times$ 2 repeated cross-fold validation (Dietterich, 1998), analysing the out-of-sample performance. This analysis ensures that the inclusion of an additional parameter does not overfit to the original training data, and instead, ensures the approach generalises well to unseen data. We present the average RMSE on the unseen data and the resulting rankings (Demšar, 2006) of each method in Table 1. The rankings account for the independence of the games, and the inability to compare errors directly across game classes. A visualisation of the resulting ranks in Fig. 3. By using these evaluation metrics, we are able to determine the effectiveness of the proposed method in comparison to existing approaches for predicting (out-of-sample) human behaviour on a range of canonical games.

The proposed quantal hierarchy method consistently performs well across the various games trialled, resulting in the best (lowest) overall rank (Table 1), as well as the most consistent (narrowest distribution of results, Fig. 3), always performing in the top 2. These results validate the modelling assumption that it becomes more difficult to reason at deeper levels of reasoning, and thus, the reasoning process becomes more erroneous. This motivates the usage of the quantal hierarchy model for capturing human decision-making in a wide-range of settings.

In the following subsections, we analyse the game results in more detail.

5.2 Simultaneous games

Market entrance

In the market entrance game, players must simultaneously decide whether to enter or stay out of a market, where the payoff depends on the decisions of the other players and market capacity c (see Sect. C.1). Experimental data show that player behaviour in market games is inconsistent with either mixed or pure Nash equilibria, although, with repeated play, players begin to approach the mixed strategy equilibrium (Duffy & Hopkins, 2005).

These deviations from equilibrium are captured well by the proposed quantal hierarchy model (Fig. 4). We see that in the beginning (before learning, Fig. 4a), the players overestimate for low c and underestimate for high c. Towards the final rounds (after learning, e.g. Fig. 4b), the behaviour approaches equilibrium, and the proposed QH model approximates this well using an increase in processing resources β and/or γ (see Table 2) to capture this player "learning". These changes highlight an important property of the QH model. If a player is learning, i.e., becoming closer to rational, this should correspond to an increase in β (and/or an increase in γ).

The level-k model fails to capture the overall trend, and is best fitted with $k = 0$, performing worse than the mixed strategy equilibrium (and all other alternatives). The reason for this is simple. Level-k ($k \geq 1$) implies a step function, where for $c > T$ where T is some threshold, the player enters with certainty, and for $c \leq T$, the player stays out with certainty. The distance to the experimental data from this step function is greater than the uniform case ($k = 0$), so the uniform case is chosen. The cognitive hierarchy model improves upon level-k, by fitting a distribution of k thinkers, able to

Table 1 Average out-of-sample (5 × 2-fold cross-validation) error

	Quantal hierarchy	Level-*k*	Cognitive hierarchy	QRE	Nash
Market entrance					
Block 1	0.565 (2)	1.503 (5)	0.725 (3)	0.564 (1)	1.242 (4)
Block 2	0.426 (1)	2.115 (5)	1.077 (4)	0.442 (2)	0.726 (3)
Block 3	0.387 (1)	2.230 (5)	1.283 (4)	0.406 (2)	0.548 (3)
Block 4	0.493 (2)	2.344 (5)	1.365 (4)	0.431 (1)	0.559 (3)
Block 5	0.489 (1)	2.513 (5)	1.387 (4)	0.519 (2)	0.574 (3)
Average rank	1.4	5	3.8	1.6	3.2
Beauty contest					
Lab	0.020 (1)	0.191 (4)	0.175 (3)	0.027 (2)	0.194 (5)
Classroom	0.042 (2)	0.188 (4)	0.162 (3)	0.019 (1)	0.190 (5)
Take Home	0.056 (2)	0.188 (4)	0.162 (3)	0.020 (1)	0.192 (5)
Internet	0.053 (2)	0.180 (4)	0.149 (3)	0.024 (1)	0.181 (5)
Newspaper	0.061 (2)	0.186 (4)	0.149 (3)	0.024 (1)	0.187 (5)
Theorists	0.071 (2)	0.171 (4)	0.135 (3)	0.040 (1)	0.172 (5)
Average rank	1.83	4	3	1.17	5
Centipede					
4-level	0.469 (1)	1.774 (4)	0.611 (3)	0.606 (2)	3.715 (5)
6-level	0.350 (1)	1.950 (4)	0.439 (2)	1.120 (3)	2.837 (5)
Average rank	1	4	2.5	2.5	5
Bargaining					
Ultimatum					
− (10, 10)	0.051 (1.5)	0.098 (3.5)	0.098 (3.5)	0.051 (1.5)	0.197 (5)
− (10, 60)	0.030 (1)	0.093 (3.5)	0.093 (3.5)	0.057 (2)	0.192 (5)
− (70, 10)	0.048 (2)	0.090 (3.5)	0.090 (3.5)	0.047 (1)	0.187 (5)
Two-stage					
$-D = 0.9$	0.040 (1)	0.096 (3.5)	0.096 (3.5)	0.084 (2)	0.198 (5)
$-D = 0.8$	0.054 (1)	0.095 (3.5)	0.095 (3.5)	0.076 (2)	0.198 (5)
$-D = 0.7$	0.048 (1)	0.099 (3.5)	0.099 (3.5)	0.075 (2)	0.197 (5)
$-D = 0.6$	0.067 (1)	0.128 (3.5)	0.128 (3.5)	0.095 (2)	0.197 (5)
$-D = 0.5$	0.037 (1)	0.111 (3.5)	0.111 (3.5)	0.054 (2)	0.190 (5)
$-D = 0.4$	0.030 (1)	0.105 (3.5)	0.105 (3.5)	0.039 (2)	0.191 (5)
$-D = 0.3$	0.024 (1.5)	0.081 (3.5)	0.081 (3.5)	0.024 (1.5)	0.192 (5)
$-D = 0.2$	0.052 (2)	0.117 (3.5)	0.117 (3.5)	0.050 (1)	0.196 (5)
Average rank	1.27	3.5	3.5	1.73	5
Overall					
Rank	1.37	4.12	3.2	1.75	4.55

Resulting ranks are indicated in brackets. In both cases, lower is better. Tied values receive the average rank between the ranks which would have been achieved had there been no ties. The overall rank is determined as the mean rank of the average rank across the game classes, meaning each game has an equal weighting in the overall rank, and the number of experiments for a game class does not affect this overall weight

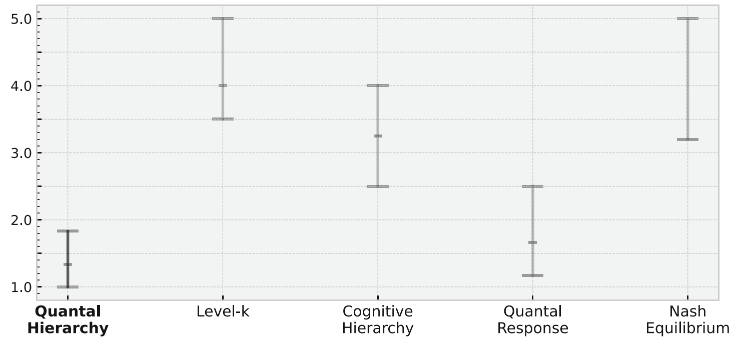


Fig. 3 Overall rankings for out-of-sample errors across the various game classes trialled. The vertical bars indicate the full range of achieved ranks across the games, with the horizontal middle bar indicating the median ranking. A lower ranking is better (with 1 = best)

"smooth" out this step function, with the line shown in Fig. 4. While this captures the qualitative trend (over entry for low c, under entry for high c, near equilibrium for mid c), quantitatively, the approach is not as strong as QRE, quantal hierarchy, or even the mixed-strategy equilibrium in most cases.

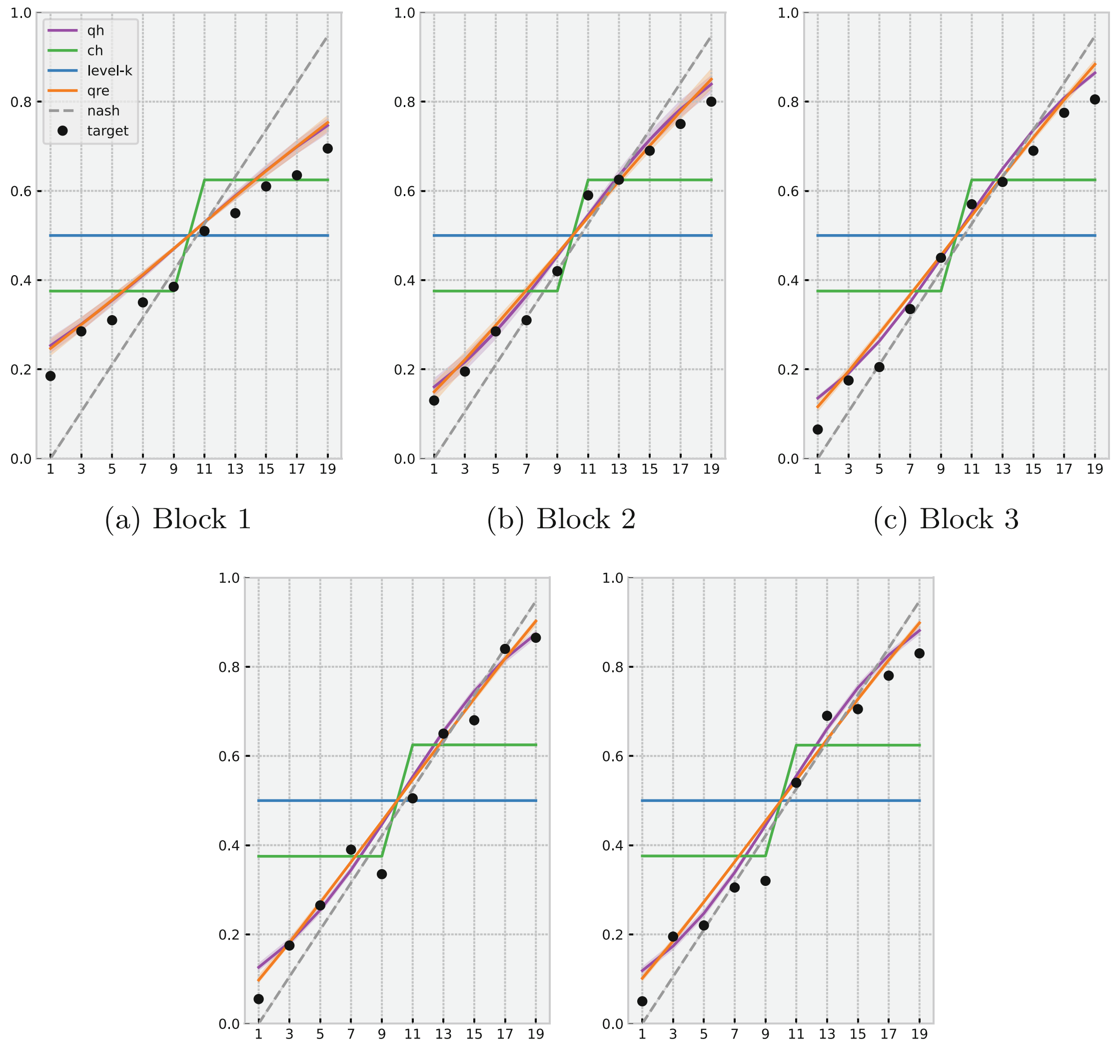
The QRE model is also a good fit here, however, due to the representation is constrained to linear lines. In contrast, the QH representation can capture such "S" shape curves, better approximating the experimental data in 3 out of the 5 blocks. QRE and QH significantly outperform the approaches which just relax mutual consistency (level-k and CH), motivating the relaxation of best-response in addition to mutual consistency.

***p*-Beauty contest**

In the p-beauty contest (Moulin, 1986), players must try and guess p times the average guess (in the range [0, 100]) of other competitors (see Sect. C.2). The Nash equilibrium is for all players to guess 0, however, experimentally, we see large deviations from this behaviour.

Analysing the experimental results (Fig. 5), we see very strong performance for the QH model when modelling the less experienced players, e.g. in the Lab experiments. The QH model fits the data well, capturing the overall distribution and achieving the lowest error rate. However, for the other experiments composed of more experienced players or players with more time (take home, newspaper), we see the distribution is better approximated by QRE.

The reason for this performance is because under the proposed approach, as a player becomes more rational, the distribution of choices narrows in to the optimal choice (see Fig. 6). However, under these experimental settings, even in the theorist case, there is bounded rational (and anti-rational) behaviour. These deviations are captured well by QRE. However, it is difficult for the QH model to capture the wide distribution of choices, as well as the bulk probability mass around the optimal case. For example, in Fig. 6 we show the proposed approach approximating level-k. As k increases, the distribution narrows. Here, this narrowing of the distribution makes it difficult to capture the entire prediction range for the more advanced subjects, due to the fact there are many sub-rational choices mixed in. As a result, we see similar fitted models for each case, despite the fact that the theorists clearly have a higher level of reasoning. If, instead, we tried to approximate the average player for each



(a) Block 1 (b) Block 2 (c) Block 3

(d) Block 4 (e) Block 5

Fig. 4 Market entrance game. The darker lines indicate the mean result from the 5 × 2 cross-validation. The shaded regions indicate ± one standard deviation. The out-of-sample data are shown as the black circles. The proposed quantal hierarchy model is the purple line. Quantal response equilibrium is the orange line, level-k is the blue line, and cognitive hierarchy is the green line. The Nash equilibrium solution is indicated as the diagonal dashed grey line

case, we could capture this increase in reasoning effort, and it would reflect an increase in β and/or γ as expected.

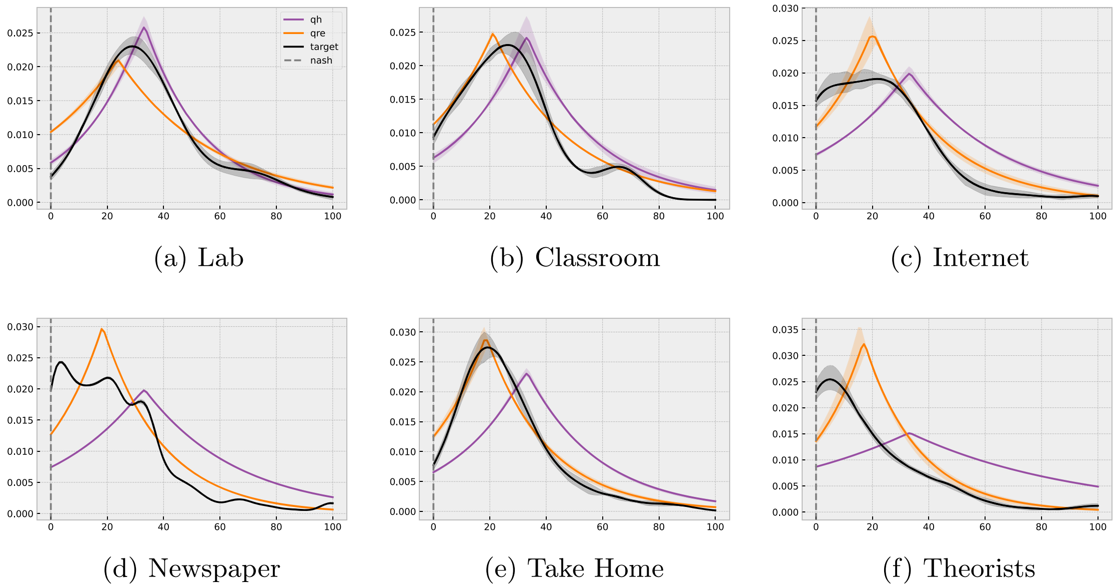
Nevertheless, in all cases, the model still significantly outperforms level-k, cognitive hierarchy, and the mixed strategy equilibrium. The level-k model predicts some of the representative spikes in the experimental data (e.g., with $k = 1$ guesses of 33, $k = 2$ of 22, etc.). However, we can see that the players do not necessarily choose according to level-k, and may make errors around the best response suggested by level-k reasoning. Level-k thinking presupposes that players will predict a multiple of p, i.e., with $p = \frac{2}{3}$, we get $p \times 50, p^2 \times 50, \ldots, p^k \times 50$, as the players at each level are best responding to lower-level players. In the proposed QH model, level-k reasoning can be recovered if $\beta_t = \infty$ for $t \leq k$ and $\beta_t = 0$ for $t > k$.

Table 2 Average fitted parameter values for each approach. Explanation of the parameters and the fitting procedure is given in Sect. A

	Quantal hierarchy β, γ	Level-k k	Cognitive hierarchy τ	QRE λ	Nash
Market entrance					
Block 1	0.43, 0.24	0.0	0.72	10.39	
Block 2	0.67, 0.18	0.0	0.68	34.56	
Block 3	0.62, 0.28	0.0	0.65	59.44	
Block 4	0.32, 0.54	0.0	0.69	83.51	
Block 5	0.75, 0.19	0.2	0.72	80.8	
Beauty contest					
Lab	0.08, 0.76	1.2	5.45	1.21	
Classroom	0.1, 0.69	1.7	5.52	1.9	
Take Home	0.06, 0.79	2.4	5.36	2.2	
Internet	0.07, 0.72	3.3	5.67	2.31	
Newspaper	0.08, 0.64	6.3	5.76	2.8	
Theorists	0.05, 0.67	5.6	5.84	3.02	
Centipede					
4-level	12.43, 0.22	0.0	1.82	2.09	
6-level	19.1, 0.14	0.3	2.3	1.09	
Bargaining					
Ultimatum					
– (10, 10)	0.08, 0.92	0.0	4.36	0.08	
– (10, 60)	0.2, 0.32	0.0	4.84	0.09	
– (70, 10)	0.06, 0.88	0.0	3.68	0.06	
Two-stage					
$-D = 0.9$	0.24, 0.13	0.0	4.07	0.04	
$-D = 0.8$	0.2, 0.22	0.0	4.29	0.05	
$-D = 0.7$	0.22, 0.28	0.0	3.73	0.06	
$-D = 0.6$	0.52, 0.2	0.0	3.89	0.08	
$-D = 0.5$	0.2, 0.36	0.0	2.97	0.11	
$-D = 0.4$	0.19, 0.38	0.0	3.78	0.1	
$-D = 0.3$	0.13, 0.49	0.0	3.69	0.08	
$-D = 0.2$	0.17, 0.52	0.0	8.02	0.11	

However, with $\beta_t < \infty$, the proposed QH model produces a distribution around these best-responding values anticipating potential errors in player reasoning, with these errors growing throughout the chain of reasoning.

The cognitive hierarchy model can improve upon the level-k approach here by weighting the “spikes“ of the level-k model differently, however, this still fails to capture the underlying distribution. A large reason for this is that certain predictions in the p-beauty contest are considered irrational, for example, any prediction over 67. However, we can see experimentally that such predictions occur, for example, in Fig. 5. If a player

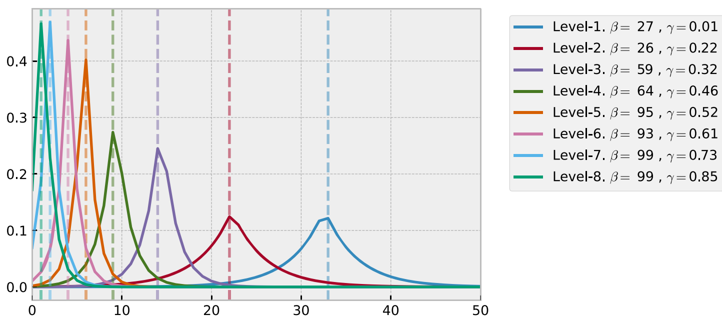


(a) Lab (b) Classroom (c) Internet

(d) Newspaper (e) Take Home (f) Theorists

Fig. 5 Beauty contest games. The darker lines indicate the mean result from the 5×2 cross-validation. The shaded regions indicate ± one standard deviation. The out-of-sample data are shown as the black line. The proposed quantal hierarchy model is the purple line. Quantal response equilibrium is the orange line. The level-k and cognitive hierarchy plots are shown in Fig. 12 due to the large difference in scales, distorting the figure. The Nash equilibrium solution is indicated as the diagonal dashed grey line

Fig. 6 Example comparing the decisions of level-k (dashed vertical lines) to the proposed QH model (solid lines) in the p-beauty contest for various settings



believes that other players would choose the maximal offer of 100, then the player should choose $\frac{2}{3}100 = \ 67$. Level-k or cognitive hierarchy models cannot capture such irrational behaviour where players choose >67. That is, there is no distribution of level-k thinkers that would predict 100. However, this feature can be captured directly under the proposed QH and QRE models due to the errors in play, again motivating the usefulness of relaxing best response in addition to mutual consistency.

In summary, we see that under the lab experiment, the QH approach is the best fit. However, QRE is a better fit in other cases of the beauty contest game.

5.2.1 Sequential games

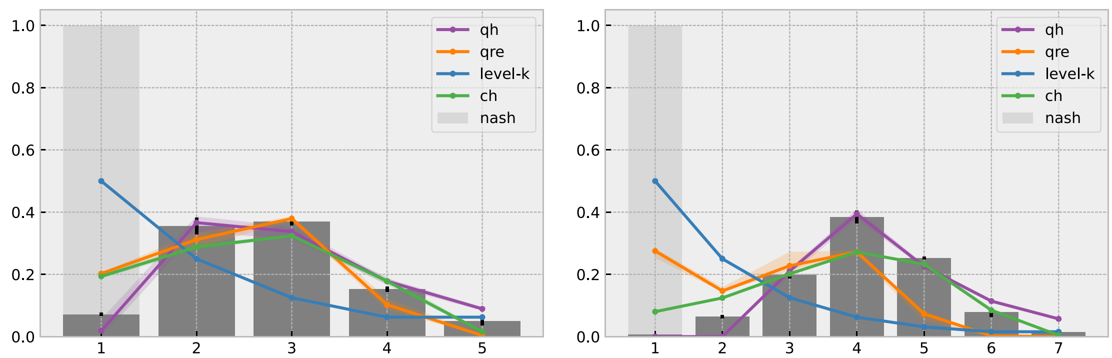
Centipede games

In the centipede game (see Sect. C.3), "two players alternately get a chance to take the larger portion of a continually escalating pile of money. As soon as one person takes, the game ends with that player getting the larger portion of the pile, and the other player getting the smaller portion" (McKelvey & Palfrey, 1992).

The subgame perfect equilibrium of the centipede game is for each player to immediately take the pot without proceeding to any further rounds, however, we see this is not the case experimentally, where players behave far from the subgame perfect equilibrium (Fig. 14). In general, many players take towards the middle of the game. The proposed QH model can capture this trend well, with QRE and cognitive hierarchy generally over-weighting the earlier nodes and under-weighting the later modes (Fig. 7).

The quantal hierarchy model provides the best fit for both the four and six-level centipede games, capturing realistic beliefs. When modelling this reasoning process, the player believes they are reasoning at a higher level than their opponent, but in addition, it is *as if* the player overestimates how noisy their own play will be when faced with a decision at later nodes. This overestimation is because once actually faced with the decision, there will be a smaller game tree for the player to consider. This reasoning process was shown to approximate the experimental results well, motivating the discounting of information processing resources for capturing future beliefs. When comparing the resulting parameters (β and γ) from the four and six-level variants (Table 2), we note that the six-level variant results in additional information processing costs for the player (larger β and γ). The additional processing costs result from the longer chain of reasoning, requiring higher processing resources.

The significantly improved performance over quantal response equilibrium on both games motivates the usefulness of relaxing mutual consistency in addition to best response. By relaxing mutual consistency, we captured the perceived "lapse" in reasoning when considering the full extensive form game tree by reducing the information processing resources the further the player tries to reason through the tree.



(a) Four Move. (b) Six Move.

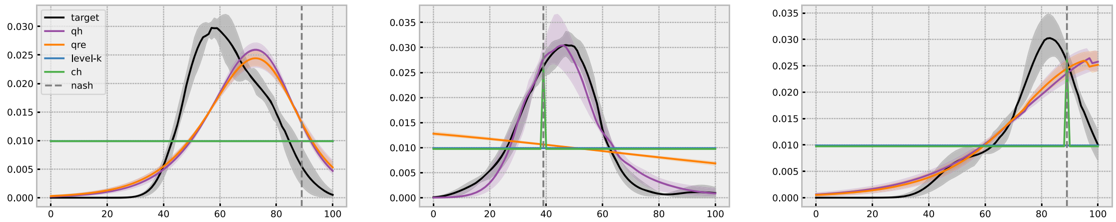
Fig. 7 Four and six-level centipede games. The darker lines indicate the mean result from the 5×2 cross-validation. The shaded regions indicate ± one standard deviation. The out-of-sample data are shown in the dark grey bars. The proposed quantal hierarchy model is the purple line. Quantal response equilibrium is the orange line, level-k is the blue line, and cognitive hierarchy is the green line. The Nash equilibrium solution is indicated as the light grey bar at the first move

Bargaining games

In bargaining games, players alternately bargain over how to divide a sum (see Sect. C.4). We examine two types, single-stage (Ultimatum) and two-stage bargaining games. These are extensions of the example game considered in Sect. 4.2.1.

Ultimatum The results are presented in Fig. 8 for the ultimatum game. We see the QH model explains important features of the observed experimental behaviour. For example, with higher V_1, Player 1 is likely to make a larger initial request (Fig. 8a vs 8c). Perfect rationality does not capture this (with the Nash equilibrium remaining unchanged), whereas the QH model suggests higher initial requests (due to higher V_1 if the request is rejected). The QRE model and the QH model behave similarly here. The reason for this similar behaviour is because the ultimatum game is nearly a single-stage decision, meaning the QH model "collapses" to QRE. This is also confirmed in the fitted parameters Table 2, with the (10, 10) and (70, 10) having almost identical values for β and λ, and QH having relatively high values of γ, meaning the differences between Player 1 and Player 2 processing resources are small. However, something interesting happens in the (10, 60) case, where QRE cannot capture the distribution. For example, if we look to the right of the rational rejection region for Player 2 (Fig. 16), we do not see any acceptances in (10, 10) or (70, 10). Whereas, if we look in the rational rejection region of (10, 60), we see several acceptances. This behaviour is irrational, because if the player rejected the request, they would have received a higher payoff. In contrast, Player 1's initial requests are relatively rational, with the peak occurring around the rational request of 40. This mismatch in player rationality is captured under the proposed model with a small γ, i.e., a large discount in processing resources. This mismatch in rationality cannot be captured with the standard QRE, which assumes a fixed β for both players. These results motivate the discounting of player resources, which can capture heterogeneous information processing resources between the two players. Alternate forms of QRE, such as Heterogenous QRE have also been proposed to deal with such dilemmas (Rogers et al., 2009), however, this is captured natively by the QH model.

The level-k model fails to capture any of the trends, with the uniform level-0 case being the best fit. The cognitive hierarchy model predicts a representative spike at the



(a) (10,10) (b) (10,60) (c) (70,10)

Fig. 8 Ultimatum game. The darker lines indicate the mean result from the 5 × 2 cross-validation. The shaded regions indicate ± one standard deviation. The out-of-sample data are shown as the black line. The proposed quantal hierarchy model is the purple line. Quantal response equilibrium is the orange line, level-k is the blue line, and cognitive hierarchy is the green line. The Nash equilibrium solution is indicated as the vertical dashed grey line rational capacity in the (70, 10) and (10, 60) cases, however, it is still clearly outperformed by both QRE and QH.

These results confirm the usefulness of QH. When both players behave with similar levels of rationality, this can be captured with $\gamma \to 1$, and the model acts the same as QRE. However, when there is a mismatch in player rationality, this heterogeneity can be captured directly with $\gamma < 1$, which became most pronounced in the (10, 60) case.

Two-stage While we saw similar behaviour between QRE and QH in the ultimatum game, under the two-stage game, the differences between the approaches become more pronounced due to the longer game tree. Under such conditions, the usefulness of discounting future paths (and relaxing mutual consistency) becomes more noticeable. The QH model convincingly outperforms QRE across the experimental results with the small disagreement penalties[3] (i.e., $D > 0.5$), and still generally outperforms QRE for the larger disagreement penalties (i.e., $D \leq 0.5$), although the two methods become closer.

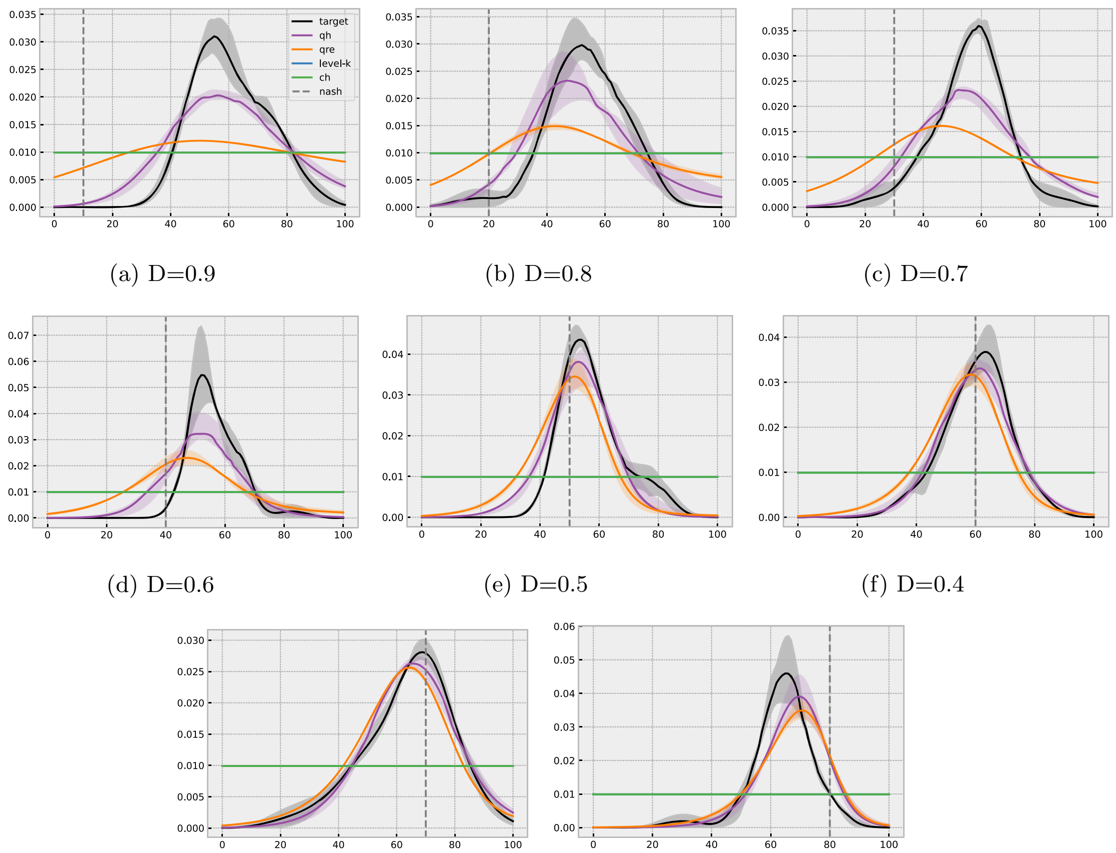
With the larger disagreement penalties (i.e. smaller D), the experimental data are closer to the perfectly rational case, as indicated with the peaks corresponding roughly to the rational request in Fig. 9. This distribution around the rational request is precisely the premise QRE is founded on, so QRE achieves adequate performance. However, with the smaller disagreement penalties (larger D, top row of Fig. 9), the distribution is not centred around the rational request, meaning QRE struggles to capture such phenomena. In contrast, the QH approach is robust to this shift due to the relaxation of mutual consistency, and is able to capture the varying distributions regardless of whether they are approximating the best-response case.

5.3 Results summary

The quantal hierarchy method consistently performed well out-of-sample in all games, ranking the best overall and achieving either the first or second position in every game. The results analysis was categorised into two game types: sequential and simultaneous games, where reasoning is represented as an extensive-form game tree with depleting information-processing resources. Although the representation worked well in both game types, it showed more improvement over alternative methods in sequential games. This improvement in sequential games can be attributed to the discount parameter that captures the heterogeneity of players, allowing for different information processing resources between the players at each stage, relaxing mutual consistency, which was crucial in bargaining games.

On the other hand, in simultaneous games, the approach aims to fit a representative distribution of the entire group, but it can struggle to capture the entire distribution of players, particularly when they exhibit widely varying levels of rationality, as in certain versions of the beauty-contest game. This highlights a potential limitation of the approach when attempting to capture multimodal distributions with varying levels of rationality, such as a bi-modal distribution with

[3] These are referred to as "discount" rates in Binmore et al. (2002). We have used the term disagreement penalties to avoid confusion with the information processing "discount" parameter γ.



(a) D=0.9 (b) D=0.8 (c) D=0.7

(d) D=0.6 (e) D=0.5 (f) D=0.4

(g) D=0.3 (h) D=0.2

Fig. 9 Two-stage bargaining game. The darker lines indicate the mean result from the 5 × 2 cross-validation. The shaded regions indicate ± one standard deviation. The out-of-sample data are shown as the black line. The proposed quantal hierarchy model is the purple line. Quantal response equilibrium is the orange line, level-*k* is the blue line, and cognitive hierarchy is the green line. The Nash equilibrium solution is indicated as the vertical dashed grey line

beginners and experts. To address this limitation, multiple versions of the model may need to be fitted, such as one for beginners and one for experts, as modelling more rational play narrows the distribution to the rational prediction and modelling less rational play widens the distribution to account for larger errors, as demonstrated in Fig. 6. However, despite this potential limitation, the method still performed exceptionally well overall.

6 Discussion and conclusions

The assumption of perfect rationality amongst players is violated in numerous experimental settings, particularly in non-repeated games. In this work, we utilised experimental datasets for several games, showing that the equilibrium behaviour is often a poor predictor of the observed actions. The proposed quantal hierarchy model offers a concise alternative representation, relaxing some traditional game-theoretic assumptions underlying rationality. The model is a good fit for experimentally observed behaviour on a range of canonical economic games, outperforming existing bounded rationality approaches on out-of-sample validation.

In the QH model, we represent higher-order reasoning as pseudo-sequential decision-making. At each level, players may reason erroneously, and this error grows the deeper one reasons (i.e., it becomes more difficult to reason about reasoning). The magnitude of the errors is governed by β, with $\beta = 0$, players do not perform any reasoning, and with $\beta \to \infty$ players reason perfectly. Parameter β, therefore, relaxes the best response assumption of players at each level of reasoning.

Decreasing β at each level of reasoning was shown to work well on a wide variety of games, reinforcing the assumption that players cognitive abilities decrease throughout the depth of reasoning. This reduction in player cognition is captured with γ, introducing an implicit hierarchy of players, relaxing the mutual consistency assumption. Representing this hierarchy of players as extensive-form game trees allowed for an information-theoretic representation, where lower-level players are assumed to make more significant playing errors (constrained by lower information processing resources). With a single-step decision, this recovers the quantal response equilibrium model. With multi-stage decisions, we recover an approximation of a generalised level-k formulation, where at each step, players are assumed to have higher resources and reasoning ability than players below themselves, but may still play erroneously.

Similar to QRE, the resource parameter β is problem dependent, and depends on the payoff magnitude (McKelvey et al., 2000). This opens an area of research analysing whether a normalised β can be used to measure problem difficulty, or whether some relationship holds between the the experimentally fitted β and the β which corresponds to the Nash solution. For example, a question arises if a normalised β can provide insights across games, and if so, can this average distance to the Nash solution be generally useful across games. A similar consideration is given to whether such payoff perturbations in QRE can be related across different games (Haile et al., 2008), and whether the boundedness parameter can be endogenised (Friedman, 2020).

There is a clear relationship between the decision-making components proposed in this work and the decision-making in multi-agent systems, such as agent-based models (ABMs) and multi-agent reinforcement learning (RL) approaches. For example, Wen et al. (2020) outline a novel framework for hierarchical reasoning RL agents, which allows agents to best respond to other less sophisticated agents based upon level-k type models. Likewise, Łatek et al. (2009) propose a recursion based bounded rationality approach for ABMs. Replacing the agents in these multi-agent approaches with the informationally constrained agents presented in our work provides a distinct area of future research, where we could examine the resulting dynamics and out-of-equilibrium behaviour from heterogeneous QH agents.

In summary, we proposed an information-theoretic model for capturing higher-order reasoning for boundedly rational players. Bounded rationality is achieved in the model by the relaxation of two central assumptions underlying rationality, namely, mutual consistency between players and best response decisions. Through relaxing these assumptions, we showed how the predictions from the proposed quantal hierarchy model align well with the experimentally observed human behaviour in a variety of canonical economic games.

Appendix A Model fitting Each model is fit to the training portion of the data (from 5 $\times$ 2 cross-fold validation), using Bayesian hyperparameter optimisation (Bergstra et al., 2013). Every model is given 1000 evaluations for a fair comparison. With level-k, due to the integer parameter, rather than Bayesian optimisation, we instead perform an exhaustive search for $k \in [0, 1, \ldots, 100]$, noting that this is an extensive range of k, easily capturing standard k's reported in the literature. The fitted parameter values with the lowest mean squared error between the predictions and the training values are selected. The out-of-sample (testing) portion is never seen by the optimisation process and is only used for evaluation after the parameter optimisation has been complete.

For quantal response equilibrium and quantal hierarchy, we sample from the range $0 \leq \beta < 100$. While β is unbounded, we find this upper bound to be more than enough with no fitted values coming close to this upper threshold. For γ, this is bounded $0 \leq \gamma \leq 1$. For cognitive hierarchy, we sample from the range $0 \leq \tau < 10$, which despite τ being unbounded, again provides a more than sufficient range for the experimental data, and covers common τ's reported in literature (Camerer, 2010; Camerer et al., 2004).

For the beauty contest games, as well as the bargaining games, due to the large action space ($a \in [0, \ldots, 100]$), rather than using the raw data directly, a fitted Gaussian kernel density estimate of the training and testing data is used to account for the large action space and the relatively small number of observations. Scott's rule is used to determine the bandwidth automatically (Scott, 2015), and we validate the robustness of this rule choice in Sect. D.2. The same kernel density estimates are used across all methods to ensure fair comparisons. For the remaining game classes, the action space is sufficiently well sampled from the observations, so no density approximation is required.

Appendix B Special cases

B.1 Backwards induction

Backwards induction can be recovered as a limiting case of the proposed model. We can see this as follows from Eq. (10) (and the expansion process from Eq. (4)), noting that softmax $e^{\beta U[a]}/Z$ converges to $\arg\max$ with $\beta \to \infty$:

$$\begin{aligned} f[a_k \mid a_{<k}] &= \frac{1}{Z_k} \times \underbrace{Z_{k+1}^{1/\gamma}}_{\text{Future Contribution}} \times \underbrace{e^{\beta\gamma^k U[a_k|a_{<k}]}}_{\text{Current Payoff}} \\ &= \underset{a_k \in A_k}{\arg\max} \left(\underbrace{U[a_k \mid a_{<k}]}_{\text{Current Payoff}} + \underbrace{V_{k+1}}_{\text{Future Contribution}} \right) \end{aligned} \tag{12}$$

where V_{k+1} is derived recursively based on choosing a_k. Backward induction assumes that all future decisions will be rational, i.e., at each stage, players choose rationally.

Appendix C Canonical games and experimental data

C.1 Market entrance/El Farol bar problem

The market entrance game was outlined in the context of cognitive hierarchies in Camerer et al. (2004), and has also been considered in prior studies, e.g., Rapoport et al. (1998). This game is fundamentally similar to the El Farol bar problem of Arthur (1994), and minority games of Challet et al. (2013). A player will profit (enjoy) in the market (bar) if less than $d \times N, d \in [0,1]$ players also enter the same market (bar).

For this work, we use the experimental data from Camerer (2011), specifically the results originally presented in Sundali et al. (1995). The payoff for staying out is fixed

$$U[\text{stay out}_k] = 1 \tag{13}$$

However, the payoffs for entering are dependent on the total demand from the other (lower-level) players and a preferential capacity $c = d \times N$:

$$U[\text{enter}_k] = 1 + 2(c - f[\text{enter}_{k+1}]) \tag{14}$$

There were $N = 20$ subjects, and various c's trialled $c \in [1, 3, 5 \ldots, 19]$.

C.1.1 Comparison methods

We can represent this pseudo-sequential structure (Camerer et al., 2004) as an extensive-form game, with each level of reasoning forming a new node in the game tree.

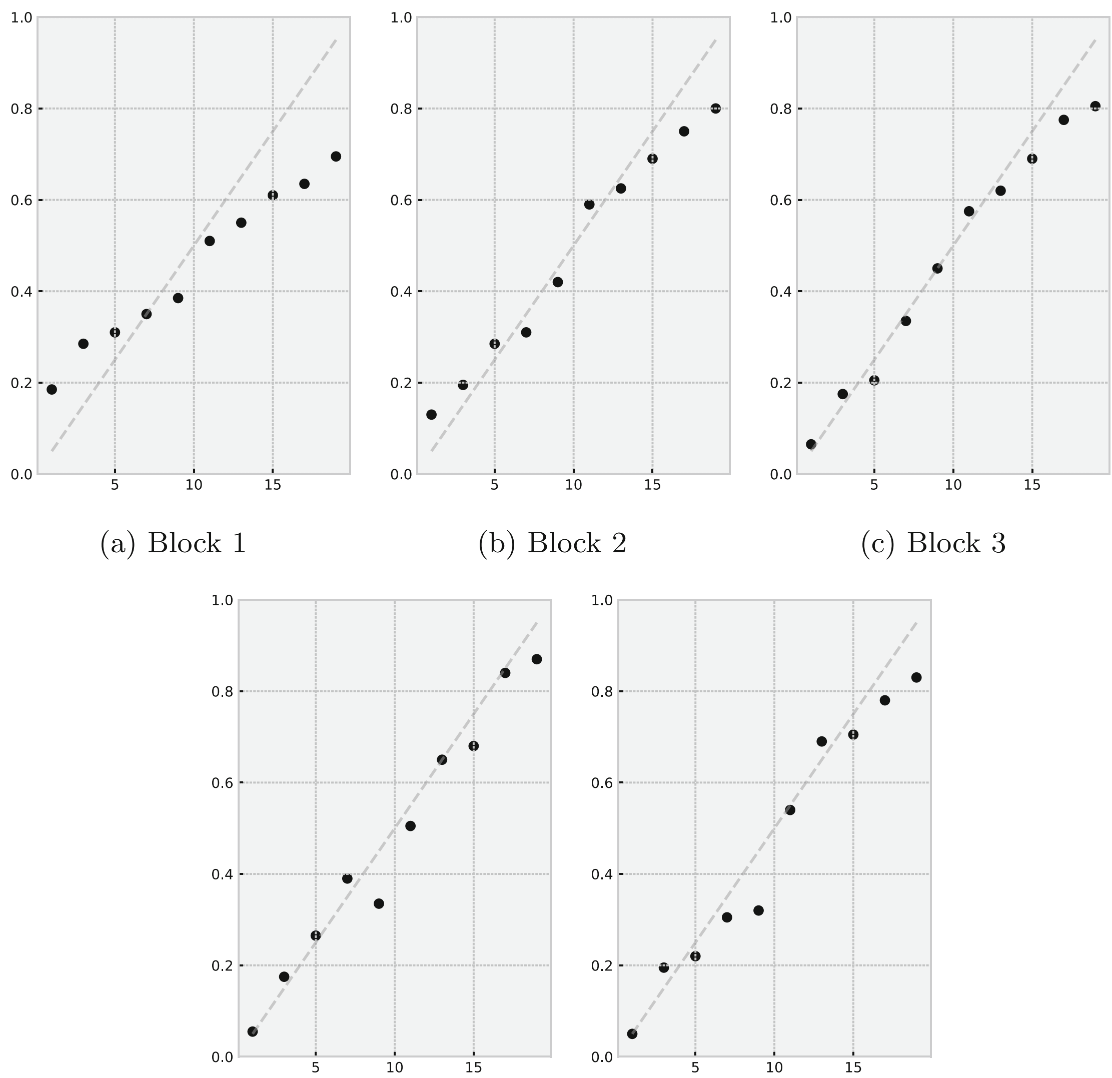
Level-k Under this configuration, a level-0 player is assumed to randomise, i.e., enter or stay out with equal probability (the same level-0 configuration is used for the naive player in the QH model). A level-1 player exploits this and attends the bar if $d > 0.5$, or stay home with $d < 0.5$, at $d = 0.5$ the player is indifferent and would attend with 50% probability. Level-2 players then base their decision assuming other players are level-1, and enter only if the level-1 players underestimated the expected capacity. Likewise, level-3 players base their decision on reasoning about level-2. Level-k behaviour necessitates step functions in the response, where players only enter at a capacity c if they believe lower-level thinkers have over or under entered.

Cognitive hierarchy Rather than assuming all players are at $k-1$, the cognitive hierarchy model fits a distribution to these k players, and best responds according to this distribution of lower level thinkers. Following Camerer et al. (2004), we use the Poisson distribution.

Quantal response equilibrium To derive the (mixed strategy) quantal response equilibrium, we use the logistic function of the differences in payoffs (between enter and stay out) as the distribution function, with numeric estimations for the fixed-point solution since no analytical solution exists, following Goeree et al. (2016, Section 2.2.2 and Section 8) and Goeree and Holt (2005) (Fig. 10).

C.2 Beauty contest

Keynes (1937, 2018) originally formulated the beauty contest game as follows. Contestants are asked to vote for the six prettiest faces out of a selection of 100. The winner is the contestant who most closely picks the overall consensus. A naive (level-0) strategy is to choose based on personal preference. A level-1 strategy is to choose as if everyone is choosing on personal preference, so the player chooses whom they think others will find most desirable. A level-2 strategy is then for players



(a) Block 1 (b) Block 2 (c) Block 3

(d) Block 4 (e) Block 5

Fig. 10 Market entrance game at various blocks using the experimental data from Sundali et al. (1995). We see over time, the experimental data becomes closer to the equilibrium of perfect attendance (dotted gray line) to choose whom they think that others will think others will choose, and so on (with Keynes believing there are players who “practise the fourth, fifth, and higher“ levels). The game was originally outlined to highlight how investors are not necessarily driven by fundamentals but rather by anticipating the thoughts of others.

An extension of the game is the p-beauty contest of Moulin (1986), where contestants are asked to guess fraction $p \in [0...1]$ (commonly $p = \frac{2}{3}$) of the average value of the other competitors guesses within the range $[0, ...100]$. The Nash equilibrium dictate that every player should choose 0. However, experimentally this is not the case, and players act boundedly rational (Nagel, 1995). Such a game shows out-of-equilibrium behaviour and motivates the modelling of such decisions in a finite-depth manner (Aumann, 1992; Binmore, 1987, 1988; Stahl, 1993).

We use the experimental data provided by Bosch-Domenech et al. (2002) with $p = \frac{2}{3}$. The resulting guesses are visualised in Fig. 11.

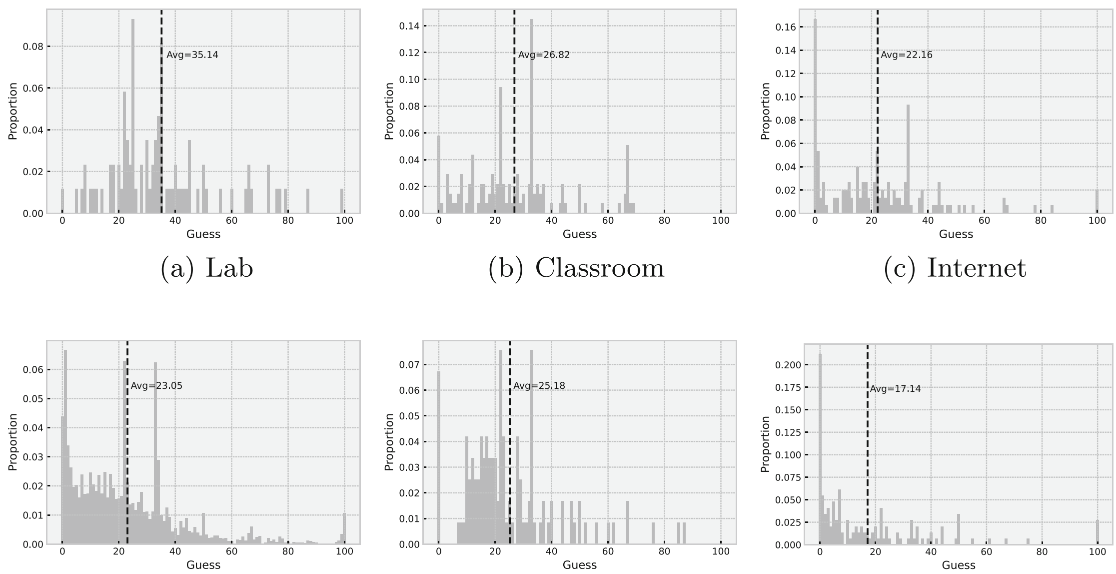
The utilities are represented as follows:

$$\begin{aligned} g_k &= p \times \frac{\sum_{a_{k+1}} a_{k+1} \times f[a_{k+1}]}{\sum_{a_{k+1}} f[a_{k+1}]} \\ U[a_k] &= |a_k - g_k| \end{aligned} \tag{15}$$

where g_k represents the predicted goal, i.e., 2/3’s of the average weighted prediction of the lower level thinkers. The utilities for each choice then become the distance to the goal.

C.2.1 Comparison methods

Level-k Level-0 competitors are assumed to guess randomly between [0, 100] (the same level-0 configuration is used for the naive player in the QH model). Level-1



(a) Lab (b) Classroom (c) Internet

(d) Newspaper (e) Take Home (f) Theorists

Fig. 11 Visualisation of various experimental p-beauty contests. The dotted vertical line indicates the average for the given dataset. Datasets source: Bosch-Domenech et al. (2002)

players then anticipate this and guess $p \times 50$ (50 being the average from the level-0 players), level-2 players then guess $p \times (p \times 50)$ and so forth. As the levels increase, the guesses, therefore, approach 0, coinciding with the perfectly rational choice.

Cognitive hierarchy Rather than assuming all players are at $k - 1$, the cognitive hierarchy model fits a distribution to these k players, and best responds according to this distribution of lower level thinkers. Again, following the convention of Camerer et al. (2004), we use the Poisson distribution.

Quantal response equilibrium For the quantal response equilibrium, we use the logit rule following Breitmoser (2012) estimated using fixed point iteration (see also Section F4.1 of the supplementary material for Anufriev et al. (2022)) (Fig. 12).

C.3 Centipede game

With perfectly rational backward induction, the subgame perfect equilibrium of the centipede game is for each player to immediately take the pot without proceeding to any further rounds. However, this is a poor predictor of what happens experimentally (Ho & Su, 2013), where players are shown to “grow“ the money pile by playing for several rounds before taking (Ke, 2019). Again, there are multiple reasons proposed to explain players deviation from the predicted unique subgame equilibrium (Georgalos, 2020; Kawagoe & Takizawa, 2012; Krockow et al., 2018).

In this work, we use the experimental data of McKelvey and Palfrey (1992) (from their Appendix C) for four and six-level centipede games. The utilities are represented as:

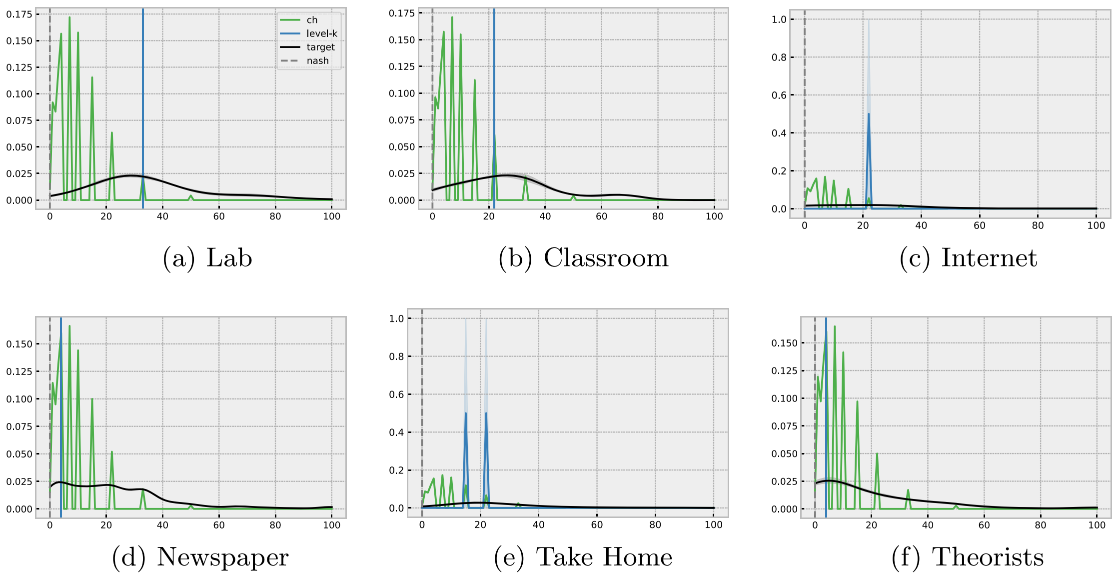


Fig. 12 Beauty contest games (extended). The darker lines indicate the mean result from 5×2 cross-validation. The shaded regions indicate ± one standard deviation. The out-of-sample data are shown as the black line. The level-k model is shown as the blue line. The Cognitive hierarchy model as the green line. The Nash equilibrium solution is indicated as the diagonal dashed grey line

$$
\begin{aligned}
U_1[\text{take}_1] &= 0.4 \\
U_1[\text{pass}_1] &= 0.2 \times f[\text{take}_2] + V_1 \times f[\text{pass}_2] \\
U_2[\text{take}_2] &= 0.8 \\
U_2[\text{pass}_2] &= 0.4 \times f[\text{take}_3] + V_2 \times f[\text{pass}_3] \\
&\;\;\ldots
\end{aligned} \tag{16}
$$

where V_1, V_2 are derived as the average expected return for the remainder of the moves. These payoffs are also visualised in Fig. 13. The conditioning on the history of decisions is implicit here, as to take an action a_n at $n > 1$, all previous actions must have been to pass (otherwise the game would have ended) (Fig. 14).

C.3.1 Comparison methods

Level-k

Under the level-k framework, we assume a level-0 agent is equally as likely to take or pass at each stage of the game (the same configuration is used as the naive player under the proposed quantal hierarchy approach). A level-1 player then takes at the node which maximises the expected utility subject to this, and so on and so forth. A full analysis of level-k framework in centipede games is presented in Kawagoe and Takizawa (2012), but this configuration used [referred to as Random Behavioral strategy (RBS) in Kawagoe and Takizawa (2012)] was shown to be the best specification for matching the experimental data (for both level-k and cognitive hierarchy).

Cognitive hierarchy Rather than assuming all players are at $k-1$, the cognitive hierarchy model fits a distribution to these k players, and best responds according to this distribution of lower level thinkers. Again, the Poisson distribution was used, which was shown to be the best experimental fit in Kawagoe and Takizawa (2012).

Quantal response equilibirum An agent-form of the QRE (McKelvey & Palfrey, 1998) is used here, where at each node, the agent choices nosily based on resource parameter β, and assuming their opponent is also operating under the same resource constraint β. This is calculated recursively following (McKelvey & Palfrey, 1998).

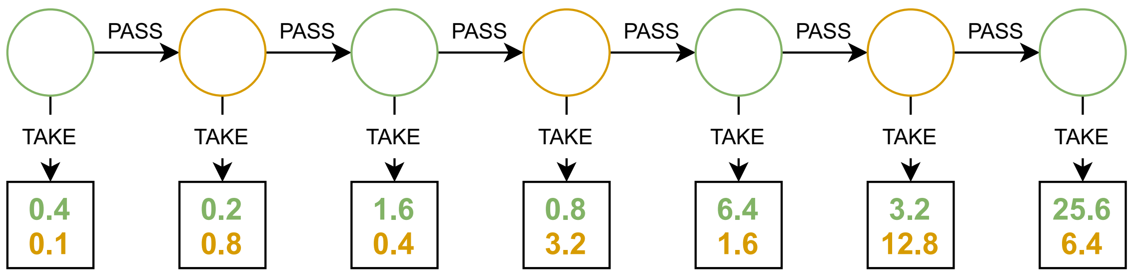


Fig. 13 The six move extensive-form centipede Game. Green (orange) circles highlight Player 1(2)’s turn, and the top (bottom) row of the boxes highlights Player 1(2)s payoff. The four-move game is equivalent until the fourth node, however, the payoffs for the fifth node become the “PASS” payoffs for Node 4

C.4 Sequential bargaining

We examine the experimental results of the ultimatum game (one-stage) and two-stage alternating-offer bargaining games (Binmore et al., 2002), which consistently demonstrate violations of backward induction (Webster, 2013), even when accounting for "fairness" in the system (Johnson et al., 2002).

C.4.1 Ultimatum game

For the ultimatum game, we use the experimental data of Game 1 from Binmore et al. (2002). In the ultimatum game, the players are faced with the following payoffs:

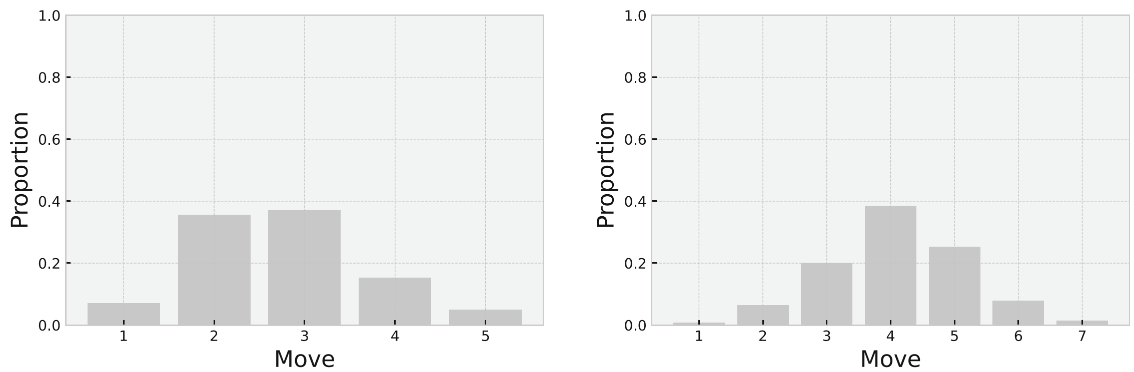
$$\begin{aligned} U_1[a_1] &= a_1 \times f[\text{accept} \mid a_1] + V_1 \times f[\text{reject} \mid a_1] \\ U_2[\text{accept}_2 \mid a_1] &= 100 - a_1 \\ U_2[\text{reject}_2 \mid a_1] &= V_2 \end{aligned} \tag{17}$$

where V_1, V_2 are the rejection payoffs for Player 1 and Player 2.

If opponents (Player 2) are rational, then Player 1, being rational, should request no more than their opponent's rejection payoff V_2. However, if opponents are not believed to be rational, then there is potential for Player 1 to exploit this fact and request higher (or lower) amounts. That is, it becomes rational for Player 1 to play as if Player 2 is not perfectly rational.

A rational opponent implies a step function, where for a_1, with payoff $100 - a_1 > V_2$ the player accepts with probability 1, and for payoffs below V_2 the player reject with certainty as shown in Fig. 16. However, from Fig. 16 we can see deviations from rationality in the observed play. This is directly shown by non-deterministic outputs, where the players may or may not accept the request based on the requested value, as well as violations where the players reject or accept with probability 1 even if a rational actor would do the opposite.

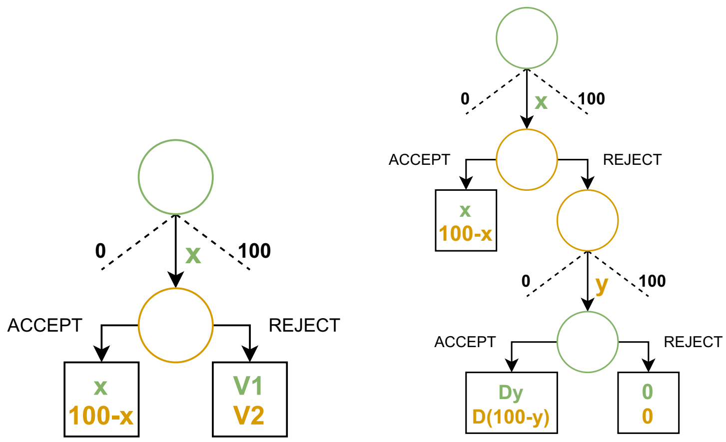
Now, knowing the opponent has potential bounds on their rationality, a rational player would respond accordingly. Figure 17 plots the distribution of Player 1



(a) Four Move.

(b) Six Move.

Fig. 14 Four and six-level centipede games using the dataset from McKelvey and Palfrey (1992)



(a) Ultimatum (One-stage) (b) Two-Stage

Fig. 15 Example extensive-form sequential bargaining games. In the ultimatum game (Fig. 15a), Player 1 makes a request $x \in [0, 100]$. If Player 2 accepts the request, Player 2 receives a payoff of $100 - x$, and Player 1 receives x. If Player 2 declines the request, they each receive the rejection payoff (V_1 or V_2). In the two-stage game, if Player 2 rejects, they can come back with a counteroffer y. Now the process repeats, and it is up to Player 1 to accept or reject. If Player 1 accepts, Player 2 gets a disagreement penalised (D) payoff of $D(100 - y)$, and Player 1 gets a payoff of Dy. However, if both decline they each get a payoff of 0

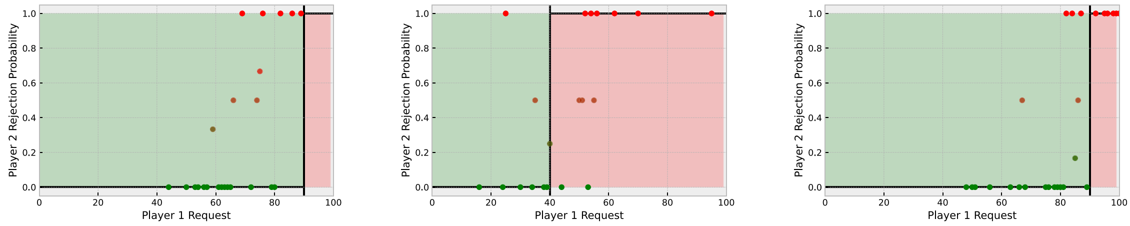
requests. We observe that Player 1's request still deviates from perfect rationality. Perfect rationality would imply a Dirac delta function with the probability mass situated at the optimal request. The observed deviation from rationality may be due to uncertainty in their opponent's abilities (reflected in the probabilities from Fig. 16), or limitations of Player 1's reasoning.

C.4.2 Two-stage bargaining

Next, we examine a two-stage bargaining game (shown in Fig. 15b). Now, if Player 2 rejects Player 1's request, they can come back with a counteroffer of their own. If the players can not come to an agreement, they both receive 0. It is, therefore, in both players best interest to reach an agreement. This is represented with the following utilities:

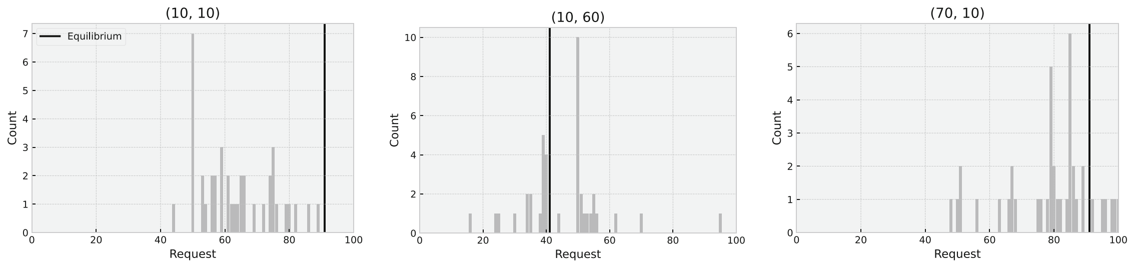
$$\begin{aligned} U_1[a_1] &= a_1 \times f[\text{accept} \mid a_1] + V_1 \times f[\text{reject} \mid a_1] \\ U_2[\text{accept}_2 \mid a_1] &= 100 - a_1 \\ U_2[\text{reject}_2 \mid a_1] &= V_2 \\ U_2[a_3] &= D(100 - a_3) \times f[\text{accept} \mid a_3] \\ U_1[\text{accept}_4 \mid a_3] &= D \times a_3 \\ U_1[\text{reject}_4 \mid a_3] &= 0 \end{aligned} \tag{18}$$

where now V_1 and V_2 are derived from the expected payoff of the rejection branch. Conditioning on past decisions are excluded from $U_2[a_3]$ as it is implicit that this can only occur when Player 2 rejects Player 1's request. We use experimental data of Game 3 from Binmore et al. (2002), considering all disagreement penalties



(a) (V_1=10, V_2=10) (b) (V_1=10, V_2=60) (c) (V_1=70, V_2=10)

Fig. 16 Observed rejection rates from experimental data of Binmore et al. (2002). A rational opponent is governed by the step function (black line), where a rational player would reject in the red area and accept in the green area. The observed points show deviations from rationality



(a) (V_1=10, V_2=10) (b) (V_1=10, V_2=60) (c) (V_1=70, V_2=10)

Fig. 17 Player 1 requests in the ultimatum game, with experimental data from Binmore et al. (2002). The black line indicates the perfectly rational choice (when assuming opponent is perfectly rational), i.e., a rational opponent would reject any lower request (left of the black line), and would accept any value above their rejection payoff (right of the black line)

$D \in [0.2, 0.3, \ldots, 0.9]$. For discussion sake, here we assume $D = 0.9$. With perfect rationality, it is in Player 1's best interest to accept any counteroffer greater than the rejection payoff of 0 (see Fig. 15b). Therefore, if prompted, Player 2 should provide a counteroffer of $y = 1$, which gives Player 2 a payoff of $D(100 - y) = 89.1$, and Player 1 a payoff of $Dy = 0.9$. Since $Dy > 0$ (the rejection payoff), Player 1 should prefer this to the alternative and accept. With this in mind, Player 1 now knows the payoff for the rejection branch for Player 2 is 89.1, so if they request $x > 10$ (assuming integer requests), Player 2 will reject this request since $100 - x < 89.1$ if $x > 10$. Therefore, the rational Player 1 requests $x = 10$, maximising their payoff, assuming Player 2 is rational.

However, from Fig. 18 which summarises results from the experiments presented by Binmore et al. (2002), we can see substantial deviations from the subgame perfect equilibrium for both players. No Player 1 requests $x < 10$ or the perfectly rational request $x = 10$. Furthermore, no Player 2 requests the rational counteroffer of $y = 1$. The distribution of initial and counteroffers is visualised in Fig. 19.

C.4.3 Comparison methods

Level-k Under the level-k model, level-0 players are assumed to be indifferent to all choices, and chooose uniformly. Level-1 players exploit this, and choose based on their opponent being a level-0 player, and so on.

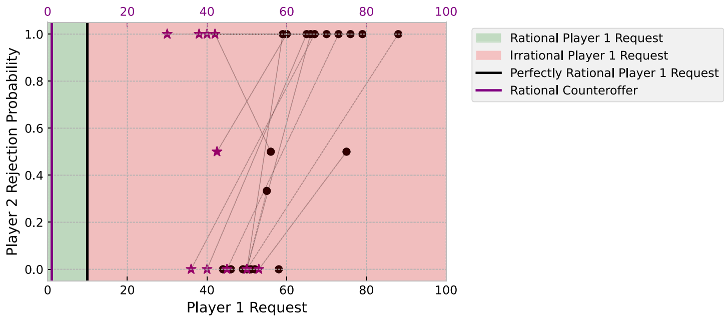


Fig. 18 Two-stage bargaining with disagreement penalty $D = 0.9$. Initial requests are shown as black circles. The y-position gives the rejection rate of these requests. For rejected requests, the counteroffers are shown as a purple star (linked to their original request by a line), where again, the y-position shows the rejection rate of the counteroffer. The perfectly rational initial request would be $x = 10$ (black line), as any requests in the red region would be rejected by a rational opponent. After a rejection, the perfectly rational counteroffer would be $y = 1$ (purple line). Deviations from the subgame perfect equilibrium are clear, with no player performing perfect backward induction

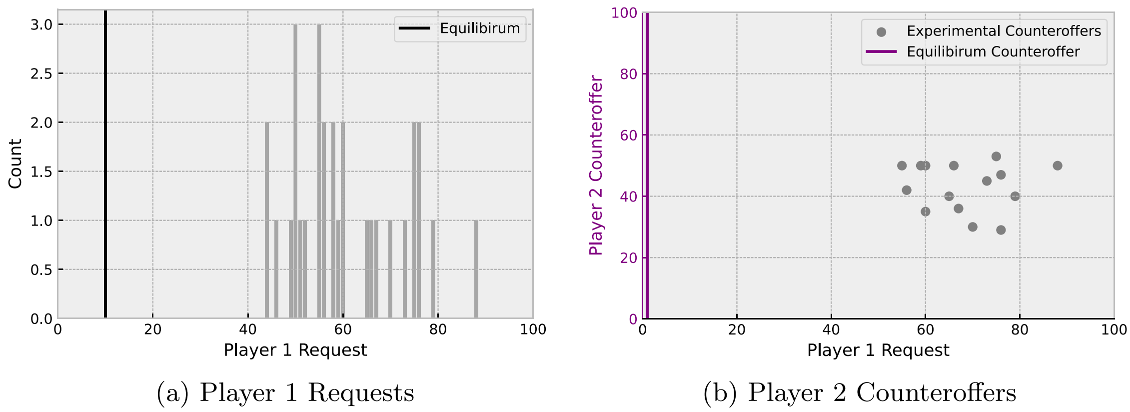


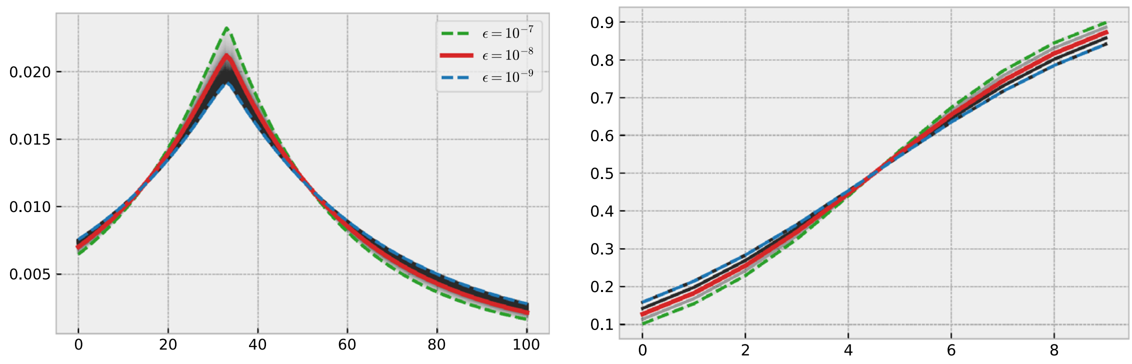
Fig. 19 Initial requests and counteroffers in the two-stage bargaining game with $D = 0.9$

Cognitive hierarchy Similar to the other game classes, again rather than assuming all players are at $k-1$, the cognitive hierarchy model fits a distribution to these k players, and best responds according to this distribution of lower level thinkers. Again, the Poisson distribution is used.

Quantal response equilibrium For the quantal response equilibrium, again, an agent form of QRE is used to account for players noisily responding at each level, which is calculated recursively from the final step.

Appendix D Sensitivity

To ensure the method's robustness, we check the sensitivity of the proposed results to various factors that may affect the outcome. Specifically, we carried this testing out with respect to the convergence/termination parameter ϵ and the fitted density estimates.



(a) Beauty Contest

(b) Market

Fig. 20 Termination parameter ϵ sensitivity. The outcome for the default value is displayed as the red line. The outcome for the upper (lower) threshold is the dashed blue (green) bar. Intermediary values are displayed as light grey lines

D.1 Termination parameter ϵ

For the bargaining and centipede games, the reasoning naturally ends at the end of the extensive-form game. However, for the market entry and beauty contest games (with no defined end point), the reasoning continues until the resources are depleted, i.e., $\beta\gamma^k < \epsilon$. We have used the threshold $\epsilon = 10^{-8}$ to determine termination. To check the robustness of the method to ϵ, here we perform sensitivity analysis across the range $10^{-7} < \epsilon < 10^{-9}$, i.e. $\pm$ one order of magnitude from the default value. We sample 1000 points uniformly from this range, presenting the results in Fig. 20.

In both cases, we can see the approach is robust to these large changes in ϵ, with an order of magnitude change only having slight effects on the resulting outcomes. These results show that ϵ does not need to be treated as a hyperparameter to optimise (as β/γ), but rather as a fixed parameter to determine “convergence“ towards 0 and termination, the choice of which depends on computational/numeric requirements. We recommend using as small value as possible (e.g. $\epsilon \leq 10^{-8}$) while still achieving reasonable convergence speed.

D.2 Density estimates

We evaluate how the resulting rankings would change with different density estimation methods. Specifically, we analyse the resulting average (out-of-sample) ranks when using Scott’s rule (Scott, 2015) (as presented), Silverman’s rule (Silverman, 2018), and the (improved) Sheather & Jones (Botev et al., 2010) for automatic bandwidth identification. While this does not provide an exhaustive list, it covers the most common rules used in literature. Density estimates are only used for the bargaining and beauty contest games, so these are the two game classes analysed here.

In Table 3, we see no change in average ranking between Scott’s and Silverman’s rules for the beauty contest games. However, when using Sheather & Jones, the rankings between Level-k and Nash change, going from 4 and 5, respectively, to 4.33

Table 3 Robustness to changes in density estimation. Average rankings are computed using the results from various automatic bandwidth determination methods. Scott's is the method we present in the paper

	Quantal hierarchy	Level-k	Cognitive hierarchy	QRE	Nash
Beauty contests					
Scott's	1.83	4.00	3.00	1.17	5.00
Silverman's	1.83	4.00	3.00	1.17	5.00
Sheather & Jones	1.83	4.33	3.00	1.17	4.67
Bargaining games					
Scott's	1.27	3.50	3.50	1.73	5.00
Silverman's	1.23	3.50	3.50	1.77	5.00
Sheather & Jones	1.23	3.50	3.50	1.77	5.00

and 4.67. This rank change does not alter any of the claims made within the paper, so we can confirm the robustness of the resulting rankings to density estimates for the beauty game.

For the bargaining games, we see slight improvement for the proposed method when comparing Scott's rule with Silverman's and Sheather & Jones–in both cases, going from 1.27 with Scott's rule to 1.23. At the same time, QREs rank worsens from 1.73 to 1.77. The remaining methods keep the same ranking. These rank changes strengthen the claims made in the paper, showing not only robustness to the rule used but also improvements for the proposed approach when utilising alternative rules for bandwidth estimation.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This work was supported by the Australian Research Council Discovery Project [Grant number DP170102927].

Data availability The experimental data analysed during the current study are publicly available from the original studies referenced. Specifically, data for bargaining games are presented in Binmore et al. (2002). Data for the beauty contest games are presented in Bosch-Domenech et al. (2002). Data for the centipede games are presented in McKelvey and Palfrey (1992). Data for the market entrance games are presented in Sundali et al. (1995), Camerer (2011). All data are subject to the original availability statements from the original authors.

Declarations

Conflict of interest No competing interests to declare. The authors have no relevant financial or non-financial interests to disclose.

References

Alaoui, L., & Penta, A. (2016). Endogenous depth of reasoning. *The Review of Economic Studies, 83*(4), 1297–1333.

Alaoui, L., & Penta, A. (2022). Cost-benefit analysis in reasoning. *Journal of Political Economy, 130*(4), 881–925.

Anufriev, M., Duffy, J., & Panchenko, V. (2022). Learning in two-dimensional beauty contest games: Theory and experimental evidence. *Journal of Economic Theory, 201*, 105417.

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *The American Economic Review, 84*(2), 406–411.

Aumann, R. J. (1992). Irrationality in game theory. *Economic Analysis of Markets and Games,* 214–227.

Ay, N., Bernigau, H., Der, R., & Prokopenko, M. (2012). Information-driven self-organization: The dynamical system approach to autonomous robot behavior. *Theory in Biosciences, 131*(3), 161–179.

Bach, C. W., & Perea, A. (2014). Utility proportional beliefs. *International Journal of Game Theory, 43*(4), 881–902.

Bergstra, J., Yamins, D., & Cox, D. (2013). Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *International Conference on Machine Learning*, pp. 115–123. PMLR.

Binmore, K. (1987). Modeling rational players: Part I. *Economics & Philosophy, 3*(2), 179–214.

Binmore, K. (1988). Modeling rational players: Part II. *Economics & Philosophy, 4*(1), 9–55.

Binmore, K., McCarthy, J., Ponti, G., Samuelson, L., & Shaked, A. (2002). A backward induction experiment. *Journal of Economic theory, 104*(1), 48–88.

Bosch-Domenech, A., Montalvo, J. G., Nagel, R., & Satorra, A. (2002). One, two,(three), infinity,...: Newspaper and lab beauty-contest experiments. *American Economic Review, 92*(5), 1687–1701.

Botev, Z. I., Grotowski, J. F., & Kroese, D. P. (2010). Kernel density estimation via diffusion. *The Annals of Statistics, 38*(5), 2916–2957.

Braun, D. A., & Ortega, P. A. (2014). Information-theoretic bounded rationality and ε-optimality. *Entropy, 16*(8), 4662–4676.

Braun, D. A., Ortega, P. A., Theodorou, E., & Schaal, S. (2011). Path integral control and bounded rationality. In *2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, pp. 202–209. IEEE.

Breitmoser, Y. (2012). Strategic reasoning in p-beauty contests. *Games and Economic Behavior, 75*(2), 555–569.

Camerer, C., Ho, T., & Chong, K. (2003). Models of thinking, learning, and teaching in games. *American Economic Review, 93*(2), 192–195.

Camerer, C. F. (2003). Behavioral game theory: Plausible formal models that predict accurately. *Behavioral and Brain Sciences, 26*(2), 157–158.

Camerer, C. F. (2010). Behavioural game theory. *Behavioural and experimental economics* (pp. 42–50). Springer.

Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics, 119*(3), 861–898.

Caplin, A., Dean, M., & Leahy, J. (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies, 86*(3), 1061–1094.

Challet, D., Marsili, M., & Zhang, Y.-C., et al. (2013). Minority games: Interacting agents in financial markets. *OUP Catalogue*.

Chong, J.-K., Camerer, C. F., & Ho, T.-H. (2005). Cognitive hierarchy: A limited thinking theory in games. In R. Zwick & A. Rapoport (Eds.), *Experimental business research* (pp. 203–228). Springer US.

Chong, J.-K., Ho, T.-H., & Camerer, C. (2016). A generalized cognitive hierarchy model of games. *Games and Economic Behavior, 99*, 257–274.

Cournot, A. A. (1838). *Recherches sur les principes mathématiques de la théorie des richesses*. L. Hachette.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research, 7*, 1–30.

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation, 10*(7), 1895–1923.

Duffy, J., & Hopkins, E. (2005). Learning, information, and sorting in market entry games: Theory and evidence. *Games and Economic Behavior, 51*(1), 31–62.

Evans, B. P., & Prokopenko, M. (2021). A maximum entropy model of bounded rational decision-making with prior beliefs and market feedback. *Entropy, 23*(6), 669.

Friedman, E. (2020). Endogenous quantal response equilibrium. *Games and Economic Behavior, 124*, 620–643.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review, 96*(4), 1043–1068.

Gächter, S. (2004). Behavioral game theory. In *Blackwell handbook of judgment and decision making*, pp. 485–503.

Georgalos, K. (2020). Comparing behavioral models using data from experimental centipede games. *Economic Inquiry, 58*(1), 34–48.

Glasner, D. (2022). Hayek, Hicks, Radner and four equilibrium concepts: Perfect foresight, sequential, temporary, and rational expectations. *The Review of Austrian Economics, 35*(1), 39–61.

Goeree, J. K., & Holt, C. A. (2004). A model of noisy introspection. *Games and Economic Behavior, 46* (2), 365–382.

Goeree, J. K., & Holt, C. A. (2005). An explanation of anomalous behavior in models of political participation. *American Political Science Review, 99*(2), 201–213.

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2005). Regular quantal response equilibrium. *Experimental Economics, 8*(4), 347–367.

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2016). Quantal response equilibrium. *Quantal response equilibrium.* Princeton University Press.

Goldfarb, A., & Xiao, M. (2011). Who thinks about the competition? Managerial ability and strategic entry in us local telephone markets. *American Economic Review, 101*(7), 3130–61.

Gottwald, S., & Braun, D. A. (2019). Bounded rational decision-making from elementary computations that reduce uncertainty. *Entropy, 21*(4), 375.

Haile, P. A., Hortaçsu, A., & Kosenok, G. (2008). On the empirical content of quantal response equilibrium. *American Economic Review, 98*(1), 180–200.

Harré, M. S. (2021). Information theory for agents in artificial intelligence, psychology, and economics. *Entropy, 23*(3), 310.

Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players. Part I. The basic model. *Management Science, 14*(3), 159–182.

Harsanyi, J. C. (1968). Games with incomplete information played by "Bayesian" players. Part II. Bayesian equilibrium points. *Management Science, 14*(5), 320–334.

Ho, T.-H., & Su, X. (2013). A dynamic level-k model in sequential games. *Management Science, 59*(2), 452–469.

Hoppe, H.-H. (1997). On certainty and uncertainty, or: How rational can our expectations be? *The Review of Austrian Economics, 10*(1), 49–78.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory, 104*(1), 16–47.

Kawagoe, T., & Takizawa, H. (2012). Level-k analysis of experimental centipede games. *Journal of Economic Behavior & Organization, 82*(2–3), 548–566.

Ke, S. (2019). Boundedly rational backward induction. *Theoretical Economics, 14*(1), 103–134.

Keynes, J. M. (1937). The general theory of employment. *The Quarterly Journal of Economics, 51*(2), 209–223.

Keynes, J. M. (2018). *The general theory of employment, interest, and money.* Springer.

Knudsen, C. (1993). Rationality and the problem of self-reference in economics! *Rationality, Institutions, and Economic Methodology, 2*, 133.

Koppl, R., Jr., & Barkley Rosser, J. (2002). All that I have to say has already crossed your mind. *Metroeconomica, 53*(4), 339–360.

Koriyama, Y., & Ozkes, A. I. (2021). Inclusive cognitive hierarchy. *Journal of Economic Behavior & Organization, 186*, 458–480.

Krockow, E. M., Pulford, B. D., & Colman, A. M. (2018). Far but finite horizons promote cooperation in the centipede game. *Journal of Economic Psychology, 67*, 191–199.

Łatek, M., Axtell, R., & Kaminski, B. (2009). Bounded rationality via recursion. In *Proceedings of Eighth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2009)*, pp. 457–464.

Levin, D., & Zhang, L. (2022). Bridging level-k to nash equilibrium. *Review of Economics and Statistics, 104*(6), 1329–1340. https://doi.org/10.1162/rest_a_00990
Lipman, B. L. (1991). How to decide how to decide how to...: Modeling limited rationality. *Econometrica: Journal of the Econometric Society,* 1105–1125.
Löfgren, L. (1990). *On the partiality of self-reference* (pp. 47–64). Gordon and Breach Science Publishers.
Mackie, J. (1971). What can we learn from the paradoxes? Part II. In *Crítica: Revista Hispanoamericana de Filosofía*, pp. 35–54.
Mattsson, L.-G., & Weibull, J. W. (2002). Probabilistic choice and procedurally bounded rationality. *Games and Economic Behavior, 41*(1), 61–78.
McKelvey, R. D., Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica: Journal of the Econometric Society,* 803–836.
McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior, 10*(1), 6–38.
McKelvey, R. D., & Palfrey, T. R. (1998). Quantal response equilibria for extensive form games. *Experimental Economics, 1*(1), 9–41.
McKelvey, R. D., Palfrey, T. R., & Weber, R. A. (2000). The effects of payoff magnitude and heterogeneity on behavior in 2×2 games with unique mixed strategy equilibria. *Journal of Economic Behavior & Organization, 42*(4), 523–548.
Mertens, J.-F., & Zamir, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory, 14*(1), 1–29.
Morgenstern, O. (1928). *Wirtschaftsprognose: Eine Untersuchung ihrer Voraussetzungen und Möglichkeiten*. Springer.
Morgenstern, O. (1935). Vollkommene voraussicht und wirtschaftliches gleichgewicht. *Zeitschrift für Nationalökonomie/Journal of Economics, 6* 337–357. https://doi.org/10.1007/BF01311642
Moulin, H. (1986). *Game theory for the social sciences*. NYU Press.
Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review, 85*(5), 1313–1326.
Nash, J. (1951). Non-cooperative games. *Annals of Mathematics,* 286–295.
Ortega, D. A., & Braun, P. A. (2011). Information, utility and bounded rationality. In *International Conference on Artificial General Intelligence*, pp. 269–274. Springer.
Ortega, P. A., & Braun, D. A. (2013). Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 469*(2153), 20120683.
Ortega, P. A., & Braun, D. A. (2014). Generalized Thompson sampling for sequential decision-making and causal inference. *Complex Adaptive Systems Modeling, 2*(1), 1–23.
Ortega, P. A., & Stocker, A. A. (2016). Human decision-making under limited time. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, R. Garnett (Eds.), *Advances in neural information processing systems*, pp. 100–108. Curran Associates, Inc.
Polani, D., Sporns, O., & Lungarella, M. (2007). How information and embodiment shape intelligent information processing. *50 years of artificial intelligence* (pp. 99–111). Berlin: Springer.
Polonio, L., & Coricelli, G. (2019). Testing the level of consistency between choices and beliefs in games using eye-tracking. *Games and Economic Behavior, 113*, 566–586.
Prokopenko, M., Harré, M., Lizier, J., Boschetti, F., Peppas, P., & Kauffman, S. (2019). Self-referential basis of undecidable dynamics: From the liar paradox and the halting problem to the edge of chaos. *Physics of Life Reviews, 31*, 134–156.
Rabin, M. O. (1957). Effective computability of winning strategies. *Contributions to the Theory of Games, 3*(39), 147–157.
Raiffa, H., & Luce, R. D. (1957). *Games and decisions: introduction and critical survey.* Wiley.
Rapoport, A., Seale, D. A., Erev, I., & Sundali, J. A. (1998). Equilibrium play in large group market entry games. *Management Science, 44*(1), 119–141.
Rogers, B. W., Palfrey, T. R., & Camerer, C. F. (2009). Heterogeneous quantal response equilibrium and cognitive hierarchies. *Journal of Economic Theory, 144*(4), 1440–1467.
Scott, D. W. (2015). *Multivariate density estimation: Theory, practice, and visualization*. Wiley.
Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.
Simon, H. A. (1976). From substantive to procedural rationality. *25 years of economic theory* (pp. 65–86). Springer.
Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics, 50*(3), 665–690.

Stahl, D. O. (1990). Entropy control costs and entropic equilibria. *International Journal of Game Theory, 19*(2), 129–138.
Stahl, D. O. (1993). Evolution of smart players. *Games and Economic Behavior, 5*(4), 604–617.
Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior, 10*(1), 218–254.
Sundali, J. A., Rapoport, A., & Seale, D. A. (1995). Coordination in market entry games with symmetric players. *Organizational Behavior and Human Decision Processes, 64*(2), 203–218.
Tishby, N., Pereira, F. C., & Bialek, W. (1999). The information bottleneck method. In *Proc. of the 37-th Annual Allerton Conference on Communication, Control and Computing*, pp. 368–377.
Tishby, N., & Polani, D. (2011). Information theory of decisions and actions. *Perception-action cycle* (pp. 601–636). Berlin: Springer.
Turocy, T. L. (2010). Computing sequential equilibria using agent quantal response equilibria. *Economic Theory, 42*(1), 255–269.
Webster, T. J. (2013). A note on the ultimatum paradox, bounded rationality, and uncertainty. *International Advances in Economic Research, 19*(1), 1–10.
Wen, Y., Yang, Y., & Wang, J. (2020). Modelling bounded rationality in multi-agent interactions by generalized recursive reasoning. In Bessiere, C., editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pp. 414–421. International Joint Conferences on Artificial Intelligence Organization. Main track.
Wolpert, D. H. (2006). Information theory-the bridge connecting bounded rational game theory and statistical physics. *Complex Engineered Systems* (pp. 262–290). Springer.
Wright, J. R., & Leyton-Brown, K. (2019). Level-0 models for predicting human behavior in games. *Journal of Artificial Intelligence Research, 64*, 357–383.

Authors and Affiliations

Benjamin Patrick Evans[1] · Mikhail Prokopenko[1]

✉ Benjamin Patrick Evans
Benjamin.Evans@sydney.edu.au

[1] Centre for Complex Systems, University of Sydney, Sydney, NSW 2006, Australia

CHAPTER 6

Endogenous emergence of stylised facts in economic markets from bounded rational agents

---

Building upon the work developed throughout this thesis; we now have the tools and techniques to analyse the formation of crises and stylised facts in economic markets under a relatively concise setting. Here, rather than making difficult modelling assumptions for a complex agent-based model, such as when to buy/sell or how much to bid, or enforcing constraints to give rise to a statistical equilibrium, we instead use Quantal Hierarchy reasoning (from Chapter 5) to model the bottom-up strategic reasoning between agents in economic markets. We show how various stylised facts can arise in this setting under generalised conditions, highlighting the usefulness of the proposed Quantal Hierarchy model for capturing decision-making in economic markets when compared to existing approaches.

ROYAL SOCIETY OPEN SCIENCE

royalsocietypublishing.org/journal/rsos

Bounded strategic reasoning explains crisis emergence in multi-agent market games

Benjamin Patrick Evans and Mikhail Prokopenko

Centre for Complex Systems, The University of Sydney, Sydney, New South Wales 2006, Australia

BPE, 0000-0002-2466-5264; MP, 0000-0002-4215-0344

Research

Check for updates

Cite this article: Patrick Evans B, Prokopenko M. 2023 Bounded strategic reasoning explains crisis emergence in multi-agent market games. *R. Soc. Open Sci.* **10**: 221164. https://doi.org/10.1098/rsos.221164

Received: 7 September 2022
Accepted: 9 January 2023

Subject Category:
Computer science and artificial intelligence

Subject Areas:
computer modelling and simulation/complexity/behaviour

Keywords:
complexity economics, bounded rationality, crisis emergence, market entrance games, multi-agent markets

Author for correspondence:
Benjamin Patrick Evans
e-mail: benjamin.evans@sydney.edu.au

The efficient market hypothesis (EMH), based on rational expectations and market equilibrium, is the dominant perspective for modelling economic markets. However, the most notable critique of the EMH is the inability to model periods of out-of-equilibrium dynamics without significant external news. When such dynamics emerge endogenously, the traditional economic frameworks prove insufficient. This work offers an alternate perspective explaining the endogenous emergence of punctuated out-of-equilibrium dynamics based on bounded rational agents. In a concise market entrance game, we show how boundedly rational strategic reasoning can lead to endogenously emerging crises, exhibiting fat tails in returns. We also show how other common stylized facts, such as clustered volatility, arise due to agent diversity (or lack thereof) and the varying learning updates across the agents. This work explains various stylized facts and crisis emergence in economic markets, in the absence of any external news, based on agent interactions and bounded rational reasoning.

1. Introduction

Economic markets have existed for millennia, with the earliest identified markets dating back to at least the Babylonian Empire [1]. Since then, the modelling of such markets has been a heavily researched topic, aiming to improve understanding of markets, increase profits and shape policy and interventions within these markets. The dominant perspective which has arisen over the last two centuries [2] is the efficient market hypothesis (EMH) [3], based on rational expectations and market equilibrium. The EMH states that prices reflect all known information about an asset, and as such, they reflect the fair (fundamental) value. However, there are multiple significant market crashes, such as the 1987 Dow Jones index crash [4], the global financial crisis in 2008 [5] and the flash crash of 2010 [6], that the EMH failed to explain purely by the arrival of external

THE ROYAL SOCIETY PUBLISHING

news. If prices reflected all information, assets matched their actual underlying value, and all agents were perfectly rational and not speculative, such bubbles (and resulting bursts) would not occur. Instead, agents perception on the market state can generate phenomena in a manner that is 'totally unrelated to economic fundamentals' [7], where asset pricing deviates from the fundamental value. For example, such phenomena can be driven by irrational exuberance of the agents [8], based on unfounded beliefs and speculation [9]. This brings into question whether pricing does in fact follow the 'rationally expected fundamental value', and highlights the need to consider alternative models of market dynamics [10].

The most serious specific critiques of the EMH include its inability to explain periods of out-of-equilibrium behaviour such as endogenous crises [11,12] or the 'stylized facts' present in economic markets such as volatility clustering where periods of high (low) volatility occur in bursts, in the absence of any significant external news [13]. Such crises are known to occur in different markets, from housing markets [14,15], to stock markets [16], and foreign exchange markets [17]. Likewise, volatility clustering is consistently observed in markets [18], with temporal correlations in volatility breaching the traditional economic assumption of heteroskedasticity. The EMH suggests that such phenomena occur due to rational agents reacting to external news in the market [19]. However, when such dynamics emerge endogenously, i.e. in the absence of external news, the traditional economic frameworks struggle to explain such stylized facts [20] or the deviation from equilibrium [21], demanding adequate models explaining the 'wildness' of market dynamics [22–24].

Thus, the endogenous emergence of crises and stylized facts are significant aspects to consider in the modelling of economic markets. An adequate model of economic markets should capture and ideally explain three desirable aspects of actual market dynamics: (i) *convergence* to a good average outcome, i.e. equilibrium, (ii) endogenous *emergence of crises* so that the resulting dynamics are categorized by fat-tailed distributions of out-of-equilibria deviations, where abrupt periods differ significantly from the 'converged' equilibrium and (iii) ability to recreate *stylized facts* of economic markets such as clustered volatility. The EMH explains (ii) and (iii) as arising only from the arrival of external news.

These notable deviations from EMH have pushed for a 'revolution' of economics [22], and encouraged a modelling focus on agents who operate under 'bounded rationality' [25]. Relaxing the perfect rationality assumption provides an alternate perspective to markets, admitting that market participants may not follow the representative *homo economicus* model [26,27], and may instead be speculative, adaptive or subject to limitations in their information processing abilities. It may initially seem desirable to modify the representative agent to account for psychological or cognitive biases directly, however, this only partially addresses the issue and still fails to explain various phenomena which crucially only arise from the interactions among heterogeneous agents [20]. These considerations have given rise to *Complexity Economics* [28], and the 'interacting agent hypothesis' [29,30] capturing the features of real-world financial markets, such as the endogenous emergence of out-of-equilibrium dynamics due to the interaction among heterogeneous boundedly rational agents [31–36], rather than purely by the rational reaction to the arrival of external news.

One of the most notable examples exploring Complexity Economics is the canonical El Farol bar problem [37]. El Farol has been called 'the most important problem' in the modelling of complex systems [38], and continues to be explored [39,40], having motivated a host of other market entrance games [41] and minority games [42] for modelling markets. In this market entrance game the agent payoffs depend on other market participants' decisions, creating complex market dynamics. Adaptive Strategies (AS) is the widely accepted solution to the El Farol bar problem [37], which generally converges to an equilibrium near the optimal resource capacity [43]. However, it is unknown to what extent deviations from this equilibrium can be captured or whether current Complexity Economics solutions to El Farol can adequately generate stylized facts similar to those observed in actual markets [36]. Hence, we need a more refined model capable of demonstrating convergence to equilibrium, punctuated by abrupt deviations, while also explaining market dynamics and stylized facts arising endogenously rather than relying on the arrival of external news.

Here, we propose an approach based on boundedly rational strategic (higher-order) reasoning agents. Each Bounded Rational AdapTive Strategic (BRATS) reasoning agent maintains a recursive model about other agents' beliefs. However, the potentially infinite chain of strategic reasoning is 'broken' at various points of recursion, following the Quantal Hierarchy (QH) model [44]. Specifically, the heterogeneous agents are limited in the amount of information processing they can perform, with reasoning resources quantified information-theoretically (in the Shannon sense). Based on the observed market outcomes, the BRATS agents learn and update their recursive beliefs (increase resources). The BRATS approach improves upon the canonical solution driven by Adaptive Strategies across all outlined criteria. Specifically, the approach not only ensures (i) general convergence towards an 'equilibrium' even with boundedly rational agents, but also allows for (ii) the explanation of abrupt endogenously emerging periods of out-of-equilibrium behaviour based on boundedly rational strategic reasoning, generating 'fat tails'. Furthermore, (iii) volatility clustering can be generated and explained endogenously due to the diversity of agent beliefs and the heterogeneous learning updates.

We provide comparisons to alternative approaches and show that the proposed approach convincingly outperforms the alternatives across the key measures in a market game, better matching dynamics observed in actual economic markets. Thus, the long-held conjecture that the AS approach provides an adequate resolution to the El Farol problem is challenged. An effective solution is shown to be provided by the agents which reason strategically (i.e. recursively) while being limited by their information processing resources.

2. Material and methods

The importance of agent irrationality on market dynamics was discussed by Keynes [45], who explained price fluctuations with the beauty contest game. Specifically, Keynes pointed out that 'I-think-you-think-they-think⋯-type of regress' better represented investor behaviour, with pricing driven by investors' beliefs about other investors, rather than by the asset fundamentals.

While this 'I-think-you-think-they-think⋯-type of regress' has been identified as a natural representation for agent behaviour in market games [38], modelling these recursive beliefs becomes problematic due to the potential infinite reasoning [46]. A recently introduced approach, the QH model [44], addressed the modelling limitations of this higher-order reasoning by 'breaking' at various points of recursion, thus accounting for boundedly rational agents and preventing infinite regress.

2.1. Background

The QH model [44] is based upon a recursive form of the variational free-energy principle, which was proposed as a thermodynamic treatment of bounded rational decision-making [47]. Agents make a decision f on which action $a \in A$ to take from the available choice set A, based upon the utility U of the choice, and the prior beliefs of the agent p. Decision-making can be represented as state changes, given by the following free-energy difference:

$$-\Delta F[f[a]] = \sum_{a \in A} f[a] U[a] - \frac{1}{\beta} \sum_{a \in A} f[a] \log\left(\frac{f[a]}{p[a]}\right), \tag{2.1}$$

which produces the equilibrium distribution for the decision function $f[a]$:

$$f[a] = \frac{1}{Z} p[a] \, e^{\beta U[a]}, \tag{2.2}$$

where Z is the partition function, and the parameter β (i.e. inverse temperature) governs the information processing available to an agent.

The QH model extended this framework to capture recursive higher-order reasoning [44] as pseudo-sequential decision-making where agents are limited in the amount of information processing they can perform. Specifically, at each level of recursive reasoning k, information processing resources are reduced by a discount parameter γ. Parameter γ modulates an agent's beliefs about other agents' decisions. Extending equation (2.1) to account for (pseudo-)sequential decision-making yields the following recursive free-energy difference:

$$-\Delta F[f] = \sum_{a_{\leq K}} f[a_{\leq K}] \sum_{k=0}^{\infty} \left(U[a_k | a_{<k}] - \frac{1}{\beta \gamma^k} \log \frac{f[a_k | a_{<k}]}{p[a_k | a_{<k}]} \right), \tag{2.3}$$

where $a_{<k}$ represents the past decisions of the agent. The overall reasoning levels are bound by the recursion depth based on γ, which ensures, given a simple computational threshold ϵ, that the recursion terminates [44]. The equilibrium solution for the agent decision function $f[a_k | a_{<k}]$ is then

given by

$$
f\,[a_k \mid a_{<k}] = \begin{cases} \frac{1}{Z_k} \underbrace{p[a_k \mid a_{<k}]}_{\text{prior belief}}, & \text{if } \beta\gamma^k < \epsilon \\ \frac{1}{Z_k} \underbrace{p[a_k \mid a_{<k}]}_{\text{prior belief}} \times \underbrace{\mathrm{e}^{\beta U[a_k|a_{<k}]}}_{\text{current utility}}, & \text{if } \gamma = 0 \\ \frac{1}{Z_k} \underbrace{p[a_k \mid a_{<k}]}_{\text{prior belief}} \times \underbrace{Z_{k+1}^{1/\gamma}}_{\text{future contribution}} \times \underbrace{\mathrm{e}^{\beta\gamma^k U[a_k|a_{<k}]}}_{\text{current utility}}, & \text{otherwise} \end{cases} \tag{2.4}
$$

The overall processing resources are governed by $\beta\gamma^k$, i.e. the information processing abilities β are discounted based on γ and the strategic reasoning depth k. When $\beta\gamma^k < \epsilon$, the recursion stops since the result simply echoes the prior belief, and no focus is placed on the payoff. In summary, the QH model represents varying levels of higher-order reasoning and captures the expectations about the processing abilities of other agents by varying parameters β and γ.

2.2. Proposed model: BRATS approach

In modelling a market, we assume that each market participant (agent) performs strategic reasoning according to the QH model, i.e. recursively but boundedly reasons about the decisions of other agents. The reasoning abilities β and γ are heterogeneously assigned among the participants. All agents start naive, with low information processing abilities ($\beta \gtrapprox 0$). Over time, agents learn and adjust their beliefs based on the observed market outcomes by suitably increasing processing abilities β (see below). We refer to this model as the Bounded Rational AdapTive Strategic reasoning (BRATS) approach.

2.2.1. Naive agent

The simplest agent is the 'naive agent'. Naive agents have no information processing abilities ($\beta\gamma^k < \epsilon$), so they make their decisions based on prior beliefs without consideration of other agents' reasoning—in other words, they perform no strategic reasoning. Naive agents simply attend if it would have been profitable to do so previously.

2.2.2. Learning

Learning in the BRATS model is represented through increasing reasoning abilities, i.e. increasing β. This is determined by a learning rate $\eta > 0$, which acts as a linear modifier for β, as follows: $\beta_{t+1} = \beta_t + \eta$. All agents start with small β_0. Agents are assigned different rates η generated from a given range ($0.1 < \eta \leq 1$), representing different learning abilities. This heterogeneous representation provides a simple configuration where varying learning rates can be captured across the population of agents at each time step.

Parameter γ is not directly modified and assumed to be fixed (but different) for all agents. However, increasing β implicitly increases the total internal resources used to reason about other agents' reasoning (through $\beta\gamma^k$).

2.3. Canonical model: Adaptive Strategies

The conventional approach to the El Farol bar problem is 'Adaptive Strategies', proposed initially in [37], and extended in [48]. Each market participant i has a set of predictors $\boldsymbol{s}_i \in S_i$, creating an 'ecology' of predictors based on the past observations that adapts over time. An individual strategy $\boldsymbol{s}_i$ is a vector of weights, where each element in the vector determines how the agent believes the historical data affects the attendance prediction for the current time [49]. By contrast to the BRATS approach, the AS model is not a strategic reasoning model. Instead, each agent makes their prediction based on the (public) history, and does not consider the reasoning of other agents at the current time step. Agents weigh the likelihood of using predictors according to the past success of this predictor based on a given history, choosing the strategy $\boldsymbol{s}^*{}_i$ which would have done the best in the previous time steps [49].

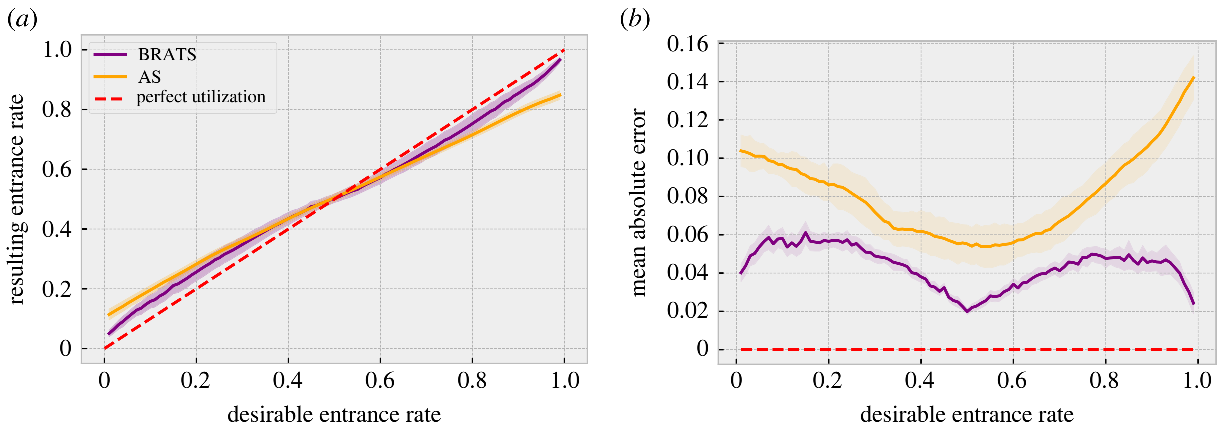


Figure 1. Resource efficiency ($N = 100$). The purple (orange) line represents the proposed BRATS model (Adaptive Strategies). The dotted red line highlights the optimal case with perfect efficiency, i.e. the mixed strategy Nash equilibrium. Figure 1*a* shows the average overall attendance rate across runs, and the shaded surrounding area shows ±1 s.d. Figure 1*b* shows the average error for each run, and the filled area shows ±1 s.d. across these runs. (*a*) Utilization and (*b*) errors.

3. Results

The basic premise of the El Farol market entrance game is as follows. Agents enjoy attending a bar, or profit entering a market, if less than 60% of other agents attend the bar; otherwise, the bar/market is deemed overcrowded/unprofitable, and they would have preferred to stay out. As a generalization, we say N agents receive payoff U_{enter} for entering if less than cN, $c \in [0, 1]$ other agents attend, otherwise they receive payoff $U_{\text{overcrowded}}$. Staying out returns a fixed payoff of U_{exit}. The inequality $U_{\text{enter}} > U_{\text{exit}} > U_{\text{overcrowded}}$ holds such that if the bar is overcrowded, the agent would have preferred to stay home. However, the preference is always to attend if it is not overcrowded.

3.1. Convergence

We first consider the 'desirable' attendance rate c, and the simulated attendance rate from the bounded rational agents. This convergence is displayed in figure 1*a*, comparing the proposed (BRATS) approach with the canonical AS solution. Agents are able to self-organize around the desired value without explicit coordination. We refer to this as the resource efficiency, i.e. the ability to operate near the optimal capacity. Figure 1*b* shows the resulting errors, with lower average errors for the proposed approach compared to AS.

The time evolution of the system is displayed in figure 5. There is an initial 'adjustment' period with both approaches in which the agents are learning appropriate beliefs. However, once some learning has taken place (e.g. several rounds have been completed), the proposed BRATS approach converges to a higher average resource efficiency than the canonical AS approach, particularly for higher or lower values of c, as confirmed in figure 1. The methods perform similarly for mid ranges of c, achieving high resource efficiency. These results show that both approaches perform well on the first criteria of *convergence* to a good average outcome, with the proposed approach outperforming AS for higher and lower entrance capacities.

It is well known that there exists a unique symmetric mixed strategy Nash equilibrium (MSE) solution to the problem, where agents attend probabilistically based on the enjoyable capacity c [50]. Such an approach offers perfect convergence and utilization (i.e. 0 error), however, this MSE solution cannot categorize learning or adaption throughout time [51], and further, cannot generate periods of endogenously emerging crises. This limitation is problematic for explaining known deviations from equilibrium in actual market settings. As we are particularly interested in out-of-equilibrium dynamics generated from bounded rational agents, we do not explore such solutions further as they cannot capture such phenomena. The following sections analyse these deviations from equilibrium in more detail.

3.2. Emergence of endogenous crises

The previous section showed the average convergence towards the desired resource capacity c. Here, we consider the endogenous emergence of crises, or self-induced shocks [52], as a result of endogenous

Table 1. The occurrence of $3^{+}\sigma$ events. The table displays the observed percentage of attendance changes falling outside three standard deviations of the mean attendance. Each cell represents the average results for the attendance rate *c*. According to the normal distribution (bottom row), the percentage of expected occurrences for $3^{+}\sigma$ events is 0.3, and any values greater than this threshold indicate a higher probability of extreme events.

	c								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
BRATS	2.3	2	1.3	0.7	1	1.1	1.3	1.9	2.3
AS	1.5	1.4	1.1	1	0.8	0.6	0.5	0.4	0.6
noise traders	0.2	0.3	0.3	0.2	0.3	0.3	0.3	0.3	0.3
normal distribution	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3

changes in agent beliefs [53]. Following [52], we consider crises to occur when the resulting attendance change is at least three standard deviations from the average historical attendance change, and refer to these as $3^{+}\sigma$ events. The EMH suggests these deviations would follow a random walk, and, thus, be normally distributed, meaning that the occurrence of $3^{+}\sigma$ events would be small (0.3%). However, in actual markets, the kurtosis of the observed return distribution can often be categorized as leptokurtic, implying fatter tails, and a higher probability of 'crisis' than the probability suggested by the random walk, casting doubts about the explanation given by the EMH [54,55].

To analyse the frequency of significant changes, we calculate the probability of $3^{+}\sigma$ events, and in addition, estimate the tail-index α of these changes using the Hill estimator [56]. The tail-index α gives a measure of the shape of the tail, with lower α's implying a slower decaying function and thus, heavier tails. It has been reported that for many markets, α can often be in the range 1 to 4 [57] with emerging markets having estimates $1 < \alpha < 2$, and developed markets $2 < \alpha < 4$ [57,58]. Thus, this section demonstrates endogenous crisis formation in terms of fat tails in attendance fluctuations, supported by estimates of the tail-indexes α.

Under the proposed BRATS approach to model market entrance, large abrupt changes in attendance occur more than one would expect from a normal distribution, as displayed in table 1. With the AS model, fatter tails than expected from the normal distribution are also observed, but to a lesser extent. By contrast, the market dynamics generated by agents who follow a random walk ('noise traders') fall entirely in line with the normal distribution. Therefore, the proposed approach demonstrates the ability to recreate an important stylized fact: fat-tailed distributions in terms of (abrupt) resource allocation changes.

We now estimate the tail-index α with a range of common tail sizes, 2.5%, 5%, 10%[1] for (i) the proposed approach, (ii) the canonical AS solution and (iii) 'noise traders'. To re-iterate, a lower tail-index α implies a heavier tail. The collated results are visualized in figure 2, and specific breakdowns are shown in table 2. The tail indices (α's) produced by the proposed approach consistently fall in the 1–3 range, as expected by crisis dynamics of emerging ($1 < \alpha < 2$) and established ($2 < \alpha < 4$) markets [57,58]. By contrast, the AS approach often generates larger α's (thinner tails), with approximately half of the estimates falling outside the actual market range for tail indices ($\alpha > 4$). With noise traders, the majority of the α's are found to be well outside this range, indicating even smaller tails.

These findings demonstrate that, unlike its alternatives, the interacting BRATS agents can consistently generate endogenous crises ($3^{+}\sigma$ events), in accordance with observed market crises (indicated by the fat tails that decay at a realistic rate), while simultaneously improving upon the average convergence.

3.3. Clustered volatility

Having analysed crisis emergence in terms of fat tails in attendance changes, we turn our attention to the temporal correlations in 'local' volatility, i.e. in the absolute percentage changes in attendance at each time step.

The temporal volatility correlations in the market entrance game are summarized in figure 3 and further analysed in figure 6.

[1]We use several tail sizes to account for estimates of α frequently depending on the tail size selection.

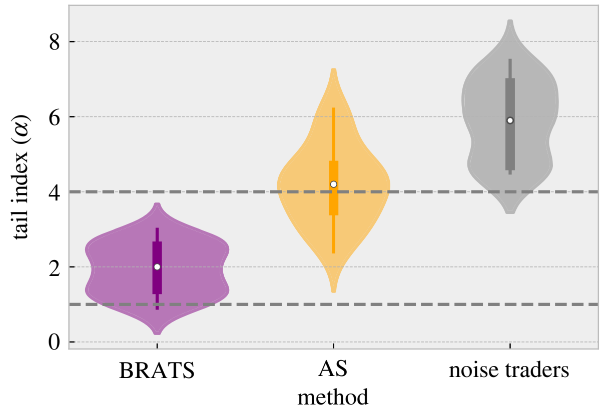


Figure 2. Tail-index estimates α computed using the Hill Estimator. The average values across tail sizes and attendance capacities are used to compute the density for the violin plots (cf. table 2). A lower α indicates a slower decay and, thus, a fatter tail. The actual range of tail indices from economic markets is shown with the dashed grey lines.

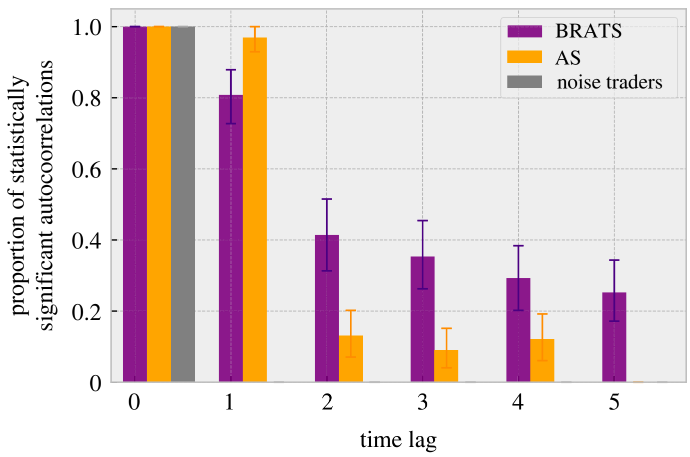


Figure 3. Proportion of statistically significant autocorrelations in the attendance volatility, showing the presence of clustered volatility. Purple represents the proposed BRATS model, orange represents the Adaptive Strategies and grey the noise traders. The y-axis shows the proportion of occurrences (across c's) which were statistically significant. The x-axis shows the temporal lag. The error bars represent the bootstrapped 95% confidence interval. Individual breakdowns for c's are presented in figure 6.

The BRATS approach shows statistically significant autocorrelations for up to five time steps. In contrast, the results produced by the AS approach generally show fewer temporal correlations in volatility, demonstrating significantly less temporal correlation beyond one time step. These findings indicate that the ability to recreate the clustered volatility observed in actual markets across multiple temporal lags is not present in the canonical AS model. Furthermore, the noise traders show no temporal volatility correlation (beyond lag zero).

The results shown in figure 3 confirm that volatility clustering can arise purely endogenously, as the observed correlations occurred without any external changes to the system. That is not to say that volatility does not also arise due to external news, but rather that periods of increased volatility can also arise purely endogenously in the absence of such news.

To explain the endogenous emergence of periods of increased volatility, we analyse whether the heterogeneity (diversity) of agents' beliefs is a leading indicator of future volatility change. In doing so, we employ a diversity measure quantified by the normalized entropy of the agent population, suitably adopted for BRATS and AS approaches (described in appendix C). When the diversity of the population decreases, this can indicate a group or herd formation and, thus, a reduction in the heterogeneity of the population. An example of a realization of these two time series is presented in figure 8.

We find that changes in the population diversity are predictive (or Granger causing) of changes in attendance (details are provided in appendix D). Crucially, these findings show that the distribution of agent beliefs is a significant predictor of volatility. To measure the effect that the diversity change has on the resulting volatility, beyond merely being predictive of this change, we use an impulse response function quantifying the impact on volatility generated by an 'impulse' of diversity (figures 4 and 7).

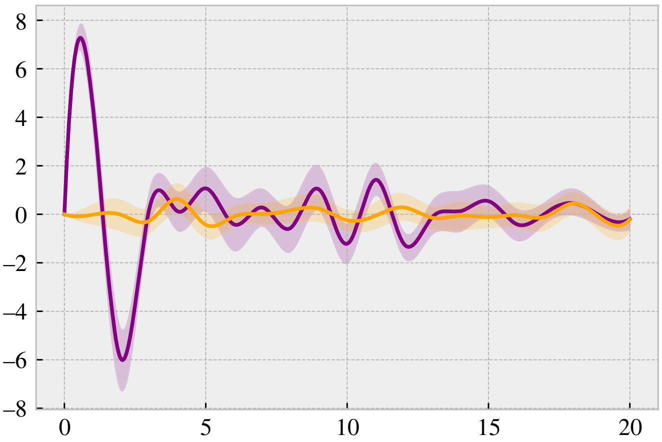


Figure 4. An example smoothed impulse response function for $c = 0.2$. Purple (orange) represents the proposed BRATS model (Adaptive Strategies). The dark line represents the mean across runs, and the shaded lighter region is the bootstrapped 95% confidence interval. The *y*-axis shows the change in volatility difference following the impulse, and the *x*-axis shows the time since the impulse. Full results (for the range of *c*) are shown in figure 7.

One unit shock in diversity change results in an initial consistent increase in positive volatility change, followed by a 'correction' to a negative change in volatility. This correction is then followed by a destabilizing period of oscillatory dynamics indicating larger changes and uncertainty in the volatility, marked by larger confidence intervals and variations. Finally, the dynamics settle down after approximately 10 periods on average, without a discernible influence beyond this interval.

When the diversity of the population is low, the system is in a 'turbulent' regime [33] where volatility is high due to the interdependence of agent decisions. Over time, agents react and modify their beliefs. Due to the agents' alternate learning updates, the belief correlation declines, and the system returns to a calm regime with less volatility, characterized by higher population heterogeneity. This reduction in the lagged autocorrelation in volatility is demonstrated in figure 3. The reduction is also reflected in the oscillatory period in figure 7 during which the impulse response flattens over time. By contrast, the results produced by the AS do not provide clear outcomes in impulse response, as shown in figure 7. These results indicate that clustered volatility cannot be explained or predicted based upon the AS approach.

4. Discussion and conclusion

The formation of crises in economic markets is a well-established phenomenon, dating back to at least the Dutch 'tulip mania' of 1636, which abruptly burst overnight [59,60]. While the EMH explains such phenomena as rational reactions to the arrival of external news, an alternative hypothesis is that booms and busts can also arise endogenously due to the irrationality of market participants [61]. In this work, exemplified by a concise market configuration, the canonical El Farol Bar Problem, we extended the QH model to capture endogenous crisis emergence in market entrance games. Crucially, this analysis showed that bounded strategic reasoning can lead to higher probabilities of crises (i.e. fat tails in attendance fluctuations), matching the dynamics of actual economic markets. Furthermore, we related the resulting volatility to the diversity of strategic reasoning resources across the agent population.

Of course, convergence to a desirable market capacity, an 'equilibrium', has been demonstrated for the El Farol bar problem before [37,43,48,62]. Moreover, it has been shown that even a large number ($N \to \infty$) of zero-intelligence agents can self-organize to the desired capacity under relatively general conditions, simply due to the law of large numbers [43]. The model proposed in this work performed at least as well as these approaches, even in systems with a relatively low number of heterogeneous agents ($N = 100$), when the agents follow boundedly rational strategic reasoning.

However, convergence to equilibrium is a necessary but not sufficient characteristic for an adequate model of actual markets, as 'a non-stationary economy must experience at least some transient moments of disequilibrium' [7]. For example, under which crises can endogenously form and volatility can cluster in time. Hence, we systemically explored the formation of crises based on abrupt deviations from an equilibrium state. Specifically, we analysed the tail indices of *endogenously* emerging crises, categorized by significant changes in attendance ($3^{+}\sigma$ events), where the system can be seen as briefly being out-of-equilibrium. This analysis demonstrated that these fat tails are significantly heavier than one would expect under a rational expectations model without the arrival of any external news. We related these resultant tail indices to actual market data, showing that the fat tails produced by the BRATS approach decay at a realistic rate.

Finally, we traced and examined the endogenous emergence of periods of volatility based on changes in agent beliefs and related this to the agent learning process. A key stylized fact of economic markets—clustered volatility—was recreated using a concise learning update of agents' reasoning resources. Notably, the change in the overall diversity of agent beliefs was predictive of the change in future volatility, indicating that the heterogeneity of agent beliefs is a crucial predictor of volatility.

These findings strengthened the conjecture that the diversity of market participant beliefs is one of the 'most important propagation mechanism[s] of economic volatility' [63]. A similar conclusion was reached by Lux & Marchesi [58], where periods of instabilities caused by the ratio of chartists and fundamentalists in the market were 'quickly brought to an end by stabilizing tendencies'. The diversity of strategies had not yet been fully explored in the context of the El Farol bar problem [64], here we offered a way to analyse the population diversity in terms of reasoning resources. In doing so, we quantified the market dynamics, confirming that during a period of high volatility, if agent beliefs are correlated, the volatility persists. As the agents attempt to alter their beliefs to recover market capacity, the correlation in the agent beliefs reduces over time—due to the different agents' learning updates. Eventually, the system returns to a 'calm' steadier state, with less volatility, supported by the increased diversity of beliefs.

In summary, the proposed approach captured three key desirable characteristics for a model of economic markets. Specifically, we have shown how the BRATS agents can (i) *converge* to a good average outcome, i.e. equilibrium, on a macro scale, which is punctuated by the (ii) abrupt *emergence of crises* on a micro-scale, categorized by spontaneous 'out-of-equilibria deviations' in accordance with actual market dynamics, and in addition, can (iii) recreate *stylized facts* of markets such as clustered volatility. Importantly, these emergent phenomena occur endogenously, without any external news to the system. The phenomena arise simply due to the interaction among heterogeneous BRATS agents.

This work adds to the growing literature on the interacting agent hypothesis. Various approaches explore this hypothesis using relatively complex and fine-grained models. For example, heterogeneous agent beliefs can model observed market price dynamics and volatility in the S&P 500 [65]. Likewise, booms and busts in various housing markets can be explained through heterogeneous expectations and beliefs [66–68]. In [34], features of financial time series can be recreated based on the dynamics of opinion formation of heterogeneous agents, and in [35,36], through the competition between local and global interaction among the agents in the market. By contrast, we aimed to capture complex market dynamics within a concise and intuitive model of agent reasoning in a general market setting. In doing so, we verified the hypothesis that the heterogeneity of interacting agents is a crucial factor behind the endogenous emergence of crisis and stylized facts in economic markets. These works can be seen as complementary, helping to support the interacting agent hypothesis.

Various approaches have been proposed to consider price changes as a result of attendance fluctuations [52,55,69]. For future work, it may be instructive to incorporate price changes into the analysis presented here, e.g. with the introduction of a market-maker. In addition, while we have examined entrance into a single market, it would be insightful to explore such findings in a generalized multiple market setting, where agents not only need to decide whether or not to enter, but also which market to enter (if any) in a congestion style game.

By contrast to the long-standing belief that AS provides an adequate solution to this class of market games, we have shown that bounded strategic reasoning can capture salient market dynamics more convincingly. This study highlights that bounded strategic reasoning ('I-think-you-think-they-think···') across heterogeneous agents can explain periods of volatility and abrupt crises emerging in economic markets, even without any external shocks.

Data accessibility. This article has no additional data.

Authors' contributions. B.P.E.: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, writing—review and editing; M.P.: conceptualization, formal analysis, investigation, methodology, resources, supervision, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. No funding has been received for this article.

Acknowledgements. The authors would like to acknowledge the referees for their valuable suggestions, which have improved the manuscript. The authors would also like to recognize the CAUL Read and Publish scheme and the Royal Society for covering the open access charge for this publication. Additionally, the authors acknowledge the

HPC service at the University of Sydney for providing HPC resources that have contributed to the research results reported in this paper.

Appendix A. Experimental setup

All experiments are run for 30 Monte Carlo simulations to account for stochasticity. To implement the AS model for our comparative analysis, we used the model presented in [49], specifically the 'El Farol' model from Chapter 3.

The explored range of attendance rates is $c \in [0.01, 0.02, \ldots, 0.99]$. Generally, the limiting cases of $c = 0$ and $c = 1$ are excluded as trivial, as these correspond to every agent going or staying out. When showing individual plots for c's, typically only $c \in [0.1, 0.2, \ldots, 0.9]$ are displayed for readability.

Appendix B. Additional results plots

B.1. Convergence

A breakdown of utilization rates for the range of c values is presented in figure 5.

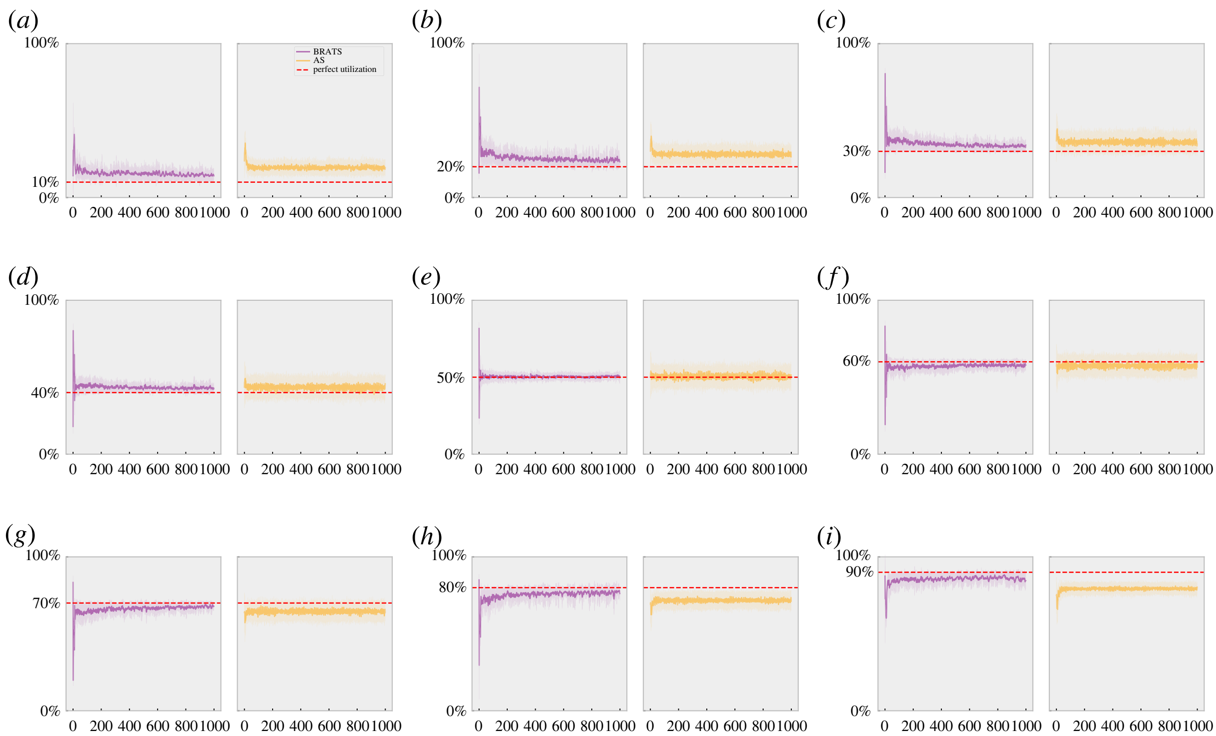


Figure 5. Average resource utilization for various enjoyable capacities c. Purple (orange) represents the proposed BRATS model (Adaptive Strategies). The mean attendance is shown as the dark line, with ±1 s.d. as the lighter bars. Perfect utilization (c) is displayed as the dashed red line. (*a*) $c = 0.1$, (*b*) $c = 0.2$, (*c*) $c = 0.3$, (*d*) $c = 0.4$, (*e*) $c = 0.5$, (*f*) $c = 0.6$, (*g*) $c = 0.7$, (*h*) $c = 0.8$ and (*i*) $c = 0.9$.

B.2. Emergence of endogenous crises

The tail-index α produced by the proposed approach falls in the 1–3 range, as expected by crisis dynamics of emerging ($1 < \alpha < 2$) and established ($2 < \alpha < 4$) markets [57,58], as shown in table 2. By contrast, the AS approach generates the α's typically in the range 2.5–6 (i.e. smaller tails). With noise traders, the α's are also found to be significantly larger, in the range 4.5–7.5, indicating even smaller tails.

Table 2. Average tail-index α computed using the Hill Estimator. Each subtable provides the results for a tail size. Each column for an attendance rate c, with the averages across c's presented as the $\mu \pm \sigma$ column. Each cell shows mean estimated tail-index α for a tail size and c. A lower α indicates a slower decay, and thus a fatter tail.

			c								
		$\mu \pm \sigma$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2.5%	BRATS	2.3 ± 0.5	1.7	1.6	2.5	2.7	2.8	2.7	2.4	1.9	2.4
	AS	4.7 ± 1.1	3.1	3.4	4.1	4.4	4.6	6.1	5.8	6.2	4.7
	noise	7.0 ± 0.3	7.1	7.0	6.8	7.2	7.5	6.6	7.2	7.0	7.0
5%	BRATS	1.9 ± 0.8	0.8	1.3	2.0	2.6	3.1	2.5	2.0	1.4	1.2
	AS	4.1 ± 1.0	2.7	2.9	3.5	3.8	4.2	4.9	5.1	5.4	4.6
	noise	5.9 ± 0.2	5.9	6.2	5.7	6.1	6.2	5.8	5.9	5.8	5.9
10%	BRATS	1.6 ± 0.6	1.1	1.0	1.4	2.1	2.6	2.0	1.5	1.1	1.2
	AS	3.5 ± 0.8	2.3	2.6	3.0	3.4	3.6	4.2	4.1	4.5	4.1
	noise	4.6 ± 0.1	4.6	4.9	4.6	4.8	4.6	4.5	4.6	4.6	4.6

B.3. Clustered volatility

The temporal auto-correlations in volatility are presented in figure 6, and the resulting impulse response functions are visualized in figure 7.

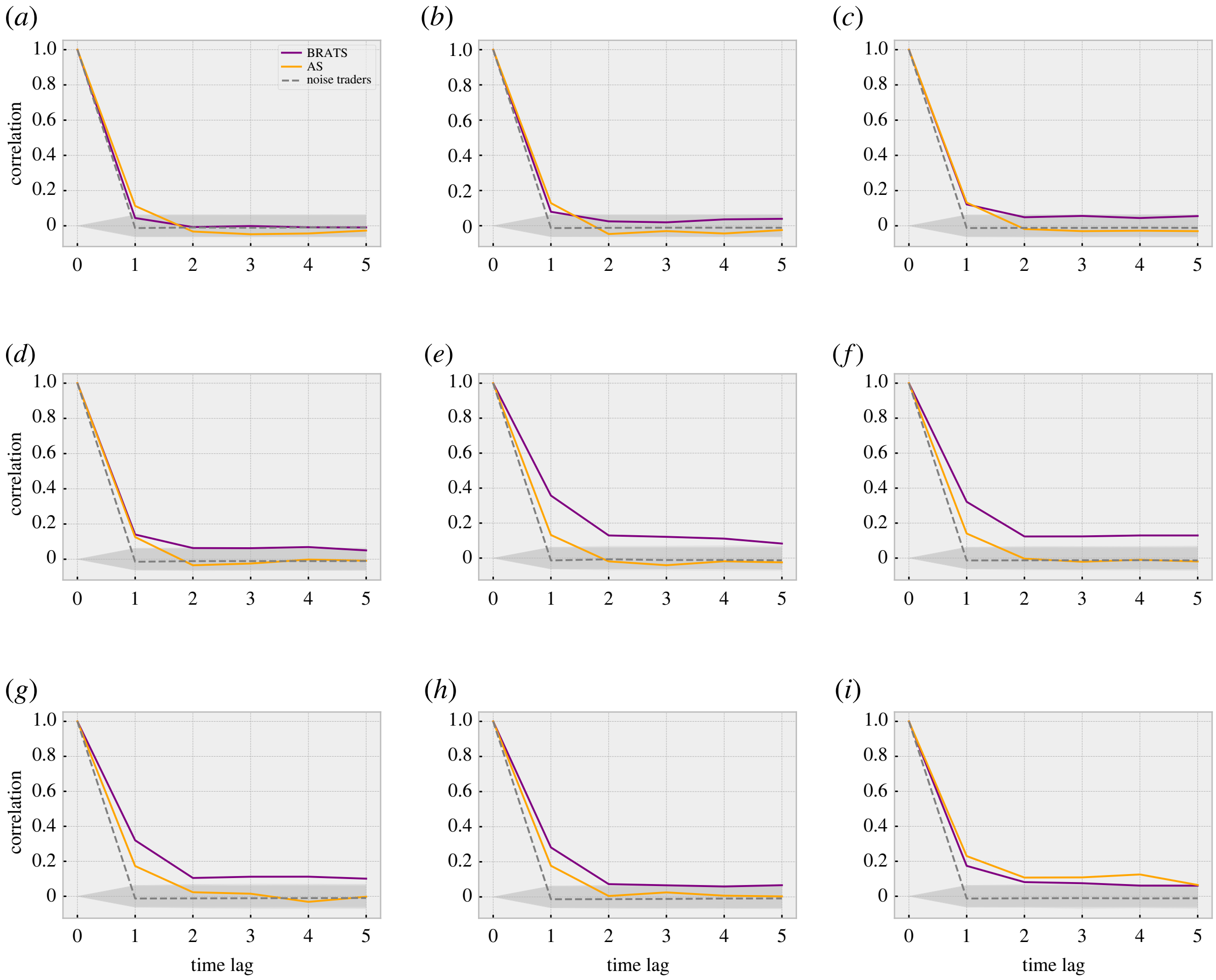


Figure 6. Autocorrelation in the attendance volatility, showing the presence of clustered volatility. Purple (orange) represents the proposed BRATS model (Adaptive Strategies). The dashed grey line represents noise traders. The filled grey region around 0 highlights the area of no statistically significant correlations. (*a*) $c = 0.1$, (*b*) $c = 0.2$, (*c*) $c = 0.3$, (*d*) $c = 0.4$, (*e*) $c = 0.5$, (*f*) $c = 0.6$, (*g*) $c = 0.7$, (*h*) $c = 0.8$ and (*i*) $c = 0.9$.

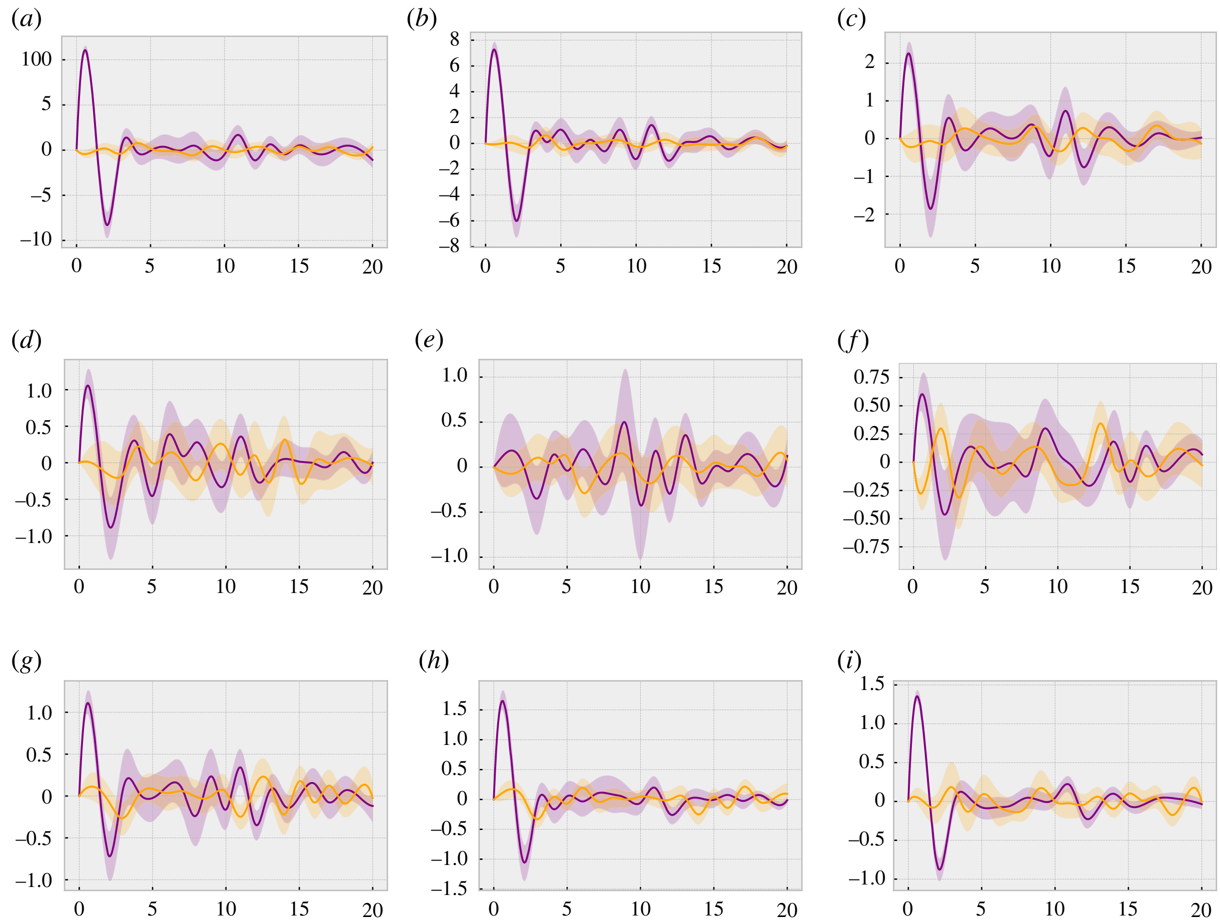


Figure 7. Smooth Impulse Response Functions. Purple (orange) represents the proposed BRATS model (Adaptive Strategies). The dark line represents the mean from across runs, and the shaded lighter region the bootstrapped 95% confidence interval. The y-axis shows the change in volatility difference following the impulse, and the x-axis shows the time since the impulse. (*a*) $c = 0.1$, (*b*) $c = 0.2$, (*c*) $c = 0.3$, (*d*) $c = 0.4$, (*e*) $c = 0.5$, (*f*) $c = 0.6$, (*g*) $c = 0.7$, (*h*) $c = 0.8$ and (*i*) $c = 0.9$.

Appendix C. Population heterogeneity

The normalized entropy of the population is used to measure the approximate heterogeneity (diversity) of the agents' beliefs. The resulting entropy is normalized by dividing by the maximal entropy distribution to provide a value between 0 and 1, with 0 indicating no diversity and 1 indicating maximal diversity. This formulation allows us to explore the relationship between the population diversity and the resulting attendance volatility, and to explain the potential emergence of periods of high volatility based on population belief changes.

For the BRATS approach (based on QH model), this corresponds to the entropy of agent resources β's, with each agent following a behaviour defined by equation (2.4) at any given time. Given the set R of all agent resources β_i across N agents ($1 \leq i \leq N$), the number of bins $|B|$ is chosen according to the Freedman–Diaconis rule for bin width W, as follows:

$$\left.\begin{aligned} W &= \frac{2\ \mathrm{IQR}(R)}{\sqrt{[3]N}} \\ \text{and} \qquad |B| &= \frac{\max(R) - \min(R)}{W} \end{aligned}\right\} \tag{C1}$$

Then, the entropy is defined across B bins:

$$\left.\begin{aligned} H_B &= -\sum_{b\in B} x_b \log_2 x_b, \\ \overline{H}_B &= -\sum_{b\in B} \frac{1}{|B|} \log_2 \frac{1}{|B|} \\ \text{and} \qquad H_\beta &= \frac{H_B}{\overline{H}_B} \end{aligned}\right\} \tag{C2}$$

where x_b is the proportion of agent resources β distributed within bin b.

royalsocietypublishing.org/journal/rsos R. Soc. Open Sci. 10: 221164

For the AS approach, the diversity is quantified by the entropy of the bitstring representation of the chosen strategies $\boldsymbol{s}^*{}_i$ across N agents $(1 \le i \le N)$ at a given time. For each agent i, in order to convert the strategy $\boldsymbol{s}^*{}_i$ into a bitstring $\boldsymbol{v}_i$, we use the following function converting an element $s_{ik} \in \boldsymbol{s}^*{}_i$ to the corresponding element $v_{ik} \in \boldsymbol{v}_i$:

$$v_{ik}(s_{ik}) = \begin{cases} 1, & \text{if } \mathrm{s_{ik}} \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{C3}$$

The bitstring $\boldsymbol{v}_i$ is then converted to its decimal representation

$$D_i = \sum_{j=1}^{|\vec{v}_i|} v_{ij} 2^{|\vec{v}_i|-j}, \tag{C4}$$

where v_{ij} is the jth element in bitstring $\boldsymbol{v}_i$ for agent i. The decimal representations D_i are then used to compute the entropy. As before, we form the set D of all the decimal representations D_i across N agents and choose the number of bins $|L|$ according to the Freedman–Diaconis rule. The entropy is computed as follows:

$$\left.\begin{aligned} H_L &= -\sum_{l\in L} x_l \log_2 x_l, \\ \overline{H}_L &= -\sum_{l\in L} \frac{1}{|L|} \log_2 \frac{1}{|L|} \\ \text{and} \qquad H_S &= \frac{H_L}{\overline{H}_L} \end{aligned}\right\} \tag{C5}$$

where x_l is the proportion of decimal representations D distributed within bin l.

We note that the used measure H_β quantifies the diversity of *beliefs* of the population, and not the diversity of *information*, as each agent observes the same past historical attendance. For the AS approach, the measure H_S quantifies the diversity of the *predictors*.

Appendix D. Granger causality

Granger causality is used to test if changes in diversity predict changes in attendance. In other words, this test is used to check if there is information in the (current) diversity about future volatility.

As non-stationarity is frequently observed in market settings [70] (as is the case in the market entrance game proposed here), to run the Granger causality test, we first consider stationary time series by taking the change in volatility and the change in diversity (first-order differencing). An example realization of these two time series for the proposed approach is presented in figure 8. The time series are confirmed to have unit-root with the Augmented Dickey–Fuller test, and stationarity with the Kwiatkowski–Phillips–Schmidt–Shin tests. A vector autoregression (VAR) model is then fitted to these time series. The lag-order L of the model is selected by minimizing the Akaike information criterion (AIC) [71]. This selection process is visualized in figure 9. VAR(L) is then used to test for Granger causality. For each c, the harmonic mean p-values [72] is given to provide an overall significance level across realizations, with a Bonferonni correction to account for multiple tests across c.

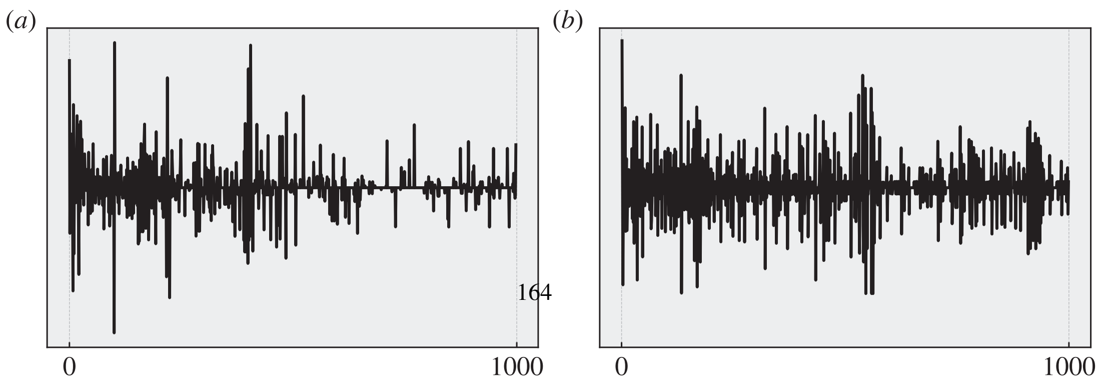


Figure 8. Example stationary transformations from one realization. (*a*) Diversity change and (*b*) volatility change.

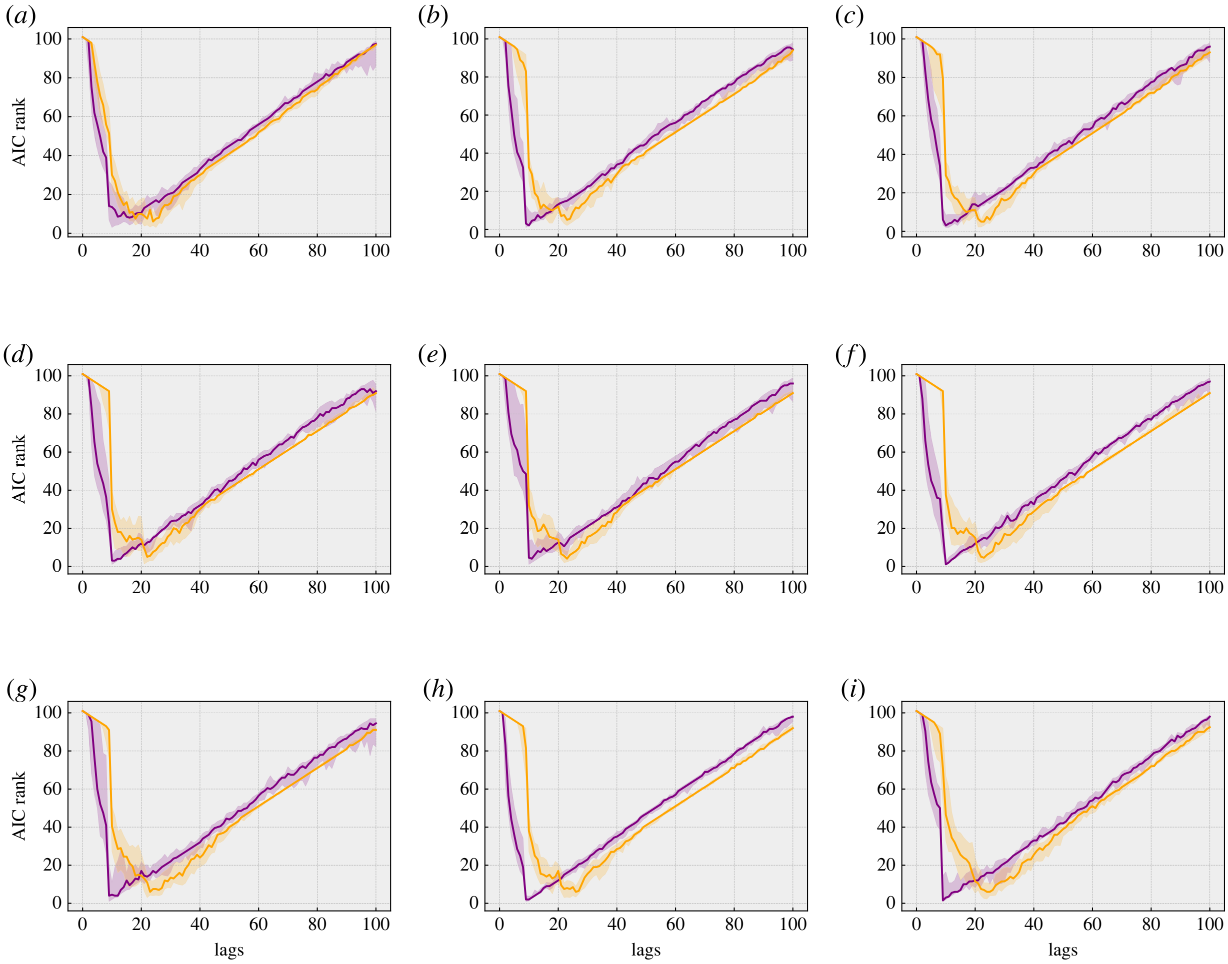


Figure 9. Lag length estimate based on expected rank of AIC, where lowest rank indicates best performance. Purple (orange) represents the proposed BRATS model (Adaptive Strategies). The darker line indicates the median rank, and the lighter filled region highlights the interquartile range. (*a*) $c = 0.1$, (*b*) $c = 0.2$, (*c*) $c = 0.3$, (*d*) $c = 0.4$, (*e*) $c = 0.5$, (*f*) $c = 0.6$, (*g*) $c = 0.7$, (*h*) $c = 0.8$ and (*i*) $c = 0.9$.

Table 3. Significance testing for Granger causality between the diversity of agent beliefs and local volatility in attendance in the proposed approach. All results are significant at the 95% confidence level.

Granger causality (diversity → volatility)									
c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
p^*	<0.01	<0.01	<0.01	<0.04	<0.02	<0.02	<0.01	<0.01	<0.01

The null hypothesis H_0 for this test is that diversity does not cause Granger volatility. Rejection of H_0 confirms if the population heterogeneity can be seen as a leading indicator for volatility in the proposed market game. Failure to reject H_0 means that there is no significant predictive information in population diversity about future attendance fluctuations. The resulting p-values from the significance tests are presented in table 3. The test for Granger causality is significant at the 95% confidence level across all entry rates, indicating we can reject H_0 and confirm that diversity change is a leading indicator of volatility in the market game. This result shows that the distribution of agent beliefs is a significant predictor of volatility in the market game.

References

1. Casson M, Lee JS. 2011 The origin and development of markets: a business history perspective. *Bus. Hist. Rev.* **85**, 9–37. (doi:10.1017/S0007680511000018)
2. Sewell M. 2011 *History of the efficient market hypothesis*. RN/11/04 London, UK: University College London.
3. Fama EF. 1970 Efficient capital markets: a review of theory and empirical work. *J. Finance* **25**, 383–417. (doi:10.2307/2325486)

4. Danielsson J, Shin HS. 2003 Endogenous risk. In *Modern risk management: A history* (ed. P Field), pp. 297–316. London, UK: Risk Books.
5. Ball R. 2009 The global financial crisis and the efficient market hypothesis: what have we learned? *J. Appl. Corp. Finance* **21**, 8–16. (doi:10.1111/j.1745-6622.2009.00246.x)
6. Kirilenko A, Kyle AS, Samadi M, Tuzun T. 2017 The flash crash: high-frequency trading in an electronic market. *J. Finance* **72**, 967–998. (doi:10.1111/jofi.12498)
7. Beja A, Goldman MB. 1980 On the dynamic behavior of prices in disequilibrium. *J. Finance* **35**, 235–248. (doi:10.1111/j.1540-6261.1980.tb02151.x)
8. Smith A, Lohrenz T, King J, Montague PR, Camerer CF. 2014 Irrational exuberance and neural crash warning signals during endogenous experimental market bubbles. *Proc. Natl Acad. Sci. USA* **111**, 10 503–10 508. (doi:10.1073/pnas.1318416111)
9. Shiller RJ. 2015 *Irrational exuberance*. Princeton, NJ: Princeton University Press.
10. Chiarella C. 1992 The dynamics of speculative behaviour. *Ann. Oper. Res.* **37**, 101–123. (doi:10.1007/BF02071051)
11. Buchanan M. 2009 Meltdown modelling: could agent-based computer models prevent another financial crisis? *Nature* **460**, 680–683. (doi:10.1038/460680a)
12. Farmer JD, Foley D. 2009 The economy needs agent-based modelling. *Nature* **460**, 685–686. (doi:10.1038/460685a)
13. Hommes CH. 2002 Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proc. Natl Acad. Sci. USA* **99**(suppl_3), 7221–7228. (doi:10.1073/pnas.082080399)
14. Evans BP, Prokopenko M. 2021 A maximum entropy model of bounded rational decision-making with prior beliefs and market feedback. *Entropy* **23**, 669. (doi:10.3390/e23060669)
15. Evans BP, Glavatskiy K, Harré MS, Prokopenko M. 2021 The impact of social influence in Australian real estate: market forecasting with a spatial agent-based model. *J. Econ. Interact. Coord.* **18**, 1–53. (doi:10.1007/s11403-021-00324-7)
16. Song R, Shu M, Zhu W. 2022 The 2020 global stock market crash: endogenous or exogenous? *Physica A* **585**, 126425. (doi:10.1016/j.physa.2021.126425)
17. Wehrli A, Sornette D. 2022 Classification of flash crashes using the Hawkes (p, q) framework. *Quant. Finance* **22**, 213–240. (doi:10.1080/14697688.2021.1941212)
18. Mandelbrot BB. 1997 The variation of certain speculative prices. In *Fractals and Scaling in Finance* (ed. BB Mandelbrot), pp. 371–418. New York, NY: Springer.
19. Inoua SM. 2020 News-driven expectations and volatility clustering. *J. Risk Financ. Manag.* **13**, 17. (doi:10.3390/jrfm13010017)
20. Hong H, Stein JC. 1999 A unified theory of underreaction, momentum trading, and overreaction in asset markets. *J. Finance* **54**, 2143–2184. (doi:10.1111/0022-1082.00184)
21. Buchanan M. 2017 Economists must broaden their horizons. *Nat. Phys.* **13**, 5–5. (doi:10.1038/nphys4011)
22. Bouchaud JP. 2008 Economics needs a scientific revolution. *Nature* **455**, 1181–1181. (doi:10.1038/4551181a)
23. Arthur WB. 1999 Complexity and the economy. *Science* **284**, 107–109. (doi:10.1126/science.284.5411.107)
24. Harré MS, Eremenko A, Glavatskiy K, Hopmere M, Pinheiro L, Watson S, Crawford L. 2021 Complexity economics in a time of crisis: heterogeneous agents, interconnections, and contagion. *Systems* **9**, 73.
25. Simon HA. 1956 Rational choice and the structure of the environment. *Psychol. Rev.* **63**, 129. (doi:10.1037/h0042769)
26. Levitt SD, List JA. 2008 Homo economicus evolves. *Science* **319**, 909–910. (doi:10.1126/science.1153640)
27. Camerer CF, Fehr E. 2006 When does 'economic man' dominate social behavior? *Science* **311**, 47–52. (doi:10.1126/science.1110600)
28. Arthur WB. 2021 Foundations of complexity economics. *Nat. Rev. Phys.* **3**, 136–145. (doi:10.1038/s42254-020-00273-3)
29. Kaizoji T. 2000 Speculative bubbles and crashes in stock markets: an interacting-agent model of speculative activity. *Physica A* **287**, 493–506. (doi:10.1016/S0378-4371(00)00388-5)
30. Lux T, Marchesi M. 1999 Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* **397**, 498–500. (doi:10.1038/17290)
31. Hommes CH. 2002 Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proc. Natl Acad. Sci. USA* **99**(suppl 3), 7221–7228. (doi:10.1073/pnas.082080399)
32. Alfarano S, Lux T, Wagner F. 2005 Estimation of agent-based models: the case of an asymmetric herding model. *Comput. Econ.* **26**, 19–49. (doi:10.1007/s10614-005-6415-1)
33. Schmitt N, Westerhoff F. 2017 Herding behaviour and volatility clustering in financial markets. *Quant. Finance* **17**, 1187–1203. (doi:10.1080/14697688.2016.1267391)
34. Zubillaga BJ, Vilela ALM, Wang C, Nelson KP, Stanley HE. 2022 A three-state opinion formation model for financial markets. *Physica A* **588**, 126527. (doi:10.1016/j.physa.2021.126527)
35. Vilela ALM, Wang C, Nelson KP, Stanley HE. 2019 Majority-vote model for financial markets. *Physica A* **515**, 762–770. (doi:10.1016/j.physa.2018.10.007)
36. Bornholdt S. 2001 Expectation bubbles in a spin model of markets: intermittency from frustration across scales. *Int. J. Mod. Phys. C* **12**, 667–674. (doi:10.1142/S0129183101001845)
37. Arthur WB. 1994 Inductive reasoning and bounded rationality. *Am. Econ. Rev.* **84**, 406–411.
38. Casti JL. 1996 Seeing the light at El Farol: a look at the most important problem in complex systems theory. *Complexity* **1**, 7–10. (doi:10.1002/cplx.6130010503)
39. Sellers MW, Sayama H, Pape AD. 2020 Simulating systems thinking under bounded rationality. *Complexity* **2020**, 3469263. (doi:10.1155/2020/3469263)
40. St Luce S, Sayama H. 2021 Network-based phase space analysis of the El Farol bar problem. *Artif. Life* **27**, 113–130.
41. Camerer CF, Ho TH, Chong JK. 2004 A cognitive hierarchy model of games. *Q. J. Econ.* **119**, 861–898. (doi:10.1162/0033553041502225)
42. Challet D, Marsili M, Zhang YC. 2001 Stylized facts of financial markets and market crashes in minority games. *Physica A* **294**, 514–524. (doi:10.1016/S0378-4371(01)00103-0)
43. Challet D, Marsili M, Ottino G. 2004 Shedding light on El Farol. *Physica A* **332**, 469–482. (doi:10.1016/j.physa.2003.06.003)
44. Evans BP, Prokopenko M. 2021 Bounded rationality for relaxing best response and mutual consistency: The Quantal Hierarchy model of decision-making. (http://arxiv.org/abs/2106.15844).
45. Keynes JM. 1937 The general theory of employment. *Q. J. Econ.* **51**, 209–223. (doi:10.2307/1882087)
46. Koppl R, Barkley Rosser Jr J. 2002 All that I have to say has already crossed your mind. *Metroeconomica* **53**, 339–360. (doi:10.1111/1467-999X.00147)
47. Ortega PA, Braun DA. 2013 Thermodynamics as a theory of decision-making with information-processing costs. *Proc. R. Soc. A* **469**, 20120683. (doi:10.1098/rspa.2012.0683)
48. Fogel DB, Chellapilla K, Angeline PJ. 1999 Inductive reasoning and bounded rationality reconsidered. *IEEE Trans. Evol. Comput.* **3**, 142–146. (doi:10.1109/4235.771167)
49. Wilensky U, Rand W. 2015 *An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo*. Cambridge, MA: MIT Press.
50. Whitehead D. 2008 The El Farol bar problem revisited: Reinforcement learning in a potential game. ESE discussion papers. 186.
51. Shubik M. 2011 El Farol revisited: a note on emergence, game theory, and society. *Complexity* **16**, 62–65. (doi:10.1002/cplx.20369)
52. Kay R, Johnson NF. 2004 Memory and self-induced shocks in an evolutionary population competing for limited resources. *Phys. Rev. E* **70**, 056101. (doi:10.1103/PhysRevE.70.056101)
53. Guzman M, Stiglitz JE. 2020 Towards a dynamic disequilibrium theory with randomness. *Oxf. Rev. Econ. Policy* **36**, 621–674. (doi:10.1093/oxrep/graa042)
54. Jefferies P, Lamper D, Johnson NF. 2003 Anatomy of extreme events in a complex adaptive system. *Physica A* **318**, 592–600. (doi:10.1016/S0378-4371(02)01743-0)
55. Johnson NF, Jefferies P, Hui PM. 2003 *Financial market complexity*. Oxford, UK: Oxford University Press.
56. Hill BM. 1975 A simple general approach to inference about the tail of a distribution. *Ann. Stat.* **3**, 1163–1174. (doi:10.1214/aos/1176343247)
57. Samanta R, LeBaron B. 2005 Extreme value theory and fat tails in equity markets. *Computing in Economics and Finance.* Society for Computational Economics.
58. Lux T, Marchesi M. 2000 Volatility clustering in financial markets: a microsimulation of interacting agents. *Int. J. Theor. Appl. Finance* **3**, 675–702. (doi:10.1142/S0219024900000826)

59. Brunnermeier MK. 2016 Bubbles. In *Banking Crises*, pp. 28–36. London, UK: Palgrave Macmillan.
60. Garber PM. 1989 Tulipmania. *J. Political Econ.* **97**, 535–560. (doi:10.1086/261615)
61. Mackay C. 1841 *Extraordinary popular delusions and the madness of crowds*. London, UK: Richard Bentley.
62. Bell AM, Sethares WA, Bucklew JA. 2003 Coordination failure as a source of congestion in information networks. *IEEE Trans. Signal Process.* **51**, 875–885. (doi:10.1109/TSP.2002.808139)
63. Kurz M, Motolese M. 2001 Endogenous uncertainty and market volatility. *Econ. Theory* **17**, 497–544.
64. Rand W, Stonedahl F. 2007 *The El Farol bar problem and computational effort: why people fail to use bars efficiently*. Evanston, IL: Northwestern University.
65. Lof M. 2015 Rational speculators, contrarians, and excess volatility. *Manag. Sci.* **61**, 1889–1901. (doi:10.1287/mnsc.2014.1937)
66. Bolt W, Demertzis M, Diks C, Hommes C, Van Der Leij M. 2019 Identifying booms and busts in house prices under heterogeneous expectations. *J. Econ. Dyn. Control* **103**, 234–259. (doi:10.1016/j.jedc.2019.04.003)
67. Zhang H, Huang Y, Yao H. 2016 Heterogeneous expectation, beliefs evolution and house price volatility. *Econ. Modell.* **53**, 409–418. (doi:10.1016/j.econmod.2015.10.039)
68. Bao T, Hommes C. 2019 When speculators meet suppliers: positive versus negative feedback in experimental housing markets. *J. Econ. Dyn. Control* **107**, 103730. (doi:10.1016/j.jedc.2019.103730)
69. Chen L. 2020 Complex network minority game model for the financial market modeling and simulation. *Complexity* **2020**, 8877886.
70. Schmitt TA, Chetalova D, Schäfer R, Guhr T. 2013 Non-stationarity in financial time series: generic features and tail behavior. *Europhys. Lett.* **103**, 58003. (doi:10.1209/0295-5075/103/58003)
71. Ivanov V, Kilian L. 2005 A practitioner's guide to lag order selection for VAR impulse response analysis. *Stud. Nonlinear Dyn. Econom.* **9**, 1–36.
72. Wilson DJ. 2019 The harmonic mean *p*-value for combining dependent tests. *Proc. Natl Acad. Sci. USA* **116**, 1195–1200. (doi:10.1073/pnas.1814092116)

CHAPTER 7

Conclusion

7.1 Summary

Throughout this thesis, we have provided methods for capturing the endogenous emergence of crises, volatility, and stylised facts in economic markets, from the interactions among bounded rational agents. This bounded rational reasoning can be captured in various ways, such as with hand-crafted heuristics, and is supported more generally by the proposed Quantal Hierarchy model for quantifying limitations in strategic reasoning. We have verified the proposed methods with actual market and underlying economic data, and demonstrated well-performing out-of-sample forecasting, outperforming a range of canonical approaches

This thesis provided essential insights, demonstrating the usefulness of multi-agent systems for capturing market dynamics from the bottom-up, and provided a contrasting perspective to the top-down efficient market hypothesis. This thesis supports the interacting agent hypothesis, explaining the endogenous emergence of out-of-equilibria phenomena through the interaction among bounded rational agents, while also providing a method of capturing the decision-making of such agents.

7.2 Contributions

The contributions of our work have been presented in four studies, addressing the three key objectives of the thesis. The key findings for each objective are presented as single-sentence summaries here.

7.2.1 Objectives

Through four studies, we have addressed our three key objectives.

Demonstrate how interactions among inductive reasoning agents can recreate key price dynamics. We developed an agent-based model which showed that interacting agents, acting according to heuristics influenced by various (not necessarily perfectly rational) factors such as social pressures, can capture key price dynamics in an economic market (Chapter 3), and supported these heuristics in more general information processing terms (Chapters 4 and 6).

Model strategic higher-order reasoning to capture decision-making in competitive markets. We developed the Quantal Hierarchy model, in which strategic reasoning can be quantified with information-processing costs, where higher-order reasoning requires additional processing resources (Chapter 5). We successfully applied this model to several canonical games.

Demonstrate how interactions among bounded strategic reasoning agents can lead to endogenous crisis emergence. We applied the Quantal Hierarchy model to the canonical El Farol bar problem and showed that bounded strategic reasoning can lead to crises emergence and the recreation of various stylised facts (such as clustered volatility) in a concise multi-agent model of economic markets (Chapter 6).

7.3 Future work

Throughout this thesis, we have outlined a novel framework for agent-based modelling, in which agents are capable of (bottom-up) higher-order strategic reasoning. Such an approach helps overcome some overly restrictive modelling assumptions made throughout traditional economics. There are still ways to go before such approaches become mainstream. Firstly, we have shown how such models can be used for out-of-sample forecasting (e.g. in Chapter 3), and other recent work is also beginning to highlight the usefulness of multi-agent methods for forecasting (Poledna et al., 2022; Gatti and Grazzini, 2020). Developing robust and generalisable models rigorously assessed through out-of-sample forecasting provides a key direction for future work on multi-agent models.

In this thesis, we used machine learning techniques for optimising models, for example, Bayesian optimisation for finding parameter configurations in our ABM and parameter values for the Quantal Hierarchy model. However, this optimisation can be taken a step further, where machine learning can be used to learn the behavioural rules of such systems (Zheng et al., 2022; Ardon et al., 2023). Combining machine learning and multi-agent systems promises to be a key direction for the next generation of market models, with AI-based simulation techniques such as reinforcement learning (Lussange et al., 2021). As an example, rather than the utility functions being manually defined, agent utility functions could be learnt automatically through revealed preference analysis with inverse reinforcement learning (Arora and Doshi, 2021).

With the developed Quantal Hierarchy model, following up on Chapter 5, the resource parameter β is problem dependent and depends on the payoff magnitude (in a similar manner to QRE, McKelvey et al., 2000). It would be interesting to analyse whether such a resource parameter can be related across different games (Haile et al., 2008), whether this parameter can be endogenised (Friedman, 2020), or whether this parameter perhaps follows some fundamental underlying law. This provides an interesting area for future research.

When the Quantal Hierarchy model is applied to sequential games, this can be seen as a type of sequential optimisation, and with $\beta \to \infty$, the Bellman optimality principle is recovered (analogous to the subgame perfect equilibrium solution). This relationship between the Quantal Hierarchy model and the Bellman equation opens an interesting potential connection with the active inference framework of Da Costa et al., 2020, and when considering beliefs about other agents beliefs, opens a possible link to the recently proposed framework of sophisticated inference (Friston et al., 2021). Investigating this connection promises to strengthen the thermodynamic foundations of the QH model.

Additionally, we have shown the usefulness of the Quantal Hierarchy representation for modelling human behaviour in a range of canonical games, and how a population of interacting

QH agents can generate key market phenomena in a concise market entrance game. Applying such an approach to a broader class of games, for example, generalised congestion games, could provide insights into a more comprehensive range of economic behaviour, e.g. capturing various asset types (or multiple markets). There is much room for applying the methods developed throughout this thesis to explain a broader range of economic phenomena not yet discussed.

7.4 Additional tool for policy-makers

This thesis showed how a concise multi-agent model can explain various stylised facts of economic markets, and the flexibility of such approaches also allows for in-depth fine-grained models to be developed that capture pricing dynamics (and downturns) of actual markets, such as the Australian housing market.

This demonstrates that multi-agent systems can capture economic phenomena unexplainable by traditional equilibrium approaches, such as the endogenous formation of stylised facts and crises. The work proposed provides an important additional tool to conventional models, which work well in general market times, but struggle to capture punctuated out-of-equilibrium phenomena, leaving policymakers "abandoned" during such times. The methods developed can provide insights into financial vulnerability (through analysis of market stability), highlight potential investment opportunities, and assist policymakers during complicated market states. In Chapter 1, we discussed some of the concerns of top policymakers with conventional approaches to market modelling, particularly under crises and uncertainty. The methods proposed here, which relax rational expectations, efficiency, and (economic) equilibrium, provide an additional tool at the disposal of such policymakers, which can be used to consider various market scenarios (for example, through counter factual analysis), even when the resulting dynamics are out-of-equilibrium. To paraphrase the opening quote of this thesis (from Herbert Simon), as economics finds it more and more necessary to understand out-of-equilibrium phenomena, the types of methods proposed here will find increased usage and importance.

Bibliography

Abergel, Frédéric et al. (2013). *Econophysics of agent-based models*. Springer.

Ackert, Lucy and Richard Deaves (2009). *Behavioral finance: Psychology, decision-making, and markets*. Cengage Learning.

Agliari, Anna, Ahmad Naimzada and Nicolò Pecora (2018). 'Boom-bust dynamics in a stock market participation model with heterogeneous traders'. In: *Journal of Economic Dynamics and Control* 91. Special Issue in Honour of Prof. Carl Chiarella, pp. 458–468.

Ardon, Leo et al. (2023). 'Phantom-A RL-driven Multi-Agent Framework to Model Complex Systems'. In: *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, pp. 2742–2744.

Arora, Saurabh and Prashant Doshi (2021). 'A survey of inverse reinforcement learning: Challenges, methods and progress'. In: *Artificial Intelligence* 297, p. 103500.

Arthur, W Brian (1994). 'Inductive reasoning and bounded rationality'. In: *American Economic Review* 84.2, pp. 406–411.

Arthur, W Brian (1999). 'Complexity and the economy'. In: *Science* 284.2, pp. 107–109.

Arthur, W Brian, ED Beinhocker and A Stanger (2020). *Complexity economics: Proceedings of the Santa Fe Institute's 2019 fall symposium (dialogues of the applied complexity network)*.

Axtell, Robert L and J. Doyne Farmer (2022). 'Agent-based modeling in economics and finance: Past, present, and future'. In: *Journal of Economic Literature*.

Axtell, Robert L et al. (2014). 'An agent-based model of the housing market bubble in metropolitan Washington, DC'. In: *Whitepaper for Deutsche Bundesbank's Spring Conference on "Housing markets and the macroeconomy: Challenges for monetary policy and financial stability"*.

Bookstaber, Richard (2017). 'The end of theory'. In: *The End of Theory*. Princeton University Press.

Bouchaud, Jean-Philippe (2008). 'Economics needs a scientific revolution'. In: *Nature* 455.7217, pp. 1181–1181.

Bouchaud, Jean-Philippe, Marc Potters et al. (2003). *Theory of financial risk and derivative pricing: from statistical physics to risk management*. Cambridge university press.

Camerer, Colin F (2010). 'Behavioural game theory'. In: *Behavioural and Experimental Economics*. Springer, pp. 42–50.

Camerer, Colin F. (2003). 'Behavioural studies of strategic thinking in games'. In: *Trends in cognitive sciences* 7.5, pp. 225–231.

Camerer, Colin F., Teck-Hua Ho and Juin Kuan Chong (2004a). 'Behavioural game theory: thinking, learning and teaching'. In: *Advances in understanding strategic behaviour*. Springer, pp. 120–180.

Camerer, Colin F., Teck-Hua Ho and Juin-Kuan Chong (2004b). 'A cognitive hierarchy model of games'. In: *The Quarterly Journal of Economics* 119.3, pp. 861–898.

Casti, John L (1996). 'Seeing the light at El Farol: A look at the most important problem in complex systems theory'. In: *Complexity* 1.5, pp. 7–10.

Chakraborti, Anirban et al. (2011a). 'Econophysics review: I. Empirical facts'. In: *Quantitative Finance* 11.7, pp. 991–1012.

Chakraborti, Anirban et al. (2011b). 'Econophysics review: II. Agent-based models'. In: *Quantitative Finance* 11.7, pp. 1013–1041.

Chen, Shu-Heng, Chia-Ling Chang and Ye-Rong Du (2012). 'Agent-based economic models and econometrics'. In: *The Knowledge Engineering Review* 27.2, pp. 187–219.

Chong, Juin-Kuan, Colin F Camerer and Teck-Hua Ho (2005). 'Cognitive Hierarchy: A Limited Thinking Theory in Games'. In: *Experimental Business Research*. Ed. by Rami Zwick and Amnon Rapoport. Boston, MA: Springer US, pp. 203–228.

Cincotti, Silvano, Marco Raberto and Andrea Teglio (2022). 'Why do we need agent-based macroeconomics?' In: *Review of Evolutionary Political Economy* 3.1, pp. 5–29.

Cont, Rama and Jean-Philipe Bouchaud (2000). 'Herd behavior and aggregate fluctuations in financial markets'. In: *Macroeconomic dynamics* 4.2, pp. 170–196.

Da Costa, Lancelot et al. (2020). 'The relationship between dynamic programming and active inference: The discrete, finite-horizon case'. In: *arXiv preprint arXiv:2009.08111*.

De Bondt, Werner FM and Richard H Thaler (1995). 'Financial decision-making in markets and firms: A behavioral perspective'. In: *Handbooks in operations research and management science* 9, pp. 385–410.

Evans, Benjamin Patrick and Mikhail Prokopenko (2021). 'A maximum entropy model of bounded rational decision-making with prior beliefs and market feedback'. In: *Entropy* 23.6, p. 669.

Evans, Benjamin Patrick and Mikhail Prokopenko (2023a). 'Bounded rationality for relaxing best response and mutual consistency: the quantal hierarchy model of decision making'. In: *Theory and Decision*, pp. 1–41.

Evans, Benjamin Patrick and Mikhail Prokopenko (2023b). 'Bounded strategic reasoning explains crisis emergence in multi-agent market games'. In: *Royal Society Open Science* 10.2, p. 221164.

Evans, Benjamin Patrick et al. (Jan. 2023). 'The impact of social influence in Australian real estate: market forecasting with a spatial agent-based model'. In: *Journal of Economic Interaction and Coordination* 18.1, pp. 5–57.

Farmer, J. Doyne (2000). 'Physicists attempt to scale the ivory towers of finance'. In: *International Journal of Theoretical and Applied Finance* 3.03, pp. 311–333.

Farmer, J. Doyne and Duncan K Foley (2009). 'The economy needs agent-based modelling'. In: *Nature* 460.7256, pp. 685–686.

Foley, Duncan K (2003). 'Statistical equilibrium in economics: method, interpretation, and an example'. In: *General Equilibrium*. Routledge, pp. 95–116.

Foley, Duncan K et al. (2017). *Crisis and theoretical methods: equilibrium and disequilibrium once again*. Tech. rep.

Foley, Duncan K (2020). 'Information theory and behavior'. In: *The European Physical Journal Special Topics* 229.9, pp. 1591–1602.

Friedman, Evan (2020). 'Endogenous quantal response equilibrium'. In: *Games and Economic Behavior* 124, pp. 620–643.

Friston, Karl et al. (2021). 'Sophisticated inference'. In: *Neural Computation* 33.3, pp. 713–763.

Garnier-Brun, Jerome, Jean-Philippe Bouchaud and Michael Benzaquen (2022). 'Bounded Rationality and Animal Spirits: A Fluctuation-Response Approach to Slutsky Matrices'. In: *arXiv preprint arXiv:2206.04468*.

Gatti, Domenico Delli and Jakob Grazzini (2020). 'Rising to the challenge: Bayesian estimation and forecasting techniques for macroeconomic Agent Based Models'. In: *Journal of Economic Behavior & Organization* 178, pp. 875–902.

Gatti, Domenico Delli et al. (2011). *Macroeconomics from the Bottom-up*. Vol. 1. Springer Science & Business Media.

Geanakoplos, John et al. (2012). 'Getting at systemic risk via an agent-based model of the housing market'. In: *American Economic Review* 102.3, pp. 53–58.

Gershman, Samuel J, Eric J Horvitz and Joshua B Tenenbaum (2015). 'Computational rationality: A converging paradigm for intelligence in brains, minds, and machines'. In: *Science* 349.6245, pp. 273–278.

Golan, Amos and Duncan K Foley (2022). 'Understanding the Constraints in Maximum Entropy Methods for Modeling and Inference'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Gontis, V. et al. (2016). 'Stochastic model of financial markets reproducing scaling and memory in volatility return intervals'. In: *Physica A: Statistical Mechanics and its Applications* 462, pp. 1091–1102.

Haile, Philip A, Ali Hortaçsu and Grigory Kosenok (2008). 'On the empirical content of quantal response equilibrium'. In: *American Economic Review* 98.1, pp. 180–200.

Harré, Michael S (2021). 'Information theory for agents in artificial intelligence, psychology, and economics'. In: *Entropy* 23.3, p. 310.

Hartley, Ralph VL (1928). 'Transmission of information 1'. In: *Bell System technical journal* 7.3, pp. 535–563.

Hommes, Cars H (2001). 'Financial markets as nonlinear adaptive evolutionary systems'. In: *Quantitative Finance* 1.1, p. 149.

Hommes, Cars H. (2002). 'Modeling the stylized facts in finance through simple nonlinear adaptive systems'. In: *Proceedings of the National Academy of Sciences* 99.suppl_3, pp. 7221–7228.

Jaynes, Edwin T (1957a). 'Information theory and statistical mechanics'. In: *Physical review* 106.4, p. 620.

Jaynes, Edwin T (1957b). 'Information theory and statistical mechanics. II'. In: *Physical review* 108.2, p. 171.

Kahneman, Daniel (2003). 'Maps of bounded rationality: Psychology for behavioral economics'. In: *American economic review* 93.5, pp. 1449–1475.

Kaizoji, Taisei (2000). 'Speculative bubbles and crashes in stock markets: an interacting-agent model of speculative activity'. In: *Physica A: Statistical Mechanics and its Applications* 287.3-4, pp. 493–506.

Kaizoji, Taisei (2006). 'An interacting-agent model of financial markets from the viewpoint of nonextensive statistical mechanics'. In: *Physica A: Statistical Mechanics and its Applications* 370.1, pp. 109–113.

Keynes, John Maynard (1937). 'The general theory of employment'. In: *The quarterly journal of economics* 51.2, pp. 209–223.

Koppl, Roger and J Barkley Rosser Jr (2002). 'All that I have to say has already crossed your mind'. In: *Metroeconomica* 53.4, pp. 339–360.

Kutner, Ryszard, Christophe Schinckus and Harry Eugene Stanley (2022). *Three Risky Decades: A Time for Econophysics?*

Kutner, Ryszard et al. (2019). *Econophysics and sociophysics: Their milestones & challenges*.

Laloux, Laurent et al. (1999). 'Noise dressing of financial correlation matrices'. In: *Physical review letters* 83.7, p. 1467.

LeBaron, Blake (2002). 'Building the Santa Fe artificial stock market'. In: *Physica A* 1, p. 20.

LeBaron, Blake, W Brian Arthur and Richard Palmer (1999). 'Time series properties of an artificial stock market'. In: *Journal of Economic Dynamics and control* 23.9-10, pp. 1487–1516.

LeBaron, Blake and Leigh Tesfatsion (2008). 'Modeling macroeconomies as open-ended dynamic systems of interacting agents'. In: *American Economic Review* 98.2, pp. 246–50.

Lewis, Richard L, Andrew Howes and Satinder Singh (2014). 'Computational rationality: Linking mechanism and behavior through bounded utility maximization'. In: *Topics in cognitive science* 6.2, pp. 279–311.

Löfgren, Lars (1979). 'Unfoldment of self-reference in logic and in computer science'. In.

Lussange, Johann et al. (2021). 'Modelling stock markets by multi-agent reinforcement learning'. In: *Computational Economics* 57.1, pp. 113–147.

Lux, Thomas (1995). 'Herd behaviour, bubbles and crashes'. In: *The economic journal* 105.431, pp. 881–896.

Lux, Thomas and Michele Marchesi (1999). 'Scaling and criticality in a stochastic multi-agent model of a financial market'. In: *Nature* 397.6719, pp. 498–500.

McKelvey, Richard D, Andrew M McLennan and Theodore L Turocy (2016). 'Gambit: Software tools for game theory'. In.

McKelvey, Richard D and Thomas R Palfrey (1995). 'Quantal response equilibria for normal form games'. In: *Games and economic behavior* 10.1, pp. 6–38.

McKelvey, Richard D and Thomas R Palfrey (1998). 'Quantal response equilibria for extensive form games'. In: *Experimental economics* 1.1, pp. 9–41.

McKelvey, Richard D, Thomas R Palfrey and Roberto A Weber (2000). 'The effects of payoff magnitude and heterogeneity on behavior in 2×2 games with unique mixed strategy equilibria'. In: *Journal of Economic Behavior & Organization* 42.4, pp. 523–548.

Nash, John (1951). 'Non-cooperative games'. In: *Annals of mathematics*, pp. 286–295.

Nyman, Rickard, Sujit Kapadia and David Tuckett (2021). 'News and narratives in financial systems: exploiting big data for systemic risk assessment'. In: *Journal of Economic Dynamics and Control* 127, p. 104119.

Nyquist, Harry (1924). 'Certain factors affecting telegraph speed'. In: *Transactions of the American Institute of Electrical Engineers* 43, pp. 412–422.

Ömer, Özlem (2020). 'Maximum entropy approach to market fluctuations as a promising alternative'. In: *The European Physical Journal Special Topics* 229.9, pp. 1715–1733.

Ortega, Pedro A and Daniel A Braun (2013). 'Thermodynamics as a theory of decision-making with information-processing costs'. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 469.2153, p. 20120683.

Palmer, RG et al. (1999). 'An artificial stock market'. In: *Artificial Life and Robotics* 3.1, pp. 27–31.

Poledna, Sebastian et al. (2022). 'Economic forecasting with an agent-based model'. In: *European Economic Review*, p. 104306.

Prokopenko, Mikhail, Fabio Boschetti and Alex J Ryan (2009). 'An information-theoretic primer on complexity, self-organization, and emergence'. In: *Complexity* 15.1, pp. 11–28.

Rosser Jr, J Barkley (2021). 'Econophysics and the entropic foundations of economics'. In: *Entropy* 23.10, p. 1286.

Scharfenaker, Ellis and Duncan K Foley (2017). 'Quantal response statistical equilibrium in economic interactions: Theory and estimation'. In: *Entropy* 19.9, p. 444.

Scharfenaker, Ellis and Jangho Yang (2020a). 'Maximum entropy economics'. In: *The European Physical Journal Special Topics* 229.9, pp. 1577–1590.

Scharfenaker, Ellis and Jangho Yang (2020b). *Maximum entropy economics: where do we stand?*

Schinckus, Christophe (2012). 'When Physicists Invade Economics'. In: *Interdisciplinary Science Reviews* 37.2, pp. 179–186.

Shannon, Claude Elwood (1948). 'A mathematical theory of communication'. In: *The Bell system technical journal* 27.3, pp. 379–423.

Sharma, Dhruv et al. (2021). 'V–, U–, L–or W–shaped economic recovery after Covid-19: Insights from an Agent Based Model'. In: *PloS one* 16.3, e0247823.

Shiller, Robert J (2015a). 'Irrational exuberance'. In: *Irrational exuberance*. Princeton university press.

Shiller, Robert J (2015b). 'The housing market still isn't rational'. In: *The Upshot. The New York Times*.

Simon, Herbert A (1957). *Models of man; social and rational.* Wiley.

Simon, Herbert A. (1959). 'Theories of Decision-Making in Economics and Behavioral Science'. In: *The American Economic Review* 49.3, pp. 253–283.

Sims, Christopher A (2003). 'Implications of rational inattention'. In: *Journal of monetary Economics* 50.3, pp. 665–690.

Stahl, Dale O and Paul W Wilson (1995). ‘On players’ models of other players: Theory and experimental evidence’. In: *Games and Economic Behavior* 10.1, pp. 218–254.

Stanley, H Eugene et al. (1996). ‘Anomalous fluctuations in the dynamics of complex systems: from DNA and physiology to econophysics’. In: *Physica A: Statistical Mechanics and its Applications* 224.1-2, pp. 302–321.

Strahilevitz, Michal Ann, Terrance Odean and Brad M Barber (2011). ‘Once burned, twice shy: How naive learning, counterfactuals, and regret affect the repurchase of stocks previously sold’. In: *Journal of Marketing Research* 48.SPL, S102–S120.

Tesfatsion, Leigh (2002). ‘Agent-based computational economics: Growing economies from the bottom up’. In: *Artificial life* 8.1, pp. 55–82.

Volcker, Paul A (2011). *Financial reform: unfinished business*.

Von Neumann, John and Oskar Morgenstern (2007). *Theory of games and economic behavior (commemorative edition)*. Princeton university press.

Yang, Jangho (2018). ‘Information theoretic approaches in economics’. In: *Journal of Economic Surveys* 32.3, pp. 940–960.

Yang, Jangho (2022). ‘Information-theoretic model of induced technical change: Theory and empirics’. In: *Metroeconomica*.

Yang, Jangho and Adrián Carro (2020). ‘Two tales of complex system analysis: MaxEnt and agent-based modeling’. In: *The European Physical Journal Special Topics* 229.9, pp. 1623–1643.

Zheng, Stephan et al. (2022). ‘The AI Economist: Taxation policy design via two-level deep multiagent reinforcement learning’. In: *Science Advances* 8.18, eabk2607.