Imperial College Business School

Imperial College London

University of London

Forecasting with Agent Games

Ross Elliot Richardson

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy of the University of London and the Diploma of Imperial College, September 2009

Abstract

This thesis investigates the use of the Minority Game and its variants in the forecasting and simulation of financial markets. We specifically consider the application of the Kalman Filter and the Bionomic Algorithm in the estimation of the strategy weight parameters, and apply this in a forecasting framework to Minority Game data and a variety of financial market data. In addition, we outline the theoretical foundation for the development of an Oil market using ideas from the Minority Game and Mix Game on a network, and perform an analysis of a financial market using concepts from the Minority Game.

Chapter 2 discusses the inadequacies of existing economic and financial modelling and advances the Minority Game as an alternate market model. Chapter 3 describes the process of building and validating the Minority Game and Mix Game. It describes the observance of a new phenomenon during simulation of a finite time-horizon Minority Game. An empirical analysis of the S&P500 Composite Index is presented, motivated by ideas regarding the conditional probability of the sign of price increments. The conceptual outline of an Oil market inspired by the ideas of the Minority Game is introduced.

Chapter 4 presents the Minority Game Forecasting Framework and the Kalman Filter, and introduces two simplifications of an existing forecast model, providing examples of their application in predicting the direction of price increments in Minority Game and financial market time-series.

Chapter 5 introduces the Bionomic Algorithm, a population heuristic specialised for the directed search of a large dimensional parameter space, and applies this to the strategy weight estimation problem. A new particularization of the Bionomic Algorithm suitable for the Minority Game and its variants is developed, and is used in the Minority Game Forecasting Framework to forecast the direction of price increments of a variety of different asset classes.

Acknowledgments

I would like to take this opportunity to thank all those who have supported me, in one way or another, throughout this undertaking. In particular, I am extremely grateful to Professor Nigel Meade and Professor Nicos Christofides for their guidance, enthusiasm, flexibility and open-mindedness during my time at Imperial. Many thanks also go to the BT Group Treasury team, especially Graham, Andy, Adam, Les, Raymond and Tom, for an interesting and enjoyable three years in Newgate Street ('the gaol'), packed with fascinating discussions, attractive auditors - an insight into the 'real world' of "sexy finance" (Les) - and the opportunity to eat so much good sushi on the steps of St. Paul's at lunchtime. Closer to my academic home, there are a few members at the Business School who have given me help, support and advice when needed, in particular Julie Paranics, Robert MacCulloch, David Lefevre, Robert Kosowski and Andrea Buraschi.

Escaping the confines of the Tanaka Building though still within touching distance, Alex LeMaison's computing advice was exceptionally helpful, as was the dedicated support from Simon Burbidge and Matt Harvey of the High Performance Computing Group. It was also a pleasure to find a safe haven from the Efficient Market Hypothesis and General Equilibrium Models just across the road in the Mathematical Institute, where an eclectic mix of fascinating research is evolving. Many thanks to Kim Christensen and Henrik Jensen, the founders of the Complexity Science group, for establishing and nurturing this cauldron of new ideas, long may it continue! With a soft spot in my heart for the old days of the Blackett Laboratory, it was great to find somebody else on campus who'd actually heard of the Minority Game in the form of my old Physics supervisor, Tim Evans. I look forward to continuing this interesting line of research with him in the near future.

Many thanks also go to Elly, Fiona, Joao, Jude, Martyn, Meurig, the Security guys, Vlad and the rest of the Garden & Weeks posse for a fantastically fun couple of years living it up in Knightsbridge, enabling me to have probably the shortest commute of anyone at Imperial. I'll miss them all dearly, but I think I've had enough Pizza, Sausage Rolls and Basics Vodka to last a lifetime.

I have made many friends during my time at Imperial, sharing many happy times together. A special mention goes to Adam, Adele, Alec, Alex, Bruno, Celina, Despina, Eduardo, Evelina, Ewan, Filip, Jair, Lu, Lucy, Maria, Marina, Martin, Nick, Paul, Pete, Poh-Choo, Richard, Saeed and Zheng, though I'm sure I'm missing a few.

It should not go without mention the unending love and support (both emotional and financial!) of my family, the battling spirit instilled by my mother, the inquisitive scientist nurtured by my father, and the creative inspiration of Zoe.

Lastly though certainly not least, I am eternally grateful to Ana for giving me encouragement and distractions in appropriate measure throughout the whole PhD journey, always with generosity, love and support.

Dedication

My family

Contents

List of Figures

List of Tables

Terminology

N number of agents

m memory length of agents

s number of strategies that each agent has

T time-horizon

τ score decay rate

L resource level

c confidence level that determines the points-threshold r

r is the points-threshold that determines whether a strategy is active or inactive ($r = (2c - 1)T$)

$\alpha = \frac{2^m}{N}$ is a control parameter that describes the 'phase' of the system - the amount of crowding across the strategy space

μ global information / state of the system

R strategy number

a_R^μ action of strategy R given the state μ

$D[t]$ excess/aggregate demand at time t

$P[t]$ price of asset at time t

$\lambda[t]$ market depth parameter

$V[t]$ trading volume of asset at time t

$S_R[t]$ is the score of strategy R at time t

$Ab[t]$ is the number of active strategies advising to buy at time t

$As[t]$ is the number of active strategies advising to sell at time t

Chapter 1

Introduction

This thesis is an investigation into the potential use of agent-based models, or more specifically the Minority Game and its variants, to simulate the statistical properties of, and to forecast the evolution of financial market time series. 'Agent-based' models in this case refer to the collection of models featuring individual, autonomous entities existing in some sort of society or ecology. Such agents in our models represent *speculative traders*, and the agents' interactions with each other through their financial transactions create a society we call *the market*.

The motivation underlying our endeavour stems from the desire to build simplistic models which resemble, at some level, the actual microscopic workings of a financial market, seen here to consist of speculating agents and a pricing mechanism relating aggregate demand to asset price changes. As Johnson states, "No matter how sophisticated one's models are for analyzing and managing risk, and pricing derivatives, it is the limitations of the underlying model describing the market dynamics that determine the applicability of the resulting predictions." [Johnson et al. 2003] The driving force behind this approach therefore, is the hope that by adding physically realistic structure - the 'human' or 'social' element of a market - to our model of a generic financial market, we may improve our understanding of its behaviour and our predictions of its future evolution. Indeed, Gupta argues that "All other things being equal, we believe that such a model may be intrinsically 'better' than a purely numerical multivariate one – and may even be preferable to many more sophisticated models such as certain neural network approaches, which also may not be capturing a realistic representation of the microscopic details of a physical market." [Gupta et al. 2005a]

There are two main aspects considered in this thesis. Firstly, the use of computational simulations of agent based models (specifically the Minority Game and the Mix Game) to reproduce the statistical properties of financial market time series, in particular the asset price increments, the trading volume, and the expected conditional probability of the sign of the next price increment. We focus on these topics in Chapters 2 and 3. Secondly, we develop techniques for calibrating these agent based models to time series using the Kalman Filter in Chapter 4 and then a population heuristic, the Bionomic Algorithm in Chapter 5. In these chapters, we subsequently apply the calibrated agent based models to financial data from a range of asset classes and examine their ability to forecast the direction of future asset price changes.

1.1 Thesis Outline

In Chapter 2 on page 26, we begin by outlining the development of the Minority Game and its variant 'the Mix Game' in the existing literature, having grown out of the fields of Statistical Physics and Complexity Science. We discuss the shortcomings of traditional models of financial time series and motivate the use of the Minority Game as a way of rectifying some of these. We explain the key concepts and variables of the Minority and Mix Games, so that we can build upon these in the later chapters.

In Chapter 3 on page 62, we discuss the building and validation of the Minority Game and the Mix Game. We display the output from some of the simulations and examine related data from financial markets. We discuss the possibility of developing similar agent based models that include network structures and suggest applying them to simulate markets featuring oligopolies such as OPEC in the oil market. The contributions of this chapter are:

- The novel empirical investigation of financial markets through the framework of the Minority Game, supporting the view that there is predictability in the S&P 500 Composite Index.
- The observation of finite-size effects of the Mix-Game model and its relation with financial data from the Shanghai Stock Exchange.
- The observation of a phase change phenomenon not discussed in the literature to our knowledge.
- The conceptual framework to build a new model of an oil market, derived from Minority Game and Mix Game concepts.

In Chapter 4 on page 88, we outline the method of reverse engineering the Minority Game model in order to train on, and subsequently forecast, time series data. We employ the Kalman Filter to calibrate the Minority Game to time series data and create a Minority Game Forecasting Framework (MGFF) that makes short-term predictions regarding the direction of future price increments of simulated Minority Game data and financial market data. The primary benefits of using the Kalman Filter in the MGFF are its dynamic ability to track the changing features of the agents' strategies and its applicability in online implementation. The computational time efficiency of the Kalman Filter could allow one to use this model to perform high-frequency forecasting, which may be useful in high-frequency trading. The contributions of this chapter are:

- A new version of the forecasting framework (Simple MGFF) that is computationally faster than existing models, and can therefore converge to an optimal solution more quickly. It can also assess patterns of greater complexity in the input data than the existing version.
- A variation of the existing forecasting framework that simplifies the number of parameters to estimate (Expanded MGFF).
- The analysis of the two forecasting frameworks over new financial data sets.

In Chapter 5 on page 144, we explore the use of population heuristics (the Bionomic Algorithm) to replace the Kalman Filter in the parameter estimation problem of the Expanded MGFF. Although the drawback of using the Bionomic Algorithm is largely related to the time necessary to perform and the assumption of static estimation parameters over the length of the time series data, the nature of the algorithm's directed search over parameter space is well suited to high-dimensional estimation problems such as in the Expanded MGFF. The Bionomic Algorithm's promising ability to escape local optima and to efficiently find the global optimum suggests that it may be possible to produce superior results. The contributions of this chapter are:

- The novelty in applying the Bionomic Algorithm (BA) to agent-based modelling.
- The particularisation of the Bionomic Algorithm to the MGFF estimation problem.
- The application of the Bionomic MGFF to the forecast of Minority Game data and financial market data.

Chapter 2

Introducing the Minority Game

"The whole is more than the sum of the parts." [Aristotle]

In this chapter, we outline the origins of the Minority Game and its variant 'the Mix Game' in the existing literature, as important examples of complex adaptive systems. We emphasize the importance of the emergent properties of such complex systems and highlight the shortcomings of traditional models of financial markets, motivating the use of the Minority Game as a way of rectifying some of these. In addition, we explain the key concepts and variables of the Minority and Mix Games, so that we can build upon these in the later chapters.

2.1 Complexity and the Minority Game

There is no unique definition of a Complex System; "Complexity is not easy to define. Worse still, it can mean different things to different people." [Johnson 2007] Whereas Computer Science has its own definition of complexity[1], the conceptual basis of the Complex Systems studied in this thesis was mostly developed at the Sante Fe Institute, New Mexico, USA, by interdisciplinary researchers

[1] also known as algorithmic complexity theory, where 'the complexity of a given object coded in an n-digit binary sequence is given by the bit length... of the shortest computer program that can print the given symbolic sequence' [Mantegna, Stanley 2000]

from a multitude of diverse disciplines such as Economics, Biology and Physics. They propose that a Complex System is a collection of interconnected components, which displays behaviour unexpected from merely an understanding of the individual components. Indeed, a description put forth by Johnson, whose book is entitled 'Two's Company, Three Is Complexity', is that complex systems are "more than the sum of their parts" [Johnson et al. 2003] or to quote Nobel Prize winning physicist Philip W. Anderson, "more is different" [Anderson 1972] - there are rich dynamics displayed by such 'many-body' systems.

The Minority Game is an example of a further specialization, a Complex Adaptive System. This is a Complex System that can learn from experience and adapt to current conditions. Formally proposed by Nobel Prize winning physicist Murray Gell-Man, John H. Holland and others at Sante Fe,

> "a Complex Adaptive System (CAS) is a dynamic network of many agents (which may represent cells, species, individuals, firms, nations) acting in parallel, constantly acting and reacting to what the other agents are doing. The control of a CAS tends to be highly dispersed and decentralized. If there is to be any coherent behavior in the system, it has to arise from competition and cooperation among the agents themselves. The overall behavior of the system is the result of a huge number of decisions made every moment by many individual agents." (Holland in [Waldrop 1994])

This sounds remarkably like a description of a capitalist, free-market economy.

One can get an impression of what we mean by a CAS when looking at systems that are said to exhibit 'Complexity'. The examples that are often cited include the dynamics of road traffic, data flow on computer networks, evolution in ecological systems, the biological immune system, and financial markets. Collectively, these systems all feature meta-stable (punctuated[2]) equilibria, are

[2]Eldredge and Gould introduced the term punctuated equilibria to describe large changes occurring sporadically, throughout a static or slowly evolving system, as found in the fossil record time-line of the existence of species [Eldredge, Gould 1972]

highly non-linear with long-memory (the Joseph effect[3]), seemingly stochastic behaviour, and whose dynamics suffer from extreme events (the Noah effect[4]) - traffic jams on roads, computer network crashes, species extinction, immune system collapse, and market crashes respectively. Indeed the dynamics are often dominated by such outliers [Jefferies et al. 2002] - the overall yearly profit or loss of a bank can mostly depend on large changes in the market that occur on just a few trading days. No one needs a reminder of just how deadly this effect can be, with the collapse of Lehman Brothers in September 2008, and the subsequent bail-outs of the American International Group (AIG) and the Royal Bank of Scotland (RBS) by the US and UK governments respectively.

In the field of finance, these shocks have traditionally been modelled as exogenous aberrations of the normal *status quo* of market equilibrium, down to external stochastic factors and thus impossible to predict [Frisch 1933, Slutsky 1937, Lucas 1978, Malkiel 1996, Taleb 2008]. However, it has been empirically established that financial market volatility is largely due to the human activity of trading itself, not solely due to external shocks or new information that becomes available [French, Roll 1986, Cutler et al. 1989, Hull 2003], nor to the change in any economic fundamentals [Shiller 1989]. Likewise, the interaction of people in a marketplace, the 'bulls' and the 'bears' and the complex dynamics that result can be seen as a defining feature of a Complex System. It has been known for centuries that it is impossible to analytically solve a 'many' (more than two)-body problem such as the orbits of the planets around the sun. It is only relatively recently however, that advances in computing power have enabled the possibility to simulate such many-body systems, and interdisciplinary institutions such as the Sante Fe Institute, have sprung up to study Complexity Science with the convergent thinking and collaborative efforts of a multitude of diverse academic disciplines. Indeed Anderson claims that agent-based modelling, the major experimental apparatus of this thesis, has become something of a trademark at Sante Fe [Anderson 2001].

[3]Proposed by Mandelbrot and Wallis to describe the unusual behaviour of water levels in the Nile river, which supposedly gave rise to the seven years of plenty followed by seven years of famine in the story of Joseph in the Book of Genesis [Mandelbrot, Wallis 1968]

[4]Proposed by Mandelbrot and Wallis to describe exceptionally large amplitudal changes observed in a process, and named after the great flood in the Book of Genesis [Mandelbrot, Wallis 1968]

The rich yet seemingly unpredictable behaviour exhibited by a Complex System is a result of the interaction of the many components or 'agents' in a collective, and the subtle temporal correlations and feedback that arises. The study of such systems from the perspective of the underlying agents with computational simulations has given encouraging evidence to suggest that large changes in the state of the system such as a crash in a financial market, can arise naturally from the internal dynamics of the system, with no need for exogenous shocks to explain disturbances to the system - the system may spontaneously find itself in an unstable state, prone to such an extreme event [Zhang 1998, Sornette 2004]. These unstable states may be encoded within the system, therefore opening up new ways to predict such extreme, internally driven events [Jefferies et al. 2002]. It has been muted that this knowledge could be used to predict extreme events in financial markets [Johnson et al. 2001] and even a theory to immunize the system against such internally driven extreme events has been proposed [Hart et al. 2002], with profound implications for state intervention in preventing the build-up of economic instabilities and the inevitable bubble-bursting that follows. Indeed, the Financial Services Authority (FSA) of the UK has recently proposed changes in the regulatory apparatus of financial markets in the Turner Review [Turner 2009], which in particular argues for better monitoring of credit conditions in the economy to examine whether a dangerous boom is being stoked up. What the FSA doesn't make clear is what it would do in such a situation. By analysing the behaviour of the Minority Game, Hart et al. advise an intervention in the market to cause 'mini-crashes' in order to prevent the system from building up to a more unstable regime that could suffer from the an even greater crash [Hart et al. 2002]. Of course, this may be difficult to implement if the Central Bank and Treasury were seen to be colluding to intervene in markets to depress asset prices, as such an intervention could prove politically unpopular amongst owners of the assets in question (the housing market is an obvious example effecting a significant proportion of the population). However, such an intervention could be seen as the complement (in fact the opposite) of what has been done recently with the quantitative easing campaigns in the UK and USA during 2009, with national governments pumping billions of dollars into the economy in order to prop-up asset prices [Flanders 2010].

It was at the Sante Fe Institute that the Economist Brian Arthur helped instigate the first Complex Systems approach to modelling financial markets, creating the Sante Fe Artificial Stock Market Model [LeBaron et al. 1998]. In doing so, he also laid the foundations for the development of the Minority Game by describing a situation of trying to predict whether it was worth going to the local bar 'El Farol' on a particular night [Arthur 1994], as an illustration of the bounded rationality[5] and inductive learning[6] that humans display when faced with a complicated situation accompanied by incomplete knowledge [Simon 1997, Shleifer 2000]. It models a system of agents competing for a scarce resource - a seat or enough space in the bar in this case, assuming that people believe it beneficial to go to the bar if it's under-crowded (so that there are enough seats to go around) but worth avoiding if overcrowded - hence the need to be in the minority. The El Farol model demonstrates the shortfall of the concepts, widespread in neo-classical Economics [von Neumann, Morgenstern 1947], of rationality and the representative, rational agent, because if there existed a correct expectational model that all rational agents would use to decide their attendance, it would automatically be the incorrect decision as everyone would turn up at the bar thus making it over-crowded, or else everyone would shun the bar, leaving it empty. Hence the system is frustrated - there are not enough seats to go around and people are forced to be heterogeneous, making their own inductive decisions in order to compete for these scarce resources, much like in financial markets. Indeed, Lamper claims that frustration[7] is an essential mechanism for modelling competition [Lamper 2002].

Arthur's Sante Fe colleague Per Bak passed the El Farol model on to the Physicists Challet and Zhang at Fribourg University who consequently developed a simpler binary version that they called the Minority Game [Challet, Zhang 1997]. This simple model captured all the fundamental properties of the El Farol

[5] the limits to the deductive rationality of models of human behaviour based on game theory and utility maximisation, see also [Levy et al. 2000] for a discussion and examples of such a breakdown of these traditional models

[6] using heuristic 'rules of thumb' to make decisions when faced with complicated problems and learning from previous experiences, see also [Orlean 2005, Nagel 1995] for a discussion and examples of experiments

[7] not all agents are able to find their preferred state, this term was originally used in a statistical physics context to describe subatomic particles not in their 'ground states'.

model, such as frustration and feedback, yet allowed easier computational representation and hence deeper investigation into the system dynamics. This initiated a race within the Complexity Science and Statistical Physics communities to explore the dynamics of the Minority Game, and over 200 papers have since been published on the topic [Econophysics website, University of Fribourg].

The Minority Game is representative of situations in which many agents must compete for scarce resources, such as traffic trying to choose the quietest of two routes from A to B, computers trying to find the least busy of two network routes, and of course people in markets, where sellers of an asset compete to find a counter-party who will buy.

2.2 Why is there a need for Complexity Science in Finance?

The random walk model developed by Bachelier and Osborne underpins much of standard finance theory [Bachelier 1900, Osborne 1959]. Advocating the use of the normal or log-normal distribution to model the probability distribution of asset-price changes, it is a reasonably good first-approximation and has allowed Option markets to flourish worldwide since the Black-Scholes-Merton model was published in 1973 [Black, Scholes 1973, Merton 1973]. However when compared with empirically observed time-series statistics of a wide range of markets, it fails to describe some of the most apparent features, or stylized facts [Cont 2001] (see Figure 2.1 on the following page to Figure 2.5 on page 37):

Excess kurtosis In the random walk model, the distribution of asset returns is supposed to be a normal distribution (with a kurtosis of three), whereas it has been known since Mandelbrot's research in the 1960s [Mandelbrot 1963, Fama 1965a, Fama 1965b] that the probability distributions of real returns is leptokurtic featuring 'fat-tails' with kurtosis greater than three [Lux 1996, Campbell et al. 1997, Farmer 1999,

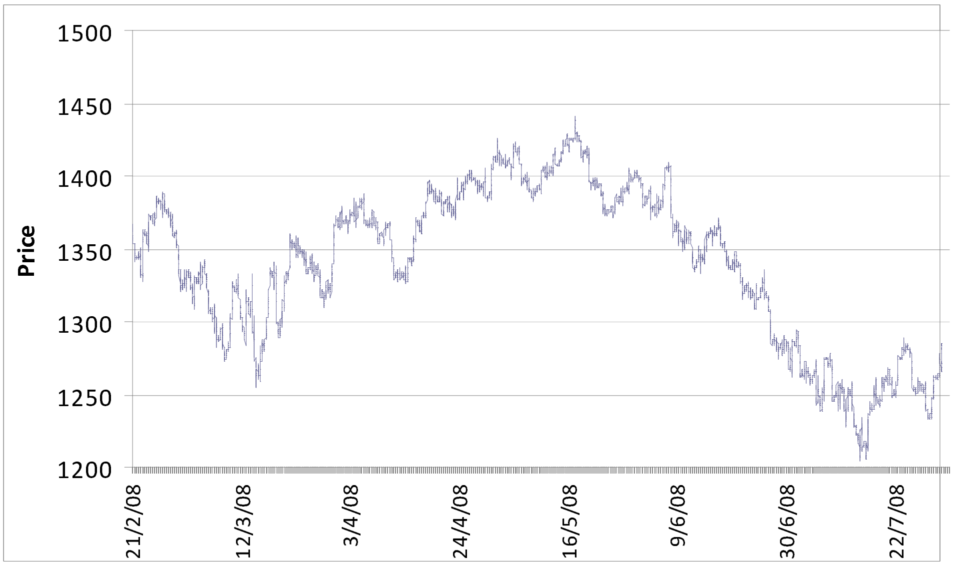


Figure 2.1: Price evolution of S&P500 Futures market at 10-minute intervals

Mantegna, Stanley 2000], even after accounting for heteroskedasticity in the data [Bollerslev et al. 1992] (see Figure 2.3).

Volatility clustering Different measures of volatility display a positive auto-correlation over several days, showing that high-volatility events tend to cluster in sporadic periods [Engle 1982, Mandelbrot 1963, Mandelbrot 1997]. See Figures 2.2 on the following page and 2.4 on page 35, which display the price increment and volume of the S&P500 market at 10-minute intervals respectively. These quantities are proxy for 'local' volatility, and illustrate the spikey behaviour of price increments and volume that tend to come in 'clusters'.

Auto-correlations The auto-correlation of linear returns decays quickly such that there is little ability to use linear auto-correlation information to make profitable forecasts, whilst the auto-correlation of higher moment returns (such as absolute returns or squared returns) decays much more slowly, with a significant auto-correlation existing for lags of the order of months or even years in some markets, see Figure 2.5 on page 37 and [Cont 2001].

Non-stationarity The price processes vary in time and are not independent

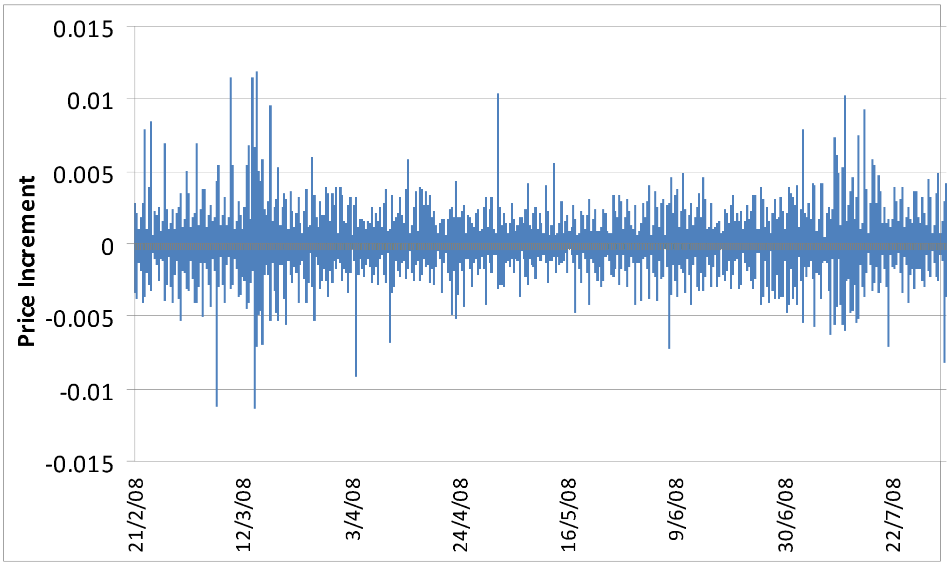


Figure 2.2: Price increments of S&P500 Futures market at 10-minute intervals

identical distributions, see for example the January effect [Keim 1983], where it has been proposed that individual investors who commonly hold small stocks sell their poorly performing investments at the end of the tax year to claim a capital loss to offset tax liabilities, before re-investing at the beginning of the next tax year. This creates a bias towards a depression in asset prices at the end of the tax year followed by an increase shortly thereafter.

The conceptual argument against any non-random walk behaviour of markets is related to the bastion hypothesis on which most standard finance theory stands – the principle of no-arbitrage. However, as Johnson et al. point out, there are various practical arguments one can make against this hypothesis [Johnson et al. 2003]:

1. In order to exploit an opportunity, you have to find it first - indeed most statistics packages used by market participants will only identify the lowest order i.e. linear correlations (see Figure 2.5). Higher order temporal correlations are extremely well hidden and hence not arbitraged away.

2. Even assuming an opportunity has been found and is being traded, arbitrage requires both a transaction cost and paradoxically an element of risk - if market conditions go against the trade and margin calls cannot be paid, no matter that eventually two asset prices must converge resulting in a profit, the position must be closed out prematurely. Indeed, such conditions almost caused the collapse of Long-Term Capital Management (LTCM)[8] in 1998 [Lowenstein 2001]. These barriers to entry mean that some arbitrage opportunities may survive in the market for a very long time.

3. From a behavioural perspective, supposing markets are effectively random at some instant in time, if there are investors such as chartists or 'noise traders' [Black 1986], who think they see patterns in this randomness, and believe they can predict the next market movement, patterns may be subsequently introduced into the dynamics merely by their collective actions. The counter-argument proposed by Black is that as long as there are 'enough' investors (or a few with the ability to invest large amounts) acting on real information as opposed to noise, they will be able to arbitrage the noise traders away, however they would still face the challenge of finding the patterns in the first place and the cost/risk or performing the arbitrage. Indeed, such profits generated by these arbitrageurs could be seen as payment for the effort necessary to find the information to trade on.

In order to estimate risk and price derivatives accurately, the model of the underlying price dynamics of an asset clearly needs to be consistent with the stylized facts of empirical data. Hence there is need for improvement over the random walk model. There have been attempts to refine the random

[8]an infamous hedge fund that pursued highly leveraged, aggressive trading strategies. When the Russian government defaulted on its sovereign debt repayments in 1998, panicked investors sold European and Japanese bonds and bought US treasury bonds, seen at that time as a safe haven. This caused bond prices to diverge unexpectedly - the opposite direction to the convergence expected in the LTCM trades - leading to massive margin calls on their trades. Thus LTCM were forced to liquidate many of their assets at heavy discounts and the intervention by a syndicate of major banks, many of whom were counter-parties to LTCM trades, was necessary to prevent LTCM from filing for bankruptcy.

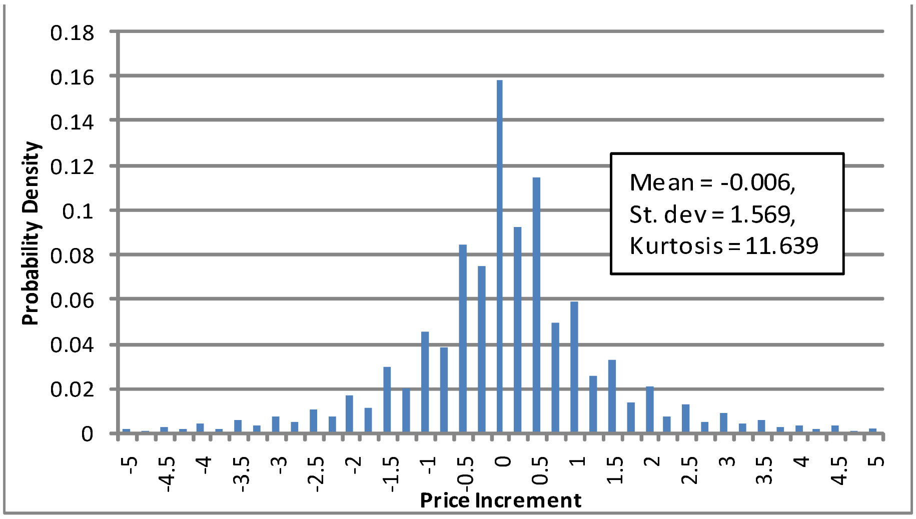


Figure 2.3: Histogram of price increments of the S&P500 Futures market at 10-minute intervals from 23/08/07 to 31/07/08

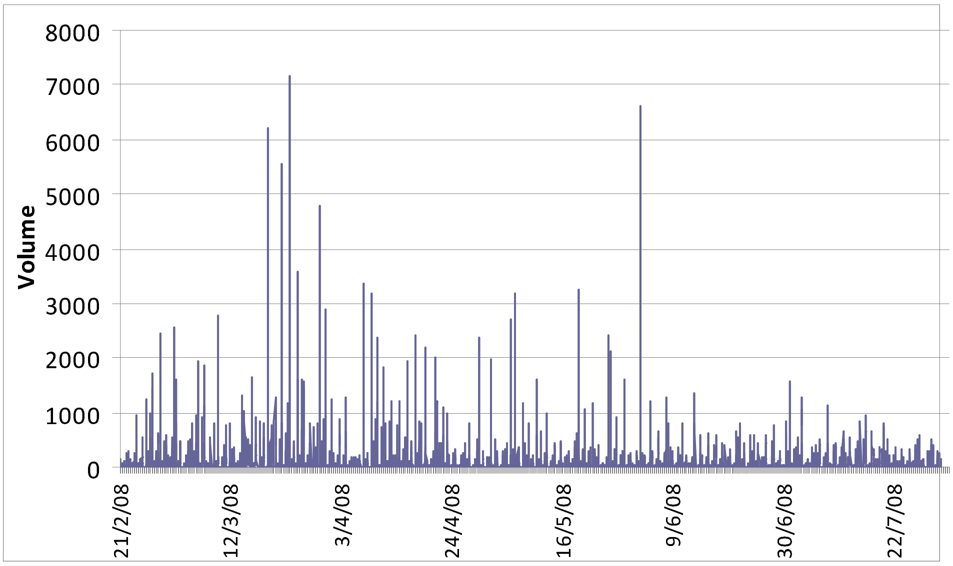


Figure 2.4: Trading volume evolution of the S&P500 Futures market at 10-minute intervals from 21/02/08 to 31/07/08

walk model by changing the shape of the probability distribution from normal or log-normal to a truncated levy-stable process which can exhibit excess kurtosis [Mandelbrot 1963, Mantegna, Stanley 2000]. There have also been attempts to include outliers by using jump-diffusion methods [Merton 1976, Cont, Tankov 2004], and time-dependent stochastic processes such as GARCH [Bollerslev 1986, Engle 2001] and stochastic volatility models [Heston 1993, Chen 1996] to model volatility clustering and non-stationarity. These adjustments are an improvement over the basic log-normal assumptions of market dynamics, however each of these approaches merely tries to replicate the price statistics from a time-series analysis point of view, rather than actually attempting to understand what is actually going on in the market at a fundamental level. Although in the long-run these statistics may be useful at describing financial market dynamics, it does not give any indication of short-term movements in asset prices. This top-down approach in which one gathers historical data, calibrates the model and implicitly assumes that history will repeat itself is dangerous, since 'past performance is no indication of future earnings', as every good investment house puts in their disclaimer [UBS Global Outlook 2008]. According to Simon Forge, this is equivalent to driving a car using the rear-view mirror – trying to judge the road ahead by what goes on behind! [Sardar, Abrams 1998] Certainly, this kind of modelling doesn't give much information or an understanding of where the road is heading next.

A more fundamental approach to understand the macroscopic behaviour of markets would be with a bottom-up perspective - to analyze the behaviour of the individual market agents, the traders, and the interactions between themselves. As Johnson claims, certainly, "the fluctuations observed in financial time-series should, at some level, reflect the interactions, feedback, frustration and adaptation of the markets' many and diverse participants." [Johnson et al. 2003] The benefit of using such agent-based models over other approaches is that not only might these models explain the aggregate statistical behaviour of the prices, but agent-based models also have the potential to mimic the micro-structure of the market. "All other things being equal, we believe that such a model may be intrinsically 'better' than a purely numerical multivariate one – and may even be preferable to many more sophisticated

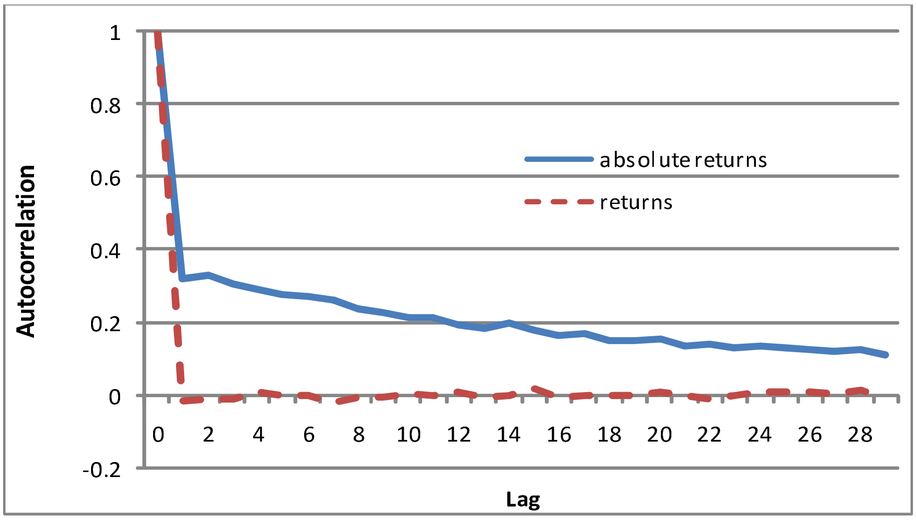


Figure 2.5: Auto-correlation of the S&P500 Futures market at 10-minute intervals from 23/08/07 to 31/07/08

models such as certain neural network approaches, which also may not be capturing a realistic representation of the microscopic details of a physical market." [Gupta et al. 2005a]

One could argue that such work is already being carried out within the traditional paradigm of finance, with agent-based modelling of microeconomics (simulating 'rational' utility-maximizing agents) or work under the guise of behavioural finance [Shleifer 2000]. However, such agent-based simulations have typically been done from the top-down, by dividing the market into a few select 'super-agents' that are supposed to be representative of a group of individuals following the same trading style, such as 'noise traders', 'value-traders' or 'momentum traders' etc. [Black 1986, Lux, Marchesi 1999, Levy et al. 2000, Bird, Casavecchia 2007]. The diversity of agents in such models is typically limited and the validity of the 'representative agent' concept is questionable. Complexity Science argues that one cannot expect the behaviour of an individual 'super-agent' (super in the sense that they have a greater weighting or market impact factor) to be accurately representative of a group of such agents interacting together. Indeed, it can be observed merely by walking down a busy street that a group of people displays a complex range of endogenously created behaviour not seen in 'independent' isolated individuals - correlations, non-linearity and feedback as the group 'swarms' down the street, which dominate and influence the behaviour of the participants (and sometimes others!). The market, like many complex systems displays *emergent phenomena* - behaviour that is not predictable simply by looking at the underlying equations, but something that can only be observed by seeing it in reality (and sometimes by doing computer simulations). Indeed as Chen points out, "generally we have no way... of inferring what would result from [such] a system except by running its course. Whatever will be will come to us as emergent properties. Neoclassical economics fails to see this..." [Chen 2002] Hence the market is certainly very different from the sum of its parts. One cannot simply equate the behaviour of a group of individuals with the behaviour of a single representational 'super-agent' and still expect to get a realistic description of a whole market, so the concept of representational 'super-agents' appears to be inappropriate.[9]

Continuing along these lines, one could also argue that there exists a limit to the potential applicability of certain types of behavioural finance experiments that focus on the fine details of the behaviour of individuals under simulated market conditions [Post et al. 2008], forgetting the importance of the complex interconnections and feedback loops that exist within the structure of a market as forces that wash away specific human attributes in such idealized settings. Therefore the inferences one can derive from this research regarding the overall market macro-structure are questionable. In addition, as behavioural finance has been developed within the paradigm of neo-classical economics, it has become preoccupied in searching for ways in which reality deviates from such theoretical - though often unobserved - concepts like market equilibrium, and has henceforth earned a reputation for being merely a collection of 'anomalies' [Kahneman et al. 1991]. There is still a lot of scope for useful contributions from behavioural finance, providing it develops as an observational science as championed by Roehner, much like sociology has developed, rather than focus on supporting or contradicting the theory of neo-classical economics

[9]The condition under it has been 'proved' that we can aggregate individual consumer demands to form a representative agent in a microeconomic (utility maximizing) setting is that the indirect utility function v of every consumer j has the Gorman form: $v^j(p, M^j) = a^j(p) + b(M)M^j$ for asset price p and income M [Gorman 1961]. This is satisfied if all consumers have either a) identical homothetic preferences, or b) quasi-linear preferences that are linear in the same good. These are fairly restrictive criteria and are unrealistic considering the diversity of real consumers.

[Roehner 2002].

An interesting line of research that has developed over the last decade, managing to draw together such considerations as the behaviour of individual market participants, as members of a heterogeneous group of 'agents' with whom they interact through the market, a place where each agent must learn and adapt to survive despite the limits of their imperfect rationality and incomplete information, is the Minority Game. Not only does this model resemble, at a basic level, the micro-mechanisms at work in a real financial market, but it also exhibits the major stylized facts observed at a macro-level as well.

2.3 What is the Minority Game?

The Minority Game (MG) is a multi-agent binary resource game - a model that simulates situations where a collection of many agents competing for a scarce resource are faced with a decision between two possibilities (encoded as 0 or 1), and have a payoff that is maximized by choosing the least popular decision. Such situations clearly arise often in nature, for example one could use the Minority game to simulate cars queuing in slow-moving motorway traffic, where agents attempt to choose the least busy of two lanes to drive on. Continuing the road analogy, it could also be used to simulate people at work trying to choose the quietest (and hence quickest) route back to the town where they live. Choosing between two lakes to go fishing, two pools to go swimming, two cinemas to watch a film, and even whether to go to the local bar or not (the El Farol bar model [Arthur 1994]) constitute the basic components of a Minority Game as in some sense the aim is to be in a smallest group (all other things being equal - the two lakes are equally populated by fish, equally easy to get to etc.). In all these circumstances, the scarce resource can be seen to be preferential access to resources - space on the road, a private fishing lake or the best seats in the cinema, and an exclusive bar. In the financial market place, the binary choices can be the action of whether to buy or sell a risky asset at the current time in exchange for cash. Here, being in the minority group represents preferential access to a counter-party to make

Figure 2.6: The price formation is a repeated discrete time process whose ordering suggests that it may be beneficial to be in the minority position when trading

a trade, so buyers are competing for sellers and vice versa, and the largest group is repeatedly frustrated - there will always be some agents in the largest group who are unable to find a counter-party, so agents in the largest group are disadvantaged on average. Indeed Savit likens the need to be in the minority as equivalent to innovation - it only pays to discover and patent a drug if you're the first to do it [Savit et al. 2004].

The reason it may be financially beneficial to be in the minority of a market rests on the ordering of events in time and how the price of an asset is derived (see Figure 2.6 and [Johnson et al. 2003]).

Assume a market-maker collates orders to buy and sell and executes the transactions at time t, with each discrete time-step representing the moment of order execution. Neglecting limit orders, if at any point in time there are more market orders to buy than sell, the market-maker will shift the price upwards (to encourage more sellers to participate in the market in future, and for the market-maker to reap a profit from being obliged to sell some of their own

time	action	cash	inventory	price	wealth
t^-	submit order to buy	100	0	50	100
t	order executed (majority sell)	51	1	49	100
$(t+1)^-$	submit order to sell	51	1	49	100
$t+1$	order executed (majority buy)	101	0	50	101

Table 2.1: Buying first in a minority decision round trip trade results in a profit

time	action	cash	inventory	price	wealth
t^-	submit order to sell	50	1	50	100
t	order executed (majority buy)	101	0	51	101
$(t+1)^-$	submit order to buy	101	0	51	101
$t+1$	order executed (majority sell)	51	1	50	101

Table 2.2: Selling first in a minority decision round trip trade results in a profit

assets in order to fulfill all market orders) before executing the orders on their book. The agents submit their market orders based on the price quoted at the last time-step, but this upwards shift means that the orders are actually executed at a higher price than when the order was placed. Agents buying the asset will have to pay a higher price than the expected price thus 'losing', whereas sellers will receive more money than they expected, thus 'winning'. Indeed, an agent who performs a round trip of buying an asset at time-step t then selling at time-step $t+1$, will accumulate a profit in real terms (increase their 'wealth' - here the sum of the value of cash and assets) if they are in the minority on both occasions (see table 2.1).

Likewise, an agent who performs a minority round trip of selling at time-step t then buying back at time $t+1$ will also accumulate a profit in real terms (see table 2.2).

Conversely, an agent who performs a majority round trip of buying at time-step t when the majority of agents are also buying, then selling at time $t+1$ when the majority of agents are selling will lose wealth in real terms (see table 2.3).

One of the main criticisms of the Minority Game is that traders usually try to profit by holding a position over a period of time, thus accumulating value

time	action	cash	inventory	price	wealth
t^-	submit order to buy	100	0	50	100
t	order executed (majority buy)	49	1	51	100
$(t+1)^-$	submit order to sell	49	1	51	100
$t+1$	order executed (majority sell)	99	0	50	99

Table 2.3: Majority decision round trip trade can result in a loss

through the change in price over many time-steps. However, as Challet et al. note, this strategy has a Minority Game component embedded within it - traders would ideally like to buy an asset at 'rock-bottom' and sell at the peak of a bubble [Challet et al. 2005]. Putting aside the fact that timing these decisions is extremely difficult to do [Buffet 2001], an agent can only do this if going against the trend, and buying before everyone else does, and selling before a bear market begins! Another consideration is that the Minority Game, under the paradigm of conventional trading strategies, resembles a fundamental view of asset pricing. Believing that on the next time-step, the price of an asset will be taken away from the fair value by a positive (negative) amount, the trader attempts to sell (buy) and pocket a small profit, thus returning the asset to the fair value again. Of course, whether such a fundamental value actually exists in some asset classes such as in the stock market is debatable, as discussed by Orlean in his self-referential hypothesis [Orlean 2005], akin to Keynes' Beauty Contest [Keynes 1936] or Soros' reflexivity [Soros 1994]. We also discuss a recent adaptation of the Minority Game that includes a section of the market with agents who 'win' by being in the majority [Gou 2005, Gou 2006a, Gou 2006 b, Gou 2007, Gou et al. 2008] and hence benefit from market trends later in this chapter.

2.3.1 Building the Minority Game - Concepts

2.3.1.1 Inductive learning and agent adaptivity

The Minority Game differs from the traditional sort of Game Theory games advanced by von Neumann & Morgenstern and Nash [von Neumann, Morgenstern 1947, Nash 1950], in that it is a repeated multi-player game with players of the order of hundreds or thousands - all N agents play at every time-step and in such a game it is impractical to expect the agents to be able to rationally deduce the equilibrium properties under which their pay-offs are maximized [Johnson et al. 2003]. Only by knowing the decisions of all the other agents would an agent be able to work out the optimal decision - and agents will be unlikely to let others know their decisions for fear of giving others an advantage, so information is generally limited and global - every trading desk with the capability of making significant market impact will have access to Bloomberg, Reuters and the same public information. Agents must therefore inductively work out their best strategy in the current state of the world. In order to do this, each agent in the Minority Game is endowed with s strategies, which tell the agent whether they should buy or sell in a particular state of the world. The Minority Game is a Complex Adaptive System as the agents generally have more than one strategy to use ($s > 1$), and will use whichever one they believe to be the most successful at that particular time. There is a score associated with each strategy, and agents will use their strategy that has made the most correct predictions over the recent past.

A variation of the MG is the Grand Canonical Minority Game (GCMG), which requires the agents to be sufficiently confident of their strategies in order to use them, leads to a variable number of agents participating in the market at each time-step (see 2.4.2 on page 52), and the ability to recreate volatility and volume clustering as observed in financial markets [Cont 2001].

In fact, there are several possible variations of how the strategy scores are maintained in the computer program. One option is to have each agent maintain their own score for each strategy, but this would involve a lot of redundant score-keeping and updating as the scores calculated by each agent should be the same, unless the initial strategy scores kept by the agents differ, or if the parameters related to how the strategy scores are updated were to be varied between agents. So, we opt for an efficient centralized scoring method, in which the strategies are scored centrally, and the strategy scores are passed to the agents for them to determine their best strategy.

2.3.1.2 The global information / state of the world μ

In financial markets, traders usually base their decisions on limited, public information. We could feature some sort of fundamental analysis in our model by declaring that the information agents have available includes the accounts of a specific company, which is thus used to derive a fundamental value of a share price. However, such information is only available at very infrequent intervals (commonly published in a company's quarterly accounts), and as highlighted on page 28, the change of an asset's price occurs far more frequently than new information comes in - on a tick-to-tick basis (the volatility in asset prices is far greater than expected from a fundamentalist point of view [French, Roll 1986, Cutler et al. 1989, Shiller 1989]). Indeed, Dacorogna et. al claim that 90% of the daily transactions in the Foreign Exchange (FX) market are down to speculative trading [Dacorogna et al. 2001].

A look at most trading room floors would probably indicate that the information most traders have on their many computer screens, and therefore use to make their investment decisions, is the price and volume of an asset. Johnson et al. therefore argue that it is appropriate to model the information used by traders in many of today's liquid financial markets, dominated by speculation and algorithmic trading, as the movement in an asset price over recent time-steps [Johnson et al. 2003]. Consequently, the state of the system μ of a Minority Game model has been defined as the history of the m most recent price changes. In this set-up, agents are assumed to be chartists, that is, they make their investment decisions based on patterns (they believe [Black 1986]) they see in financial time-series. Whether the Minority Game, with these embedded assumptions, can capture the observed market dynamics more successfully for some markets rather than others will be examined in Chapter 4.

The MG is binary not only in that the agents are required to choose between two possibilities of buying or selling a traded asset, but moreover that the state of the system is updated at each time-step with one of two outcomes. We use the convention to label the most recent state update as '1' if the majority of the agents choose to buy (meaning a positive aggregate demand leading to a price increase of the traded asset), and ‘0’ if the majority of the agents choose to sell (negative aggregate demand and a resultant price decrease), although the symmetry of the system means that it could equally be re-interpreted the other way around, as long as the price-formation process (see 2.3.1.4 on page 49) is also re-interpreted in the reverse fashion. Encoding this information as a binary code is a major innovation of the Minority Game over the El Farol bar model [Arthur 1994]. This binary nature of the global information, the state of the system, is the simplest yet representative way of encoding the price movement history and it greatly reduces the number of possible states the Minority Game can be in, making it a lot more tractable for analysis and simulation. The small, finite value of m leads to there only being 2^m different possible histories or states of the system μ, and therefore only 2^{2^m} distinct strategies on which the agents base their investment decisions (whether to buy or sell at the next time-step, given the state of the system). Moreover, Challet and Zhang, the creators of the Minority Game, found that the space of strategies could be further reduced from the unique Full Strategy Space (FSS) of 2^{2^m} strategies to a degenerate Reduced Strategy Space (RSS) of 2^{m+1} by only using strategies that are anti-correlated or uncorrelated with each other - analogous to creating a statistically orthogonal, or uncorrelated, basis in strategy space [Challet, Zhang 1998]. This removes strategies that are very closely correlated, which would lead to similar investment decisions being taken most of the time. Indeed, this speeds up computation without effecting the qualitative nature of the results.

See Figures 2.7 on the following page and 2.8 on page 47 for examples of a full and reduced strategy space respectively for the case where $m = 2$. The $2^2 = 4$ states of the world written in binary form are either $\{00\}$, $\{01\}$, $\{10\}$ or $\{11\}$, which represent cases where the two most recent aggregate majority decisions were sell-sell, sell-buy, buy-sell, or buy-buy respectively (and price changes were ‘down-down’, ‘down-up’, ‘up-down’ and ‘up-up’ respectively). These states can also be written in decimal form, with the usual interpretation of binary into decimal such that $\{00\}$, $\{01\}$, $\{10\}$ and $\{11\}$ are equivalent to $\mu = 0$, $\mu = 1$, $\mu = 2$ and $\mu = 3$ respectively. The action proposed by strategy R^* under a particular state of the world μ^* can be written as $a_{R^*}^{\mu^*}$. In the tables, the actions of all strategies $\{R\}$ in a particular state of the world

Strategy \ State	{00}, $\mu = 0$	{01}, $\mu = 1$	{10}, $\mu = 2$	{11}, $\mu = 3$
$R = 1$	-1	-1	-1	-1
$R = 2$	-1	-1	-1	+1
$R = 3$	-1	-1	+1	-1
$R = 4$	-1	-1	+1	+1
$R = 5$	-1	+1	-1	-1
$R = 6$	-1	+1	-1	+1
$R = 7$	-1	+1	+1	-1
$R = 8$	-1	+1	+1	+1
$R = 9$	+1	-1	-1	-1
$R = 10$	+1	-1	-1	+1
$R = 11$	+1	-1	+1	-1
$R = 12$	+1	-1	+1	+1
$R = 13$	+1	+1	-1	-1
$R = 14$	+1	+1	-1	+1
$R = 15$	+1	+1	+1	-1
$R = 16$	+1	+1	+1	+1

Figure 2.7: Full Strategy Space table for $m = 2$, which shows the actions stated by strategy R (a row) for each possible state of the world μ (a column). A '+1' entry indicates an agent using this strategy should buy at the next time-step, and '-1' indicates the agent should sell, for example, strategy $R = 12$ in the state $\mu = 1$ stipulates to sell at the next time-step. There are 2^{2^m} strategies in the Full Strategy Space.

$\mu = \mu^*$ are represented by a specific column vector in the strategy space tables ($\underline{a^{\mu^*}}$), whilst the actions of a particular strategy $R = R^*$ for every state of the world μ are represented by a specific row vector in the strategy space tables ($\underline{a_{R^*}}$). For example the first row ($R = 1$) of both the Full and Reduced Strategy Spaces is to sell ($a^{\mu}_{R=1} := a^{\mu}_1 = -1$) for every state of the world μ, in effect regardless of what happened to the price. The second row ($R = 2$) of the Full Strategy Space represents the strategy that advises to sell for all states of the world except in the state {11} (if the last two price changes were 'up-up'), in which case the strategy advises the agent to buy ($a^{\mu=3}_{R=2} := a^3_2 = +1$).

Like many Complex Systems, the Minority Game exhibits feedback - the choice made by each agent aggregates up at every time-step to define the winning outcome. This is consequently used to define the new state of the system, which can be written as a binary string of length m. For example if $m = 3$,

Strategy \ State	{00}, $\mu = 0$	{01}, $\mu = 1$	{10}, $\mu = 2$	{11}, $\mu = 3$
$R = 1$	-1	-1	-1	-1
$R = 2$	-1	-1	+1	+1
$R = 3$	-1	+1	-1	+1
$R = 4$	-1	+1	+1	-1
$R = 5$	+1	-1	-1	+1
$R = 6$	+1	-1	+1	-1
$R = 7$	+1	+1	-1	-1
$R = 8$	+1	+1	+1	+1

Figure 2.8: Reduced Strategy Space for $m = 2$, which shows the actions stated by strategy R (a row) for each possible state of the world μ (a column). A '+1' entry indicates an agent using this strategy should buy at the next time step, and '-1' indicates the agent should sell, for example, strategy $R = 8$ in the state $\mu = 0$ stipulates to buy at the next time step. There are 2^{m+1} strategies in the Reduced Strategy Space.

and the previous state of the system was {010} or $\mu[t-1] = 2$ in decimal form, if there are more agents using strategies that advise to buy ($a^2_{R\ni 1}[t^-]$) rather than sell ($a^2_{R\ni -1}[t^-]$), the system state is updated to become {101} (or $\mu[t] = 5$ in decimal form), with the most recent state update added to the right, and the oldest digit on the left 'forgotten'. As most agents decided to buy, the minority group were selling and it is this minority group that is defined to have 'won' (this is why it's called the Minority Game), meaning that within the strategy space, strategies predicting to sell (action $a^2_{R\ni -1}$) are awarded a point, whilst those predicting to buy (action $a^2_{R\ni 1}$) are deducted a point. This occurs regardless of whether the strategies were actually used or not, so these are referred to as virtual points, but do not represent any real wealth that is accumulated by the agents. In fact, these points are instead accumulated by the strategies, and it is the (relative) score of the strategies - a measure of how successful the predictions of the strategies have been over recent history - that determines which strategies are used to predict future market behaviour by the agents at the following time-step. In most versions of the MG in the literature, the feedback is perfect in that no external factors such as shocks to the aggregate demand or price are included, and the market just feeds off itself. In this case, the state of the world μ is the only (global) information available to the agents - the only input into the system - which consequently dictates

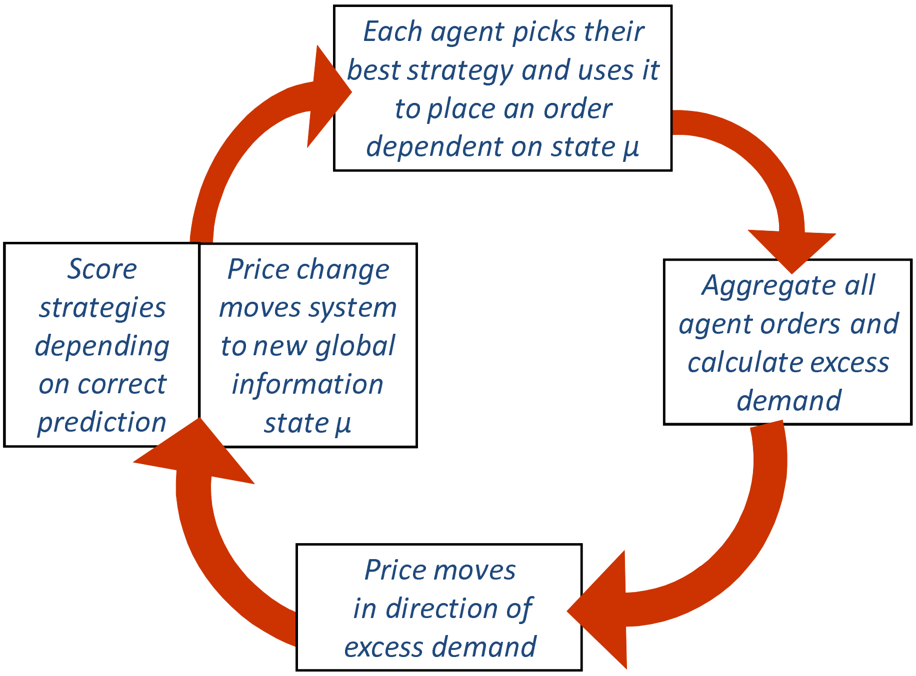


Figure 2.9: The feedback loop exhibited in the Minority Game. The agents' trading decisions are collected by the market maker who adjusts the price in the direction of the excess demand. The change in price of the asset leads to an updates to the state of the world μ and the strategy scores. In this new state, the agents pick their highest scoring strategy to decide how to make their next trading decision at the next time-step.

the agents' investment decisions at the next time-step. These decisions create new aggregate demand, define a new state of the world and the new global information that acts as the input into the system at the next time-step. This process then repeats in successive time-steps *ad infinitum* (see Figure 2.9).

Currently the agents are only interested in the price movements to make their investment decisions, though it could be extended to include other information too, with trading volume or recent price volatility used to capture changes in agent behaviour when markets become more volatile.

2.3.1.3 Stochasticity

There are two situations that need to be resolved by a random coin toss:

1. When the number of agents in both groups (i.e. buy and sell) are equal - meaning that there is no well-defined minority group, hence no outright winning decision.
2. A tie in an agent's strategy scores - where an agent has more than one strategy with the same score, so cannot inductively work out an outright optimal strategy.

Both situations inject stochasticity into the simulations, meaning that, given the same starting conditions but different random number generator seed, no two runs of a simulation will result in the same time evolution.

2.3.1.4 The price formation process

In order to construct a price time-series, we use the empirically supported theory that an excess in aggregate demand (e.g. the number of agents choosing to buy (action 1) is greater than the number of agents choosing to sell (action 0) leads to an approximately linear change in the asset price [Cont, Bouchaud 2000, Farmer 2002]. This can be modeled as

$$P[t] = P[t-1] + \frac{D[t^-]}{\lambda} \tag{2.1}$$

where

$$D\left[t^-\right] = N_1\left[t^-\right] - N_0\left[t^-\right] \tag{2.2}$$

represents the excess aggregate demand in the market just prior to time t, while time t represents the time when the new price $P[t]$ is set and the market orders are executed. (The time just prior to execution at time t is denoted by t^-). λ is a scale parameter of 'market depth' that reflects how sensitive the price change is to excess demand.

We can assume the financial market made by a market-maker is inherently discrete in time, as $N_1[t^-]$ and $N_0[t^-]$ are the number of buy and sell orders that have accumulated since the last order execution at time $t-1$ and just before the new order execution at time t. It is debatable whether the discreteness is still valid as computer interfaces and algorithms become more commonly relied upon to make quotes in markets (the only market left in the City of London that still uses the 'open-outcry' method is the London Metal Exchange during the 'Ring trading' sessions), or if there are very many market-makers operating all at the same time as in the 24-hours a day global FX market to make it appear continuous. The discrete time-step can be interpreted in two ways - the time it takes for a market-maker to collate and execute new orders, or the time in which it takes the agents to update their investment decisions and decide whether to change their inventory of assets, though the latter interpretation would be a much more diverse variable amongst the wide range of investors from day traders to pension fund managers.

The scale parameter λ reflects the 'market depth' - that is, how sensitive the market is to an order imbalance, and should be an increasing function of the liquidity, or number of traders participating in the market:

$$\lambda = f(N_0[t^-] + N_1[t^-]) \approx \theta \cdot (N_0[t^-] + N_1[t^-]) + \rho \tag{2.3}$$

where the linear constant of proportionality θ and constant ρ can be estimated from a linear regression.

If there are only a few agents, the impact on the price of each agent's order is likely to be higher, and a market-maker will be attempting to 'scalp' a similar amount of money from a smaller number of agents. Indeed we could specifically include the inventory $M[t]$ of a market-maker [Jefferies et al. 2001] by modifying equation 2.1 thus:

$$P[t] = P[t-1] + \frac{(D[t^-] - M[t^-])}{\lambda} \tag{2.4}$$

though this then introduces unrealistic outcomes where the agents arbitrage the market-maker, which need to be resolved by further elaboration of the model.[10]

[10]Remarkably, the agents very quickly learn that they can arbitrage the market maker by having enough agents using the same strategy consisting of 'buy-wait-sell' or vice versa. This can be prevented by allowing the market-maker to raise a bid-ask spread that is proportional

It should also be noted that most MG models in the literature feature the case where agents can only buy or sell one unit of the asset, which is an unrealistic assumption. There have been extensions to the MG that relax this assumption [Johnson et al. 2003], though they merely resulted in more complicated models with fairly similar simulation results. For our investigations into the forecasting ability of the MG framework, we do not have to make any assumptions about the purchasing power of the agents. Indeed, the weightings of the strategies available to agents is generally assumed to vary over time, thus implicitly allowing for a heterogeneity in market order sizes, and we specifically try to estimate the relative weight of these strategies, which resemble an aggregation of the order sizes of all the agents using each particular strategy at each time-step.

2.4 The development of the Minority Game

There have been many adaptations of Challet & Zhang's original Minority Game model (see [Econophysics website, University of Fribourg] for an extensive list of the many papers based on the Minority Game) to include increasingly more realistic features of financial market trading, thus demonstrating the adaptability and robustness of the model. However we have to remember what our goals are when tailoring an agent-based model to suit a particular purpose. In the case of Johnson et al., it was to understand "what gives rise to the observed high volatility in today's crowded markets?" [Johnson et al. 2003] In this thesis, it will be to investigate the possibility of making forecasts using agent games based on the Minority Game paradigm. In this case, assuming we will need to parametrize our model to real-world market data, we should try to keep the number of parameters small by having the model as simple as possible on condition that it provides a suitably realistic representation of the empirical market data (it can recreate the stylized facts). To quote Johnson et al., it might be "impossible to answer these types of question if the model we use to simulate the market is of comparable complexity to the market itself.

to the average wealth of the market maker divided by the average volume of transactions. See [Jefferies et al. 2001] for more details.

Instead, we need to focus on a minimal set of underlying assumptions in order to make sense of what is, after all, a very 'complex' system."

2.4.1 The Basic Minority Game

The Basic Minority Game, as introduced by Challet and Zhang in 1997 originally involved only four parameters - the number of agents N, the memory of the agents m, the number of strategies available to each agent s, and the resource level of the game L [Challet, Zhang 1997]. With just these four parameters, the Basic Minority Game succeeded in reproducing many of the stylized facts such as the excess kurtosis of the distribution of asset returns, yet as all agents participate at every time-step, the 'market' volume is a constant 100% of the market. The simple nature of the model allowed much study through simulations and even the foundations for the development of an analytic mathematical theory of the Minority Game [Coolen 2005], a rare occurrence in the field of agent-based computer modelling.

2.4.2 The Grand Canonical Minority Game

The Basic Minority Game suffers from the unrealistic property that all agents participate at every time-step. Three groups independently introduced the Grand Canonical Minority Game (GCMG), so called after the term's use in statistical physics to indicate a variable number of agents participating in the system [Slanina, Zhang 1999, Johnson et al. 2000, Challet et al. 2000]. In this version, the agents have an additional three parameters, the finite time-horizon of the strategy scores T, the strategy score (exponential) decay factor τ, and a confidence threshold c. See section 2.5 on page 54 for more information on the parameters and their meanings. Strategies with scores below a specific threshold $r = r(c)$ which is determined by the parameter c are not used. If all strategies held by an agent have scores below this threshold, the agent won't participate in the market at that time-step, thus creating a variation in the number of agents and consequently an evolving volume of trade that changes from one time-step to the next.

The GCMG is one of the most interesting developments of the MG as it can recreate the clustered volatility as observed in financial markets, in addition to reproducing the other stylized facts highlighted in Section 2.2 on page 31 when the parameters are of suitable values. Research by Lamper et al. and Johnson et al. indicates that the GCMG shows some promise in forecasting hourly data of the USD-JPY exchange rate during the 1990s, achieving a 54% success rate at predicting the next-step ahead directional changes [Lamper et al. 2001, Johnson et al. 2001]. This may have practical significance, assuming one can establish a trading method to exploit this predictive power. Indeed, as highlighted by Doyne Farmer of the Sante Fe Institute, "The nice thing about markets is that you don't really have to predict very much to do an awful lot." [Johnson et al. 2003]

2.4.3 Mix Games

We also introduce the Mix Games of De Martino et al. [de Martino et al. 2004a] and Gou [Gou 2005, Gou 2006a, Gou 2006 b, Gou 2007, Gou et al. 2008]. Gou highlights the limited diversity of the agents in the GCMG as a weakness - all agents in most versions of the GCMG have the same general parameters:- memory length m, number of strategies s, time-horizon T, score-decay rate factor τ and confidence threshold c. In addition, she criticized the failure of the MG to take account of trend-following behaviour of the agents. Although there has been some work considering Minority / Majority Games as early as 2001 [Marsili 2001, de Martino et al. 2003, de Martino et al. 2004b, Tedeschi et al. 2005], agents still had limited diversity, generally switching with a specific probability between the two types of game (one where the agents win by being in the minority, the other by being in the majority) but with the same parameters. Consequently Gou introduced a new version dubbed the 'Mix Game' where two groups of agents, one playing a minority game and the other playing a majority game, are coupled together and simultaneously decide to buy or sell, based on their predictions of the next winning outcome. In this case, the majority game agents 'win' by choosing the action that is in the majority. Each group has its own set of parameters - $\{N_{Min}, m_{Min}, s_{Min}, T_{Min}, \tau_{Min}, c_{Min}\}$ and $\{N_{Maj}, m_{Maj}, s_{Maj}, T_{Maj}, \tau_{Maj}, c_{Maj}\}$

for agents playing the minority game and majority game components respectively - thus allowing more diversity of agents than the traditional GCMG or the minority/majority game put forward by Marsili, de Martino et al.

Gou discovers that there are parameter inequalities that are required to ensure the stability of the Mix Game:

$$N_{Maj} \leq N_{Min} \tag{2.5}$$

(for $\alpha = \frac{2^m}{N} \leq 0.3$)

$$m_{Maj} \leq m_{Min} \tag{2.6}$$

$$T_{Maj} \leq T_{Min} \tag{2.7}$$

Like the GCMG, there are certain parameter regimes that more accurately reproduce the statistics displayed by most financial markets. By tailoring the 'local volatility' (the volatility of the time-series over a short moving window of around 5 time-steps) and using a Generic [*sic*] Algorithm to optimize the other parameters, Gou manages to obtain up to around a 61% hit-rate for the Shanghai Stock Market (the success at predicting the directional change of the market at the next time-step), as opposed to a poor (as good as tossing a coin) 50% for the Minority Game [Gou 2005]. However, previous research reported a 54% success rate in predicting hourly US Dollar/Japanese Yen exchange rate data over the 1990s [Johnson et al. 2001]. See section 3.4 on page 76 for work on validating the Mix Game model.

2.5 The parameters of the Minority Game

In this section, the main parameters of the (Grand Canonical) Minority Game will be defined and discussed, specifying the typical parameters values that are used:

N number of agents (in the market simulation - this is not used in the forecasting procedure of Chapters 4 and 5)

m agents' memory length (the number of previous time-steps over which the price changes are used to derive the global state)

s number of strategies available to each agent

T time-horizon (the number of previous time-steps over which strategies are scored)

τ score decay-rate (exponential factor of score decay)

c confidence threshold (level that controls the threshold at which strategies are determined to be 'active' i.e. that they are sufficiently successful for the agents to confidently use)

L resource level

2.5.1 Typical parameter values in the literature and their effects

$N \sim 101$ represents the order of the number of large trading institutions (banks, pension funds, corporate treasuries etc.) whose trading activities can have significant impact on the markets. Papers generally have a range of values from 51 to 1001, and the number of agents has traditionally been an odd number to prevent an equal number of agents in both groups, so encouraging a well-defined minority group and negate the reliance on the random coin toss to resolve the winning decision in the case where there's a tie in the number of agents. This is less of an issue after the introduction of the more widely used Grand Canonical Minority Game, since the number of agents participating at any time is variable and generally less than 100%.

$m \sim 3$ determines the length of pattern that the agents can see and use to trade, hence if there are enough agents in the simulation, there should be no predictability using patterns of length $\leq m$ as predictive information, as agents arbitrage any predictability away, see validation work in Section 3.1 on page 67 and [Savit et al. 1999]. The number of different states of the world (2^m), and the number of distinct strategies (in the reduced strategy space - 2^{m+1}) increases exponentially with the value of m, thus for computational feasibility it is important to keep this fairly low.

The interesting Interactive Minority Game [Website, University of Fribourg], where users can actually play the Minority Game against computer generated agents, reports that humans can generally only perform a deductive 'rational' analysis of the game if the system has a complexity corresponding to memory size of $m \leq 3$ [Laureti et al. 2004]. When exceeding this threshold, humans tend to exhibit repeated, biased behaviour, characteristic of the inductive learning [Arthur 1994]. Indeed, Li et al. and Savit et al. claim that in an evolutionary setting, where agents can drop their worst strategy and pick a new one of different memory length m, the strategies that improve agents' average winnings have $m \leq 6$ [Li et al. 2000b, Savit et al. 2004]. In fact, the number of agents holding strategies of varying m in the long-run tends to a step-function - strategies with $m \geq 6$ are rarely held by agents and perform poorly in comparison to smaller m, though this does depend on the number of agents (they use $N = 401$ evolutionary agents). Interestingly, this contradicts earlier studies by Challet & Zhang, where agents with 'bigger brains' (larger values of m) are found to always outperform agents with smaller memory, supposedly as they can see observe longer patterns - information not observed by the smaller memory agents [Challet, Zhang 1997, Challet, Zhang 1998]. Savit believes Challet & Zhang's inferences from their simulations are too general, especially being non-applicable to high m / low N regimes (in fact, high α regimes - see 3.1 on page 63) [Savit et al. 2004]. Gou [Gou 2005] henceforth restricts her studies to $m \leq 6$.

$s \sim 2$ If $s = 1$, agents are 'frozen' in that they only have one strategy in which to participate in the market, and the result is a behaviour with trivial periodic structure [Savit et al. 1999]. As long as $s > 1$, agents are able to adapt their behaviour, picking the most successful strategy that they possess (as measured over recent history of length T - see next point). Research has shown that the behaviour of $s = 2$ and $s > 2$ are qualitatively similar, however the volatility of the system - an inverse measure of system efficiency and how well the agents can co-ordinate themselves to reach optimality - increases as s increases [Marsili et al. 2000]. Thus, the agents generally perform more poorly - that is, the average success they have in predicting the winning decision decreases with increasing s - owing to them being more likely to switch strategies more often, which creates more 'noise' in the system. Therefore, the value of s at which the agents can co-ordinate optimally is 2, and most studies in the literature use this value. This also has the benefit of being computationally simpler to simulate.

$T \sim 100$ Introduced as a modification to the original Minority Game (which only contained parameters N, m and s and simply accumulated the success of agents' strategies from the start of the simulation), the 'time-horizon' T represents a rolling window over which agents judge their strategies' successes. For the most recent T time-steps, a strategy is awarded a point if it predicted the correct outcome of the next time-step, or is deducted a point if it predicted an incorrect next outcome - this is called the 'score increment' at each time-step. The score of the strategy is then the sum over the all score-increments in the rolling window of T time-steps (albeit sometimes with an exponential weighing factor τ to mimic the likely assumption that the older score increments are less important - for more details, see the definition for τ in this section). As is the nature of the Minority Game and Complex Adaptive Systems in general, the strategies in play vary over time, as success of any strategy is intricately interdependent with the success of the other strategies. T is therefore an estimate to the length of time agents believe a market 'ecology' exists without evolving significantly, and is therefore believed to be an applicable time period over which to judge their strategies. In the literature, the period over which scores are kept is typically of the order of a few multiples of the memory length m, certainly we need $T \gg m$ for it to make any physical sense. (As a minor point, if T

is too large, it requires more computer power to process and may slow the computer program somewhat.) Thus the literature usually has T at around 50 to 200 time-steps. Occasionally in the Mix-Game simulation studies (see 3.4 on page 76) however, Gou calibrates the model to data from the Shanghai Stock Exchange with $T = 12$.

Another variation of updating the strategy scores - with an iterative updating procedure (see information about the score decay rate on the current page for more details including the equations) makes T as a control parameter obsolete and causes the system to act as if $T \longrightarrow \infty$. (However, the value of T in the Grand Canonical Minority Game still has a secondary effect if it is used to derive the score threshold at which agents are confident enough to participate in the market, $r = r(c, T)$.) Lamper claims that a Fourier transform of the simulation with a finite time-horizon shows that there is information in the time-series corresponding to the time-horizon length-scale (wavelength), whereas none appears with the iterative score-update [Lamper 2002]. Interestingly, when validating the Mix-game model, we note that there are finite size effects associated with the T parameter that we recover only if using the finite T version (without iterative score updates) - and this is also seen empirically in markets, thus implying that in reality, people use a finite time-period over which to judge their strategies (see 80 for more details).

$\tau \sim 0.99$ The score decay rate τ, is used as a weighting factor in the procedure to update strategy scores. With $\tau = 1$, all score increments during the recent time-horizon window T are equally weighted, whereas with $\tau < 1$, agents consider older score increments as being less important, or less valid an indication of strategy use in the current market conditions. In fact, with $\tau > 1$ the converse is true - agents believe historic data is more representative of the true 'market ecology' currently in existence. As this is unlikely to be the case, we generally have the following restriction $\tau \leq 1$. This enters into the model through the following equations, where $S_R[t]$ represents the score of strategy R at time t and $dS_R[t]$ represents the score increment of strategy R at time t

$$S_R[t] = \textstyle\sum_{i=t-T}^{i=t} \tau^{t-i} dS_R[i] \tag{2.8}$$

for finite time-horizon, T. If we use an iterative score updating implying an infinite time-horizon ($T \to \infty$), the relationship becomes:

$$S_R[t] = (S_R[t-1] \cdot \tau) + dS_R[t] \tag{2.9}$$

$c \sim 0.53$ An essential part of the Grand Canonical version, c is a 'confidence threshold' which mimics the amount of confidence the agents need in order to use their strategies to participate in the market. Specifically, the parameter c determines the score needed by a strategy for it to be defined as 'active' r, and thus available to be used by the agents, through the equation:

$$r = (2c-1)T \tag{2.10}$$

If a strategy's score is below this, it will not be used by the agents, and if the agents have no strategy above this points-threshold, they will abstain from trading in the market - thus causing the volume of trading to vary over time. Note that $c = 0.5$ implies $r = 0$, meaning that if a strategy R is active, the anti-correlated strategy $\overline{R}$ is always inactive and vice versa. We can measure whether the strategy is active using the Heaviside Step-Function to create the active strategy indicator function:

$$\mathcal{H}\{S_R - r\} \tag{2.11}$$

which returns 1 if strategy R is active and 0 if inactive. The total number of active strategies is thus

$$\sum_{R=1}^{R=2^{m+1}} \mathcal{H}\{S_R[t] - r\} \tag{2.12}$$

is the total number of active strategies at time t.

$L \sim 0.5$ is the resource level, the threshold of the proportion of agents that defines the winning outcome. If N_0 and N_1 are the number of agents submitting sell orders and buy orders respectively, then

$$Winning\,Threshold[t] = L \cdot (N_0[t] + N_1[t]) \tag{2.13}$$

Under the convention used in the literature and my own model, it is the number of buyers (agents choosing option 1) which is used to define the winning outcome:

$$N_1[t] < WinningThreshold[t]$$

then the winning action is 'buy' (the buying agents win, sellers lose) and the binary digit on the right of the global information state is updated with a 0 e.g. $\{111\} \rightarrow \{110\}$, otherwise if

$$N_1[t] > WinningThreshold[t]$$

then the winning action is 'sell' agents win (the selling agents win, buyers lose) and the binary digit on the right of the global information state is updated with a 1 e.g. $\{000\} \rightarrow \{001\}$. In the unlikely event of an equality

$$N_1[t] = WinningThreshold[t]$$

a random coin toss is used to define the winning action and consequently the new binary digit on the right of the global information state.
For example, in the situation where $L = 0.6$, if over 60% of the active agents choose to buy, the winning action is to sell, whereas if less than 60% of active agents choose to buy, the winning action is to buy. For it to be truly a minority game, we need $L = 0.5$ so that the group which wins is always the minority group.

L has the effect of creating a bias in the system. From the convention used here, if $L = 0.6$, there is generally an increase in the simulated asset price on average, as if we assume that initially the number of agents buying is roughly equal to the number selling on average (since we have an unbiased method for distributing the possible strategies across the agent population), it is more likely that there will be less than 60% of agents buying at any time-step. Thus the buying agents win, the price formation process registers an upwards pressure on the asset price, and the buying strategies' scores improve at the detriment to the selling strategies. The MG is (assuming an equally distributed coverage of the strategy space) unbiased for $L = 0.5$. This parameter is useful in generating long-term trends in an asset price - acting perhaps like a macroeconomic variable which could be used to model, for example, the growing strength of the Chinese Renminbi since 1994. For the Mix-Game version (2.4.3 on page 53), where a proportion of agents 'win' by being in the majority, Gou [Gou 2006a] doesn't investigate the L parameter, setting $L = 0.5$.

2.6 Concluding Remarks

In this chapter, we have discussed the motivations behind the Minority Game model, from the importance in recognizing the emergent properties of complex systems such as financial markets, to the simplistic rules and small number of parameters necessary to build a MG simulation. In Chapter 3, we present the results from a sample of simulations and highlight important characteristics that describe the behaviour of simulations, which can also be applied to financial markets.

Chapter 3

Model Building

> *"'Think simple' as my old master used to say - meaning reduce the whole of its parts into the simplest terms, getting back to first principles."* Frank Lloyd Wright

In order to run our own simulations and build upon the work of previous researchers, we must first be sure that the model we produce is actually a properly working Minority Game, hence the need to validate our model. Just from an observation of the time-series output (see Figure 3.1 on page 64), we would be unable to distinguish what was really the underlying engine of such seemingly stochastic behaviour. Indeed, just performing a random coin-toss, or using a stochastic process of a normal or log-normal distribution can look similar to the time-series of the price of an asset, although without the clustered volatility appearing as in the Figure 3.2 on page 65 of 'trading volume'.

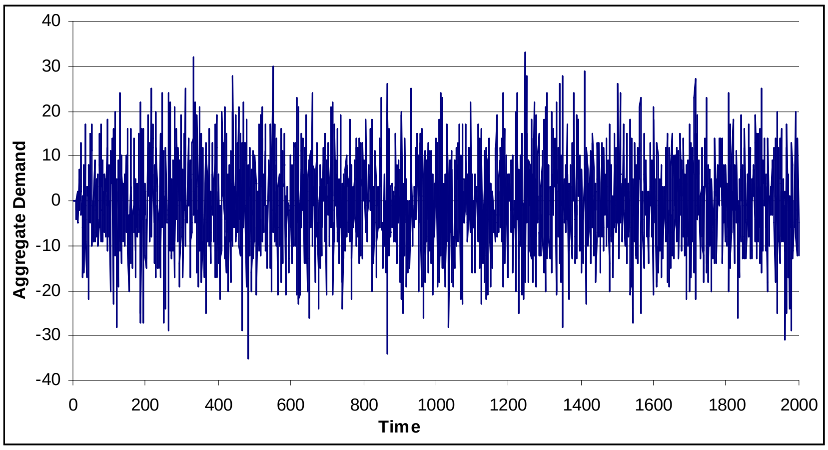
This chapter covers the steps taken to ensure the models we have built behave as expected, by presenting results from a sample of simulations that reproduce features of the different versions of the Minority Game and Mix-Game. We also highlight the important characteristics of the behaviour of both models, and relate them to the statistical behaviour of S&P500 Composite Index and Shanghai Stock Exchange data. In addition, we discuss the possibility of adding further structure in the form of networks of agents to allow for private shared information and co-ordination amongst a subset of the agent population, and it's potential applicability in simulating oligopoly dominated markets such as the oil market.

3.1 (Grand Canonical) Minority Game

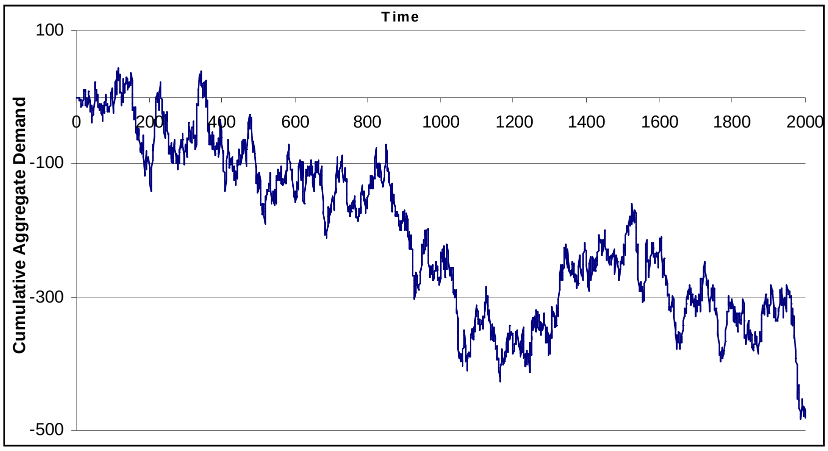
We display examples of the output of simulations of the Grand Canonical Game with parameters $N = 201$, $m = 6$, $s = 2$, $T = 100$, $\tau = 0.99$, $c = 0.53$, $L = 0.5$ in Figures 3.1 on the next page and 3.2 on page 65. Figure 3.1 shows the time evolution of aggregate demand, cumulative aggregate demand, and a price formation process derived from the underlying MG engine (the Cumulative [Aggregate Demand/Volume] - see equation 2.1 on page 49. Figure 3.2 shows the evolution of the number of active agents at each time-step, akin to a measure of trading volume. Compare these to Figures 2.2 on page 33 and 2.4 on page 35, and one can see the clustered volatility in the aggregate demand and the trading volume, as desired.

In numerical simulations we find results to corroborate the work in [Savit et al. 1999] that demonstrate the variation of $\frac{\sigma_{N_1}}{\sqrt{N}}$, where σ_{N_1} is defined as the standard deviation of the number of agents choosing option 1 at time t^-, over a range of values for m and N. As you can see in Figure 3.3 on page 66, for small values of m, there is a relatively high volatility, which decreases with increasing m up to the minimum, which occurs at $m = 6$. After this, volatility increases slowly, tending asymptotically to the value expected in a random coin-toss experiment (where agents flip a coin to decide which option they will choose).

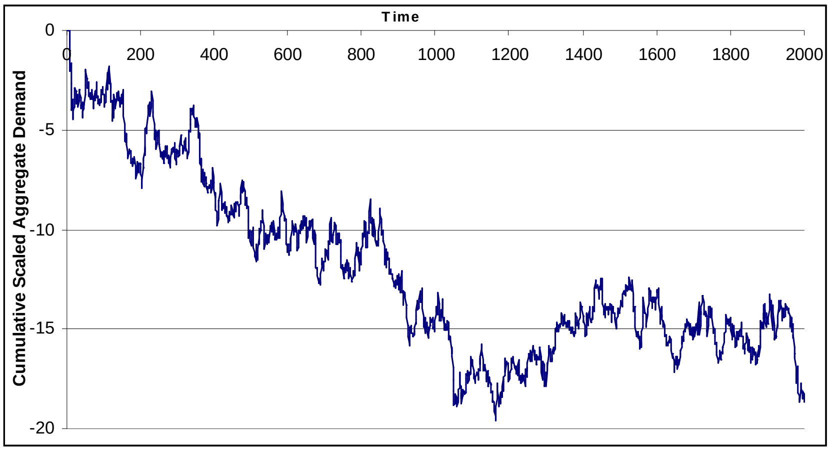
Through these studies, it was discovered that an important ratio of parameters $\alpha = \frac{2^m}{N}$, characterises the dynamics of the system [Challet, Marsili 1999, Savit et al. 1999]. Crowd-anticrowd theory, as developed by Hart et al. explains that this is because α specifies the degree to which the strategy space is well covered by the agents [Hart et al. 2001]. This can more easily be seen by realising that the number of different strategies (in the Reduced Strategy Space - which only includes the statistically orthogonal / uncorrelated and



(a) Aggregate Demand time-series generated in a MG simulation



(b) Cumulative Aggregate Demand time-series generated in a MG simulation



(c) Price time-series determined from Aggregate Demand and 'Volume' generated in a MG simulation

Figure 3.1: The evolution of aggregate demand during a Minority Game simulation with parameters $N = 201$, $m = 6$, $s = 2$, $T = 100$, $\tau = 0.99$, $c = 0.53$, $L = 0.5$

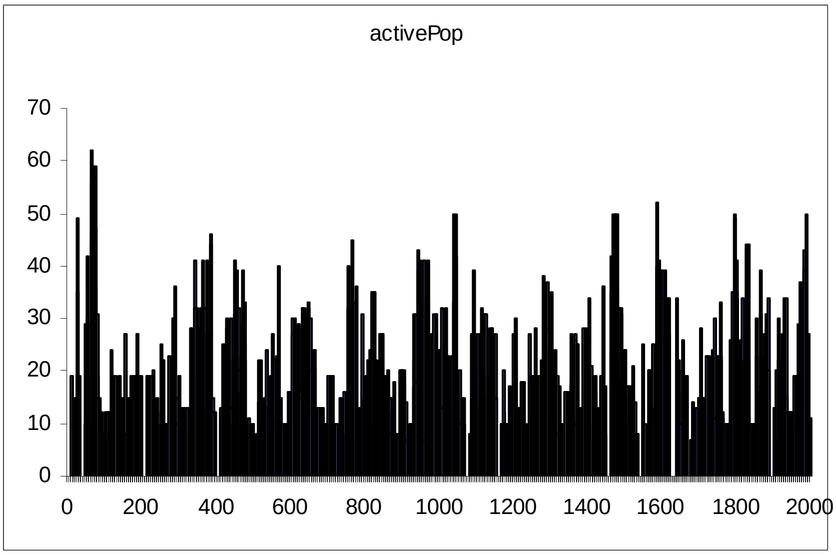


Figure 3.2: Trading 'volume' of Minority Game simulation with parameters $N = 201$, $m = 6$, $s = 2$, $T = 100$, $\tau = 0.99$, $c = 0.53$, $L = 0.5$. The volume is calculated as the number of active agents (those that have at least one strategy with a score above the confidence threshold).

anti-correlated strategies) is 2^{m+1}, therefore α is inversely related to the average number of agents holding each strategy.

For small α (low m or high N), the strategy space is well-covered (crowded regime) - many agents are likely to hold each strategy, though not necessarily use it as they may have a better strategy. In most versions of the GCMG, there is centralized scoring - meaning that all agents agree on the relative merits of each strategy, by using the same technique to measure the success of each strategy they hold. This means that if one strategy R^* accumulates more points than a subset of others at a particular point in time, all agents holding these strategy pairs will use strategy R^* at the next time-step, analogous to rushing to the market at the same point in time with the same investment decision - usually the cause mooted by financial commentators for the sudden change in the price of an asset. Normally, this would be attributed to the arrival of new information regarding the asset, but it is also possible that it is internally caused, as in our model, by heterogeneous traders drawing the same conclusions on the value of an asset based on patterns observed in the market. As m increases, the strategy space becomes less crowded, and σ_{N_1} decreases

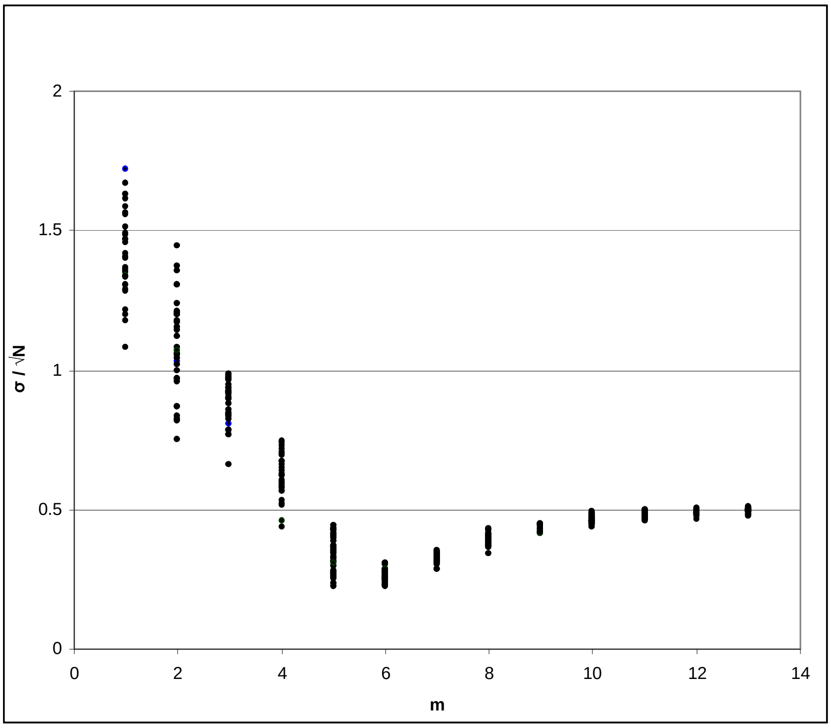


Figure 3.3: The volatility per agent ($\frac{\sigma_{N_1}}{\sqrt{N}}$) in the MG simulations as a function of m for $N = 101$ and $s = 2$. Here, σ_{N_1} is the standard deviation of the number of agents choosing option 1 over a simulation run, and we perform 32 independent runs of 10,000 time-steps for each value of m. The low m (crowded) region displays high-volatility often prevalent in financial markets, whilst the large m (sparse) region asymptotically approaches the limit of independent agents (who decide their investment decisions by flipping a coin). There is a minimum in the volatility per agent at $m = 6$ (when $N = 101$ and $s = 2$), suggesting this is when agents manage to 'co-operate' optimally and the global efficiency is highest.

as a smaller proportion of the total population hold the same strategy, and hence there is a smaller proportion of agents rushing to the market to make the same trade.

At the other extreme, when α is large (high m or low N), the strategy space is sparse - few agents, if any, hold each strategy. As agents are very sparsely distributed across strategy space, there are large gaps of available strategies that are not being used, therefore patterns are often not arbitraged away and agents are unable to co-ordinate themselves to take advantage of information in the history bit-string, thus minimising the system wastage. In fact with increasing m, σ tends asymptotically to the value as if the agents were playing the game independently, and tossing a coin to decide whether to buy or sell.

The difference between the two regimes of α can be seen by assessing the information content (or predictability) of the Minority Game in both phases, which we do by conditioning the probability that the excess demand at the next time-step will be negative (the winning action at the next time-step is 1) on the past k outcomes or 'bits' as in [Savit et al. 1999]. The results are displayed in Figures 3.4 on page 70 to 3.6 on page 71 for $N = 101$ and $s = 2$.

The flat Figure 3.4 with $k = m = 4$ is indicative of the crowded (large N) region, where all information contained in the price patterns less than or equal to the agents' memory length m is exploited by the agents. As there are enough agents to 'arbitrage' away any biases for the winning action at the next time-step at every state μ, there is no predictive information left that the agents can exploit. Thus, given we are in a specific state of the world μ (or as Savit describes it, an 'm-bit past'), the predictability of the next market movement, measured as the probability of the next market movement being a price increase (that the new bit on the right of the binary representation of the state of the world is '1'), is such that $\mathbf{P}(1|\mu) = \frac{1}{2}$. The situation is symmetric, there is no predictability given the global information. Note that this is similar to the idealised case of the efficient market hypothesis, where all information is instantaneously aggregated to asset prices such that no more predictive information exists.

For the same case, we look at conditional probabilities using an extra bit of information compared to that available to the agents such that $k = m + 1$ in Figure 3.5. This shows that predictability (biases) are still prevalent in the time-series of winning outcomes, however as the agents only look at the past m time-steps, they can't see any patterns longer than this and are therefore unable to exploit these biases when making investment decisions.

Figure 3.6 on page 71 displays the situation where the conditional probabilities exhibit biases even when $k = m = 6$. In comparison to Figure 3.4, when $k = m = 4$, despite the fact that in both cases the agents see exactly the same length of patterns that we condition the probabilities to ($k = m$), the system is now in a regime where not all possible strategies are held by the agents. This is because the strategy space is not well covered - number of agents ($N = 101$) is comparable to the number of strategies ($2^{m+1} = 128$), and as there is a random distribution of strategy pairs amongst agents (for $s = 2$), there is a significant chance that some strategies will not be held by any agents. These missing strategies result in biases in the conditional probabilities that are not able to be exploited by the agents, despite them being able to analyze such pattern lengths. Therefore, in this case $\mathbf{P}(1|\mu) \neq \frac{1}{2}$ meaning that, given we are in a specific state of the world μ, there is a greater chance of the next market movement being in a specific direction. Therefore, there is predictability given the global information.

In Section 3.2 on page 72, we assess the conditional probabilities of the direction of closing price changes in the S&P500 Composite Index using an analogous procedure and find evidence for such biases over a range of values of k. This could therefore be explained by the unlikely possibility that strategies used by traders are of lower complexity (shorter bit-length) than the length of conditioning that we use ($m < k$), or by the more plausible argument that the strategies being used by traders do not fully cover the space of all possible strategies.

The volatility of excess demand $\sigma_D = \sigma(N_1 - N_0)$ is an inverse measure of how effective the system is at distributing resources, and how well the agents can co-ordinate their behaviour to reduce wastage by maximizing the number of people in the minority (the number of winners), thus minimizing the number of people in the majority (the number of losers). If 49 out of 101 agents pick the consequent 'winning' action, there are 49 winners and 52 losers. This is a much more equitable outcome compared to 1 agent picking the resulting winning option, and 100 agents losing - in this case, there are a lot of opportunities for agents to 'win' that are not utilised. In fact, some versions of the minority game use a score system that in effect divides up the scarce resource by the number of winners, so if the number of winners is small, they benefit greatly, as opposed to the minimum amount when divided by the maximum amount of winners (49%).

The optimal outcome would have the winning group being as close to $\frac{N-1}{2}$ as possible, with conversely the losing group at just over $\frac{N}{2}$ - and the more often the excess demand is further away from the optimum, the greater is the volatility. It has been found [Challet, Marsili 1999, Savit et al. 1999, Li et al. 2000a] (indeed, we verify this in simulations and display the results in 3.3 on page 66) that for $s = 2$, the optimal value is $\alpha_c \sim 0.5$, where the subscript c refers to criticality, a concept from Statistical Physics. This gives rise to a minimum in the volatility of excess demand per agent of the system, a tendency for the size of the winning group to be near the maximum possible value of $\sim \frac{N-1}{2}$, and consequently leading to a minimum in the wastage of resources. It seems that in order to reproduce most of the stylized facts of financial markets, α needs to be around the 'critical' value α_c, suggesting the possible 'self-organized criticality' of Financial Markets (see discussions on page 69).

The volatility minimum at $m = 6$ in Figure 3.1 is robust for many values of N, and represents a global optimization - the agents have managed to coordinate themselves such that volatility of excess demand is even less than a simple random coin-toss. Indeed, we demonstrate, as in [Savit et al. 1999], that there is a transition at the global optimum between a symmetric phase of $E(p(excess\ demand \mid \mu)) = \frac{1}{2}$ for low α regimes (see Figure 3.4), and an asymmetric phase of a non-flat $E(p(excess\ demand \mid \mu))$ for high α regimes (see Figure 3.6 on page 71). This is analogous to the concept of dynamical symmetry breaking[1] in Statistical Physics [Challet, Marsili 1999].

[1] where the symmetry is the frequency equivalence between the winning action (1 or 0, or analogously a negative or positive excess demand) in the region $\alpha < \alpha_c$, which is 'broken' for $\alpha > \alpha_c$ where there is a bias in the frequency of agent choices.

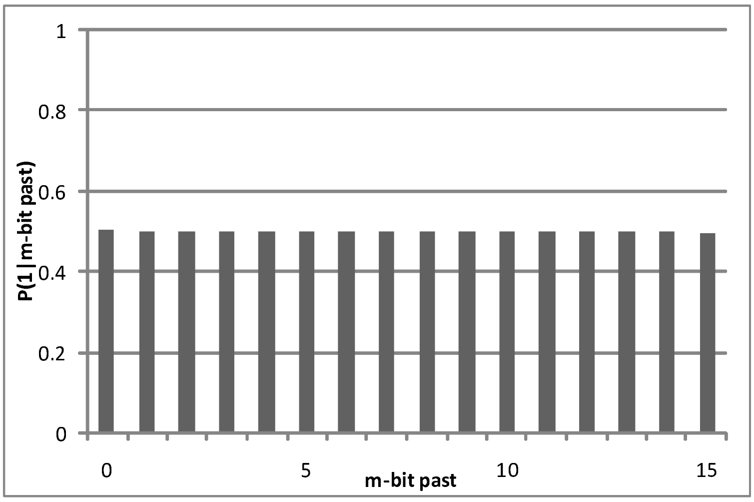


Figure 3.4: The conditional probability $P(1|k\ bit\,past)$ of winning action '1' with $k = 4$ for the game played with $m = 4$, $N = 101$. The flat histogram is indicative of the crowded region - there are enough agents to 'arbitrage' away any biases for the winning action at the next time-step, given that we are in state μ. As the strategies used by the agents are in effect looking for patterns of length $m = 4$, and there are enough agents participating in the market, there is no information left that the agents can exploit.

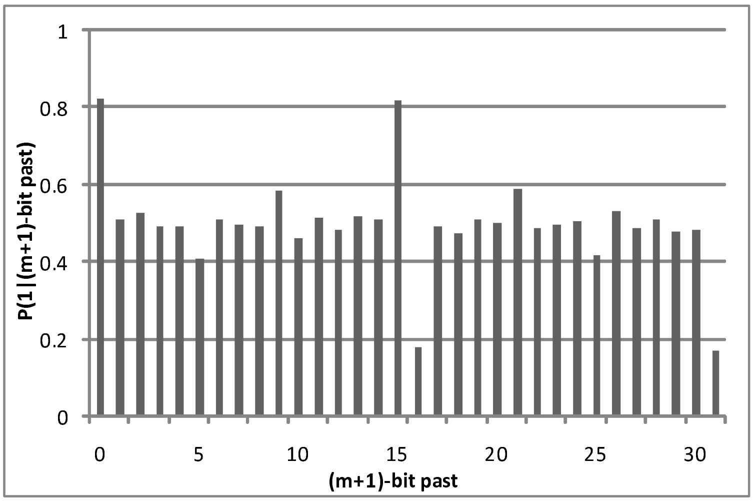


Figure 3.5: The conditional probability $P(1|k\ bit\,past)$ of winning action '1' with $k = 5$ for the game played with $m = 4$, $N = 101$. We use an extra bit of information (compared to Figure 3.4) to assess the conditional probabilities ($k = m + 1$), showing biases are still prevalent in the time-series of winning outcomes. As the agents only look at the past m time-steps, they can't see any patterns longer than this and are therefore unable to exploit these biases when making investment decisions.

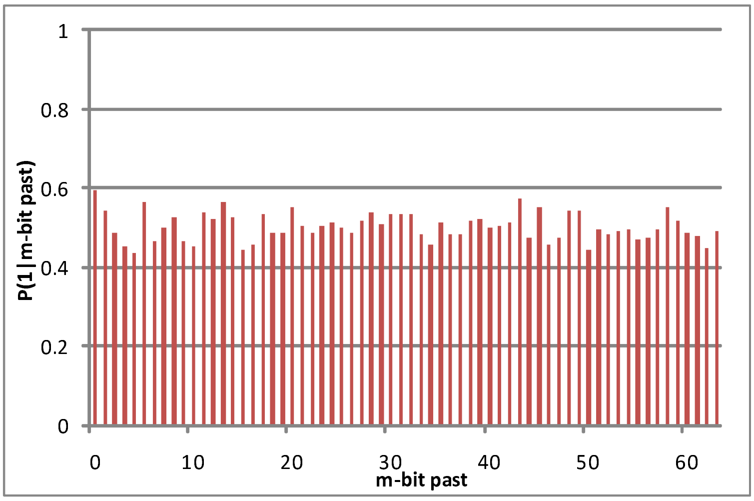


Figure 3.6: The conditional probability $P(1|k\ bit\, past)$ of winning action '1' with $k = 6$ for the game played with $m = 6$, $N = 101$. Although we assess the directional probabilities conditional on the same bit-length of past winning actions as the agents do to make predictions ($k = m$ as in 3.4), we find that there are biases in the system that the agents could in theory exploit. In this case, not all of the $2^{m+1} = 128$ strategies are held by the (101) agents, therefore the patterns that would have been exploited by these missing strategies remain.

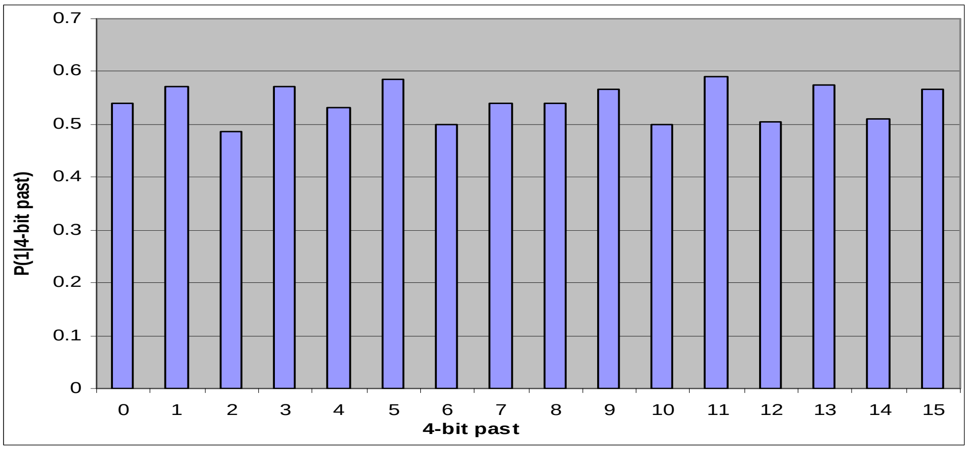
Some propose that there is 'self-organization' [Johnson et al. 1999, Kauffman 1993] or even further 'Self-Organized Criticality' (SOC) in markets [Challet, Marsili 2003, Challet et al. 2005, Bak, Chen 1993, Bak, Paczuski 1995, Bak 1997] in which the market over time naturally tends to a critical point where the 'correlation length' of the system (a measure of how far the correlations between the actions of market participants spread across the market) diverges to infinity - one of the possible causes of the power-law behaviour observed in financial markets [Lux 1996, Farmer 1999, Mantegna, Stanley 2000, Bouchaud, Potters 2003] [Sornette 2004].

If this global optimization occurred by itself, it would support such an argument [Challet et al. 2005]. As imagined in [Challet et al. 2001], one could imagine the implications this has for the development of emerging markets: when there are plenty of opportunities (a high α region - low N / high m), more agents are drawn to participate in the market leading to an increase in N evolving the system to a lower α regime, until a global co-operation is maximized (volatility is minimized). One could argue though, that upon reaching this point, there would appear to be nothing to immediately stop the momentum of newcomers entering the market seeking such apparent opportunities to make a profit. In the long-run, agents may realize that the available profits aren't what they used to be, and that the number of people in the market is unsustainable considering the diminishing profits they can reap. In due course, this could lead to an instability that results in lot of the agents losing all their money and therefore unable to participate in the speculative markets, hence cutting the number of participating players in the market to a sustainable level. However this analysis assumes that the resources / opportunities in the market remain constant overtime, as opposed to the widely held view that innovation and investment brings forth further opportunities that can be exploited. In addition, there is a market 'ecology' in the sense that the success of an agent's strategies depends largely on the strategies of every other agent (remember Keynes beauty contest [Keynes 1936]), and it is likely that as time continues and even after the strategy space is well covered by newcomers to the market, as agents have the possibility to switch between strategies, the system may never settle down to an 'equilibrium'.

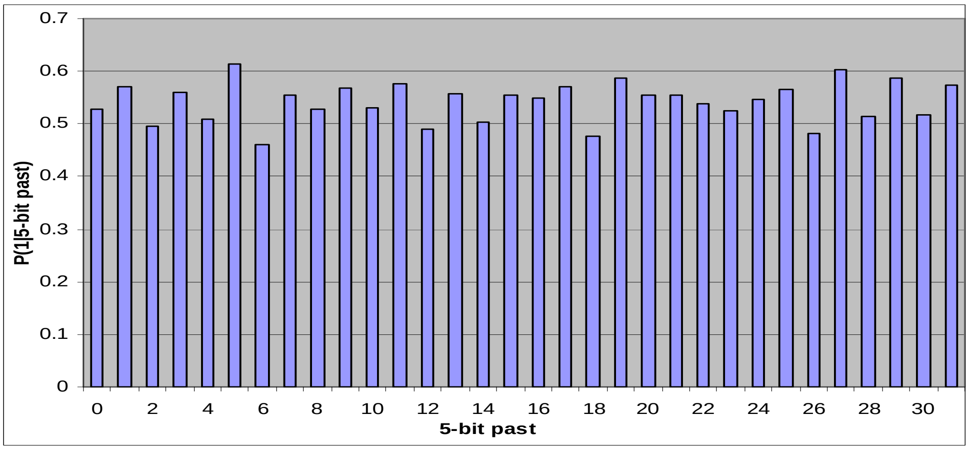
3.2 Anti-symmetric phase of the S&P 500 Index

We perform a new analysis of the S&P 500 Composite Index, similar to that done for hourly USD-JPY data in [Jefferies et al. 2001], by assessing the conditional probability of the daily closing price increments from 31st December 1963 to 10th December 2008. The analysis is motivated by the investigations in [Savit et al. 1999] and discussed on page 67. We find that the conditional probability of the directions are not flat at $\mathbf{P}(1|m - bit\,past) = 0.5$ for the values of m that we analyze - see Figure 3.7 on the next page, similar to Figure 3.6 on the preceding page and different from Figure 3.4 on page 70. This is in accordance with the results presented in [Jefferies et al. 2001], where the conditional probabilities also do no have a flat $\mathbf{P}(1|m - bit\,past) = 0.5$.

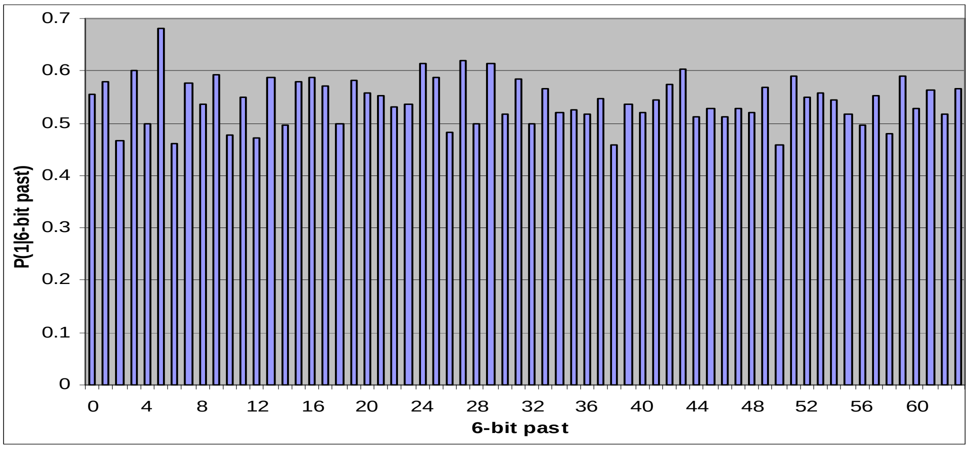
This suggests that, if we were to consider the behaviour of the S&P 500 Index



(a) $k = 4$ bit past



(b) $k = 5$ bit past



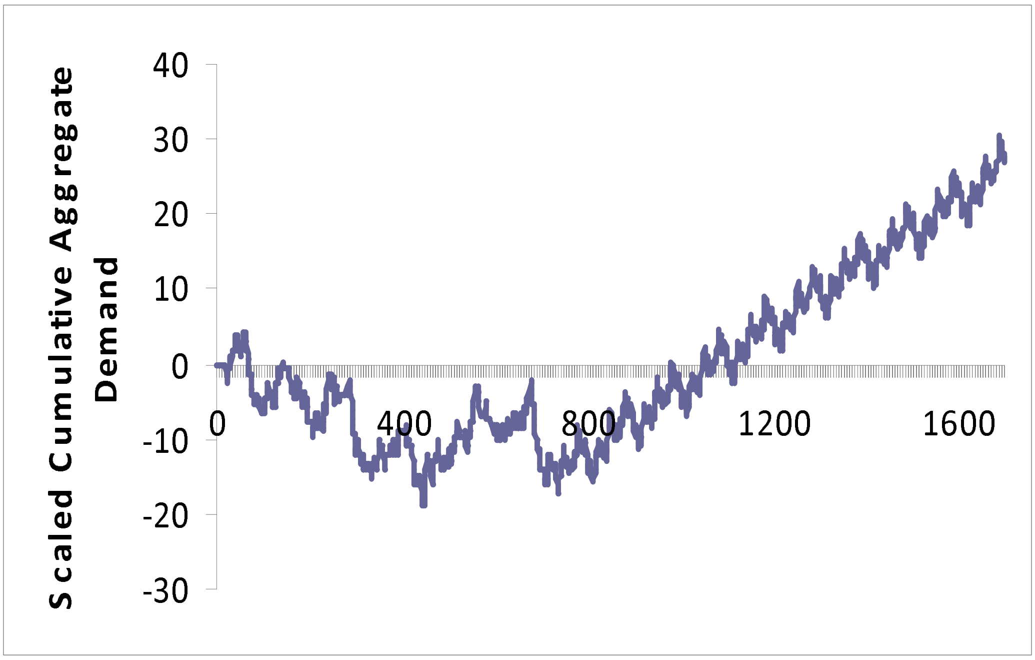
(c) $k = 6$ bit past

Figure 3.7: Histograms of the conditional probability $P\{(\Delta p > 0)|k\ bit\, past\}$ for the daily closing price increments (Δp) of the S&P 500 Index from 1963 to 2008. The fact that the conditional probabilities exhibit biases for the range of k we analyse suggests that there are patterns in the price increments that could be exploited by a Minority Game style forecasting procedure.

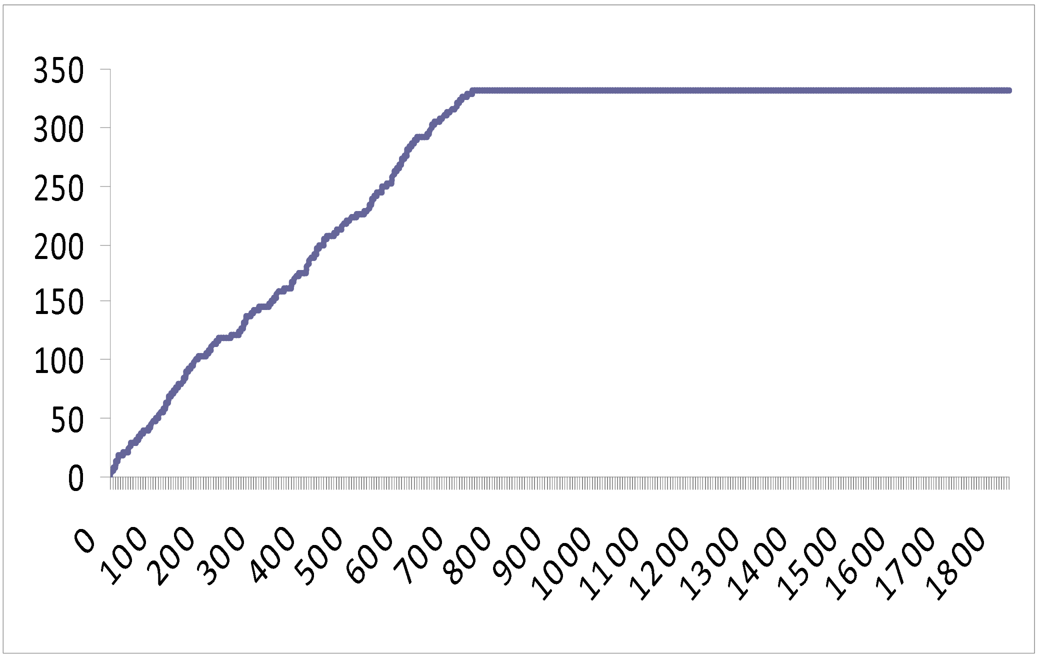
from the point of view of the Minority Game, the market represented by the S&P 500 Index - the 500 largest market capitalised companies in the United States of America - exists in the anti-symmetric phase, the meaning of which was discussed on page 67. This anti-symmetric phase corresponds to the sparse strategy space, with there being more trading strategies possible than those currently being used by the agents. This would more closely correspond to reality because there isn't such a well defined concept of a 'strategy space' for real traders to occupy, as noted in [Lamper 2002]. Indeed, there are many more strategies than can easily be defined in the same binary form that we are using, and these must be responsible for the biases in the choices. However it is still surprising that such a simple record of patterns are not arbitraged away and also that the deviation from the symmetric level of 0.5 can be as large as 0.68 in Figure 73c. This provides some promising evidence that a forecasting framework built around the concepts of the Minority Game - that of expected conditional probabilities - may be able to capture this predictability to make forecasts of the direction of asset prices.

3.3 Periodic and stochastic phases of the Minority Game simulations

In our simulations of the Minority Game, we observe a new phenomenon that we believe has not been discussed in the existing literature, to the best of our knowledge. For the parameter set we use, our Minority Game simulations appear to have two phases - starting with an initial phase, where the evolution is dominated by stochastic 'coin-tosses' to resolve frequent ties in the population of agents choosing each action (to $N_{buy} = N_{sell}$) and the underlying ties in strategy scores. This can only happen in the Grand Canonical Minority Game where the number of agents participating in the market varies at each time-step - in the original Minority Game all agents are forced to participate, however as the total population of agents is an odd integer there can be no tie in agent numbers (though there will still be stochasticity to resolve ties in the strategy scores).



(a) Scaled Cumulative Aggregate Demand time-series, akin to a price formation process from a MG simulation. This illustrates the change of behaviour between an initial stochastic phase (dominated by 'coin-toss' resolutions) into an irregular pattern of increments that are repeated in cycles of length T.



(b) The cumulative number of agent population ties ($N_{buy} = N_{sell}$) that are resolved by 'coin-tosses'

Figure 3.8: The two phases of a GCMG with finite Time-horizon (parameter set $N = 101$, 3, $s = 2$, $T = 100$, $\tau = 0.99$, $c = 0.53$, $L = 0.5$). The phase change occurs after 800 time-steps.

After this initial stochastic phase, the system settles down into a phase where no ties in the population choosing action N_{buy} and N_{sell} occur, giving rise to a cumulative aggregate demand pattern as seen in Figure 3.8, where we see cycles of period T within which are embedded the same stochastic pattern. Figure 3.8a shows the cumulative aggregate demand (the price formation process of equation 2.1 on page 49), and Figure 3.8b shows the cumulative ties in agent population, where $N_{buy} = N_{sell}$. Both figures show a change in phase at the same time-step (around time-step 800), when no more ties in agent population occur. Before this change, the rate of agent population ties - the average number of ties per time-step - is just under 50%. After this phase change in Figure 3.8a, the linear trend occurs, as each repeated period results in the same aggregate demand at each time-step, and thus also over the whole period T, giving rise to an increasing (or sometimes decreasing) price time-series.

This cyclical phase only occurs when a finite Time-horizon is used - if we use the GCMG with a iterative score update (equivalent to letting $T \to \infty$, see 2.5.1 on page 58 for more details), this does not happen. In our example, the phase changes around time-step 800, however in further studies this change can occur much later, though usually no more than after around 20,000 time-steps. We believe that, given enough time-steps, the finite Time-horizon system will eventually find a period of length T time-steps within which there are no ties in agent choice to be stochastically resolved, and thus the system becomes deterministic. This evidence supports the findings of Lamper that a Fourier Transform of the aggregate demand time-series finds evidence of periodicity [Lamper 2002].

3.4 Mix Game

We continue by developing the computer program to simulate Gou's Mix-game model, with the resultant simulated price, volatility of excess demand and trading volume of the Mix-Game with parameters $N = 201$, $N_{maj} = 40$, $m_{maj} = 3$, $T_{maj} = 12$, $N_{min} = 161$, $m_{min} = 6$, $T_{min} = 60$, $s = 2$, $\tau = 1$, $c = 0.53$ displayed in Figures 3.9, 3.10 and 3.11 respectively. These are analogous to

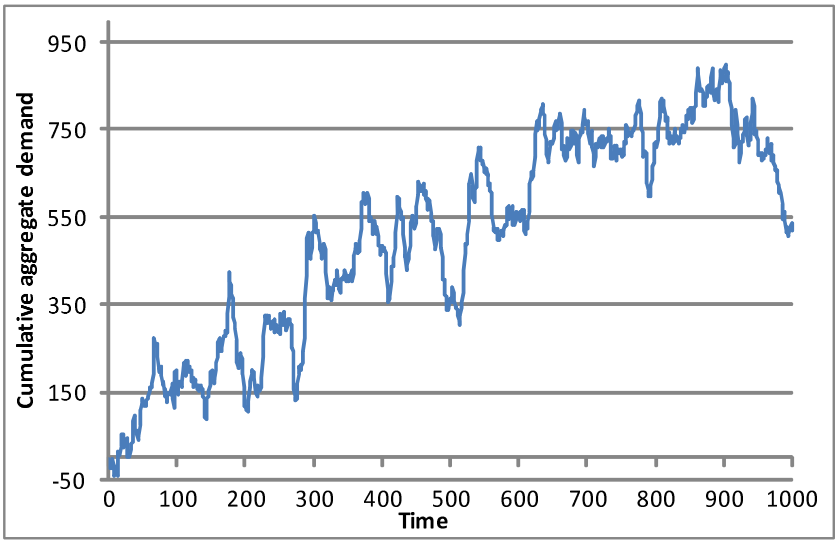


Figure 3.9: Cumulative Aggregate Demand time-series of Mix-Game with parameters $N = 201$, $N_{maj} = 40$, $m_{maj} = 3$, $T_{maj} = 12$, $N_{min} = 161$, $m_{min} = 6$, $T_{min} = 60$, $s = 2$, $\tau = 1$, $c = 0.53$

those in Figure 8b of [Gou 2006a] - indeed, with the same parameters used by Gou, we find similar orders of magnitude of the measures.

However, simply by looking at these graphs would not be enough to distinguish the underlying process, therefore we reproduce additional Figures from Gou's work. Figure 3.12 on page 79 recreates Figure 3 in [Gou 2006 b], which shows the variation of the mean of local volatility of excess demand $< \sigma_D >$ ('local' meaning that it is calculated over a short moving window of 5 time-steps) whilst fixing the time-horizon of the minority-game agents $T_{min} = 36$ and varying the time-horizon parameter of the majority-game agents $T_{max} = 36$, and vice versa. Although the behaviour of the constant T_{min} blue line is slightly different to Gou's, this line is fairly volatile over an ensemble of simulations and our line remains within the (1 standard deviation) error bars and is therefore not a significant difference. Gou fails to show any error bars on her graphs so it is impossible to tell the underlying behaviour of her graph over such an ensemble of simulations. As can be seen, the mean local volatility grows steadily as T_{min} increases keeping $T_{maj} = 36$ constant, whereas mean local volatility is a lot higher for small T_{maj} and constant $T_{min} = 36$ and decreases as T_{maj} is increased. Although the theoretical crossover should be at $T_{maj} = T_{min} = 36$,

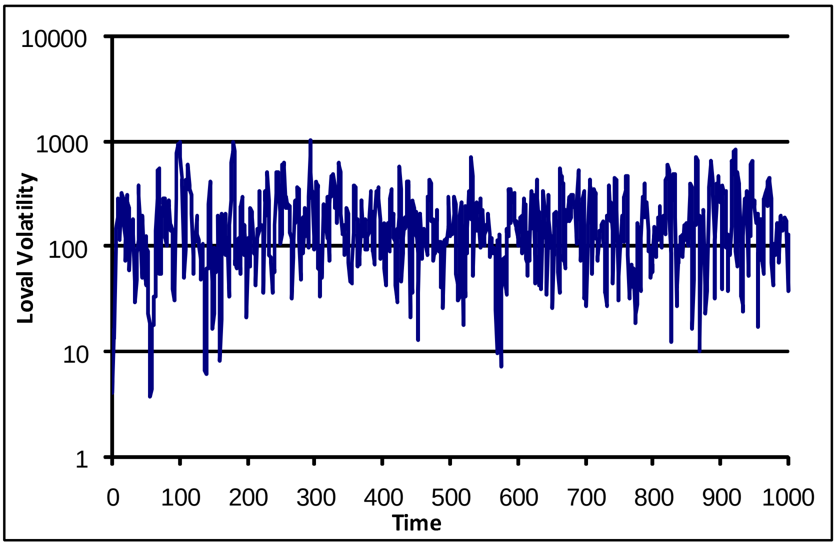


Figure 3.10: Local volatility σ_D (standard deviation of excess demand over a moving window of 5 time-steps) time-series of Mix-Game with parameters $N = 201$, $N_{maj} = 40$, $m_{maj} = 3$, $T_{maj} = 12$, $N_{min} = 161$, $m_{min} = 6$, $T_{min} = 60$, $s = 2$, $\tau = 1$, $c = 0.53$

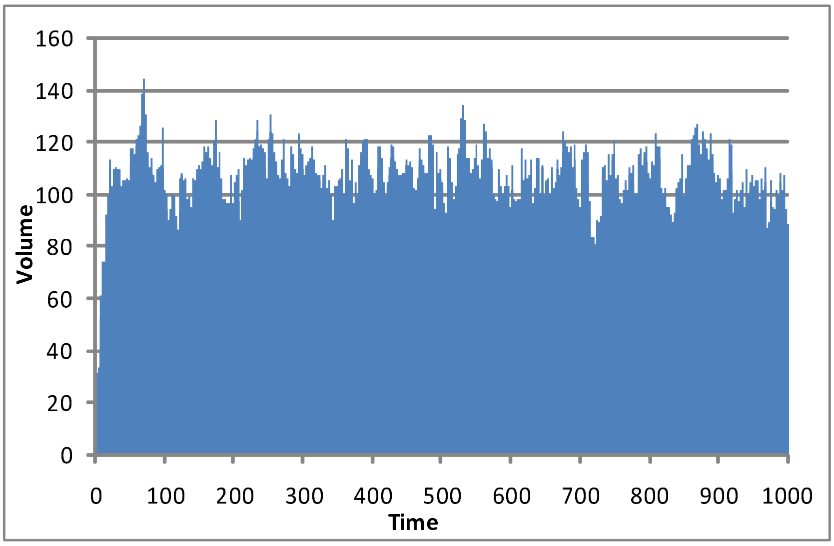


Figure 3.11: 'Volume' (number of active agents) time-series of Mix-Game with parameters $N = 201$, $N_{maj} = 40$, $m_{maj} = 3$, $T_{maj} = 12$, $N_{min} = 161$, $m_{min} = 6$, $T_{min} = 60$, $s = 2$, $\tau = 1$, $c = 0.53$

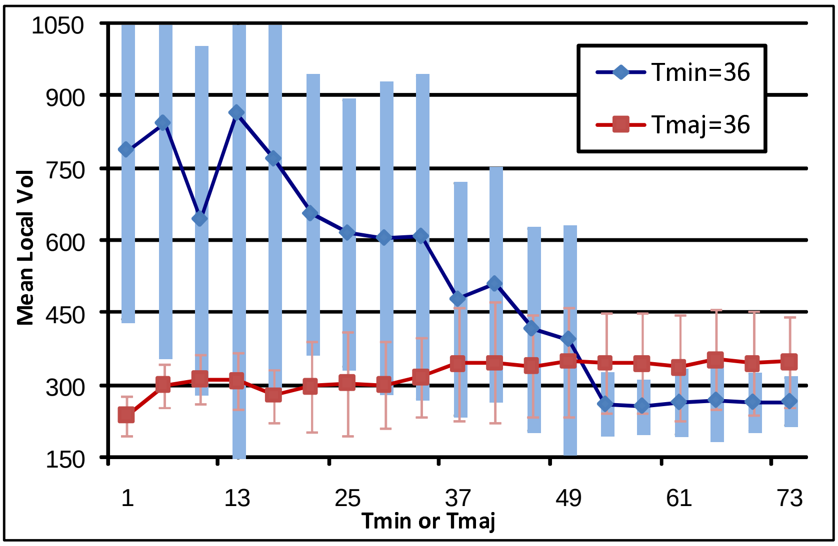


Figure 3.12: Variation of the mean local volatility of excess demand $< \sigma_D >$ over time horizon T_{min} (T_{maj}) keeping $T_{maj} = 36$ ($T_{min} = 36$) fixed, with $m_{maj} = 3$, $m_{min} = 6$, $N = 201$, $N_{maj} = 72$ and $s = 2$. The blue line corresponds with a variable T_{maj} and constant T_{min}, whilst the red line corresponds with a variable T_{min} and constant T_{maj}.

this does not happen in the sample mean of the simulations, though the sample error bars are large enough to show that this would be acceptable from these results within 1 standard deviation.

To analyse the effect of varying the proportion of majority-game agents, we recreate Figure 2a in [Gou 2006a] below in 3.13 on the following page, in the case where the memory parameters and time-horizon parameters are equal ($m_{maj} = m_{min} = 2$, $T_{maj} = T_{min} = 12$). It can be seen that the system becomes more efficient (less wastage / volatility decreases) when agents playing the majority-game are added to the system. This continues until they represent half the population, above which the majority-game agents dominate the system and cause the simulated asset price (and excess demand) to diverge. The shape of the local volatility curves are sensitive to the memory parameters and time-horizon, with the case being that for $m_{maj} = 6 > m_{min} = 5$ and $T_{maj} = 60 > T_{min} = 12$, the local volatility increases, as observed in [Gou 2006a].

The following Figure 3.14 on page 81 reproduces Figure 7 in [Gou 2006 b],

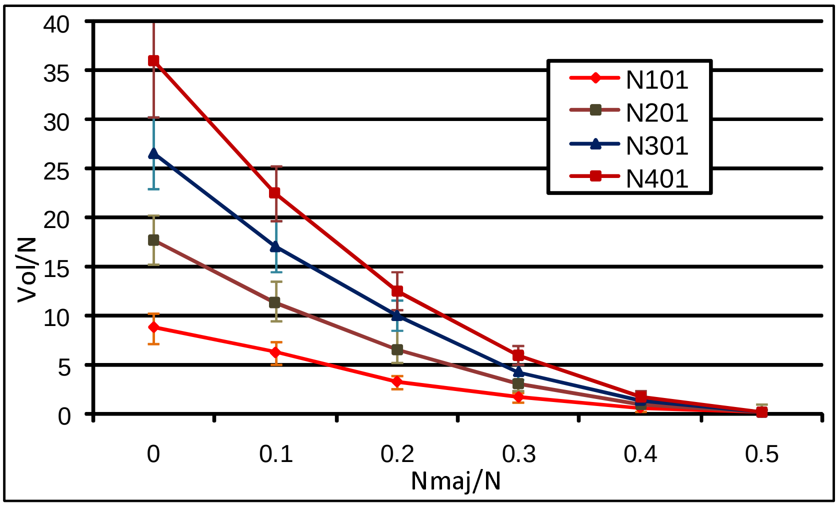


Figure 3.13: Variation of the mean local volatility per agent $\left(\frac{<\sigma_D>}{N}\right)$ over the relative proportion $\left(\frac{N_{maj}}{N}\right)$ with $m_{maj} = m_{min} = 2$, $T_{maj} = T_{min} = 12$ and N varying from 101 to 401

which aims to mimic the daily returns histogram of the Shanghai Index from 1992 to 2004 displayed on a log-log plot. All of Gou's work involves using the finite time-horizon procedure to score and update agents' strategy scores. Additionally, we build a novel version of the mix-game that uses a recursive score updating procedure that, in effect, leads to an infinite time-horizon as discussed in Section 2.5.1 on page 58, and compare the two different set-ups.

Compare the curvature of the two cases for small return/high frequency points, where the infinite time-horizon is a lot more like a power law (straight line) and has a higher frequency of small returns and lower frequency of fat tails when compared to the finite time-horizon version. These are likely to be finite size effects owing to the finite time-horizon, which are also present in Gou's simulations and indeed the actual market data, supporting the view that agents in reality also have a finite time-horizon over which they analyse the success of their strategies. This suggests that the finite time-horizon version is the most suitable version of the Mix-Game in the simulation and forecasting of financial markets.

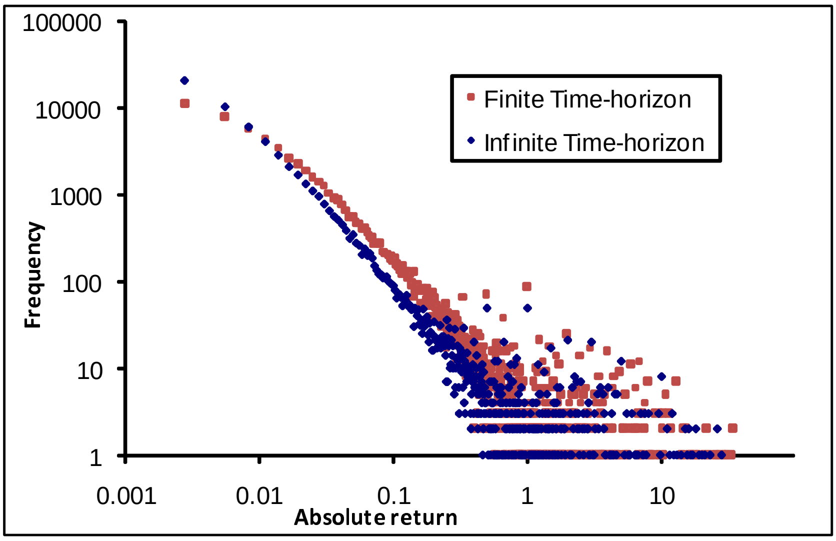


Figure 3.14: Log-log histogram of the Mix-Game absolute returns with parameters $N = 201$, $N_{maj} = 40$, $m_{maj} = 3$, $T_{maj} = 12$, $N_{min} = 161$, $m_{min} = 6$, $T_{min} = 60$, $s = 2$, $\tau = 0.99$, $c = 0.53$

3.5 The Mix Game and Agent Networks - an oil market model

As a prospective future development of the Mix Game in relation to financial markets, we propose the conceptual framework to build a new version of the Mix Game model adapted to mimic the structure of the Oil Market in order to garner some understanding of the features of such a market via simulation. In particular, we discuss the possibility of adding further structure in the form of networks of agents to allow for private information to be shared and there to be co-ordination amongst a (dominant) subset of the agent population.

The Oil Market is an interesting example of a market featuring an oligopoly, with a small collection of entities - oligopolists - united to form a powerful influence in the market. Whilst the Organization of the Petroleum Exporting Countries (OPEC's) "ability to manage the market has long been a subject of intense interest," [Stevens 05, Adelman 1980, Seymour 1980, Gately 1984, Griffin 1985, Mabro 1992, Parra 2004] the media often highlights the significance of OPEC statements regarding the co-ordination of their efforts to exert influence on the market price of oil, either by increasing of decreasing the supply rate [Herrick 2001].

For the purposes of our model, we adapt the ideas from [Anghel et al. 2004, Gourley et al. 2004, Choe et al. 2004] that explores the possibility of adding structure to Minority Game systems with the addition of a network (see [Evans 2004] for discussion of the variety of networks and important concepts). In these models, the agents of the Minority Game are placed on a network (for example a random graph [Erdos, Renyi 1960], a small-world network [Watts, Strogatz 1998], or hierarchical structure [Foldy et al. 2003]), and are allowed to communicate with their neighbours. In particular, they borrow the strategies of their nearest neighbours, thus increasing the number of strategies available to them. The dynamics introduced by the network include a reduction in volatility, owing to the increased level of local co-ordination.

3.5.1 The Oil Producers - OPEC

We can picture OPEC members as agents forming a collective of a small, dense network (with every agent linked to many of the other OPEC agents), sharing their strategies in order to co-ordinate their market participation to influence the price of oil. The concept of 'producers' from the perspective of the Minority Game framework has been introduced in [Challet et al. 2000, Challet et al. 2005] by suggesting that producers have less adaptability in their market transactions - they simply wish to sell their products and most likely do so the moment they have them. This is likely to occur in fairly regular, periodic cycles, as expected in the behaviour of farmers for example, who have specific times of the year when their crops will be in harvest. This sort of behaviour can be captured in our model by setting the number of strategies available to each OPEC agent $s^{OPEC} = 1$, so that they are unable to adapt using their own strategy alone. Through their network, the OPEC agents are able to borrow the most successful strategies being used in the group, thus they do have some adaptability in co-ordinating their efforts to maximise their profit or utility. In some ways the fact that the network of OPEC agents can still collectively use a variety of different strategies may reflect the possibility of some countries 'cheating' the production limits set for each country, or perhaps that they can co-ordinate themselves overall to produce an steady average output but don't need to use exactly the same strategies individually.

Although it could be argued that OPEC may want to maintain a high price of oil, it will usually claim the reason for their actions is the need to maintain stable prices [Herrick 2001], favoured by the oil exporters as it guarantees a steady stream of income, and desired by oil importers so that they can reliably price their industrial operations. Indeed, this stability is one of the reasons attributed to the massive investment in capacity by the Saudi Arabia government in the mid-1980s so as to encourage energy consumers to return to oil as a primary fuel source [Stevens 05]. Indeed, as can be seen in recent years with the global financial crisis since the autumn of 2007, such high oil prices may have been a primal cause of the inflationary pressures that led to rising interest rates and mortgage repayments, resulting in the default of many home-owners on their mortgage repayments in the USA, thus precipitating the record numbers of defaults on collateralized debt obligations containing such 'toxic debt', the mistrust and subsequent failure of many banks and finally, the global downturn. So in the quest for such price stability, the mean-reverting nature of the agents playing the Minority Game is ideally suitable. We can therefore picture OPEC as a small, dense network of agents, each with a single strategy, playing the Minority Game. In simulations, we could analyse the effects of changing the network structure of the agents - from a random or small-world network with a few countries acting as major hubs of information, to a hierarchical structure dominated by one of two major oil exporters such as Saudi Arabia.

3.5.2 The Oil Consumers

Sharing this market with OPEC is the rest of the market participants - the energy consumers - independently acting, speculative agents, able to adapt their behaviour in order to maximise their profit. We therefore model these agents much like the traditional Minority Game agents, having more than one strategy to choose from. This 'diffuse' group of independent agents features a myriad of different market participants, some trading using fundamental pricing ideas such as the fair-value of oil (the macro-economically justifiable price of oil based on global demand for crude oil or the marginal cost, or based on the budget assumptions of OPEC members or oil companies in order to balance their accounts) and thus, like OPEC playing the mean-reverting style Minority Game. Other, more speculative participants, may use technical trading methods to profit from market trends and will be playing the majority game. Consequently we decide to model this group as independent agents playing the Mix Game.

3.5.3 Conditions for stability

In order to guarantee stability in the model, we require equations 2.5, 2.6 and 2.7 on page 54 in Chapter 2 to be fulfilled. These constraints are global requirements, not to be applied to the distinct OPEC network or the diffuse group separately. Constraint 2.5 applies to a Mix Game where each agent chooses to buy or sell one unit of an asset at each time-step. This necessitates the introduction of another concept - that of the impact factor I of the agents which corresponds to the buying power of the agents. Up until now, we had assumed this to be constant and equal amongst the agents, although we note that a variable buying power has been previously discussed in [Johnson et al. 2003]. Without even considering the massive impact that the whole of OPEC could muster if fully co-ordinating its actions such that the might of several state run companies simultaneously entered into the market with the same action, an individual OPEC state would still be quite a formidable market player. We compare this to a proprietary trader in the City, whose market impact most likely pales in comparison. Therefore we should factor into the model a rough estimate of the impact each agent can have on the market. We can do this by creating a variable for each agent that determines the number of units with which they buy or sell in the market place which could be dependent on other new measures such as their wealth or inventory (in particular, the massive production capacity of Saudi Arabia dominates several other OPEC members such as Venezuela, who cannot even supply their own export quota limits [Stevens 05] - indeed there is even debate over the relative importance of OPEC versus Saudi Arabia

[Cremer, Isfahani-Salehi 1991, Dahl, Yucel 1991, Al-Turki 1994, Gately 1995, Salehi-Isfahani 1995, Gulen 1996, Al-Yousef 1998]). Alternatively, we could consider one high-impact agent as being the aggregate of several identical traders (with similarities to the concept of the 'representative agent' as discussed in Chapter 2, although these aggregated agents would still be a minor component of the market population and not be representative of the whole market). In fact, we could allow an infinite combination of different impact factors, however to simplify the matter somewhat, we propose to initially use two different impact factor values - that of an OPEC agent I^{OPEC} and that of the diffuse population I^{diff} such that:

$$I^{OPEC} > I^{diff} \tag{3.1}$$

and reserve the possibility for a more complicated distribution of market impact values to be investigated afterwards.

We can now specify the updated version of constraint 2.5 on page 54 for our Oil market game as follows:

$$\sum_{i}^{N_{maj}^{diff}} I_{maj}^{diff} < \sum_{j}^{N_{min}^{diff}} I_{min}^{diff} + \sum_{k}^{N^{OPEC}} I^{OPEC} \tag{3.2}$$

where $I_{maj}^{diff} = I_{min}^{diff}$ as the market impact of the diffuse agents playing the majority game is equal to that of the agents playing the minority game. Here the summation limits are the number of diffuse agents playing the majority game N_{maj}^{diff}, the number of diffuse agents playing the minority game N_{min}^{diff}, and the total number of OPEC members N^{OPEC}, also playing the minority game. These values are parameters in our model, however we could simplify the investigation by fixing N^{OPEC} and allowing N_{maj}^{diff} and N_{min}^{diff} to vary. As of the time of writing, the actual number of OPEC members amounts to 13 [OPEC], so we might choose for $N^{OPEC} = 10$. Also, the size of the network - the number of agents embedded on it - may have an effect on its properties, and so we need to verify whether this number is suitable for our simulations.

3.5.4 Other factors - market shocks

If we really wanted to increase the realism, we could also include unpredictable market shock factors. These commonly include hurricanes over Texas, engineering problems in major oil refineries, political instabilities and terrorism in western Africa's oil exporters such as Nigeria and Equatorial Guinea. However, these stochastic factors would merely add noise to our system and hamper understanding of the behaviour of our model - unless analysing the response of the market to such stochastic shocks, which would also be an interesting issue to investigate. Therefore we initially decline to feature stochastic market shock factors.

3.5.5 Potential applications and questions

We suggest running simulations with this structure, investigating the behaviour over the space of parameters, to see if we can garner an understanding of particular phenomena observed specifically in oil markets or other oligopolies, in particular the price volatility and its relation to the network effects and stability. We could also consider using such a model as the basis for a forecasting framework, analogous to that of the Minority Game used in Chapters 4 and 5. In particular, we could see if it improves the accuracy of forecasts of oil markets, oil-based derivatives markets or even the share price of oil companies such as BP (the logic being that the profitability of oil companies tends to be positively correlated with the price of oil), over the conventional Minority Game Forecasting Framework (MGFF).

3.6 Concluding Remarks

In this chapter, we have discussed several aspects of the simulation of Minority Game and Mix Game models. We presented examples of the time series that can be generated in these simulations, discussed the different phases we can observe and the statistical properties such as the conditional probabilities of the sign of the price increment for the MG. We then applied similar ideas to financial data from the S&P500 Composite Index, observing evidence to suggest predictability that may be open to exploitation from a forecasting technique built around Minority Game principles. We have also compared results from our Mix Game to work in *[Gou 2006 b]*, in particular the presentation of the Shanghai Stock Exchange data, noting the finite-size effects in both. Lastly, we discussed the possibility of adding further structure to the models by introducing networks of agents to allow for private shared information and co-ordination amongst a subset of the agent population, and highlight it's potential applicability in simulating an oligopoly dominated market such as the oil market.

Chapter 4

The Forecasting Framework

> *"...but in the world nothing can be said to be certain except death and taxes."* Benjamin Franklin

The (Grand Canonical) Minority Game is an interesting candidate for a generic market model as the simulations we presented in Chapter 3 demonstrate that it can reproduce some of the most obvious stylised facts of financial markets, despite containing a small number of parameters. From this standing point, the next logical question to ask is whether it can be used to forecast these financial markets.

There is considerable evidence in the financial econometric literature that excess stock returns are predictable [Fama, French 1989, Campbell, Shiller 1988b, Menzly, Santos, Veronesi 2004, Lettau, Ludvigson 2001, Lioui, Rangvid 2009]. However, many studies have suggested that only the direction of asset returns is predictable [Breen et al. 1988, Hong, Chung 2003, Christoffersen, Diebold 2006]. Nyberg suggests that this is possibly because the noise in the observed returns is too high for the accurate forecasting of the overall return. Indeed, he highlights [Leitch, Tanner 1991] who "find that the direction of the change is the best criterion for predictability because traditional statistical summary statistics may not be closely related to the profits that investors are seeking in the financial market. Directional predictability is also important for market timing, which is crucial for asset allocation decisions between stock and risk-free interest rate investments" [Nyberg 2008].

In this vein, we outline a method of reverse engineering the Minority Game model in order to train on, and subsequently forecast, the directional change of time series data. A Kalman Filter is employed to calibrate the Minority Game to time series data and create a Minority Game Forecasting Framework (MGFF) that makes short-term predictions regarding the direction of future increments of a time series. The primary benefits of using the Kalman Filter in the MGFF are its dynamic ability to track the changing features of the agents' strategies and its applicability in online implementation. The computational time efficiency of the Kalman Filter could allow one to use this model to perform high-frequency forecasting, which may be useful in high-frequency trading. We present the results of applying the MGFF to time series data form MG simulations and a range of financial assets.

4.1 The General Methodology

A general methodology for forecasting the Minority Game was originally outlined in [Johnson et al. 2001] and [Lamper 2002], the main steps of which are described in this section.

4.1.1 Scoring the strategy space, determining the state μ and making next-step predictions

Specifically, this involves 'reverse engineering' the Minority Game so that a time series of asset prices is used as inputs that are fed into the MGFF. This data is then converted into binary data representing the direction of price changes ('1' for a positive price change, '0' for a negative one, see 4.1.5 on page 106 for discussion of the event of no price change). At each time-step, the data is used to determine the score of each strategy over the previous period equivalent to the time-horizon T (discussed in 2.5 on page 54). From this, we can determine the strategies that have performed well enough that their scores are greater than the necessary points-threshold to be considered 'active' strategies, and therefore be used by the agents to make their investment decision for the next time-step. The past m price changes are used to determine the current state of the system μ, and from this the action of all the strategies can be evaluated.

Owing to the symmetry of the strategy space (for every strategy there is an anti-correlated strategy which will predict the opposite action at the next time-step for every state of the world), the aggregate demand arises from a bias in the prediction of the 'active' strategies. See tables 4.1 and 4.2 on the next page for examples of when the activity of the strategy space is unbiased and biased respectively for the case $\mu = 0$. In the first case, the number of active strategies suggesting to sell at the next time-step

$$As[t] = \sum_{R=1}^{R=2^{m+1}} \mathcal{H}\{S_{a_R^0 \ni -1}[t] - r\} = 2 \tag{4.1}$$

and the number of active strategies suggesting to buy at the next time-step

$$Ab[t] = \sum_{R=1}^{R=2^{m+1}} \mathcal{H}\{S_{a_R^0 \ni 1}[t] - r\} = 2 \tag{4.2}$$

are equal with two strategies each. Thus if we make the simplifying assumption made originally in the literature [Lamper 2002], that all strategies are held by equal numbers of agents and hence weighted equally, this would create an aggregate demand of 0 at the next time-step (see 4.1.4 on page 94 for discussions on relaxing the equal strategy distribution assumption). In the latter case, the number of active strategies to sell at the next time-step $As[t] = 3$, whilst the number of active strategies to buy at the next time-step $Ab[t] = 1$, this leading to a bias of $Ab[t] - As[t] = -2$, implying a negative aggregate demand at the next time-step if all strategies are equally weighted, and consequently a negative price movement at the next time-step.

If we assume an equally weighted strategy space, the number of agents choosing to sell (buy) will be proportional to the number of active strategies predicting to sell (buy) at each time-step. Thus, the aggregate demand in the Minority Game is caused by a bias in the strategy space, and the expected excess demand

Strategy	Active status, $\mathcal{H}\{S_R - r\}$	Action at next time-step (given μ), a_R^μ
$R = 1$	Active	-1
$R = 2$	Inactive	-1
$R = 3$	Active	-1
$R = 4$	Inactive	-1
$R = 5$	Active	+1
$R = 6$	Inactive	+1
$R = 7$	Active	+1
$R = 8$	Inactive	+1

Table 4.1: Unbiased activity of strategy space for $m = 2$ and $\mu = 0$. As the number of active strategies stipulating '+1' (buy) at the next time-step equals the number of strategies stipulating '-1' (sell), the strategy activities are unbiased. Assuming each active strategy is being used by an equal proportion of agents, this would suggest a zero aggregate demand over the next time-step.

Strategy	Active status, $\mathcal{H}\{S_R - r\}$	Action at next time-step (given μ), a_R^μ
$R = 1$	Active	-1
$R = 2$	Active	-1
$R = 3$	Active	-1
$R = 4$	Inactive	-1
$R = 5$	Active	+1
$R = 6$	Inactive	+1
$R = 7$	Inactive	+1
$R = 8$	Inactive	+1

Table 4.2: Biased activity of strategy space for $m = 2$and $\mu = 0$. As there are three active strategies ($R = 1, 2\,3$) stipulating '-1' (sell) and only one active strategy $R = 5$ stipulating '+1' (buy), the strategy activities are biased. Assuming there are an approximately equal number of agents using each active strategy, this would suggest a negative aggregate demand at the next time-step.

is:

$$\boldsymbol{E}\{D[t^-]\} = N_1[t^-] - N_0[t^-] = \psi \cdot (Ab[t-1] - As[t-1]) \tag{4.3}$$

where ψ is a constant of proportionality to be determined by regression.

The variance in this estimate will depend on the number of agents active at each time-step - the more active agents there are, the greater the number of possible orders that are submitted and aggregated at each time-step. Indeed, the number of possible ways to get a particular value of excess demand can be calculated from the binomial co-efficients. Extensive statistical analysis by Lamper has shown that a linear relationship of $\phi \cdot (As[t] + Ab[t])$ for the variance of the forecast is broadly appropriate, where ϕ is a constant of proportionality [Lamper 2002]. Lamper then models the excess demand as a normal distribution:

$$D[t^-] \sim \mathrm{N}\{\psi \cdot (Ab[t-1] - As[t-1]),\ \phi(Ab[t-1] + As[t-1])\} \tag{4.4}$$

arguing that "we seek the simplest stable distribution as a density forecast, whilst acknowledging that the true distribution of D is indeed fat-tailed".

As the price increment is believed to be approximately linearly related to aggregate demand, we have the equation for the expected price change $dP[t]$ at time t:

$$\boldsymbol{E}\{dP[t]\} = \mathbf{E}\left\{\frac{D[t^-]}{\lambda}\right\} = \gamma \cdot \{\frac{(Ab[t-1] - As[t-1])}{(Ab[t-1] + As[t-1])}\} + \beta \tag{4.5}$$

where γ and β are regression co-efficients which can be determined from historic data.

4.1.2 Multi-step forecasts

In order to make forecasts of the price changes several time-steps ahead, the forecast model is run ahead on the assumption that the expected next price increment is realised. The strategy scores and state are updated with the expected price increment direction, and a similar estimate is made for the following time-step.

4.1.3 Recovering parameters of a black-box MG

In order to test the forecasting ability of the methodology, we start by using data generated by a 'black box Minority Game' - one in which we pretend not to know the parameters used in the simulation to generate this data. By scanning over the parameter space (i.e. m, T, c, τ) and measuring the correlation between the predictions and the actual outcome of the black box minority game simulation, Lamper found that he was able to recover the parameters of the 'black-box' MG (which gave a maximum correlation or hit-rate measure) [Lamper 2002]. Note that our forecasting framework focuses on the strategies in play and their relative weights, and not on any real notion of the agent, unlike the original derivation of the Minority Game. This has the simplifying effect that we need not consider the parameters N and s in our parameter search. The total population parameter N in our MGFF, would simply be absorbed as a coefficient of all the strategy weights, and in this chapter we rely on a Kalman Filter estimation procedure to increase or decrease the strategy weights as necessary as in [Lamper 2002] (see Section 4.1.4 for further information). This is different to the work by Gou, whose generic [*sic*] algorithm still considers the individual agents, thus requiring a search for N (and setting $s = 2$) in order to deduce the distribution of strategies across agents [Gou 2006c].

Here we take the correlation of predictions measure to be the usual 2-point correlation function between the price increment $dP(t)$ at time t and the prediction of the price increment $\widehat{dP}(t)$ calculated at time $t-1$ by the MGFF. The sample is taken over a window of size M:

$$Corr = \frac{\sum_t^M (dP(t) - \overline{dP}) \cdot (\widehat{dP}(t) - \overline{\widehat{dP}})}{\sqrt{\sum_j^M (dP(j) - \overline{dP})^2} \sqrt{\sum_k^M (\widehat{dP}(k) - \overline{\widehat{dP}})^2}} \tag{4.6}$$

The hit-rate measure, which measures the number of times the sign of the price increment is correctly predicted, is defined as:

$$Hit\,rate = \frac{|\ dP(t)\,\widehat{dP}(t) > 0,\ t = 1, 2, \ldots, M\ |}{|\ dP(t)\,\widehat{dP}(t) \neq 0,\ t = 1, 2, \ldots, M\ |} \tag{4.7}$$

Here, the norm of the set is the cardinality, which is the number of elements in the set.

Another measure of predictive ability is the relative hit-rate, defined as the ratio of the hit-rate measure and a naive predictor:

$$Relative\ Hit\,rate = \frac{Hit\,rate}{Naive\ Predictor} \tag{4.8}$$

$$Naive\ Predictor = \frac{\mid dP(t)\,dP(t-1) > 0,\ t = 1,\,2,\,\ldots,\,M \mid}{\mid dP(t)\,dP(t-1) \neq 0,\ t = 1,\,2,\,\ldots,\,M \mid} \tag{4.9}$$

Here the naive predictor measures the success of a prediction based on the assumption that the sign of the price increment at time t will be the same as the sign of the price increment observed at time $t-1$.

We additionally investigate the hit-rate of e time-steps ahead using the multi-step forecasting discussed in Section 4.1.2:

$$e\ step\ hit\,rate = \frac{\mid dP(t+e)\,\widehat{dP}(t+e) > 0,\ t = 1,\,2,\,\ldots,\,M \mid}{\mid dP(t+e)\,\widehat{dP}(t+e) \neq 0,\ t = 1,\,2,\,\ldots,\,M \mid} \tag{4.10}$$

and the relative hit-rate of predictions e time-steps ahead using:

$$e\ step\ relative\ hit\,rate = \frac{e\ step\ hit\,rate}{e\ step\ naive\,predictor} \tag{4.11}$$

Whilst the set of Grand Canonical Minority Game parameters $\{m,\,T,\,c,\,\tau\}$ will be obtained using a simple trial-and-error search for a parameter set that maximizes the average hit-rate measure, the next section outlines a general method for estimating the popularity of the Minority Game strategies in play.

4.1.4 Estimating the strategy weightings distribution

Initial models had the simplifying assumption that all strategies were equally popular and hence equally weighted. Further work considered the relaxation of this assumption to improve the flexibility of the forecasting methodology, and various methods such as Kalman Filtering [Lamper 2002] and Generic [*sic*] (Genetic) algorithms [Gou 2005] have been applied to determine the strategy weightings.

4.1.4.1 Possible estimation techniques

There are several possible techniques that can be used to estimate the strategy distribution weights. These include the simple least squares (batch) regression [Bar-Shalom et al. 2001], genetic algorithms [Beasley 1999] or recursive estimation such as the Kalman Filter [Gershenfeld 1999]. The least squares regression, performed with the dimension of the regression being the number of strategies in the strategy space (2^{m+1} in the reduced strategy space), would be particularly inefficient for on-line implementation of the MGFF, as when new data is observed, the regression has to go through largely the same data all over again. Genetic algorithms also suffer from this problem, although the superior performance of genetic algorithms, especially for high-dimensional estimation problems can sometimes compensate for this extra computational time and effort. Indeed Gou employs such techniques to calibrate the mix-game [Gou 2005] and we consider it in chapter 5 on page 144.

Recursive estimation procedures such as the Kalman Filter or recursive least squares are much more computationally efficient at handling on-line estimation. Parameter estimates are calculated and updated when new information is observed without the need to store and process all the old data again at each time-step. Indeed, this approach is much faster than the batch regressions or genetic algorithms, and is suitable if the MGFF were to be used for high-frequency / intra-day trading. It is possible that superior results may be achieved with genetic algorithms, though this may be restrictive for practical use, if for example it takes a day or more to process all the information and find optimal estimates - however this may be good enough for trading over the frequency of a few days or a week.

4.1.4.2 The Kalman Filter

> "[The Kalman Filter] must rank as one of the most important contributions to control theory in the twentieth century." [Hansen, Snyder 1997]

The Kalman Filter is a particular form of recursive estimation procedure that was developed in the early 1960s to handle the problem of tracking bodies with well known dynamical equations of motion such as aircraft and the Apollo spacecraft [Kalman 1960, Kalman, Bucy 1961]. During the 1980s it started being applied to economic and business estimation problems [Harvey 1989]. Not only does it have the benefits of being recursive and hence computationally fast and efficient for large amounts of repeated data, but another strength is its ability in estimating unobserved internal states of a system, such as the position of an aircraft (which is in fact unobserved, but estimated from radar time returns or a global positioning satellite receiver). In our case the unobserved internal state of the system is the (MG) strategies in play in the market-place.

The procedure features two main equations, the first of which is the dynamic (also known as plant) equation:

$$x[t+1] = F[t] \cdot x[t] + G[t] \cdot u[t] + \varepsilon[t] \tag{4.12}$$

where x is the n_x-dimensional Kalman state vector - in our case the weightings of the MG strategies, $F[t]$ is the n_x by n_x transition matrix that describes the (linear) deterministic part of the transition from $x[t]$ to $x[t+1]$, u is an n_u-dimensional input/control vector representing a (deterministic) input into the system that we might like to make, $G[t]$ is a n_x by n_u response matrix that reflects the sensitivity of the Kalman state vector to the input vector, and $\varepsilon[t] \sim \mathbf{N}(0, Q[t])$ is a n_x-dimensional zero-mean and white (no auto-correlation) process noise reflecting the unpredictable noise or model mismatch errors between time t and $t+1$ where $Q[t]$ is the n_x by n_x covariance matrix of this noise.

The input vector $u[t]$ in our model is assumed to be zero (and henceforth we neglect the response matrix $G[t]$) as we are attempting to estimate the state vector with no knowledge of any deterministic change to the system. There may at times be additional knowledge of our system to suggest that $u[t']$ is non-zero at time t'; perhaps for some reason we have information that there will be a large increase in one strategy over all others. For example if a Central Bank decides to enter into the market to prop up its currency (or weaken it as in the case of China in recent years), regardless of any state of the world μ, this could be modeled by an increase in the strategy advocating '+1' in every column i.e. the bottom strategy number 8 in Figure 2.8 on page 47. Alternatively, the case in 2008 when the UK's Financial Services Authority imposed rules banning the short selling of banking stocks in an effort to reduce share price volatility could be modeled by a reduction in the strategy proposing to sell in every state of the world, i.e. a reduction in the strategy proposing '-1' in every column (the top strategy number $R = 1$) in Figure 2.8 on page 47. However, this would be difficult to model in the sense that we would need to calibrate the effect of the intervention in the market, plus model issues such as the timing of the intervention. This may be easier when the model frequency is based on daily data, but would be extremely difficult for high-frequency modeling. In the general case where we have no knowledge of this, we take $u[t]$ and $G[t]$ to be zero.

The second main equation in the Kalman Filter system is the measurement equation:

$$z[t] = H[t] \cdot x[t] + \omega[t] \tag{4.13}$$

where $z[t]$ is the n_z-dimensional measurement vector, $H[t]$ is the n_z by n_x measurement matrix that projects the state vector onto the measurement space, and $\omega[t] \sim \mathbf{N}(0, J[t])$ is the n_z-dimensional zero-mean white measurement noise vector reflecting the error in taking measurements where $J[t]$ is the n_z by n_z covariance matrix at time t.

There are several good resources [Gershenfeld 1999, Bar-Shalom et al. 2001, Grewal 2008, Gupta, Hauser 2007] that explain the theory of the Kalman Filter and its implementation in more detail and we refer the reader to these, whilst outlining the salient points below:

1. Use the 'best' estimate of the preceding state $x[t-1]$ which we shall label

$\hat{x}_{t-1|t-1}$ to predict the current state $x[t]$, calculated using the expectation operator on the dynamic equation (4.12):

$$\hat{x}_{t|t-1} = F[t-1] \cdot \hat{x}_{t-1|t-1} \tag{4.14}$$

The current state prediction is labelled $\hat{x}_{t|t-1}$, reflecting that only information up to and including time $t-1$ has been used in determining this value. Notice that the predicted current state is not influenced by the process noise term $\varepsilon[t]$ as this has an expected value of zero. We also assume that there is no deterministic input control of the system as discussed on page 96, hence $u[t] = 0$.

The difference between the current state prediction $\hat{x}_{t|t-1}$ and the true state $x[t]$ is the state prediction error $\widetilde{x}_{t|t-1}$:

$$\widetilde{x}_{t|t-1} = x[t] - \hat{x}_{t|t-1} \tag{4.15}$$

with a related covariance matrix called the state prediction error covariance $P_{t|t-1}$

$$P_{t|t-1} := \mathbf{E}\{\widetilde{x}_{t|t-1} \otimes \widetilde{x}'_{t|t-1}\} = F[t-1] \cdot P_{t-1|t-1} \cdot F[t-1]' + Q[t] \tag{4.16}$$

where $P_{t-1|t-1}$ is state estimate error covariance defined in step 4 at time $t-1$, A' indicates the transpose of a vector or matrix A, and $a \otimes b$ means the outer product between two vectors a and b.

2. Use the predicted state $\hat{x}_{t|t-1}$ to predict the measurement $z[t]$ taken at time t, calculated using the expectation operator on the measurement equation (4.13), which takes the predicted state into the measurement space:

$$\hat{z}_{t|t-1} = H[t] \cdot \hat{x}_{t|t-1} \tag{4.17}$$

The predicted measurement is labelled $\hat{z}_{t|t-1}$to indicate that this is determined only using information up to and including time $t-1$. Again, notice that the predicted measurement is not influenced by the measurement noise term $\omega[t]$ as this has an expectation of zero.

3. Information from the current time t now flows into the estimation process as we compare the predicted measurement $\hat{z}_{t|t-1}$ with the actual measurement outcome $z[t]$:

$$v[t] = z[t] - \hat{z}_{t|t-1} \tag{4.18}$$

where $v[t]$ is the measurement residual that we want to minimise in this procedure, with a related covariance matrix called the measurement residual covariance B_t:

$$B_t := \mathbf{E}\{v[t] \otimes v[t]\} = H[t] \cdot P_{t|t-1} \cdot H[t]' + J[t] \tag{4.19}$$

4. We update the predicted state $\hat{x}_{t|t-1}$ with the new information at time t by adding the linearly weighted measurement residual $v[t]$ to get the best linear state estimate $\hat{x}_{t|t}$ for time t:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K[t] \cdot v[t] \tag{4.20}$$

where $K[t]$ is the *Kalman Gain* weighting factor, which takes the measurement residual from the measurement space back into the internal state space (where $x[t]$ lies) and weighs it in a way that reflects how much we 'believe' this new information over our existing predicted state $\hat{x}_{t|t-1}$. The larger the Kalman Gain, the more we believe the new information than our existing state predictor, so we perturb the state prediction by a large amount in the 'direction' of $v[t]$.

The difference between the true state $x[t]$ and the state estimate $\hat{x}_{t|t}$ is the state estimate error $\widetilde{x}_{t|t}$:

$$\widetilde{x}_{t|t} = x[t] - \hat{x}_{t|t} \tag{4.21}$$

with a related covariance matrix called the state estimate error covariance

$$P_{t|t} := \mathbf{E}\{\widetilde{x}_{t|t} \otimes \widetilde{x}'_{t|t}\} = (I - K[t] \cdot H[t]) \cdot P_{t|t-1} \cdot (I - K[t] \cdot H[t])' + K[t] \cdot J[t] \cdot K[t]' \tag{4.22}$$

where I is the identity matrix of the appropriate dimensions (n_x by n_x in this case), and this form of $P_{t|t}$ is the 'Joseph form' of the state estimate error covariance - useful in the implementation of the filter for preserving the positive-definite and symmetric structure of the covariance matrices.

The optimal Kalman Gain $K[t]$ is derived from the state prediction and measurement residual covariance matrices such that the mean square state estimate error $\mathbf{E}\{| \widetilde{x}_{t|t} |^2\}$ is minimised - this is also equivalent to minimising the trace of $P_{t|t}$ - which consequently minimises the measurement residual $v[t]$ and hence optimizes our model's explanatory power. After some algebraic manipulation, the optimal Kalman Gain is shown to be:

$$K[t] = P_{t|t-1} \cdot H[t]' \cdot B_t^{-1} \tag{4.23}$$

The state prediction error covariance $P_{t|t-1}$ and state estimate error covariance $P_{t|t}$ equations (4.16 and 4.22) show that they are not explicitly dependent of the state and measurement vectors and hence can be calculated separately, however they do enter into the process of updating these vectors via the calculation of the optimal Kalman Gain.

4.1.4.3 Covariance Matching

As we don't know the state and measurement noise covariance matrices $Q[t]$ and $J[t]$ *a priori*, we need to estimate these parameters during the implementation of the Kalman Filter. We follow the procedure in [Lamper 2002] and [Gupta 2005a], who both use the technique of covariance matching as described in [Maybeck 1979], suitable for on-line implementation, by taking time-averages of quantities calculated by the filter. Solving equation 4.19 for $J[t]$, we can see the following relationship:

$$\widehat{J}[t] = \frac{1}{W-1} \sum_{i=t-W}^{t-1} (v[i] \otimes v[i]' - H[i] \cdot P_{i|i-1} \cdot H[i]') \tag{4.24}$$

In practice, we use the better conditioned but more computationally intensive version of this estimate based on the updated state estimates advocated by

Lamper:

$$\widehat{J}[t] = \frac{1}{W-1}\sum_{i=t-W}^{t-1}\{(z[i]-H[i]\cdot x_{i|i})\otimes(z[i]-H[i]\cdot x_{i|i})'+H[i]\cdot P_{i|i}\cdot H[i]'\} \quad (4.25)$$

To calculate $Q[t]$, we follow Gupta's method using equations 4.16 and 4.19 to get:

$$B_t = H[t]\cdot(F[t-1]\cdot P_{t-1|t-1}\cdot F[t-1]' + Q[t])\cdot H[t]' + J[t] \quad (4.26)$$

which can be re-arranged to form:

$$H[t]\cdot Q[t]\cdot H[t]' = B_t - H[t]\cdot F[t-1]\cdot P_{t-1|t-1}\cdot F[t-1]'\cdot H[t]' - J[t] \quad (4.27)$$

and then solved for $Q[t]$ to become:

$$Q[t] = (H[t]'H[t])^{+}H[t]'(B_t - H[t]F[t-1]P_{t-1|t-1}F[t-1]'H[t]' - J[t])H[t](H[t]'H[t])^{+} \quad (4.28)$$

where A^{+} refers to the (Moore-Penrose) pseudo-inverse of matrix A.

Gupta supports the use of setting the off-diagonal terms of $Q[t]$ to zero, in order to keep $Q[t]$ diagonal, as it has the benefits of making the covariance matrices 'more positive-definite' and is motivated by the fact that having cross-correlation in the process noise is generally rare. We concur and also follow Lamper's technique of using a lower bound on the diagonal elements of $Q[t]$ of 0.0001 to prevent the filter from being unduly optimistic regarding how certain it is of finding the correct values of $x_R[t]$. Using pessimistic (larger) covariance matrices are suitable in this case as we acknowledge the limitations of our ability to capture reality with our model - although the agent-based games may capture a degree of market behaviour, there is still likely to be significant model mismatch.

In addition, we also follow the method of Gupta, who claims that taking a sample average of past measurements of the residual process to calculate B_t:

$$\hat{B}_t = \frac{1}{W-1}\sum_{i=t-W}^{t-1} v[i]\otimes v[i]' \quad (4.29)$$

could be considered more accurate than using equation 4.19.

4.1.4.4 Applying the Kalman Filter to the MG framework

The components of the strategy weightings vector x can be considered to be probabilities - the probability of a specific strategy being held by the population of agents. Indeed Gupta et al's early work [Gupta 2005a, Gupta et al. 2005a, Gupta, Hauser 2007] considers this set-up and the need to constrain the state vector such that all the components of the state vector are positive and sum to 1:

$$\sum_R x_R = 1 \tag{4.30}$$

However, the computational efforts necessary to maintain the equality constraint and the difficulty of maintaining the positive definiteness of the relevant covariance matrices lead us to relax this constraint as discussed in [Gupta et al. 2006]. Indeed, if one were to constrain the weightings to sum to 1, the model would then need to undergo a regression to estimate the sensitivity of the price increments to the predicted excess demands anyway, thus scaling the weights to potentially be greater than 1. Therefore, we relax this equality constraint acknowledging that it is the relative weights of the strategies that are important in determining future price changes. There is however, still the inequality constraints that the components of $x[t]$ must be greater than zero at all times

$$x_R[t] > 0 \tag{4.31}$$

There are a few options of how to ensure this - the simplest being by setting the weighting to zero if the Kalman Filter estimates a component to fall below zero, or keeping the component unchanged from the time-step before - an idea that would be particularly valid if strategy weightings change only slowly over time. Another method employed by Gupta et al. [Gupta et al. 2007a, Gupta 2007] is to combine the Kalman Filter dynamic and measurement equations (eq.4.12 and eq.4.13) with the inequality constraint eq. 4.31 (Gupta also allows the possibility of including equality constraints such as eq. 4.30) into an augmented pseudo-measurement space.

For our use and the way it appears in the literature [Lamper 2002] [Gupta et al. 2005a], the transition matrix $F[t]$ is taken to be the identity matrix, reflecting that there is no deterministic way in which the strategy weightings evolve. As Gupta claims "it would be a tough modeling problem to choose another matrix... however, if desired we could incorporate a more complex transition matrix into the model, even one that is dependent on previous outcomes." [Gupta et al. 2007a] Because the deterministic part of the dynamic equation is treated as the identity matrix in our model, this means that we rely on the evolution of the state vector $x[t]$ being made through the stochastic influences $\varepsilon[t]$ and $\omega[t]$. The Kalman Filter in effect implements a 'least-squares' approach to minimising the measurement residual $z[t]$, and it is the comparison of the real price increments to the predicted price increments that forces the state vector $x[t]$ to change over time, to minimise this measurement residual. In effect, it is the measurement residual that drives the estimation, rather than any deterministic force. Owing to the stochastic nature of the price increments and measurement error, the state vector components (the estimate of the weightings of the strategy space) also evolve stochastically.

The question remains that, if we model the strategy space with just a vector of size 2^{m+1} - the number of strategies available (in the reduced strategy space) - is this able to adapt fast enough to capture the sudden strategy weight changes caused by some strategies overtaking others in their strategy score, thus being picked by agents to make predictions at the next time-step? We explore this question by developing what we call the Simple Minority Game Forecasting Framework (Simple MGFF) and applying it to both Minority Game simulated data and financial data in Section 4.3 on page 111. We also explicitly capture the sudden changes due to strategies becoming inactive by using the Heaviside step function in the measurement matrix $H[t]$, yet the first simple version of the Kalman Filtering methodology cannot model the potentially sudden change in the agents' optimal 'dominant' strategies. A procedure to remedy this problem would be to explicitly use information about the strategy scores in the transition matrix $F[t]$, so it becomes conditionally deterministic rather than static. However we choose to attack this problem directly by expanding the state vector space and creating a model to estimate the weights of the

$\binom{2^{m+1}}{2} + 2^{m+1} = 2^{2m+1} + 2^m$ distinguishable strategy pairs in the Reduced Strategy Space (strategy pairs are unordered so the pair $\{R, R'\}$ is equivalent to $\{R', R\}$ and repetition is allowed ($R' = R$), so we can model the number of possible strategy pairs using a binomial coefficient with an additional 2^{m+1} to account for the repetitions). We call this version the Expanded Minority Game Forecasting Framework (or Expanded MGFF) and explore its ability to forecast Minority Game simulated data and financial data in Section 4.4 on page 125. In this way our model is then not reliant on the Kalman Filter's ability to track the changing usage of single strategies, but the weight of each strategy pair - the proportion of agents that hold each strategy pair, which should, if we believe the Minority Game setup to be representative of the real world, change at a slower rate. Gupta discusses estimating strategy pairs in [Gupta et al. 2007a], but in the context of using the Full Strategy Space that necessitates the random choosing of a subset of strategy pairs to make the problem computationally tractable and averaging over several runs, however this would be more time-consuming and has the potential to create biases in the strategy choice if some of the chosen strategies are similar [1]. We therefore believe that our method of estimating the weights of all strategy pairs in the Reduced Strategy Space, which only contains uncorrelated and anti-correlated strategies, to be logically preferable.

As originally described by Lamper [Lamper 2002], the measurement vector $z[t]$ is our time-series input data at time t; the first vector component is the price increment of the asset, and we will also extend it to investigate the effects of using a second component, specifically the trading volume of the asset. The measurement matrix $H[t]$ in our model varies with time and is determined by the state $\mu[t]$ and strategy scores $S_R[t]$. In most implementations of the Kalman Filter, and certainly for the original intentions of the use of the Kalman Filter to track bodies in motion, the measurement matrix $H[t]$ would be deterministic, known from the start or even static. In our case however, although we don't know $H[t]$ until after time $t-1$ when all the strategy scores are updated with information from $t-1$, we can treat $H[t]$ as constant for the next iteration of the Kalman Filter.

[1] close to each other on a de Bruijn graph

In our model, where the price increment is the first component of the measurement vector, the first row of the measurement matrix H is the transpose of the strategy space column that refers to the current state $\mu[t]$ multiplied by the active strategy indicator function $\mathcal{H}\{S_R[t]-r\}$ (eq. 2.11 on page 59):

$$H_{1R}[t] = a_R^{\mu[t]} \cdot \mathcal{H}\{S_R[t]-r\} \tag{4.32}$$

which is the action of strategy R if active.

Thus, the expected price increment can be written as a weighted sum over all strategies:

$$z_1[t] = C(t) \cdot (H[t] \cdot x_R[t] + \mathbf{E}(\omega_1[t])) = C(t) \cdot \left(\sum_R a_R^{\mu[t]} \cdot \mathcal{H}\{S_R[t]-r\} \cdot x_R[t] \right) \tag{4.33}$$

where $C(t)$ is a constant of proportionality that may depend on time. Thus the expected direction of the next price movement can be written as:

$$Predicted\ direction = sgn\,(z_1[t]) = sgn \left(\sum_R a_R^{\mu[t]} \cdot \mathcal{H}\{S_R[t]-r\} \cdot x_R[t] \right) \tag{4.34}$$

For the second data source (the trading volume of the asset), the analogous equations are:

$$H_{2R}[t] = \mathcal{H}\{S_R[t]-r\} \tag{4.35}$$

and the expected trading volume measurement:

$$z_2[t] = \sum_R \mathcal{H}\{S_R[t]-r\} \cdot x_R[t] \tag{4.36}$$

It is hoped that by employing a second data source, this will increase the speed of convergence of the strategy weight estimates and lead to improved results, as discussed (though not explicitly implemented) in [Lamper 2002].

4.1.5 Cleaning data from repeated price quotes

As the Minority Game and consequently the forecasting framework are primarily binary models - the strategies at each time-step predict either a price increase or decrease - the question of what to do with stationary price quotes i.e. prices that are repeated must be addressed. One way would be to allow the model to receive such stationary price updates, and then use an unbiased method such as a random coin toss to decide whether the price 'moved up or down' as in [Jefferies et al. 2001]. However this would introduce unnecessary noise into the calculation of the state and the strategies scores, and considering that the forecasting methodology focuses on determining the direction of future asset price changes, the simplest method is simply to remove such repeated prices, with an understanding that the next-step prediction we are making may not be the next price quote, but in fact the next time the price does actually move.

This also has an impact on the way we have to handle data, such as trading volume, when used in addition to the asset price time series. With tick data, there are often repeated price quotes corresponding to trades that have no effect on the price, presumably as they are small enough not to exert a significant influence on the price. The volume for these trades at the same price are summed up before being fed into the Kalman Filter equations along with a single realisation of the repeated price.

4.2 Statistical Significance of Results

A simple way we can assess the statistical significance of our hit-rate measurements is to use the composite hypothesis testing procedure outlined in [Bar-Shalom et al. 2001]. We compare two hypotheses, the null hypothesis $\mathbf{H_0}$ that our measurements are due to random chance, with hypothesis $\mathbf{H_1}$ that the MGFF has some ability in predicting the direction of price movements of a specific financial market.

The null hypothesis $\mathbf{H_0}$ can be imagined as a system making predictions with no information, i.e. we make a sequence of predictions about the direction of the next price change by performing a sequence of Bernoulli trials with the probability of success $\mathbf{P}(success) = \mathbf{P}(failure) = 0.5$. In effect, we are tossing an unbiased coin to make predictions with no information whatsoever, and we perform the coin toss at each time-step, over the sample window M. The 'null hit-rate', the proportion of successful predictions achieved by the repeated coin-toss experiment purely by random chance is then described by the Binomial distribution, whose first two central moments are:

$$\mathbf{E}\left(\#\, successful\, predictions\right) = \mathbf{P}\left(success\right) \cdot M = \frac{M}{2} \tag{4.37}$$

and

$$variance\ \left(\#\, successful\, predictions\right) = \mathbf{P}(success) \cdot \mathbf{P}(failure) \cdot M = \frac{M}{4} \tag{4.38}$$

We can obtain the expected null hit-rate κ_0 and its standard deviation σ_0 by dividing the mean and standard deviation of the number of successful predictions by the sample window M, thus we would expect the null hit-rate to be distributed with mean value $\kappa_0 = 0.5$, and standard deviation

$$\sigma_0 = \frac{1}{2} M^{-\frac{1}{2}} \tag{4.39}$$

thus we formulate our problem as a decision between the null hypothesis $\mathbf{H}_0$, written as

$$\mathbf{H}_0 : \quad \kappa_0 = 0.5 \tag{4.40}$$

and the alternative hypothesis $\mathbf{H}_1$:

$$\mathbf{H}_1 : \quad \kappa_1 > 0.5 \tag{4.41}$$

For us to reject the null hypothesis $\mathbf{H_0}$ that our hit-rate has been obtained by chance, and so support the belief $\mathbf{H_1}$ that there is predictive power in our method, we require our hit-rate measurement to be above the bounds of the null hypothesis. For us to reject the null hypothesis with 95% confidence, we require the difference between the observed hit-rate κ and the expected null hit-rate κ_0 to fulfil the following condition:

$$\kappa - \kappa_0 = \kappa - 0.5 > 1.64 \cdot \widehat{\sigma} \tag{4.42}$$

that is, the null hit-rate is far enough away from our observed hit-rate, given the standard deviation of the observed hit-rate $\widehat{\sigma}$. The factor 1.64 comes from calculating the one-sided alternative hypothesis at the 95% confidence level, assuming the observed hit-rate to be normally distributed around the 'true' hit-rate.

If we were to assume that the standard deviation of the null hit-rate σ_0 and the observed standard deviation $\widehat{\sigma}$ were the same, we would have the following equation:

$$\kappa > 1.64 \cdot \sigma_0 + 0.5 = 0.82 \cdot M^{-\frac{1}{2}} + 0.5 \tag{4.43}$$

The standard deviation of the null hit-rate over the moving time window $M = 200$ is

$$\sigma_0 = \frac{1}{2} M^{\frac{-1}{2}} = \frac{1}{2 \cdot \sqrt{200}} = 0.035$$

Therefore, if we are using a hit-rate with a moving window size of 200 time-steps for example, in order to accept the observed hit-rate as significant, we require a hit-rate of at least:

$$\kappa > \frac{0.82}{\sqrt{200}} + 0.5 = 0.558 \tag{4.44}$$

We generally measure the statistics of the time-series such as hit-rate and correlation of prediction over a moving window of $M = 200$ time-steps for the Simple and Expanded MGFFs in this Chapter, see equations 4.6 and 4.7 on page 93 (note that in Chapter 5 we set $M = 100$). Therefore the MGFF generates a time-series of hit-rates (see Figure 4.3 on page 115), each calculated over a moving window of size M. From this we calculate a sample average to get the average hit-rate over the whole period of the financial data sample.

As an alternative to using the Binomial estimate of the standard deviation as a proxy for the standard deviation of the measured hit-rate κ, we could directly estimate the standard deviation of the hit-rate time-series, by taking non-overlapping samples of the hit-rate across the data sample. If our sample

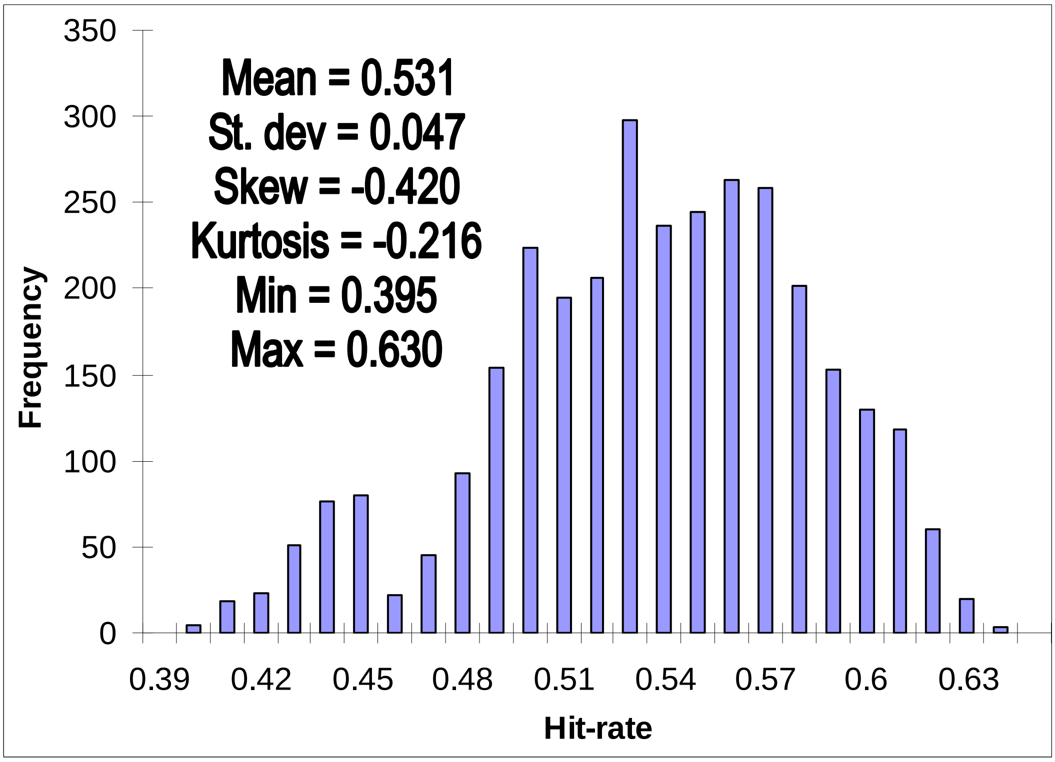


Figure 4.1: Histogram of overlapping hit-rates from Simple MGFF of daily NASDAQ data (3200 moving average hit-rates of window $M = 200$, equivalent to 16 independent hit-rate measurements)

data is of a size several times greater than the window M used to calculate the statistics, we can record several non-overlapping hit-rate estimates and calculate the mean and standard deviation of the distribution. However, in order to be able to form a reasonable amount of non-overlapping hit-rate estimates, we would require long data sets. This would require a longer computing time and also enough financial data - sometimes not always possible when considering the closing price and volume of a recently introduced asset class that has only a few years of daily financial data available, such as credit default swaps. One hit-rate estimate would be equivalent to almost 1 year worth of daily data, and so require us to have many years worth of data. In order to increase the data set, we could always decrease the moving window M to 100 time-steps for example, thus doubling the number of hit-rate estimates to use in the determination of the sample standard deviation.

Another possibility is to simply replace $\widehat{\sigma}$, which is a measure of the variability of the hit-rate, with a proportion of the spread of the maximum to minimum observed hit-rate of the (overlapping) moving average. This rectifies the difficulties of computing a large enough data set to estimate $\widehat{\sigma}$. Figure 4.1 on the previous page shows the histogram of 3200 hit-rates from a moving sample window of size $M = 200$ obtained by applying the MGFF to daily NASDAQ data (closing prices and trading volume) from 1995 to 2008. We can see that the half-spread of hit-rates is

$$\frac{\kappa_{max} - \kappa_{min}}{2} = \frac{0.235}{2} = 0.1175 \tag{4.45}$$

, which would correspond more closely with the 99% confidence interval of a normally distributed variate, rather than the 95% confidence interval, because of our inclusion of all the tail of the distribution. The 99% confidence interval of a 2-sided normally distributed random variable corresponds to random variable taking a value within the range $\pm 2.58 \cdot \sigma$ about the mean value. Therefore we can estimate $\widehat{\sigma}$ with the equation:

$$\widehat{\sigma} = \frac{\kappa_{max} - \kappa_{min}}{2 \cdot 2.58} = \frac{\kappa_{max} - \kappa_{min}}{5.152} = 0.046 \tag{4.46}$$

giving us very close agreement with the standard deviation of 0.047 measured from the 3200 overlapping hit-rates, and good agreement with the 16 independent (non-overlapping) hit-rate measures of 0.052. We therefore choose to use the maximum-minimum range technique ('min-max' technique) as the favoured method to estimate the standard deviation of κ.

4.3 Simple Minority Game Forecasting Framework

We begin with our simple Kalman estimation method - the Simple Minority Game Forecasting Framework (Simple MGFF) - unique in that it estimates single strategy weights as opposed to strategy pairs, and relies on the Kalman Filter to track changes in the dominance of strategies. This is different to the work of Gupta and Gou who focus on strategy pair sets, however it means that if we assume the data we use is generated by a system analogous to the Minority Game, we have a model mismatch as an integral part of the Minority Game is the adaptability of agents from their multiple strategies. To compensate for this, we use two data sources - the asset price and the trading volume - to allow for faster convergence of the Kalman Filter.

In an attempt to simplify the computation and increase the speed of the process, we focus on using the much smaller Reduced Strategy Space of size 2^{m+1}, whereas [Gupta et al. 2005a] and [Gou 2006c] focus on the Full Strategy Space version (size 2^{2^m}) of the Minority Game (see 2.3.1.2 on page 45 for more details). In simulations of the Minority Game, there was not found to be any significant changes in the general behaviour of the output, and so in our 'reverse engineering' forecasting methodology we also expect it not to make much difference in the forecasting power. By using the RSS, we are removing the strategies that are correlated and only estimating the weights of strategies that are statistically orthogonal to each other. The Kalman Filter adapts the weights of these statistically orthogonal strategies over time, and can thus *in theory* recreate the appearance of the FSS correlated strategies.

The ideas underlying the Simple MGFF, where the Kalman Filter is used to track the changing popularity of single strategy weights rather the weights of strategy pairs, may have some justification. If investors were to behave exactly as a Minority Game agent, they would keep current scores of the success of their investment strategy, and if one of their strategies did slightly better than another, they would switch immediately. If the situation reversed at the next time-step, they would switch immediately back to the original strategy. In reality, we would expect real market investors to exhibit some form of inertia in changing their strategies (and their opinions) - indeed we could model this inertia by requiring agents to switch investment strategies only after a certain trial time-period has passed to see if the new strategy maintains its greater success rate, or alternatively requiring an upcoming strategy to accumulate a certain number of 'buffer-points' before switching from the incumbent. Or we can just assume that this slow adaptability of investor opinion can be successfully captured by the Kalman Filter, hence the Simple MGFF may be able to produce reasonably good results, despite its simplicity. This is not an absurd assumption, as one would expect the rate of incoming new information (such as in the form of price and volume data) about an asset to be of higher frequency than the rate at which investors change their investment decisions or strategies. This would be especially true when we train our model to tick data, although depending on the type of investor being modelled, it would also be correct over timescales of several orders of magnitude greater if we consider typical institutional investors such as pension fund and asset managers, who may take years to change their investment strategies.

The benefits of only estimating single strategy weights instead of strategy pairs is that is reduces the size of the estimation problem significantly; there are only 2^{m+1} single strategies, whereas there are $\binom{2^{m+1}}{2} + 2^{m+1} = 2^{2m+1} + 2^m$ possible strategy pairs. As the number of estimation parameters - the strategy (pair) weights - increases exponentially as we increase the agent memory size m, the computational feasibility of running this estimation creates a practical upper bound on the value of m. As the number of estimation parameters of our Simple MGFF is not so sensitive to m, the value of m can be relatively higher than the existing method, allowing us to search a wider parameter space of m. In addition, the reduction in time taken to execute the program means that we can search more widely or finely the space of the other parameters $\{T$, c,

$\tau\}$ to find the optimal parameter set for the input data. The larger the value of m, the larger the number of time-steps we use to define the state of the system μ, so that more complicated patterns in the data can be handled by the system. This should logically increases the likelihood of being able to spot regularities in the market, of the sort observed with the S&P 500 Composite Index in Section 3.2 on page 72.

Whether the inherent model mismatch will effect the ability of the Simple MGFF to replicate the promising results of Lamper and Gupta will be assessed, comparing our results with the output from the Expanded Minority Game Forecasting Framework (Expanded MGFF) in Section 4.4 on page 125, which expands the number of estimation parameters, to return to estimating strategy pairs rather than single strategies. We note however, that our Expanded version is also different to the literature, in that it is derived using the Reduced Strategy Space of 2^{m+1} strategies rather than the Full Strategy Space of 2^{2^m} to further simplify the model to make it computationally more amenable.

4.3.1 MG data

As in [Lamper 2002], we begin by running the MGFF using a data set generated by a 'black box' Minority Game, i.e. one in which we assume not to know *a priori* the parameters, to see if we can recover the correct parameter set. The parameter set of our black box model is chosen to be $\{m = 3, N = 101, s = 2, T = 100, c = 0.53, \tau = 1\}$ for consistency with Lamper's choice as it is seen to produce dynamics similar to financial markets (the stylized facts), and whilst we estimate the strategy weights using the Kalman Filter, we also need to scan through the parameter space of m, T, c and τ to search for the optimal objective function value, in our case the best (average) hit-rate over the data set.

Figure 4.2 on the next page demonstrates the sensitivity of the MGFF to the initial values of the covariance matrices Q, J, $P_{t|t}$ of the Kalman Filter in the simple state vector estimate. For the matched parameter set $\{m = 3,$

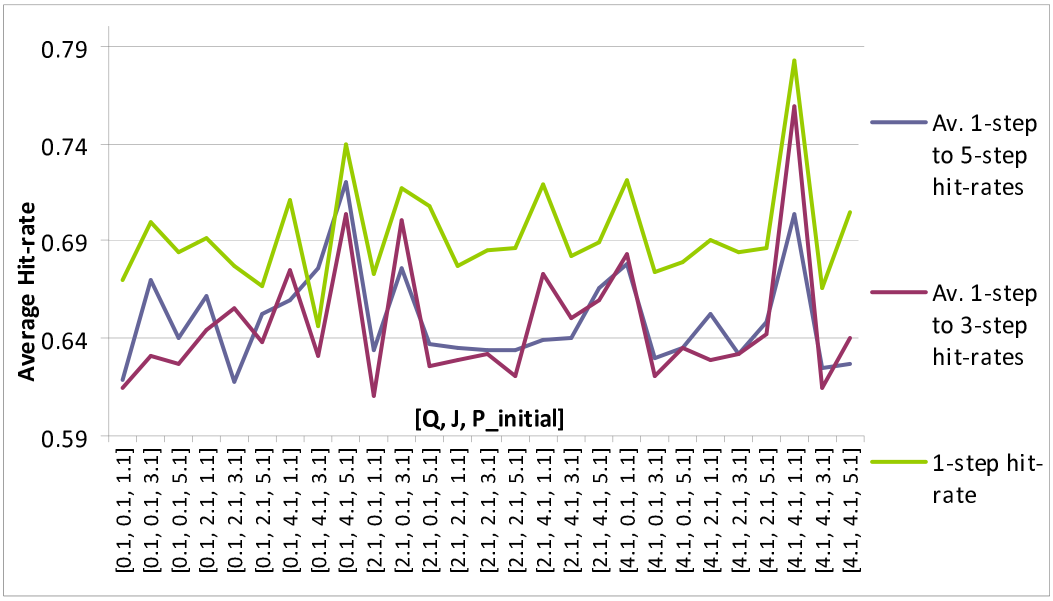


Figure 4.2: Variation of the average hit-rate to changes in the initial covariance parameters ($Q_{initial}$, $J_{initial}$, $P_{initial}$) of matched MGFF (where the MGFF is applied to data created from a MG simulation using the same parameter set $\{m, T, c, \tau\}$ as the MGFF)

T=100, $c = 0.53$, $\tau = 1\}$ we find the optimal hit-rate occurs with the initial values $Q_{initial} = 4.1$, $J_{initial} = 4.1$ and $P_{initial} = P_{0|0} = 1.1$. We do not see this sensitivity for the the expanded state vector estimation in Section 4.4 on page 125, so this is either due to the model mismatch in the Simple MGFF version, or the sparsity of the Expanded MGFF version's expanded covariance matrices (which have 2^{4m+2} elements, as opposed to the Simple MGFF setup that has 2^{2m+2} elements) where the errors in one strategy estimate are less likely to influence the other strategy estimates and the system as a whole.

Figure 4.3 on the next page shows the convergence over time of the 1-step hit-rate for the matched parameter set and optimal initial covariance matrices, taken over a sample window of $M = 200$ time-steps, towards the optimal estimates of the strategy weight parameters. Note that, despite the actual matched parameter set $\{m, T, c, \tau\}$ being used, this simplified version of the MGFF can only achieve an optimal average of 79% (see Figure 4.3) and for these parameters at most a peak hit-rate of around 89% (see Figure 4.2). This is lower than the rate achieved with the expanded MGFF in Section 4.4 on

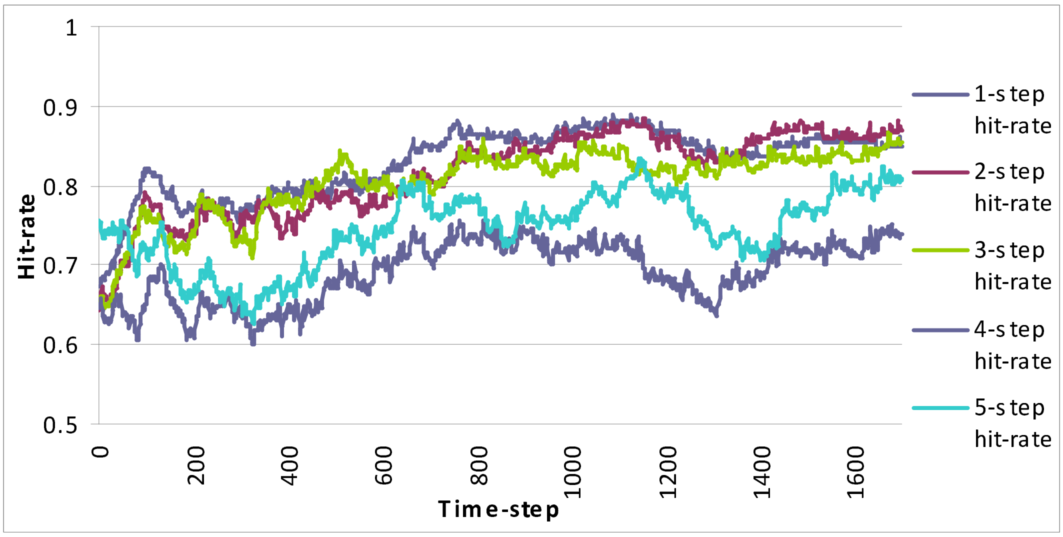


Figure 4.3: Time-series of e-step hit-rates (where $e = 1, 2, \ldots, 5$) for matched parameters (where the MGFF parameter set $\{m = 3, T = 100, c = 0.53, \tau = 1\}$ is the same as in the MGFF) and with optimal initial Kalman Filter parameters $Q_{initial} = 4.1$, $J_{initial} = 4.1$, $P_{initial} = 1.1$) .

page 125 where we realise a 100% hit-rate for the periodic (no population ties) phase of the Minority Game. We can therefore conclude that this reduction in forecasting ability is down to the known model mismatch between the Simple MGFF, which is not able to capture the behaviour of the adaptive Minority Game agents that can switch investment decisions between their strategy pairs.

We show the range of hit-rates produced across the space of different time-horizons T and confidence threshold parameters c respectively in Figure 4.4 on the next page. Through scanning across the parameter space, we find the optimal hit-rate approaches the value of 1 at $\{m = 3, T = 100, c = 0.54, \tau = 1\}$ with the only incorrect parameter c which is actually 0.53 in the 'black-box' Minority Game used to generate the time-series. The matched parameters generated an average hit-rate of around 79%. The increased hit-rate with a larger confidence threshold may be down to the removal of the 'weaker' strategies - those with lower points from the active strategy set used to specify the measurement matrix H of the Kalman Filter (see equations 4.32 and 4.35 on page 105). In the actual 'black-box' MG system, agents with only weak strategies in their set will be more prone to switching between strategies,

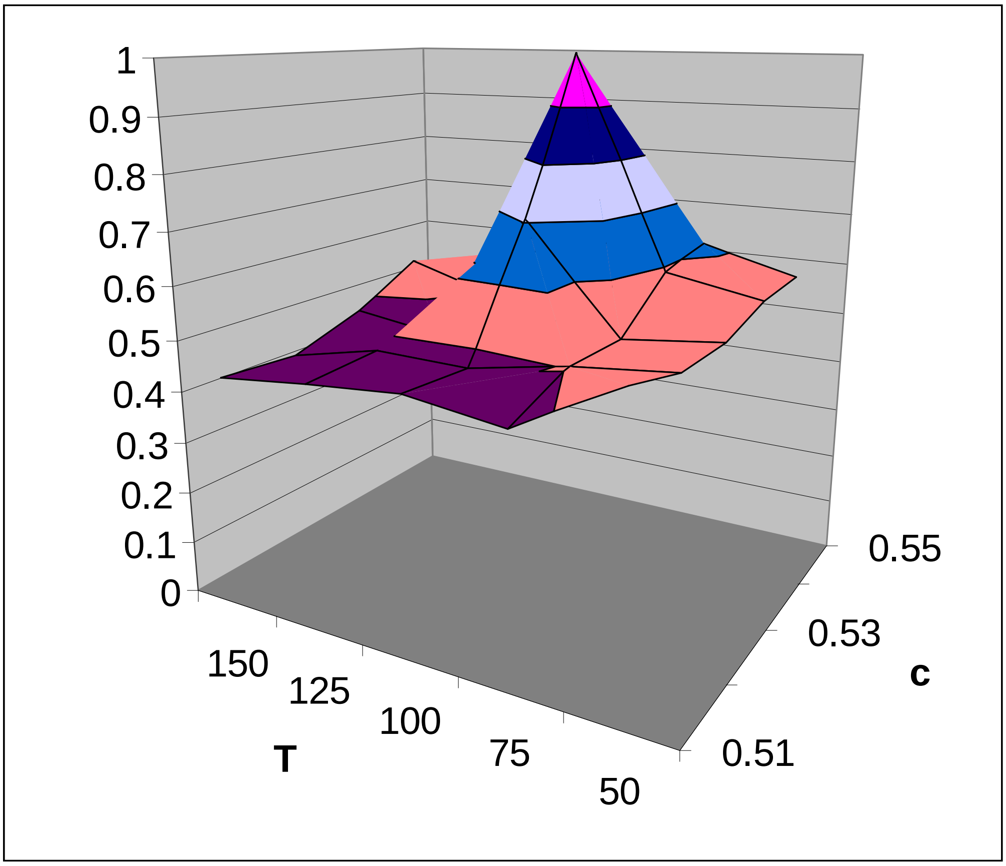


Figure 4.4: The variation of the average hit-rate of MGFF with the T and c parameters (fixing $m = 3$, $\tau = 100$ and the initial covariance matrices $Q = 4.1$, $J = 4.1$, $P_{initial} = 1.1$). This hit-rate surface shows a clear peak at the correct time horizon parameter $T = 100$ and (nearly) the correct confidence threshold $c = 0.54$ (the actual value is $c = 0.53$). This demonstrates that we can recover the original parameters $\{T, c\}$ used in the simulation, and we similarly use this method to find the correct values of m and τ.

or additionally between participating in the market or not. However the Simple MGFF, which relies on the Kalman Filter to track such changes in strategy usage, is unlikely to be able to keep up with changes in agent choice or activity. Assuming such weak strategy switching has an aggregate effect of zero on the aggregate demand (not unreasonable as over large periods of time there will be on average as many switches from buy-to-sell as from sell-to-buy, and more fundamentally, there will be as many states of the world μ specifying a strategy to buy as there are to sell), by increasing c we reduce this 'switching noise' from the estimation by removing the weaker strategies from the measurement matrix H. Thus, the optimal average hit-rate is observed using a higher than expected parameter $c = 0.54$.

4.3.2 Financial data

With the possibility that the Minority Game may be able to capture, to a certain degree, the complex dynamics of a financial market, we strive to test this hypothesis by applying the MGFF to real market data. In choosing our data sets, we look for a variety of time periods, data frequencies and asset classes, and in particular, choose data sets that have not yet been examined from a Minority Game forecasting perspective in the literature.

Using Bloomberg as a resource for much of our data, we pick a range of asset classes and periods of time when both asset price and volume data are available so that we can use two streams of simultaneous data in the Kalman Filter to estimate the strategy weights. We generally focus on daily data, but also have some tick data from the S&P500 Futures Index. The data sample size ('Data size') is calculated after the removal of repeated price quotes (zero price increment) as discussed in Section 4.1.5 on page 106 and excludes the initial period of time-horizon length T necessary to obtain the correct strategy scores. We list the data sources that we use and the asset class in Table 4.3 on page 119, where the LME is the London Metal Exchange, NYMEX is the New York Mercantile Exchange, and TSE is the Tokyo Stock Exchange.

The average values of the e-step ahead hit-rate κ_e calculated using the optimal parameter set and their standard deviations $\widehat{\sigma_e}$ are presented in Table 4.4a on page 120. Whilst we search the parameter space of $\{m, T, c, \tau\}$ for the best one-step ahead hit-rate κ_1, we also present the two-step κ_2 and three-step ahead hit-rate κ_3 calculated using equation 4.10 on page 94, by running the MGFF forward assuming the predicted steps are realised, as discussed in Section 4.1.2 on page 92. We also present the relative hit-rate κ_e^{rel} for predictions e steps ahead in Table 4.4b on page 120, calculated using equations 4.10 and 4.8 on page 94. For assets where we have repeated the analysis over different time-periods or using different frequencies, we label them with a lower case letter.

From Table 4.4a on page 120, we can see that as e increases, the hit-rates κ_e decrease across most assets. This is analogous to weather forecasting - the further into the future one looks, the less likely our predictions will be correct. This is not only down to the stochastic nature of financial markets and the weather, but also reflects the uncertainty in our model, the accumulation of measurement errors and the effect of factors we have neglected to include in our model. Indeed, any predictability several time-steps into the future would enable traders to enter deals to exploit this, thus arbitraging away these opportunities. In Figure 4.5 on page 121, we illustrate this behaviour, omitting error bars for clarity. The cases such as S&Pa which appear to have increasing predictability are within the bounds of statistical uncertainty, suggesting that such increases are not statistically significant.

Table 4.4b on page 120 shows that all the 1-step, 2-step and most of the 3-step ahead relative hit-rates have values greater than 1, suggesting that our model, on average, achieves a better prediction success rate than the naive prediction. The size of the standard deviations however, are such that only S&P(c)'s 1 and 2-step hit-rates, and Copper Futures 1-step hit-rates are statistically significant.

We note that the optimal parameter sets, the values $\{m, T, c, \tau\}$ that produce the best 1-step ahead hit-rate κ_1 for each asset, are distributed fairly evenly across the possible strategy space that we searched. This may indicate that across the range of markets we investigate and within the limit of the parameter space we search, there is no particular bias towards any parameter values. This may reflect that different classes have market participants with different attributes, captured by the parameters such as the length of time

Asset Class	Asset Label	Bloomberg Ticker	Unit	Exchange
Commodity Futures	West Texas Intermediate Crude Oil Futures	CL1 <Cmdty>	US$/barrel	NYMEX
Commodity Futures	Copper Futures	LMCADY <Cmdty>	US$/tonne	LME
Commodity Futures	Nickel Futures	LMNIDY <Cmdty>	US$/tonne	LME
Commodity Index	Topix Real Estate Index	TPREAL <Index>	Index points	TSE
Stock Index	NASDAQ Composite Index	CCMP <Index>	Index points	NASDAQ
Equity Futures	S&P500 Futures Index	SP <Index>	Index points	CME

Table 4.3: List of assets analyzed using the Simple MGFF

Asset Label	Time Period	Size	κ_1	$\widehat{\sigma}_1$	κ_2	$\widehat{\sigma}_2$	κ_3	$\widehat{\sigma}_3$	$\{m, T, c, \tau\}$
Real Estate(a)	30/07/93 - 28/08/08	3600	**0.519**	0.034	**0.511**	0.032	**0.505**	0.036	{3, 50, 0.52, 0.9}
Nickel Fut	19/05/06 - 22/08/08	500	**0.574**	0.012	**0.557**	0.014	**0.528**	0.022	{3, 100, 0.51, 0.9}
Copper Fut	19/05/06 - 22/08/08	500	**0.581**	0.030	**0.524**	0.040	**0.490**	0.044	{4, 50, 0.52, 0.9}
Oil Fut(a)	10/05/00 - 22/04/08	1900	**0.542**	0.041	**0.527**	0.037	**0.537**	0.068	{4, 100, 0.52, 0.9}
S&P Fut(a)	30/06/08 - 01/07/08	1400	**0.542**	0.044	**0.577**	0.046	**0.522**	0.040	{2, 100, 0.52, 1}
S&P Fut(b)	02/07/08 - 03/07/08	1000	**0.557**	0.024	**0.516**	0.022	**0.469**	0.035	{2, 50, 0.52, 1}
S&P Fut(c)	30/07/08 - 31/07/08	300	**0.651**	0.015	**0.627**	0.033	**0.477**	0.023	{4, 100, 0.52, 0.9}
NASDAQ(a)	03/01/95 - 19/08/08	3400	**0.531**	0.047	**0.505**	0.052	**0.519**	0.051	{2, 50, 0.52, 0.9}
NASDAQ(b)	17/05/02 - 19/08/08	1500	**0.541**	0.034	**0.514**	0.028	**0.502**	0.022	{5, 50, 0.52, 1}

(a) Average hit-rates achieved by applying Simple MGFF to Financial Data

Asset Label	Freq.	Time Period	Size	κ_1^{rel}	$\widehat{\sigma^{rel}}_1$	κ_2^{rel}	$\widehat{\sigma^{rel}}_2$	κ_3^{rel}	$\widehat{\sigma^{rel}}_3$	$\{m, T, c, \tau\}$
Real Estate(a)	daily	30/07/93 - 28/08/08	3600	**1.01**	0.10	**1.00**	0.10	**0.99**	0.11	{3, 50, 0.52, 0.9}
Nickel Fut	daily	19/05/06 - 22/08/08	500	**1.12**	0.08	**1.08**	0.05	**1.03**	0.04	{3, 100, 0.51, 0.9}
Copper Fut	daily	19/05/06 - 22/08/08	500	**1.28**	0.09	**1.16**	0.13	**1.09**	0.14	{4, 50, 0.52, 0.9}
Oil Fut(a)	daily	10/05/00 - 22/04/08	1900	**1.14**	0.09	**1.14**	0.13	**1.18**	0.23	{4, 100, 0.52, 0.9}
S&P Fut(a)	tick	30/06/08 - 01/07/08	1400	**1.12**	0.10	**1.21**	0.12	**1.11**	0.13	{2, 100, 0.52, 1}
S&P Fut(b)	tick	02/07/08 - 03/07/08	1000	**1.14**	0.09	**1.04**	0.07	**0.95**	0.08	{2, 50, 0.52, 1}
S&P Fut(c)	tick	30/07/08 - 31/07/08	300	**1.18**	0.03	**1.16**	0.08	**1.05**	0.69	{4, 100, 0.52, 0.9}
NASDAQ(a)	daily	03/01/95 - 19/08/08	3400	**1.05**	0.15	**1.01**	0.17	**1.04**	0.17	{2, 50, 0.52, 0.9}
NASDAQ(b)	daily	17/05/02 - 19/08/08	1500	**1.13**	0.10	**1.08**	0.10	**1.05**	0.08	{5, 50, 0.52, 1}

(b) Data frequency, average relative hit-rates by applying Simple MGFF to Financial Data

Table 4.4: Average hit-rates achieved using the Simple Minority Game Forecasting Framework applied to Financial Data. κ_e is the observed average hit-rate e-steps ahead, $\widehat{\sigma}_e$ is the observed standard error, κ^{rel} is the observed average relative hit-rate and $\widehat{\sigma^{rel}}$ is the observed standard error of the relative hit-rate. $\{m, T, c, \tau\}$ is the optimal parameter set used in the Simple MGFF that produced the highest hit-rate values.

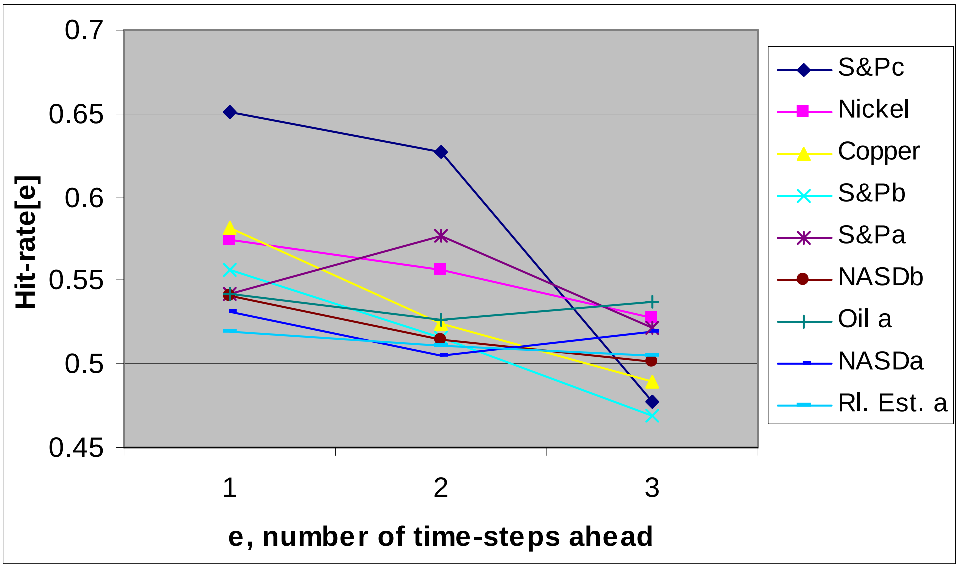


Figure 4.5: Variation of the average hit-rate with the number of time-steps e into the future that we forecast

they measure their strategy successes, or the confidence they require in order to trade. However, in order to establish any reliable conclusions, we would wish to perform a more exhaustive examination using longer data samples, more asset classes, a wider parameter space search and a finer search - with smaller step size increments across the parameter values.

We present examples of the resultant time-series of hit-rates produced by our Simple MGFF, with the Copper Futures (LMCADY) in Figure 4.6, the S&P500 Futures(b) in Figure 4.7, the WTI Crude Oil Futures(a) in Figure 4.8 on page 123 and the Topix Real Estate Index(a) in Figure 4.9 on page 124.

Although the data set used to generate Figure 4.6 on the next page is only over a fairly short time-period of 400 time-steps, it manages to maintain an average hit-rate of 58.1% and is statistically significant. As we extend the period of time over which we run the MGFF, the hit-rate tends to decrease, as can be seen in Table 4.4a on the preceding page between S&P500 Futures(a,b,c) data of rows 5, 6 and 7, and the NASDAQ(a,b) data of rows 8 and 9. In general, the shorter the sample size, the less reliability can be placed on the results. It is possible that, assuming we have a model mismatch between our simple MGFF and the real world, any (well-behaved) unexplained time-dependent variables

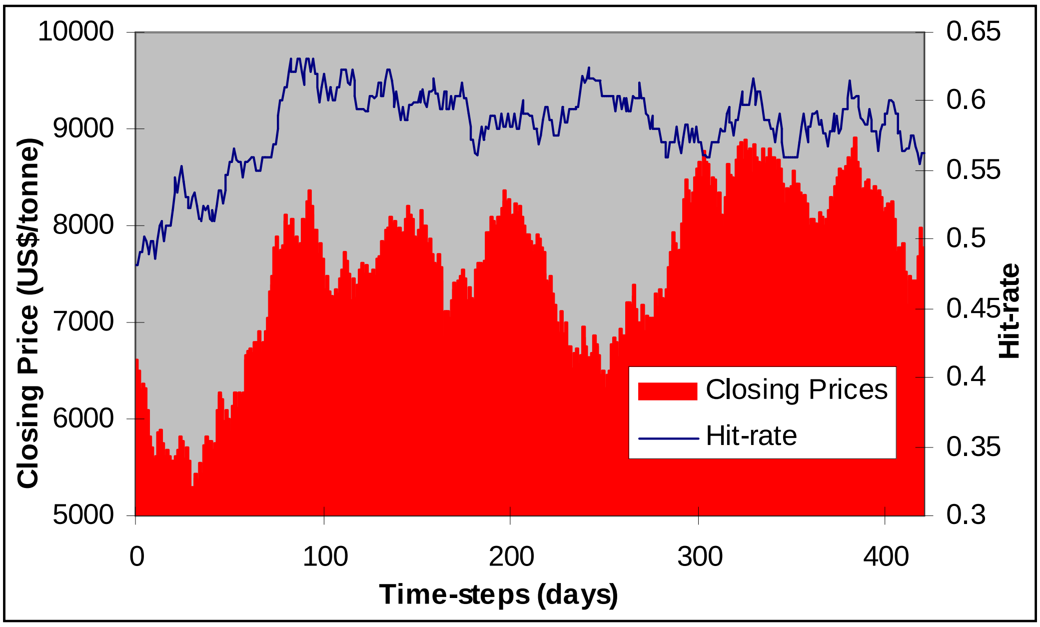


Figure 4.6: Closing prices and hit-rate time-series using Simple MGFF on Copper Futures (LMCADY) daily data

that we have not captured in our model will have changed by a smaller amount, causing less unexplained error in our calibration, thus an improvement in the hit-rate measures.

Figure 4.9 is over the longest time period, and experience the lowest optimal average hit-rate (52%) amongst the examples. However there are certain periods of the hit-rate time-series when the MGFF manages to achieve a fairly high moving hit-rates of around 57%-58%, perhaps corresponding to the 'pockets of predictability mentioned in [Johnson et al. 2001, Andersen, Sornette 2005, Gupta et al. 2005a].

We also present the correlation between the hit-rate and the price data ($\rho_{\kappa-price}$), and the correlation between the hit-rate and trading volume data ($\rho_{\kappa-volume}$) for the figures we have presented in Table 4.5 on page 124. Although the correlation between the hit-rate and price data for Copper Futures appears to be fairly high at 0.55, there doesn't appear to be any systematic bias in correlations when compared with the other asset classes, all of whom have the opposite sign. This can also be compared with the correlations for the Expanded MGFF in Table 4.8 on page 131 for other asset classes.

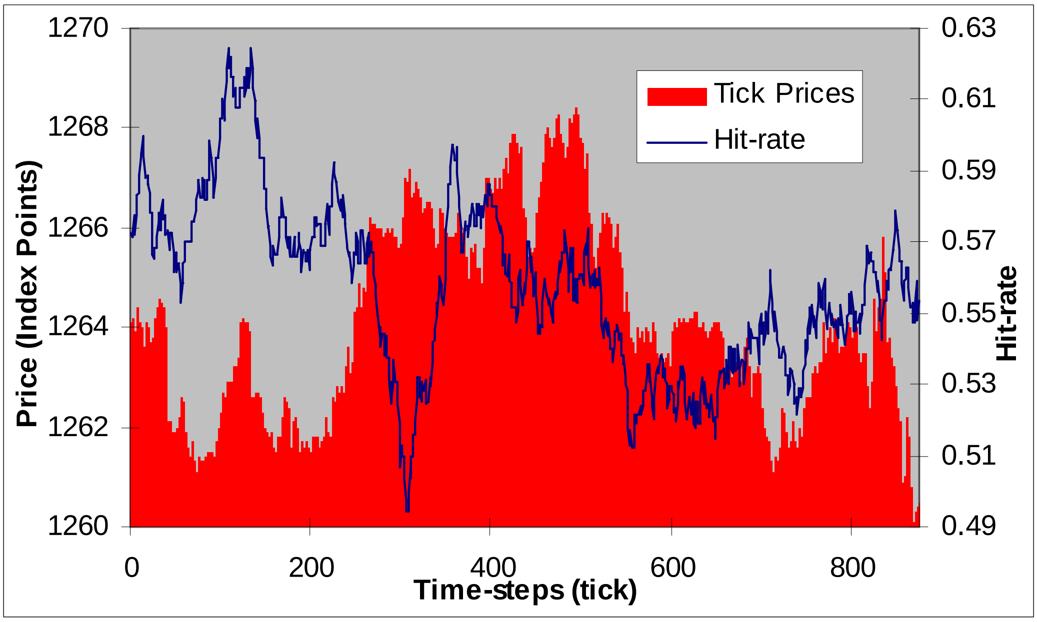


Figure 4.7: Prices and hit-rate time-series using Simple MGFF on S&P500(b) futures tick data

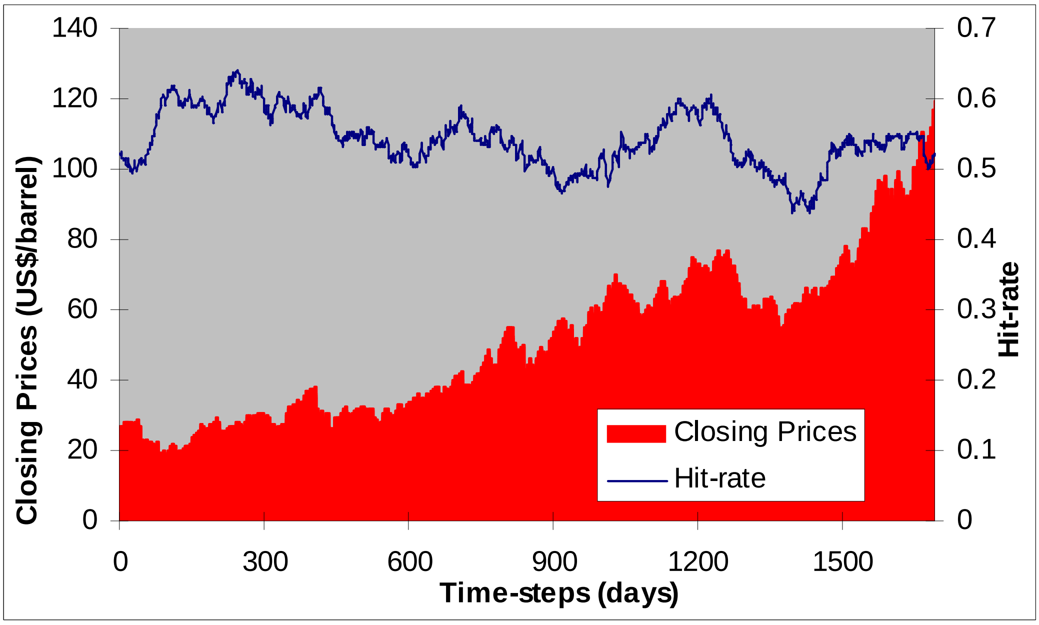


Figure 4.8: Closing prices and hit-rate time-series using Simple MGFF on Crude Oil Futures(a) (CL1) daily data

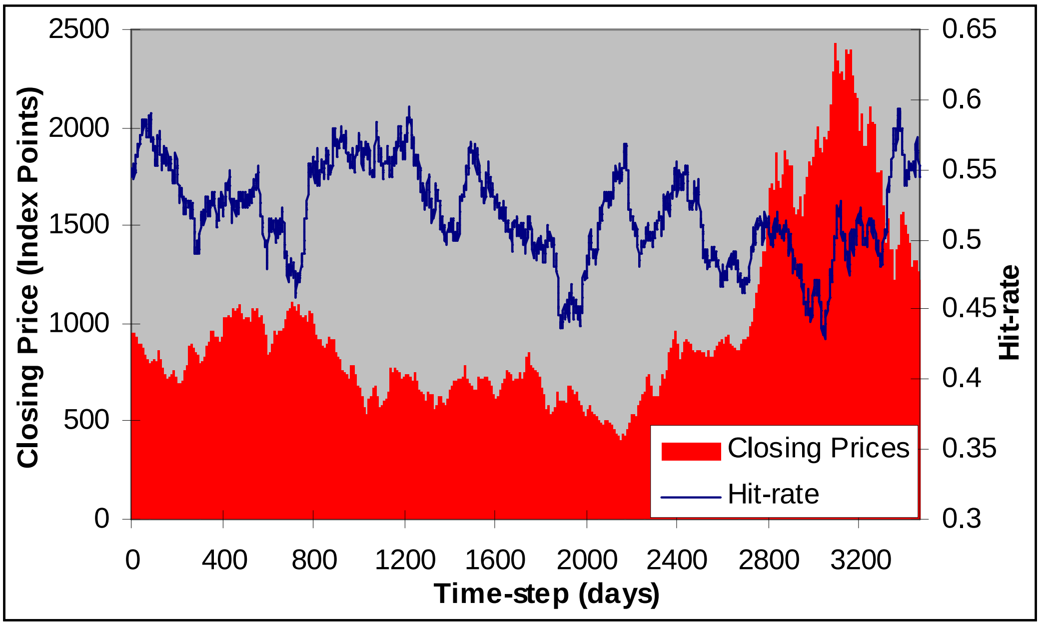


Figure 4.9: Closing prices and hit-rate time-series using Simple MGFF on daily Japanese Real Estate(a) data

Asset	$\rho_{\kappa-price}$	$\rho_{\kappa-volume}$
Copper Futures	0.55	0.08
S&P500 Futures(b)	-0.17	-0.16
Crude Oil Futures(a)	-0.34	-0.32
Japanese Real Estate(a)	-0.29	-0.09

Table 4.5: Correlations between the hit-rate, and the price ($\rho_{\kappa-price}$) and volume ($\rho_{\kappa-volume}$) data for certain assets using the Simple MGFF.

4.4 Expanded Minority Game Forecasting Framework - strategy pair estimation

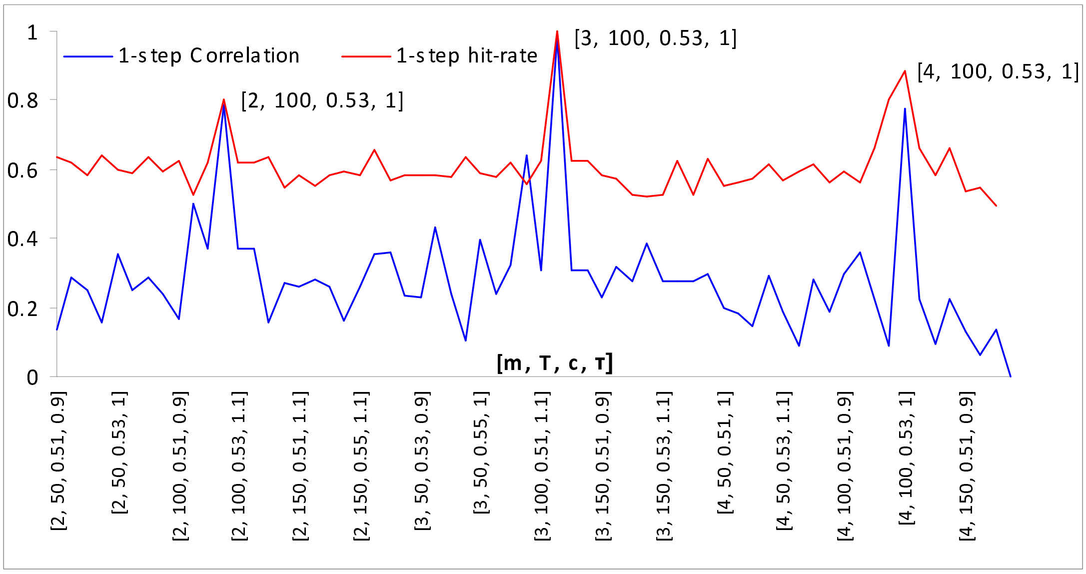
We now implement the expanded formulation of the MGFF, such that we estimate strategy pair weightings, as considered in [Gupta et al. 2005a], to see if it improves the hit-rates over our simple MGFF of Section 4.3 on page 111.

4.4.1 Applications to MG data

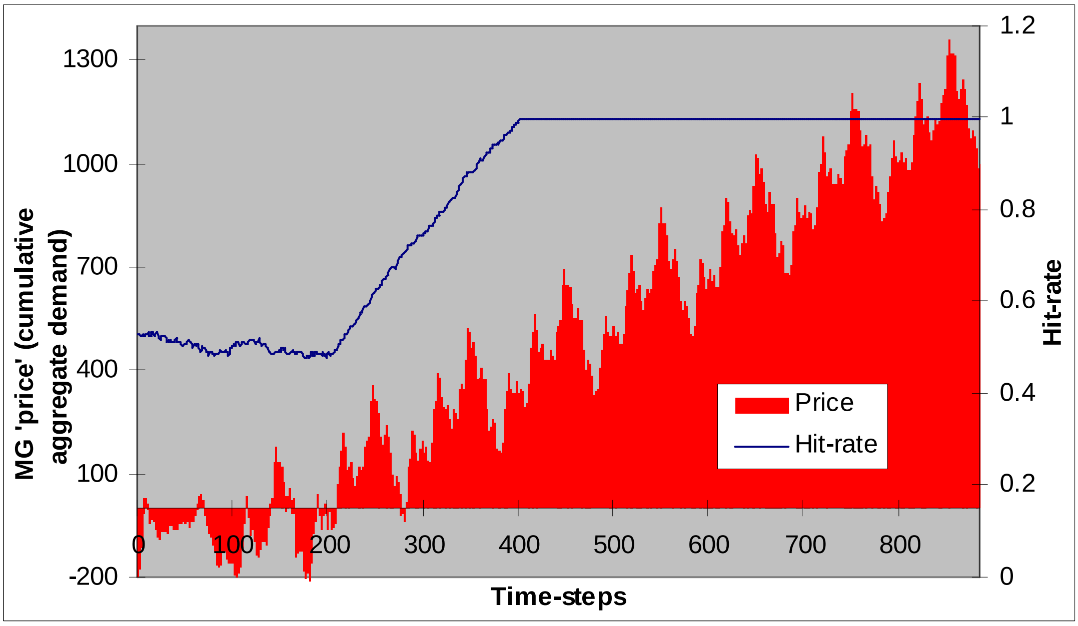
We would expect the Expanded MGFF to capture the behaviour of data created by a 'black-box' more successfully than the Simple MGFF as we remove a known model mismatch. Indeed, we confirm this as can be seen in Figure 4.10 on the following page, where both the hit-rate and the correlation of predictions with outcomes[2], taken over the sample window of $M = 200$, approach the value of 1 at the correct parameter set of $\{m = 3,\ T = 100,\ c = 0.53,\ \tau = 1\}$. Also note in Figure 4.10a on the following page that there are two other parameter sets that also produce high hit-rates and correlations to the perfectly matched set. These other sets have the same values of T, c and τ, but differing values of m, suggesting the relative insensitivity of the Expanded MGFF towards the value of the parameter m. This is encouraging as m is the parameter that determines the complexity of the MGFF, with an increase in m creating an exponential increase in the number of strategies / strategy pairs to be estimated, and consequently an exponential increase in the computational time needed to run. Therefore, there is some justification, not just in terms of practicality, to restrict our parameter searches to lower values of m. In fact, we generally restrict our values of m in the extended model such that $2 \leq m \leq 4$.

In Section 3.3 on page 74, we discussed the observation that our simulations of the finite-time horizon version of the Grand Canonical Minority Game exhibited two different phases, the initial one dominated by ties in the population of agents choosing to buy or sell at each time-step, resolved by a random coin-toss. The latter phase, having no population ties and thus no obvious

[2] as given by equation 4.6 on page 93



(a) The variation of the average hit-rate over the parameter set $\{m, T, c, \tau\}$ when applying the Expanded MGFF to MG simulation data $\{m = 3, T = 100, c = 0.53, \tau = 1\}$ in the periodic phase. We can clearly see peaks in the hit-rate when T, c and τ are correctly matched. The fact that the average hit-rate has significant peaks suggests the relative insensitivity of the MGFF procedure to the m parameter. This demonstrates that our MGFF procedure can indeed be used to recover the appropriate parameter values from a time-series that is well described by a Minority Game.



(b) Convergence of the optimal hit-rate time-series using matched parameters $\{m = 3, T = 100, c = 0.53, \tau = 1\}$ on MG data in the periodic phase.

Figure 4.10: The Expanded MGFF tracking data from a 'black-box' MG in the periodic phase

coin-tosses, experiences periodicity - a repeated stochastic pattern of length T.

Whilst the periodic phase of the Minority Game data can easily be tracked by the Expanded MGFF (see Figure 4.10 on the preceding page) with the best hit-rate tending to 1, the stochastic phase has a much lower hit-rate (see Figure 4.11 on the next page). In fact, this also explains the reason why the hit-rate of Figure 4.10b starts at around 0.5, before the system settles into the periodic, predictable phase at around 200 time-steps when the hit-rate begins to rise towards 1.

Whereas one might think that some resemblance of the features of the Minority Game could be captured by the Expanded MGFF, the unpredictability introduced by the 'coin-toss' resolutions dominates the aggregate demand time-series, ensuring that any sort of estimation of strategy weights, and consequently also reproducing the correct parameter set is much more difficult. As seen in Figure 4.11, the optimal hit-rate achieved over the initial coin-toss phase was with parameter set $\{m = 2,\ T = 70,\ c = 0.53,\ \tau = 1\}$, with the matched parameter set achieving a hit-rate of only 0.48. Note that the hit-rates appear to be higher for smaller m. This may be because the population ties in the stochastic period of the Minority Game that generated the time-series occurred approximately every two time-steps, as can be seen in Figure 3.8b on page 75, and so the shorter memory m of the MGFF results in a smaller number of time-steps being used to determine the state of the world μ and hence the strategies' predictions at the next time step. In effect, the smaller memory value m means the MGFF has a 'greater resolution' to make predictions in between the coin-tosses and so the predictions are less corrupted by the random coin-toss 'noise'.

4.4.2 Applications to Financial data

Similarly to Section 4.3.2 on page 117, we now test the Expanded MGFF to analyse financial market data. In particular, we attempt to assess the validity of earlier claims regarding the use of the Minority Game to forecast

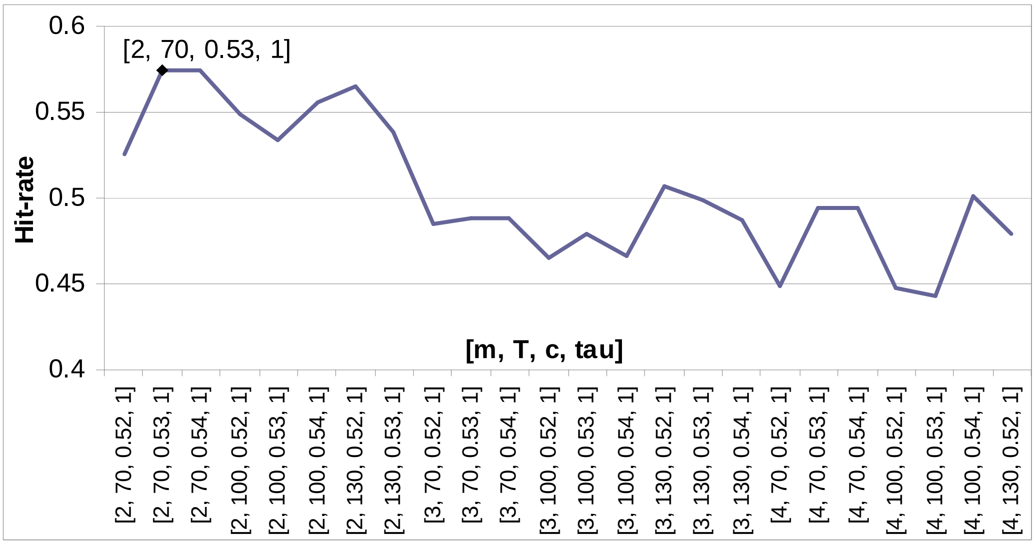


Figure 4.11: The variation of the average hit-rate over the parameter set $\{m, T, c, \tau\}$ when applying the Expanded MGFF to MG simulation data $\{m = 3, T = 100, c = 0.53, \tau = 1\}$ in the stochastic 'population tie' phase. We have difficulty distinguishing unique optimal parameter sets (compare with Figure 4.10a), especially over T, c and τ, whilst the procedure generally favours small m.

Asset Class	Asset Label	Bloomberg Ticker	Unit	Exchange
Equity	BP	BP/LN <Equity>	GBP£	LSE
Currency	USD-JPY	USDJPY <Crncy>	Yen per US$	OTC
Currency	USD-EUR	USDEUR <Crncy>	Eur per US$	OTC

Table 4.6: Additional assets tested with the Expanded MGFF

hourly data in the USD-JPY currency market but using daily data instead. Note that owing to the much greater computing time required to process the expanded strategy pair space, we reduce the sample size of the data due to time constraints and also only analyse $m \leq 4$ (whilst the Simple MGFF used $m \leq 5$). We also extend our analysis to encompass three new assets, described in Table 4.6. Note that for the tests using Foreign Exchange data (USD-EUR and USD-JPY), no trading volume is available as currencies are traded globally over the counter (OTC), and not on any central exchange. The hit-rate results of applying the Expanded MGFF to financial data are presented in Table 4.7a on the next page, with the relative hit-rate results displayed in Figure 4.7b.

Asset Label	Time Period	Size	κ_1	$\widehat{\sigma}_1$	κ_2	$\widehat{\sigma}_2$	κ_3	$\widehat{\sigma}_3$	$\{m, T, c, \tau\}$
Real Estate(b)	11/03/02 - 27/09/05	800	**0.518**	0.031	**0.517**	0.0545	**0.527**	0.047	{4, 50, 0.51, 0.9}
Nickel Fut	19/05/06 - 22/08/08	500	**0.563**	0.040	**0.579**	0.063	**0.548**	0.065	{2, 50, 0.51, 0.9}
Copper Fut	19/05/06 - 22/08/08	500	**0.579**	0.041	**0.542**	0.017	**0.512**	0.059	{4, 50, 0.52, 0.9}
Oil Fut(b)	22/10/92 - 28/02/95	500	**0.589**	0.017	**0.559**	0.020	**0.508**	0.042	{2, 50, 0.51, 0.9}
Oil Fut(c)	01/05/06 - 26/08/08	500	**0.540**	0.024	**0.522**	0.041	**0.517**	0.029	{3, 50, 0.51, 1}
S&P Fut(d)	01/08/06 - 31/07/08	400	**0.559**	0.018	**0.518**	0.011	**0.469**	0.027	{4, 100, 0.51, 1}
S&P Fut(e)	28/07/08 - 29/07/08	300	**0.552**	0.028	**0.564**	0.016	**0.483**	0.043	{2, 50, 0.51, 1}
S&P Fut(c)	30/07/08 - 31/07/08	300	**0.611**	0.011					{4, 100, 0.52, 0.9}
NASDAQ(c)	22/08/06 - 19/08/08	450	**0.564**	0.028					{2, 50, 0.52, 0.9}
BP	13/07/89 - 06/06/97	1800	**0.529**	0.046	**0.517**	0.038	**0.527**	0.041	{3, 50, 0.52, 1}
USD-EUR	26/04/05 - 26/08/08	800	**0.537**	0.037	**0.564**	0.059	**0.575**	0.057	{2, 100, 0.52, 1}
USD-JPY	29/11/91 - 17/03/94	550	**0.539**	0.024	**0.437**	0.032	**0.469**	0.029	{2, 50, 0.52, 1}

(a) Average hit-rates κ_e and observed standard error $\widehat{\sigma_e}$ for forecasts e-steps ahead.

Asset Label	Freq.	Time Period	Size	κ_1^{rel}	$\widehat{\sigma^{rel}}_1$	κ_2^{rel}	$\widehat{\sigma^{rel}}_2$	κ_3^{rel}	$\widehat{\sigma^{rel}}_3$	$\{m, T, c, \tau\}$
Real Estate(b)	daily	11/03/02 - 27/09/05	800	**0.98**	0.14	**0.96**	0.13	**0.91**	0.09	{4, 50, 0.51, 0.9}
Nickel Fut	daily	19/05/06 - 22/08/08	500	**1.10**	0.12	**1.12**	0.12	**1.04**	0.09	{2, 50, 0.51, 0.9}
Copper Fut	daily	19/05/06 - 22/08/08	500	**1.27**	0.10	**1.23**	0.06	**1.16**	0.18	{4, 50, 0.52, 0.9}
Oil Fut(b)	daily	22/10/92 - 28/02/95	500	**1.17**	0.06	**1.17**	0.09	**1.09**	0.20	{2, 50, 0.51, 0.9}
Oil Fut(c)	daily	01/05/06 - 26/08/08	500	**1.14**	0.07	**1.05**	0.09	**1.06**	0.06	{3, 50, 0.51, 1}
S&P Fut(d)	daily	01/08/06 - 31/07/08	400	**1.19**	0.08	**1.14**	0.07	**1.07**	0.06	{4, 100, 0.51, 1}
S&P Fut(e)	tick	28/07/08 - 29/07/08	300	**1.00**	0.05	**1.03**	0.04	**0.90**	0.10	{2, 50, 0.51, 1}
S&P Fut(c)	tick	30/07/08 - 31/07/08	300	**1.10**	0.02					{4, 100, 0.52, 0.9}
NASDAQ(c)	daily	22/08/06 - 19/08/08	450	**1.22**	0.08					{2, 50, 0.52, 0.9}
BP	daily	13/07/89 - 06/06/97	1800	**0.99**	0.12	**0.97**	0.11	**0.97**	0.12	{3, 50, 0.52, 1}
USD-EUR	daily	26/04/05 - 26/08/08	800	**1.17**	0.07	**1.23**	0.15	**1.25**	0.16	{2, 100, 0.52, 1}
USD-JPY	daily	29/11/91 - 17/03/94	550	**1.13**	0.06	**0.91**	0.11	**0.99**	0.09	{2, 50, 0.52, 1}

(b) Data frequency, average relative hit-rates κ^{rel} and observed standard error $\widehat{\sigma^{rel}}$ of the relative hit-rate for forecasts e-steps ahead.

Table 4.7: Applying the Expanded MGFF applied to financial data. $\{m, T, c, \tau\}$ is the optimal parameter

We observe similar behaviour to the Simple MGFF in that as we make predictions further into the future (increasing e) the hit-rate measure of predictability declines. We also manage to achieve a comparable hit-rate to that found in [Johnson et al. 2001] for the USD-JPY market during the 1990s, although we use daily data and they observe a hit-rate of 54% for hourly data. The relative hit-rates we observe are greater than 1, except for the Real Estate(b) Index and the BP stock analysis, and USD-JPY analysis for $e > 1$ and S&Pb for $e > 2$. As in the Simple MGFF, we observe a fairly even distribution of optimal parameter sets across the parameter space $\{m, T, c, \tau\}$.

We present examples of the resultant time-series of hit-rates produced by our Expanded MGFF, with the NASDAQ(c) index hit-rates shown in Figure 4.12 on the following page. We show the Nickel Futures (LMNIDY) in Figure 4.13 on page 132, featuring both the hit-rate and relative hit-rate time-series, which although having different values, generally display a similar evolution over time. As seen in equations 4.8 and 4.9 on page 94, a relative hit-rate greater than 1 corresponds to a better forecasting technique than the naive prediction of assuming that the next price increment will have the same sign as the previous price increment. As can be seen, the relative hit-rate does remain above 1 most of the time, suggesting that our method is, at the very least, better than the naive prediction method. In Figure 4.14 on page 132, the hit-rates for 1-step, 2-step and 3-step predictions are displayed for the daily Crude Oil Futures(c).

In Table 4.8 on the following page, we display the correlations between the hit-rate and the input data of price and volume for the Expanded MGFF (compare this with Table 4.5 on page 124 for the Simple MGFF). Although the correlation between hit-rate and price for the NASDAQ(c) index appears very high at 0.81, the analogous measure for the Nickel Futures price has the opposite sign at -0.64 and the Crude Oil Futures(b) measure is only 0.20. This suggests no systematic correlation in the methodology. The data is taken over a relatively short period of 250 data points, and it is possible that the time-series during this period is such that certain states of the system μ occur far more often than others - there is a (short-term) bias in the states of the system. If the input data is such that the Kalman Filter estimates a larger proportion

Asset	$\rho_{\kappa-price}$	$\rho_{\kappa-volume}$
NASDAQ(c)	0.81	0.03
Nickel Futures	-0.64	0.16
Oil Futures(c)	0.20	-0.08

Table 4.8: Correlations between the hit-rate time-series, and the price ($\rho_{\kappa-price}$) and volume ($\rho_{\kappa-volume}$) data for Expanded MGFF

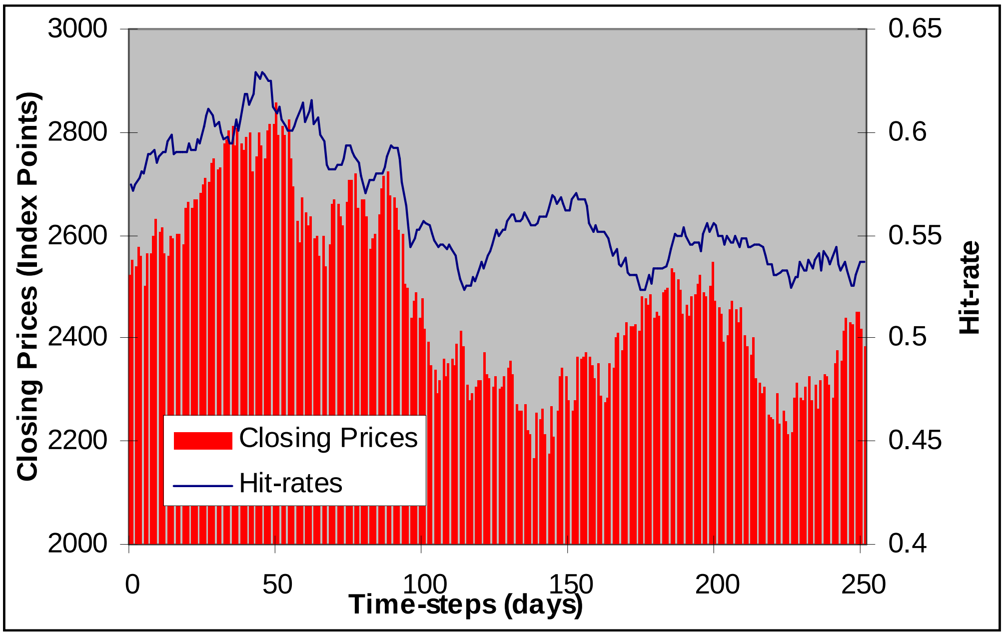


Figure 4.12: Closing prices and hit-rate time-series of Expanded MGFF on daily NASDAQ index data (NASDAQ c)

of weight to strategies that forecast price increases more frequently when the system is in the recurring subset of states μ of the short-term biased system (and assuming that the strategy pair weightings do not change so much over the data sample), there will be a bias in the predictions that will create such a positive correlation between the hit-rate measure and the price data. We also note that the Oil Futures(a) in the Simple MGFF and Oil Futures (c) in the Expanded MGFF are of similar magnitudes but opposing signs, despite the Expanded MGFF price data being a subset of the data used in the Simple MGFF. This suggests no systematic bias in the overall MGFF method.

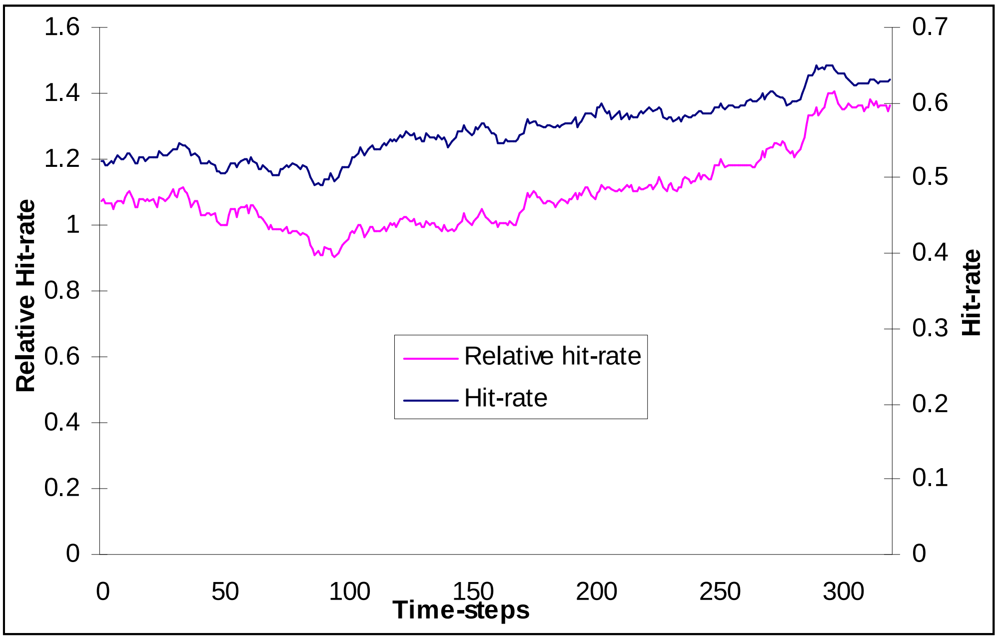


Figure 4.13: Hit-rate and relative Hit-rate time-series using the Expanded MGFF on daily Nickel Futures (LMNIDY) data

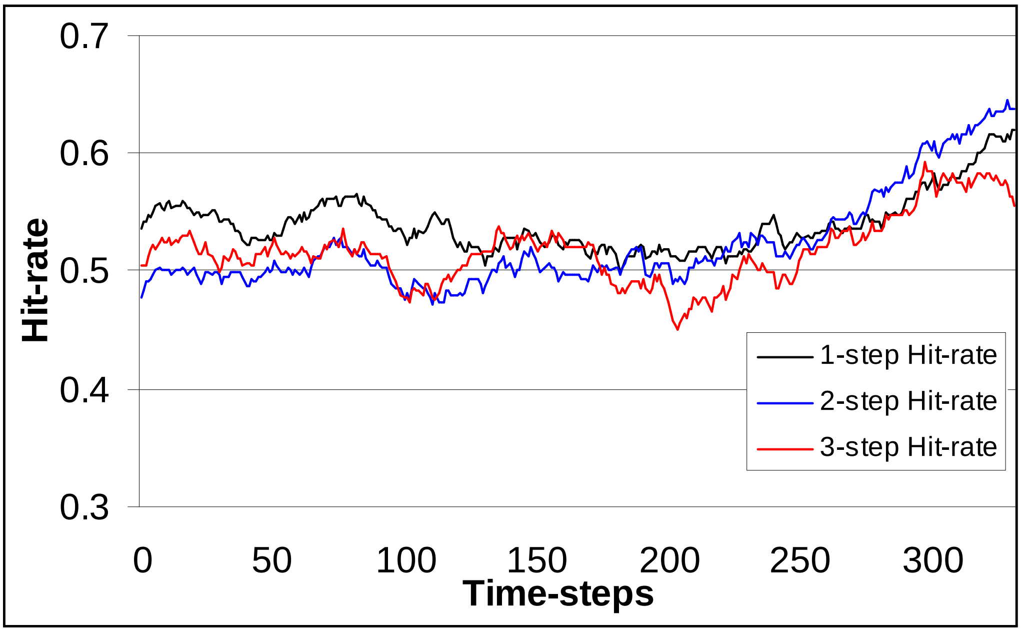


Figure 4.14: Hit-rate time-series for $e = 1, 2, 3$ time-steps ahead using Expanded MGFF on daily Crude Oil Futures(c) (CL1) data

4.5 Discussions

4.5.1 Testing for significance

As discussed in Section 4.2 on page 106, we need to test these results to determine how likely they are the result of real predictive power, or to evaluate if they are more likely to be the result of chance. Here, we use the composite hypothesis testing to choose between the null hypothesis that it is likely that we have achieved these results by chance:

$$\mathbf{H}_0 : \ \kappa_0 = 0.5 \tag{4.47}$$

and the alternative hypothesis that it is likely that there is predictive power in our method:

$$\mathbf{H}_1 : \ \kappa_e > 0.5 \tag{4.48}$$

where we use the one-sided composite hypothesis test, in that we focus our attention on testing for the optimal hit-rate - which if $\mathbf{H}_1$ is true, should be greater than the null hit-rate. We decide between the two hypotheses by using equation 4.42 on page 108 to measure the number of standard deviations that our null hit-rate of $\kappa_0 = 0.5$ is from our observed hit rate κ_e (for e-step ahead predictions). For us to reject the null-hypothesis with 95% confidence, we require:

$$\frac{\kappa_e - \kappa_0}{\widehat{\sigma}} > 1.64 \tag{4.49}$$

In the case where this criterion is not fulfilled, we also use the p-value ('p-val') to evaluate the significance of our result in 4.9b on page 135. The p-value is defined as the probability of achieving the observed hit-rates or higher, given that the null hypothesis $\mathbf{H}_0$ is true:

$$p\ value = 1 - \Phi\left(\frac{\kappa_e - \kappa_0}{\sigma_0}\right) \tag{4.50}$$

where Φ(x) is the standard normal (cumulative) distribution for the random variable X = x. Note that the separation of the observed hit-rate and the null hit-rate are measured in relation to the null standard deviation σ_0 in this calculation, whilst equation 4.49 uses the observed standard error $\widehat{\sigma}$. If the p-value is greater than 0.05, then we can say that the observed hit-rates are not statistically significant at the 5% level. Ideally, we wish for equation 4.49 to be satisfied and for the p-value to be less than 0.05 in order for us to have confidence in the significance of our results.

We present these measures in Tables 4.9a and 4.9b on the next page, with the results that fulfil equation 4.49 highlighted in bold lettering, in which case we also embolden the p-value if less than 0.05.

4.5.2 Observations

4.5.2.1 Comparing the Simple and Expanded MGFF

Table 4.9 on the following page shows that for both the Simple MGFF and the Expanded MGFF, around half of all the data samples assessed appear to have significant hit-rates - far enough away from the null hit-rate to reject the null hypothesis. Therefore, in these particular cases, the hypothesis that the MGFF has some predictive power is supported for data size less than or equal to 500. This would suggest that, in the cases analyzed, the Simple MGFF can compete with the Expanded MGFF, and provides the benefit that it is quicker, owing to the smaller number of strategy weights to be estimated and we can thus search over a larger parameter space to find the optimal parameter sets.

Interestingly for the S&P500 Futures(c) results from tick data between 30/07/08 - 31/07/08 we find that both models have equal significance for the one-step ahead predictions, and indeed find the same optimal parameter set {4, 100, 0.52, 0.9}. We also find the Copper futures results for daily data between 19/05/06 - 22/08/08 to be statistically significant in both the Simple MGFF and Expanded MGFF, with the same optimal parameter set {4, 50, 0.52, 0.9}. The NASDAQ results from daily data are not statistically significant in the Simple MGFF but are in the Expanded MGFF, although in this case the data size is a lot smaller. We perform the analysis twice in the Simple MGFF case, where we have a larger and smaller data set to compare the effects of sample

Asset Label	Freq.	Time Period	Size	$\frac{\kappa_1-\kappa_0}{\widehat{\sigma_1}}$	p-val	$\frac{\kappa_2-\kappa_0}{\widehat{\sigma_2}}$	p-val	$\frac{\kappa_3-\kappa_0}{\widehat{\sigma_3}}$	p-val	$\{m, T, c, \tau\}$
Real Estate(a)	daily	30/07/93 - 28/08/08	3600	0.56	0.30	0.34	0.38	0.14	0.44	{3, 50, 0.52, 0.9}
Nickel Fut	daily	19/05/06 - 22/08/08	500	**6.17**	**0.02**	**4.07**	**0.05**	1.27	0.21	{3, 100, 0.51, 0.9}
Copper Fut	daily	19/05/06 - 22/08/08	500	**2.7**	**0.01**	0.6	0.25	-0.23	0.61	{4, 50, 0.52, 0.9}
Oil Fut(a)	daily	10/05/00 - 22/04/08	1900	1.02	0.12	0.73	0.22	0.54	0.15	{4, 100, 0.52, 0.9}
S&P Fut(a)	tick	30/06/08 - 01/07/08	1400	0.95	0.12	**1.67**	**0.01**	0.55	0.27	{2, 100, 0.52, 1}
S&P Fut(b)	tick	02/07/08 - 03/07/08	1000	**2.38**	**0.05**	0.73	0.33	-0.89	0.81	{2, 50, 0.52, 1}
S&P Fut(c)	tick	30/07/08 - 31/07/08	300	**10.1**	**0**	**3.85**	**0**	-1	0.74	{4, 100, 0.52, 0.9}
NASDAQ(a)	daily	03/01/95 - 19/08/08	3400	0.66	0.19	0.10	0.44	0.37	0.30	{2, 50, 0.52, 0.9}
NASDAQ(b)	daily	17/05/02 - 19/08/08	1500	1.21	0.12	0.5	0.35	0.09	0.48	{5, 50, 0.52, 1}

(a) Simple MGFF significance tests

Asset Label	Freq.	Time Period	Size	$\frac{\kappa_1-\kappa_0}{\widehat{\sigma_1}}$	p-val	$\frac{\kappa_2-\kappa_0}{\widehat{\sigma_2}}$	p-val	$\frac{\kappa_3-\kappa_0}{\widehat{\sigma_3}}$	p-val	$\{m, T, c, \tau\}$
Real Estate(b)	daily	11/03/02 - 27/09/05	800	0.58	0.31	0.31	0.32	0.57	0.22	{4, 50, 0.51, 0.9}
Nickel Fut	daily	19/05/06 - 22/08/08	500	1.58	0.04	1.25	0.01	0.74	0.09	{2, 50, 0.51, 0.9}
Copper Fut	daily	19/05/06 - 22/08/08	500	**1.93**	**0.01**	**2.47**	0.12	0.20	0.37	{4, 50, 0.52, 0.9}
Oil Fut(b)	daily	22/10/92 - 28/02/95	500	**5.24**	**0.01**	**2.95**	**0.05**	0.19	0.41	{2, 50, 0.51, 0.9}
Oil Fut(c)	daily	01/05/06 - 26/08/08	500	**1.67**	0.13	0.54	0.27	0.59	0.32	{3, 50, 0.51, 1}
S&P Fut(d)	daily	01/08/06 - 31/07/08	400	**3.28**	**0.05**	1.64	0.31	-1.15	0.81	{4, 100, 0.51, 1}
S&P Fut(e)	tick	28/07/08 - 29/07/08	300	**1.86**	0.07	**4**	**0.04**	-0.40	0.68	{2, 50, 0.51, 1}
S&P Fut(c)	tick	30/07/08 - 31/07/08	300	**10.1**	**0**					{4, 100, 0.52, 0.9}
NASDAQ(c)	daily	22/08/06 - 19/08/08	450	**2.29**	**0.04**					{2, 50, 0.52, 0.9}
BP	daily	13/07/89 - 06/06/97	1800	0.63	0.21	0.45	0.32	0.66	0.22	{3, 50, 0.52, 1}
USD-EUR	daily	26/04/05 - 26/08/08	800	1.00	0.15	1.08	0.04	1.32	0.02	{2, 100, 0.52, 1}
USD-JPY	daily	29/11/91 - 17/03/94	550	1.63	0.13	-1.97	0.96	-1.07	0.81	{2, 50, 0.52, 1}

(b) Expanded MGFF significance tests

Table 4.9: Significance tests of hit-rates showing the separation of means criterion $\left(\frac{\kappa_e-\kappa_0}{\widehat{\sigma_e}}\right)$ for forecasts e-steps ahead, the p-values (the probability of achieving this hit-rate or better given that the null hypothesis is true) and the optimal parameter set $\{m, T, c, \tau\}$

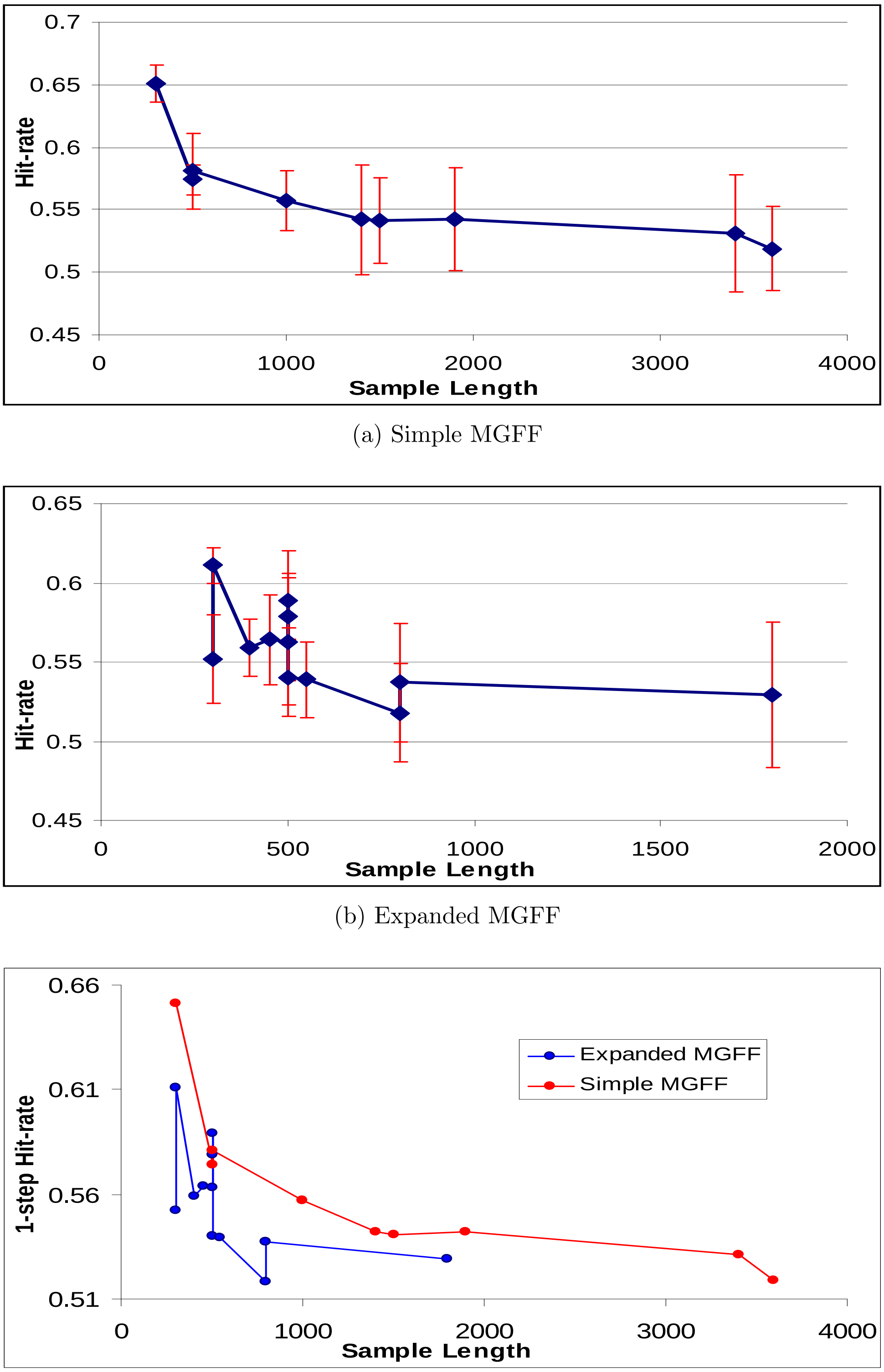
length. We find the shorter sample length of NASDAQ(b) from 17/05/02 - 19/08/08 of 1500 samples to have more significance than the longer sample length of NASDAQ(a) from 03/01/95 - 19/08/08 of 3400 samples. Both cases find similar parameter sets, with only the score decay rate τ different. In fact, the parameter set for the longer sample length is the same as the optimal set found by the Expanded MGFF, although in the latter case the results are deemed to be statistically significant.

For most of the samples analyzed however, the sample length (the number of data points over which we run the MGFF) of the Expanded MGFF is smaller than in the Simple MGFF, making direct comparison difficult, especially if we consider the sample length has an impact on the hit-rate measures, as discussed in Section 4.5.2.2. The Expanded MGFF has a greater number of strategy weight estimates to make, and consequently takes longer computational time to run, therefore we leave extending the investigation of longer sample lengths of the Expanded MGFF for future work.

4.5.2.2 The Sample Length

From the hit-rate results tables 4.4a on page 120 and 4.7a on page 129, and the significance tests in Table 4.9 on the preceding page, it can be observed that the significance of results depends highly on size of the sample length - the number of data points over which we run the MGFF. We illustrate this by plotting figures to show the relationship. Figure 4.15 on the next page shows the variation of the hit-rate with the sample length, whilst figure 4.16 on page 138 shows the significance criterion of equation 4.49 with the sample length.

Despite our relatively small sample size, this suggests that as we increase the sample length size, the MGFF produces less significant results, or equivalently doesn't perform as well. For a procedure geared towards on-line implementation such as the Kalman Filter, this is disappointing as one would hope that the procedure would be able to adapt to changing market conditions, so that the MGFF would perform consistently well over such longer time horizons. The reliability of any performance measurements are naturally improved if



(a) Simple MGFF

(b) Expanded MGFF

(c) Simple and Expanded MGFF

Figure 4.15: The variation between the average 1-step hit-rates κ_1 and the sample length of the financial data set, comparing the Simple MGFF results and Expanded MGFF results.

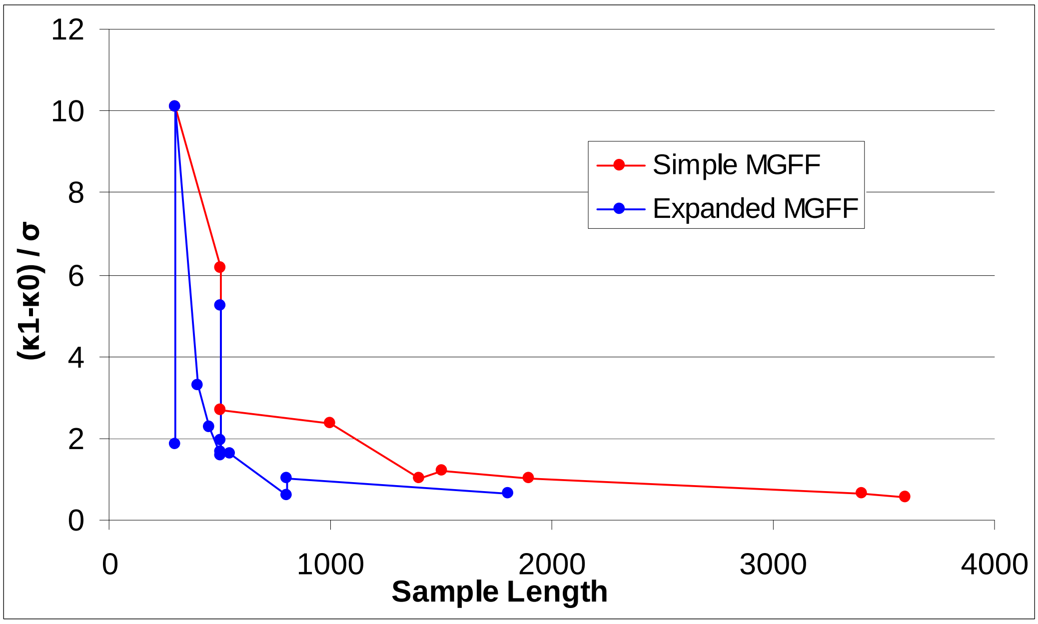


Figure 4.16: The variation between the separation of means criterion for 1-step hit-rates and the sample length of the financial data set used, comparing results of the Simple MGFF and Expanded MGFF.

we increase the sample length over which we measure, therefore the current reliability of our forecasting is questionable. However, Figure 4.16 does demonstrate that, given a specific data size, the Simple MGFF can compete with, or indeed produce more significant results than the Expanded MGFF. Alternatively, the Simple MGFF produces statistically significant results over a greater range of data size (those with values above 1.64 in Figure 4.16).

4.5.3 Comparison with other studies

The study of directional change prediction blossomed only relatively recently, "Although there exists a vast number of articles addressing the predictability of stock market return as well as the pricing of stock index financial instruments... most of the proposed models rely on accurate forecasting of the level (i.e. value) of the underlying stock index or its return... Depending on the trading strategies adopted by investors, forecasting methods based on minimizing forecast error may not be adequate to meet their objectives. In other words, trading driven by a certain forecast with a small forecast error may not be as profitable as trading guided by an accurate prediction of the direction of movement... Therefore, predicting the direction of change of the stock market index and its return is also significant in the development of effective market trading strategies." [Leung et al. 2000]

In the same paper, Leung et al. present a comparison of different multivariate classification techniques and level estimation counterparts, applying them to forecast the monthly directional change of the S&P500 Composite Index between January 1991 and December 1995. They report a range of hit-rates for the 60 months of their out-of-sample forecasts using the classification techniques:- linear discriminant analysis[3] (0.57), the binary choice models[4] logit (0.60) and probit (0.60), and probabilistic neural network models[5] (0.63). All techniques are statistically significant at the 95% level of a one-sided hypothesis test except for the discriminant analysis. The hit-rates achieved by the out-of-sample forecasts using their level estimation models are:- adaptive exponential smoothing (0.48), vector auto-regressions with a Kalman Filter (0.53), multivariate transfer functions (0.53) and multi-layered feed-forward neural networks[6] (0.63). None of these hit-rates are significant at the 95% level except for the feed-forward neural network model. Indeed, a similar study on daily Dow Jones Industrial Average (DJIA) data using a simple technical trading strategy based on non-parametric methods (created by a feed-forward neural network) was performed in [Gencay 1998]. It shows a range of hit-rates between 0.57 and 0.61 for five different periods of out-of-sample daily forecasts between 1963 and 1988, all statistically significant at the 95% level.

[3] "a multivariate statistical technique that investigates the differences between two or more groups of observations with respect to a set of independent (input) variables. These independent variables, called discriminator variables, are used to distinguish the characteristics among different groups." [Leung et al. 2000] In our case, the two groups are the predicted direction - up or down, and a third group (no change) could also in theory be added.

[4] used when trying to model dependent variables that can take on only binary values (e.g. 0 or 1).

[5] conceptually built on the Bayesian method of classification

[6] which use the explanatory variables (inputs) to control the dependent variable target (output)

In comparison for Stock Indices, our MGFF looks at the S&P500 Composite Index over tick and daily data, and NASDAQ daily data, and it would be interesting to test it on monthly S&P500 data or daily DJIA data in future work to make a direct comparison with Leung et al. and Gencays' studies. For the S&P500, our Simple MGFF achieves a hit-rate of 0.54 over 1400 ticks (S&P500a), 0.56 over 1000 ticks (S&P500b) and 0.65 over 300 ticks (S&P500c) of data from the three different periods we tested in middle of 2008. The latter two hit-rates are deemed to be significant at the 95% level of a one-sided hypothesis, though the hit-rate from the longer sample period of 1400 ticks is not. The Expanded MGFF achieves a hit-rate of 0.60 over 400 data points of the daily S&P500(d) data, whilst it obtains 0.55 and 0.61 for the samples of 300 tick data points in S&P500(e) and S&P500(c) respectively. All the Expanded MGFF hit-rates are statistically significant at the 95% level. For the NASDAQ data, our results are not statistically significant over 3400 (NASDAQa) and 1500 (NASDAQb) data point samples using the Simple MGFF, but are statistically significant at the 95% level with a hit-rate of 0.56 over 450 data points (NASDAQc) using the Expanded MGFF. We therefore conclude that our Stock Index results are comparable in magnitude of hit-rate and significance with the results achieved by Gencay and the directional forecasting techniques in Leung et al.'s study, despite being over a longer sample size than Leung et al. and using higher frequency data. In addition, our results for the S&P500 Composite Index appear to be better than those they achieved with the level estimation methods, except for the feed-forward neural network.

Feed-forward and recurrent neural networks[7] are also applied to foreign exchange rates in [Kuan, Liu 1995], who analyzes daily data of USD-JPY and USD-DEM (US dollar-Deutschmark, which can be seen as a proxy for the USD-EUR rate) among other currencies between 1980 and 1985. They find mixed success in forecasting the currencies. For the USD-JPY, there are no feed-forward networks and only two recurrent networks that have statistically significant hit-rates of the direction (what they call 'market timing ability') at the 95% level - in these cases they are 0.61 and 0.60 and are made over only 100 out-of-sample observations. For the USD-DEM, neither the feed-forward

[7] recurrent neural networks feed the outputs back into the inputs, which is useful in modelling feedback behaviour

nor recurrent network models have statistically significant hit-rates at the 95% level. We only analyse FX markets with the Expanded MGFF, and although achieving hit-rates of around 0.54, we have similar difficulties in finding statistically significant results at the 95% level, although we have performed the analysis over a larger data set (800 observations for the USD-EUR and 550 for the USD-JPY). Therefore our results for the FX markets are comparable with that found using neural networks in Kuan and Liu's study.

4.6 Future work

In order to gain more confidence in any conclusions drawn from our investigations, we propose to analyse a greater number of assets and moreover longer data sets, to increase the reliability of the results. It may also be that we have not found the optimal parameter set to maximise the hit-rate, owing to our search over a relatively small subset of parameters. Ideally, we would increase both the size of the search, i.e. the lower and upper bounds of parameters m, T, c and τ, and also the resolution of our search, i.e. by using smaller interval steps for the discrete variable T and the continuous variables c and τ.

Alternatively it may be that the Minority Game Forecasting Framework using the Kalman Filter, at least in the version we have built, cannot capture the markets investigated here to a degree necessary to allow short-term predictions of the market direction, in contradiction to the work in [Jefferies et al. 2001, Johnson et al. 2001]. Or perhaps the Kalman Filtering techniques we have used need refining. Indeed, Gupta presents several alternative versions of Kalman Filtering methods [Gupta 2007, Gupta, Hauser 2007], in particular by constraining the state estimates such that they sum to 1, in order to decrease the measurement residuals and ensure better tracking. This may improve the convergence of the estimates. Indeed, there are also alternate ways of implementing the Kalman Filter such as the Square Root Filter[8] [Bar-Shalom et al. 2001], that can double the improve the accuracy and improve numerical stability (reduce round-off errors), consequently improving the

[8] achieved by sequentially calculating each state vector component and then factorizing the covariance matrix P by Cholesky decomposition in order to evolve the covariance

symmetry and positive-definiteness of the state covariance matrices $P_{t|t}$. Such techniques will be investigated at a later date to assess whether they improve performance.

We could reduce the size of our parameter search by removing the time-horizon parameter T in developing an (exponentially weighed) 'iterative' score update version of the MGFF (as discussed in 2.5.1 on page 55). The determination of strategy scores would be thus controlled by the score decay-rate parameter τ, with a decrease in τ effectively reducing the relevance of older strategy score increments, similar to the effect of decreasing T, but in a smoother fashion. Generally, the benefit would be to speed up the parameter search process as we would only be searching the parameter set $\{m, c, \tau\}$.

Another alternative to improve the adaptability of our MGFF would be to simultaneously run multiple versions of the MGFF, each with a different parameter set, to make an ensemble of forecasts. We would therefore choose the forecast from a subset of all the forecasts, such as the best performing version at that current time as measured by its hit-rate, or by taking an average prediction of several high performing versions. This is a technique favoured in [Lamper 2002] and also by the MET Office when making weather forecasts[Met Office]. We endeavour to create such an ensemble MGFF platform in the future. Initial studies should consider the hit-rate dynamics of several competing parameter sets - when one parameter set starts becoming less successful at predicting the time-series, does another parameter set start working better?

Although we have merely focused on the hit-rate as we have been primarily interested in using this measure to find the optimal parameters of our model, if our framework is to have any use for trading in the markets, another important consideration is the monetary gain possible using our MGFF technique. Though this is heavily dependent on the kind of trading strategy employed given the information provided by the MGFF, a relatively simple simulation experiment would be to use a buy-hold-sell strategy, where we buy the asset when the MGFF predicts a price increase, and then sell the asset when the MGFF next predicts a price decrease. We would then be able to show the increase (or decrease) in wealth over time by using the MGFF and our trading strategy in the market place (see [Jefferies et al. 2001, Lamper 2002] for examples). We note, however, that the monetary gain of any trading strategy is highly dependent on externalities such as transaction costs, which may prove to be prohibitive for trading strategies that require frequent transactions in the market.

4.7 Concluding Remarks

In this chapter, we have outlined a method to reverse engineer the Minority Game model so that we can calibrate it to time series data and subsequently make forecasts of the future directions of the time series, calling it the the Minority Game Forecasting Framework. We employed a Kalman Filter to estimate the weights of the strategies held by the agents, presenting a novel version that focuses on estimating single strategy weights (the Simple MGFF), and secondly a version that deals with strategy pair weights (the Expanded MGFF). Both these MGFFs are simpler than existing models in the literature as they only estimate strategies from the Reduced Strategy Space. We applied them to data generated from the MG simulations and also from a range of financial asset classes, finding comparable results in both versions. Considering the results are comparable, this would suggest the benefit of using the Simple MGFF, which is computationally quicker than the Expanded MGFF.

Chapter 5

The Bionomic Framework

> *"If I were reincarnated, I would wish to be returned to Earth as a killer virus to lower human population levels."* Prince Phillip, His Royal Highness The Duke of Edinburgh

In this chapter, we explore the use of a population heuristic, the Bionomic Algorithm, to replace the Kalman Filter in the parameter estimation problem of the Expanded MGFF in Section 4.4. Although the main drawbacks of using the Bionomic Algorithm are similar to every population heuristic - namely the relatively large amount of time needed to perform the optimization compared to a single solution optimization, and the additional assumption of static estimation parameters over the length of the time series data - it is the nature of the algorithm's directed search over parameter space that makes it well suited to high-dimensional estimation problems such as in the Expanded MGFF. The Bionomic Algorithm's promising ability to escape local optima and to efficiently find the global optimum suggests that it may be possible to produce superior results.

We begin by outlining the basic ideas of population heuristics and presenting the Bionomic Algorithm, and then move on to developing a particularization that can be specifically applied to the Expanded MGFF estimation problem, or more generally to any equality constrained (or conserved quantity e.g. the length of a vector) optimization problem. We apply this particularization of the Bionomic Algorithm to the calibration of the Minority Game Forecasting Framework, therefore creating the Bionomic MGFF. We apply this to time series data generated by a Minority Game and also from a range of financial assets and present our results.

5.1 Population heuristics and the Genetic Algorithm

Population heuristics, the methodology of generating several solutions to a specific problem as a basis for making further improvements, is an enticing alternative to the more traditional optimization techniques such as Newton-Raphson, Levenberg-Marquardt [1], and Simplex[2] methods that aim to improve a single solution (see [Gershenfeld 1999] for an introductory overview to optimization). Motivated by the ideas of sexual reproduction and Darwinian evolution as the respective method and force that enables nature to adapt and 'improve' the fitness of living organisms in a changing environment, John H. Holland of the Sante Fe Institute pioneered the use of population heuristics known as Genetic Algorithms in the 1970s [Holland 1975]. Rather than spending a large amount of time trying to search the neighbourhood of one solution and being vulnerable to getting stuck in local optima, a Genetic Algorithm generates an ensemble of potential solutions and then combines and possibly mutates these in order to produce future 'generations' of solutions, in analogy with the evolutionary update of the genome [Forrest 1993]. Indeed Beasley claims, "there is something about a population heuristic which means that it is often capable of producing better quality solutions than single solution heuristics."[Beasley 1999]

Gershenfeld outlines four main steps for this update [Gershenfeld 1999]:

[1] which employs a combination of gradient search methods when far away from an optimum and the Newton method when close. This non-linear search needs the 1st and 2nd derivatives of a reasonably smooth objective function (both derivatives and smoothness are usually not satisfied in computational simulations), with starting parameter values near the desired extremum of the cost function.

[2] useful for multi-dimensional search of the local area around one solution

Fitness - the fitness step evaluates the objective function, the measure which we are interested in optimizing, for every member (in our case a set of parameters of our model) of the population.

Reproduction - members of the population are selected probabilistically, based on a weighting derived from their fitness. A parameter set with a high fitness (high optimality of the objective function), is more likely to be used to generate the next generation than a parameter set with low fitness. We can also choose how much we depend on the fitness in determining the relative rates of reproduction - with high selectivity, one solution is forced to dominate, whereas in low (or zero) selectivity we accept any solution.

Crossover - after two 'parents' are chosen in the reproduction step, an 'offspring' (a potential new member of the next generation of solutions) is generated by some kind of random mix of components of the parent solutions, for example the first k components of the parameter set of parent A and the rest of the parameter components from parent B. This process enables the possibility of jumping around to new parts of the parameter space without needing to make a series of discrete moves to get there, as in the simplex method.

Mutation - this step changes some values of the solution (i.e. some of the parameters) either randomly, or preferably by taking advantage of what is known about how to generate good moves for the problem.

Although these methods will not guarantee the global optimal solution, they do generally provide good solutions in a reasonable amount of time, often the desired result for many practical applications of optimization.

5.2 A new Genetic Algorithm for the Minority Game Forecasting Framework

5.2.1 Considerations

In subsection 4.1.4.1 on page 95 we discussed a few possible techniques to estimate the relative weightings of the strategy space and followed the path of Lamper and Gupta in applying Kalman Filtering methods to the problem. The benefit of using the Kalman Filter was the applicability of its use in on-line forecasting, when we recursively generate a new estimate of the strategy weightings in play whenever new information, such as the new price increment and volume data is made available. By doing this, we can estimate how the strategy weightings evolve over time throughout a period of time-series data.

The success of our forecasting methodology is measured by the hit-rate (or correlation) of our predictions against the directional change of a real asset price-time series. It is the hit-rate or correlation that we want to optimize - this is our objective function. Assuming the Minority Game (and Mix-Game) setup is somehow representative of a real market structure that produces the asset price time-series we observe, one could argue that as the success of strategies is highly interdependent on the ecosystem of other strategies in play, it is likely that the objective function forms a complicated search landscape (non-convex and non-linear in the language of the field of Optimization). Indeed, it is this 'rough' type of search landscape that is particularly suited for investigation by Genetic Algorithms, although we also note that if the search landscape is such that there are multiple optimal parameter sets with similar qualities of solution, the Simulated Annealing method introduced in [Metropolis et al. 1953] and later developed in [Kirkpatrick et al. 1983, Kirkpatrick 1984] may also prove to be useful (see discussion for applicability of optimization methods in [Gershenfeld 1999]).

The drawback with using a Genetic Algorithm to solve the strategy weighting problem is the time expenditure necessary - whereas the Kalman Filter technique gives us a quick on-line solution, the Genetic Algorithm takes a lot longer as it necessitates the generation of an ensemble of possible solutions, and several iterations to this population to obtain an improved solution. In addition, although theoretically possible, it would be computationally very expensive to calibrate to multiple data series such as trading volume in addition to price increments, as the dimensionality of the search landscape would exponentially increase for every new data input, thus taking even longer to find an optimal solution. (Gershenfeld [Gershenfeld 1999] remarks that Genetic Algorithms are most useful when there is a rough search landscape with several smaller local optima and one distinctly best global optimum, as it is better to spend less time improving one particular solution and more time on several solutions that can cover the search-space to find the distinctly best solution.)

Owing to the time cost of running a Genetic Algorithm search, any sort of on-line implementation would be computationally impractical, so we resort to performing a batch optimization where we execute the Genetic Algorithm over a sample of price time-series data lasting a few hundred time-steps. We note however, that if there is a distinct global optimum, it would still be theoretically possible to observe the evolution of this optimal set of strategy weights over time by using a moving window of overlapping asset price data. In fact, Gou [Gou 2006c] demonstrates a similar methodology by splitting a time-series of daily closing prices of the Shanghai Index into non-overlapping periods, with each period being judged to have distinctly different price trends - either increasing, decreasing or stationary.

Our work is novel in that Gou performs her search over the Full Strategy Space, whereas we hope to simplify and speed up the search procedure (exponentially so) by using the Reduced Strategy Space of strategies. In addition, we employ a unique version of a Genetic Algorithm called the Bionomic Algorithm (BA) - see 5.2.3.2 on page 150 for a detailed description - that, as far as we are aware, has not been applied to calibrating agent-based models of financial markets before. Also, Gou primarily investigates the calibration of the Mix-game model whilst we only focus on the Minority Game, though an obvious next step in future research would be to apply the BA to the Mix Game too.

5.2.2 The calibration methodology

We wish to find the strategy weightings that maximise the hit-rate measure of equation 4.7 on page 93. Following on from a methodology similar to that used by Gou in [Gou 2006c], we use the Minority game to do a series of one-step predictions of an asset price time-series. Just as in the Kalman Filter calibration methodology of Section 4.1.1 on page 89, we extract the sign of the price increments of a target financial time-series and use this information to derive the Minority Game state μ at every time-step and score the strategies to determine the active strategies, given a set of model parameters $\{m, T, c, \tau\}$. Using the knowledge of the active strategy scores and an estimate of the (relative) strategy weightings, we then make a prediction of the next price increment and use this in our hit-rate measure. By choosing parameters that maximise the hit-rate, we can deduce the most likely configuration of strategy weightings. The strategy weighting space will be searched using the Bionomic Algorithm, whilst the set of Grand Canonical Minority Game parameters $\{m, T, c, \tau\}$ will be obtained using a simple trial-and-error search.

For the same reasons as in the Kalman Filter forecasting method (see 2.5.1 on page 56 for discussion), we fix the parameter $s = 2$, so that the number of strategies available to each agent is only two. Whilst Gou assumes a value for the number of agents in the simulation N, such as $N_{total} = 201$ and $N_{maj} = 40$, we remove the necessity of estimating N by again focussing on the relative strategy weights as we did for the Kalman Filter forecasting method in Chapter 4, assuming equation 4.34 on page 105 holds where x_R is the proportion of the population holding strategy R, rather than using any absolute weighting measure.

Because the Mix-Game can produce behaviour unlike that seen in financial markets (such as the case when Majority-Game agents dominate the Minority-Game agents), [Gou 2006c] stipulates the need to restrict the parameter set $\{m_{maj}, m_{min}, T_{maj}, T_{min}, N_{maj}, N_{min}\}$ for the Mix-Game model to ones that produces statistics similar to the stylized facts of financial time-series (see for example the necessary parameter inequality relations starting with equation 2.5 on page 54 to ensure stability of the simulations). In addition she claims it is necessary to match the median of the 'local volatility of the Mix-Game and the target financial time-series, and make sure that the log-log plot of absolute returns in both cases looks similar. As the Minority Game is less prone to such unrealistic (divergent) behaviour, and moreover as the size of the volatility is dependent on the value of the N parameter that we have removed from our model, we choose to ignore these specifications.

5.2.3 The Bionomic Algorithm

5.2.3.1 Introducing the Bionomic Algorithm

The Bionomic Algorithm (BA) is an intelligent probabilistic search algorithm combining ideas from genetic algorithms and scatter search[3]. It was first presented in [Christofides 1994] and has since been applied to the optimal design of telecommunications networks [Christofides 1996], corporate tax structures [Christofides et al. 2003], and scheduling aircraft landings [Pinol 2006]. One of the main strengths of the algorithm is its requirement to maintain a population of child solutions with high distinctiveness, thus preventing the search over the parameter space being restricted to a narrow subset of possible parameter sets. This generally improves the quality of the solutions possible and is particularly good at searching the space of high-dimensional problems such as our quest to estimate the probability weightings of the strategy pairs available to the agents playing the Minority Game, where the dimension (number of estimation parameters) is of the order of 100 (see Section 5.2.4.1 on page 153.

5.2.3.2 An overview of the Bionomic Algorithm

The following description is directly adapted from [Christofides et al. 2003] and [Karavias 2006]:

We use y and x to be solutions to a general abstract problem, and $\eta(y)$ and $\rho(y)$ for the objective value and $\hat{m}$-dimensional infeasibility vector of $\hat{m}$ problem

[3]this method uses linear combinations of parent solutions to form the next generation of child solutions (see [Glover 1994, Glover 1995, Glover 1997])

constraints, respectively. $\rho_m(y) = 0$ if constraint m is feasible. Given two solutions y and x, we define a choice function f, so that $f_{\eta\rho}(y, x) = y^*$, where $y^* \in \{y, x, u\}$ is the preferred solution between y and x, if $y^* \in \{y, x\}$, and undefined (i.e. no choice is made) if $y^* = u$. Note that the subscripts $\eta\rho$ are meant to indicate that y and x enter f only through their properties η and ρ. The function f is required to be transitive and defines a partial ordering on the set of solutions. The following outline (algorithm 5.1 on the following page) is described with respect to a maximisation problem, such as finding the MGFF strategy weight parameters that give the maximum hit-rate or correlation compared to a real market price time-series:

Algorithm 5.1 The Bionomic Algorithm

(1.1) **Generate an initial population of solutions.**
The initial solutions may be generated at random or by any other algorithm, such as maximally avoiding points (see Section 5.2.4.3 on page 160 and Figure 5.2 on page 161).

(1.2) **Improve each solution in the initial population.**

repeat: for each solution y

Compute the solution value $\eta(y)$ and infeasibility vector $\rho(y)$.
(a) If a solution is infeasible minimise a measure of $\rho(y)$ by applying a local-optimization algorithm. If the solution has become feasible go to the following step (b) otherwise if it is still infeasible proceed to the next solution.
(b) else if a solution is feasible maximise $\eta(y)$ by a local optimization algorithm, maintaining feasibility.

until: all solutions are examined.

repeat: generating new solutions.

(2) **Select a parent solution set.**
A parent set can be constructed by sequentially choosing from the population and adding to the set, solutions y with high 'dominance' over the population and high '*distinctiveness*' from the other members already in the set.

(3) **Combine parents to produce a child solution.**
(a) A combining procedure is used to produce a pool of child solutions for this parent set.
(b) Compute the dominance between all solutions in the pool and keep only the 'best' one.
(c) Improve the child solution using the same procedure as in [1.2a] and [1.2b] above.

(4) **Add the child to the population.**
(a) **If** the child solution dominates a sub-list of the population, replace the most dominated member of this sub-list with the child.
(b) **else if** a member of the population (excluding the currently best solution) has reached a critical *age*, replace this member with the child.
(c) **else** reject the child.

until: a stopping criterion is satisfied.

(5) **Report the best solution encountered.**

5.2.4 Applying the BA to the MG

Like all heuristic optimization techniques, the BA is general and has to be adapted to the problem of interest, and it is the particularisation of the algorithm to the specific problem in question that can also influence the success or failure of the algorithm [Christofides et al. 2003]. Indeed, we present here a new particularisation of the BA suitable for searching the strategy space parameter weightings of such agent-based problems as the Minority Game, Mix-Game and other such variants.

5.2.4.1 The dimensionality of the problem, n

In Chapter 4 we continued the work of Lamper and Gupta to develop and apply the Kalman Filter to estimate the relative strategy weights of the Minority Game being used in a financial market setting, with an underlying assumption that the Minority Game can, to a certain degree, capture some behaviour of the complex systems that are financial markets. The benefit of using the Kalman Filter was its suitability in on-line implementation and thus its inherent ability in capturing the evolution of strategy weights. Although one of the Minority Game's primary features is the adaptability of the agents as they hold more than one strategy, we were able to reduce the computational complexity of the problem in the Simple MGFF by only estimating single strategy weights and hoping that the model mismatch would be compensated for by the adaptability of the estimation procedure, with the Kalman Filter able to adapt fast enough to the changing weights.

In the Bionomic Algorithm framework, we use a technique that involves batch optimization - taking a static data set and finding parameter estimates that provide the highest value of our objective function - in our case the hit-rate of predictions, i.e. the proportion of successful predictions of the one-step ahead market movement. The optimal strategy weightings we find remain static over the data sample period. Thus, in order to capture the adaptability of the agents in this batch optimization setup, we are forced to consider the original Minority Game setup where agents have at least two strategies to choose from and are able to switch between them - the simplest case would be to estimate the relative weight of strategy pairs, such as strategies {1, 2} and {1, 3} etc. Our problem becomes the estimation of the relative weight of these strategy collections, just like the Expanded MGFF using the Kalman Filter in Section 4.4 on page 125. This is also supported in [Marsili et al. 2000] by Marsili's observation that the Minority Game agents are most successful when their strategy collection size is only two, as greater collection sizes lead agents to switch too often, as highlighted in Section 2.5.1 on page 56.

In order to make our optimization procedure as efficient as possible, we need to minimise the dimensionality of our problem. There are a few simplifications we can make, the first being that we use the Reduced Strategy Space of only uncorrelated and anti-correlated strategies which has size 2^{m+1}, rather than the Full Strategy Space of size 2^{2^m} as used in [Gou 2006c]. Another one is that the pair set is unordered i.e. we do not distinguish between the pair {strategy R, strategy R'} with pair {strategy R', strategy R}, thus we only need one parameter weight to capture both possibilities. In addition, we note that the strategy space contains anti-correlated pairs such as the pair {strategy $\overline{R}$, strategy $\overline{R'}$} which advocates the opposite action to the pair {strategy R, strategy R'} in every state of the system. Rather than model these as separate parameters, we reject the inequality constraint used in the Kalman Filtering, where all strategies had weight $x_R > 0$ and extend the parameter space to the negative domain to model the aggregate weight of anti-correlated strategy pairs:

$$x^{agg}_{R,R'} := \begin{cases} x_{R,R'} - x_{\overline{R},\overline{R'}} & \forall R \neq \overline{R} \\ x_{R,\overline{R}} & \forall R = \overline{R} \end{cases} \tag{5.1}$$

where $x^{agg}_{R,R'}$ is the aggregate weight of the strategy pair {R, R'} which is what we want the BA to estimate.

This is possible because we restrict the optimization to just one data series[4] - the price of the asset - and do not consider the volume data as well (compare

[4]The ability of the Kalman Filter to handle multiple data series is another feature that is not practical in many other optimization techniques including the Bionomic Algorithm, however we hope that the alternative strengths of the Bionomic Algorithm, such as its efficiency in searching large dimensional estimation spaces will compensate.

this with equation 4.36 on page 105), therefore there is no volume constraint such as:

$$z_2 = \sum_{R,\,R'} \left(\mathcal{H}\{S_R - S_{R'}\} \cdot \mathcal{H}\{S_R - r\} + \mathcal{H}\{S_{R'} - S_R\} \cdot \mathcal{H}\{S_{R'} - r\} \right) \cdot \mid x_{R,\,R'} \mid \tag{5.2}$$

where the time variable has been removed for clarity. There is therefore no unique solution to equation 5.1 - there are an infinite combination of weights $x_{R,R'}$ and $x_{\overline{R},\overline{R'}}$ that satisfy this.

We represent the estimation parameters we require on the strategy space matrix for the case $m = 2$ in table 5.1 with the convention that an element labelled by a lower case letter (i.e. 'u') represents a strategy pair $\{R,\, R'\}$ that needs to be estimated, whilst the italicized lower case with an over-line ($\overline{u}$) represents the related anti-correlated pair $\{\overline{R},\, \overline{R'}\}$ that can be obtained from its respective partner pair and doesn't need to be estimated separately. For example, strategy pair $\{2,\, \overline{1}\}$ is labelled by the letter 'i', whilst its anti-correlated partner pair $\{\overline{2},\, 1\}$=$\{1,\, \overline{2}\}$ is labelled by $\overline{i}$. Note this is a symmetric matrix, so the lower triangle would be the transpose of the upper triangle shown, however we omit this for clarity (we do not need to estimate this as well).

As strategies in the Reduced Strategy Space are uncorrelated or anti-correlated across the states of the system μ, this means that strategy score increments for each strategy must also either be uncorrelated or anti-correlated. Therefore it is reasonable to expect that each strategy in a pair spends approximately an equal amount of time on average as the dominant strategy of the pair. Therefore, the strategy pair $\{R,\, R'\}$ is uncorrelated with $\{R,\, \overline{R'}\}$ as long as $R \neq R'$, and we therefore have to treat the weight of these two strategy pairs as different parameter weights to be estimated.

We can cover all elements to be estimated if we consider the strategy pair $\{R,\, R'\}$ such that R is the row number and R' is the column number of the pair element. In this convention, strategy $R \in 1, 2, \ldots, 2^m$ and strategy $R' \in 1, 2, \ldots, 2^{m+1}$ with $R' \geq R$. Also note that every strategy pair (lower case letter) has a respective anti-correlated partner pair (italicized over-line) except for the special case of four pairs 'e', 'j', 'o' and 't', which correspond to

$R\backslash R'$	1	2	3	4	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
1	a	b	c	d	e	$\overline{i}$	$\overline{m}$	$\overline{q}$
2		f	g	h	i	j	n	$\overline{r}$
3			k	l	m	n	o	$\overline{s}$
4				p	q	r	s	t
$\overline{1}$					$\overline{a}$	$\overline{b}$	$\overline{c}$	$\overline{d}$
$\overline{2}$						$\overline{f}$	$\overline{g}$	$\overline{h}$
$\overline{3}$							$\overline{k}$	$\overline{l}$
$\overline{4}$								$\overline{p}$

Table 5.1: Strategy Pair arrangement of the distinct strategy pairs $\{R,\ R'\}$ and their anti-correlated counterparts $\{\overline{R},\ \overline{R'}\}$. In our estimation problem, it is the aggregate weight between the distinct strategy pairs minus their anti-correlated counterparts that we need to estimate. The matrix is symmetrical and the lower triangle is omitted for clarity.

the case where the pairs are their own respective anti-correlated partner pairs, i.e. $\{R,\overline{R}\}=\{\overline{R},R\}$. Thus we only need to estimate 2^m+1 columns over 2^m rows so that we can reduce our estimation problem to the size of only

$$n := 2^{2m} + 2^m \tag{5.3}$$

dimensions, much less than the total number of possible strategy pairs or elements in the strategy space matrix which numbers $2^{m+1} \cdot 2^{m+1} = 2^{2m+2}$ and still just over half the number of unordered strategy pairs that we estimated in Chapter 4 for the Expanded MGFF, which uses $2^{2m+1} + 2^m$ estimation weights. Table 5.2 on the following page illustrates the exponential increase in the number of dimensions of our Bionomic estimation problem over the smallest values of the agent memory parameter m.

5.2.4.2 The estimation vector space

Now that we have established the dimensionality - the number of different free parameters to estimate - of our problem, we need to think about how best to represent it. The flexibility of the population heuristic style of optimization now comes into play as we can relax the restrictions imposed by the linear matrix representation necessitated by the formulation of the Kalman Filter-

m	n
1	6
2	20
3	72
4	272
5	1056
6	4160
7	16512

Table 5.2: Hyperspace dimension values n for small m. The hyperspace dimension corresponds to the number of estimation parameters, and increases (super) exponentially and we are therefore restricted to small values of m to ensure computational tractability.

ing estimation problem. Specifically, we can readily consider the estimation weights to be directly related to a probability measure, and embed the equality constraint that all weights sum up to 1, if we represent the state of the system - the set of all strategy pair weights that we need to estimate - as a state vector $\underline{V}$ lying in an n-dimensional (normed) vector space Υ.

In this case, the uncorrelated aggregate strategy pairs are represented by axes on the orthonormal basis of the vector space. As all strategy pair aggregate weights x_u^{agg} sum up to 1:

$$\sum_{u=1}^{u=n} x_u^{agg} = 1 \tag{5.4}$$

we require the state vector to have length (2-norm) equal to 1:

$$\| \underline{V} \|_2 := \left(\sum_{u=1}^{u=n} | V_u |^2 \right)^{\frac{1}{2}} = 1 \tag{5.5}$$

From equation 5.4 and 5.5, we can clearly see that the aggregate weight x_u^{agg} of strategy pair 'u' can be represented by the square of the vector component V_u^2, the square of the projection of the state vector $\underline{V}$ onto axes 'u'. Thus, the state vector components V_u correspond to aggregate strategy pair 'square root-weights' $\sqrt{x_u^{agg}}$ and by Pythagoras' theorem, the sum of squared vector components, in effect the aggregate strategy pair weights, are constrained to sum to 1.

In fact, this formulation has close parallels to the state vector representation in Quantum Mechanics, where the state vector represents a physical system (at the quantum level), and the probability that the system is in a specific state is given by the square of the vector component related to that state, such as the quantum spin of a subatomic particle.

Acknowledging that we cannot determine the absolute weight of a strategy pair x_u and its anti-correlated partner pair $x_{\overline{u}}$, only the relative weighting between them x_u^{agg}, we use the following convention to derive the strategy pair weightings from the aggregate strategy pair weightings:

If a state vector component V_u is positive, we give to the strategy pair $\{R, R'\}$ the full weight by setting $x_u = V_u^2$, and the associated weight that represents the anti-correlated strategy pair $\{\overline{R}, \overline{R}'\}$ $x_{\overline{u}} = 0$.

Else, if the state vector component is negative, we set the strategy pair $\{R, R'\}$ weight $x_u = 0$, and give the associated anti-correlated strategy pair $\{\overline{R}, \overline{R}'\}$ the full estimated weight by setting $x_{\overline{u}} = V_u^2$.

5.2.4.3 Population of solutions

As the Bionomic Algorithm is a search heuristic of the 'population type', we need to generate an initial set of state vectors (estimates of the strategy pair weightings) of length (2-norm) equal to 1 which we write as $\mathbf{V}$. From each state vector $\underline{V} \in \mathbf{V}$, we can calculate the strategy pair weightings and thus use in the Minority Game Forecasting Framework to calculate the objective value (the hit-rate). Each state vector $\underline{V}$ is equivalent to a point on the surface of a n-dimensional hypersphere (also called a n-sphere by geometers or a $(n-1)$-sphere by topologists) of radius 1, with the vector from the origin of the hyperspace to these points defining the state vector $\underline{V}$, as can be seen in Figure 5.1.

We therefore need to generate an initial population of points on an n-dimensional hypersphere. There are several techniques that can be employed to generate

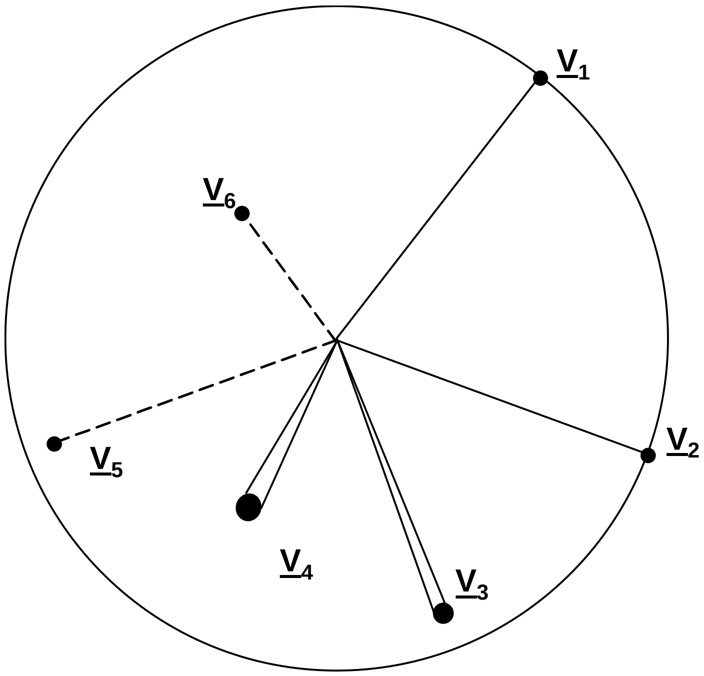


Figure 5.1: State vector solutions ($V_{1,2,3...}$) as points on an n-sphere. Each state vector corresponds to one possible set of n aggregate strategy pair weights, whilst each axis corresponds to the domain of each aggregate strategy pair weight that needs to be estimated. The domain of each aggregate strategy pair weight is from '+1' (in which case every agent only uses the strategy pair $\{R,\ R'\}$ corresponding to this axis and no other strategy pair) to '-1' (in which case every agent only uses the anti-correlated strategy pair $\{\overline{R},\ \overline{R'}\}$ corresponding to this axis and no other strategy pair). Thus, all state vectors are of length (2-norm) equal to 1 and all state vectors lie on an n-sphere in n-dimensional hyper-space. The projection of the state vector onto each axis equals the square root of the aggregate strategy pair weight represented on that particular axis.

points on a sphere [Marsaglia 1972], however we need one that can be generalized to an arbitrary number of dimensions - if we vary the agent memory parameter m, we change the number of dimensions as $n = 2^{2m} + 2^m$. A simple method that utilises the spherical symmetry of the multi-variate Gaussian distribution to generate such random points on a hypersphere was proposed by Muller [Muller 1959]. This involves generating n Gaussian distributed random variables ε, one for each component of the state vector, and normalising by dividing by the 2-norm of the Gaussian inputs:

$$\underline{V} = (\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_{n-1} + \varepsilon_n)^{\frac{-1}{2}} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{n-1} \\ \varepsilon_n \end{bmatrix} \tag{5.6}$$

We can use the simple and widely known Box-Muller method to transform uniformly distributed random variates into normally distributed variates [Box, Muller 1958]. In fact, the larger the surface area of the n-sphere that we explore (the larger the coverage of the n-sphere that our population of solutions spans) the more likely it is that we will find ‘good’ solutions. Ideally, we would therefore wish to start by generating an initial population of solutions with well-dispersed points across the surface.

Following on from the work of [Niederreier 1992], Karavias develops a ‘low-dispersion’ algorithm specifically designed to generate maximally avoiding points on a search space, which would provide optimal coverage of the search space by the initial population [Karavias 2006]. We, however, use possibly the simplest technique of generating uniformly distributed numbers using the rand() function in the cmath library for C++, which seemed to produce good coverage across the surface of the hypersphere in initial studies. Figure 5.2 highlights the difference in the expected outcome from the two techniques, and we note that it would be interesting to see in further work if we can improve our search using such a ‘low-dispersion’ algorithm to generate the uniformly distributed random variable set to undergo Box-Muller transform to be turned into normally distributed random variables.

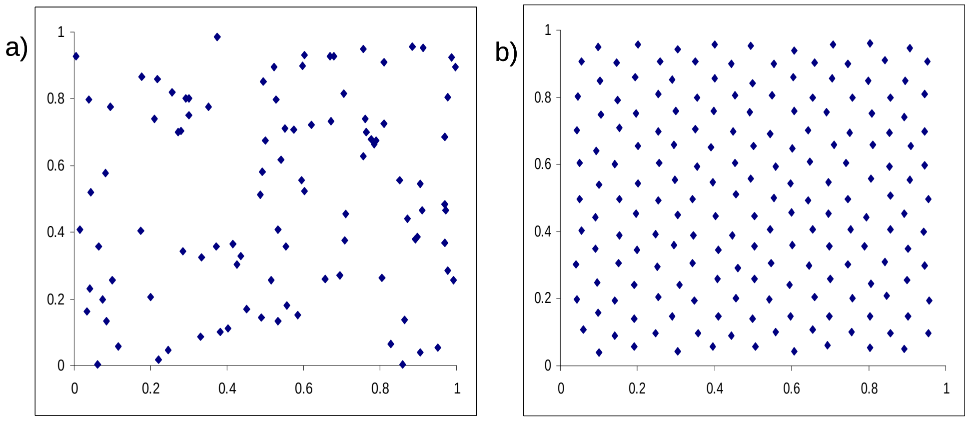


Figure 5.2: Two possible methods for searching the estimation space (the surface area of the n-sphere). (a) uses randomly produce points whilst (b) using a low-dispersion algorithm to generate maximally avoiding points. The latter method provides a more even coverage of the estimation space.

5.2.4.4 Local Search Optimization

We are required to produce a local optimization algorithm for use in the Bionomic Algorithm's steps 1.2a, b and 3c. The benefit of formulating the problem on the hypersphere is that by restricting our search for solutions on the surface of a hypersphere rather than inside or outside the hypersphere, we generate only feasible solutions, i.e. strategy pair weights that satisfy:

$$x_u^{agg} \geq 0 \ , \ \forall u \tag{5.7}$$

$$\sum_{u=1}^{u=n} x_u^{agg} = 1 \tag{5.8}$$

therefore we do not need to perform step 1.2a of the BA (see Algorithm 5.1 on page 152. Our local optimization procedure therefore becomes a search across the surface of the sphere and we do this by performing 'small' angular rotations around an axis perpendicular to two dimensions (or state vector components or basis vectors in hyperspace) at a time as shown in Figure 5.3, rejecting the rotation if it doesn't lead to any improvement in the objective function (the hit-rate in our case), or accepting the rotation and continuing in the same angular direction until it stops improving the objective function. This is then repeated in the other possible dimension pairs.

From Figure 5.3, it can be seen that the change in the state vector components are related in the following fashion:

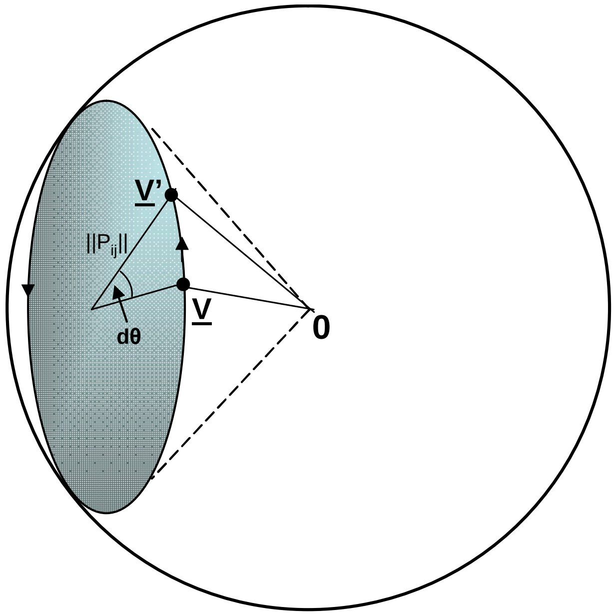
$$\delta i = -2 \cdot \parallel p_{ij} \parallel \cdot \sin(\frac{\delta\theta}{2}) \cdot \sin(\theta + \frac{\delta\theta}{2}) \tag{5.9}$$

$$\delta j = 2 \cdot \parallel p_{ij} \parallel \cdot \sin(\frac{\delta\theta}{2}) \cdot \cos(\theta + \frac{\delta\theta}{2}) \tag{5.10}$$

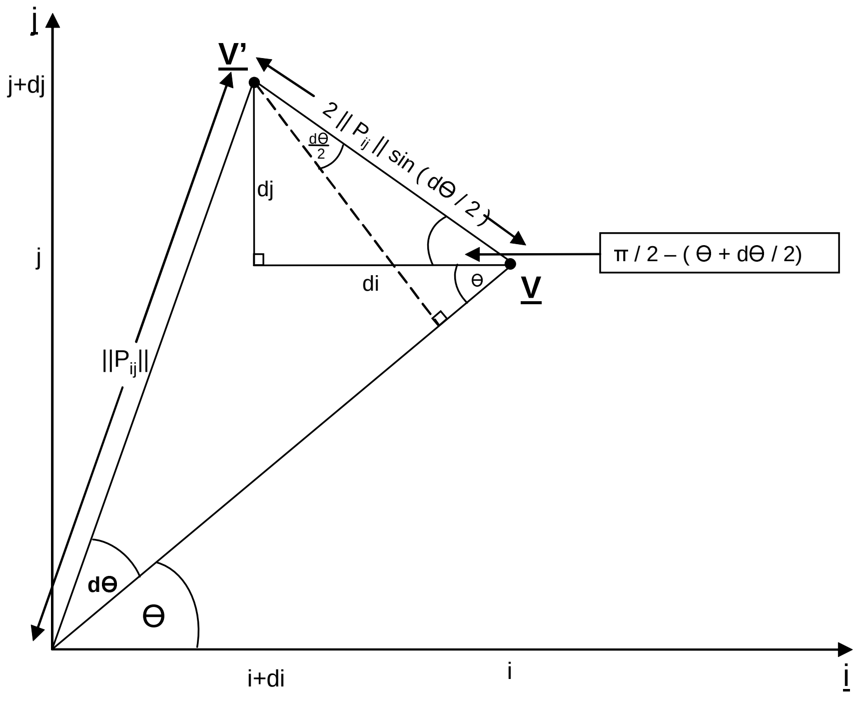
where δi and δj are the change in state vector component i and j respectively, p_{ij} is the projection of the initial state vector onto the plane defined by the $i - j$ axes before the local optimization procedure is executed,$\parallel p_{ij} \parallel$ is the length of this projection which remains constant during the rotation under the local optimization procedure, θ is the angle between the initial state vector projection p_{ij} and the i axis, and $\delta\theta$ is the angular step size that we rotate p_{ij} in both angular directions, keeping all orthogonal components of the state vector constant.

Initial studies with $\delta\theta = \frac{\pi}{50}$ radians showed very little change in the objective function (hit-rate) with only a small minority of dimension searches displaying an improvement, however this was improved by increasing the angular step size to $\delta\theta = \frac{\pi}{10}$ radians.

Originally, we scanned through all possible dimension pairs, however this method achieved approximately six state vector optimizations per hour for $n = 72$ ($m = 3$) and a data set of size 500 points. Considering it is recommended [Christofides et al. 2003] that the population of solution state vectors should be of the order of the dimensionality n, and that we also need to perform this optimization for the best child solutions produced from each parent set (BA step 3c) - early executions of the computer program created around 20 parent sets per iteration of the child creation procedure, with around 70 iterations in total - this would require around 1500 optimizations, taking around 10 days to complete. If one were interested in training up the Minority Game Forecasting Framework in this way on tick data in order to make investment



(a) 3 dimensional representation of hypersphere



(b) 2 dimensional cut through the hypersphere

Figure 5.3: The local optimization procedure involves rotating the state vector V around the axis perpendicular to the plane defined by basis vectors $\hat{i}$ and $\hat{j}$ (each corresponding to a square root aggregate weight of a strategy pair). The procedure rotates the state vector V to V' in small angular steps $d\theta$ such the projection $i \rightarrow i + di$ and $j \rightarrow j + dj$, whilst maintaining a constant state vector projection $\parallel P_{ij} \parallel$ onto the i-j plane.

decisions over the near time-horizon, this may prove to be an unreasonable amount of time. Therefore we simplify the method and rather than running through all dimensions, we randomly pick two dimensions and attempt the local optimization method, and repeat this until a maximum number of times, which we chose to be equal to the number of dimensions n.

5.2.4.5 Generate parent solution sets

Step 2 of the BA requires us to construct a set of existing state vectors from which we will produce children to be accepted or rejected into the population. In most Genetic Algorithms, parent sets only have a cardinality (set size) $k = 2$, however the Bionomic Algorithm can have an arbitrary cardinality $k \geq 2$, subject to the criteria that the set elements have high 'distinctiveness' between them. This larger cardinality and 'distinctiveness' parameter is a strength of the BA over Genetic Algorithms, as it means that we can enforce a wider search area when generating new children than conventional Genetic Algorithms. As Christofides et al. argues, this spanning property is "perhaps the most important component of the BA... [as] the larger the part of the solution space spanned by the parent sets, the less likely it is for the BA search to be 'stuck' at a local optimum." [Christofides et al. 2003] Beasley concurs, claiming that with children produced by parents too near each other, we are effectively exploring the same space that led to the maturation of the current parents, adding that the 'distinctiveness' parameter controls the degree of diversification [Beasley 1999].

Whereas in Euclidean space we usually use the length (2-norm) between two points to measure the distance or 'distinctiveness' of two state vectors solutions, it is natural to think of the distance between two points on a hyper-sphere using the arc length. As the arc length has a monotonically increasing, injective (one-to-one) relationship with the traditional Euclidean distance, the Euclidean distance is still a valid way of describing distances on the hyper-sphere. However, as we would like to allow for the possibility of linear additivity of our 'distinctiveness' measure, and so that the step size of our local optimization procedure is also linearly additive (if we double the step size, this is equivalent to moving the initial step size twice), we choose to use the arc length measure. The equation to calculate the arc length l between two points on the surface of a sphere is:

$$l = r \cdot \phi \tag{5.11}$$

where r is the radius of the hypersphere and ϕ is the angle between the two points (measured on the plane defined by the two state vectors and the origin). Our hypersphere has radius equal to 1, so the arc length becomes simply $l = \phi$. As seen in Figure 5.4, we can form a plane defined by the two state vectors and the origin, and use trigonometry to determine the angle ϕ in the following way:

$$\phi = 2 \arcsin \left(\frac{\parallel \underline{d} \parallel}{2} \right) \tag{5.12}$$

where $\underline{d}$ is the vector difference between the two state vectors $\underline{V_A}$ and $\underline{V_B}$ ($\underline{d} = \underline{V_A} - \underline{V_B}$) and $\parallel \underline{d} \parallel$is the standard Euclidean metric (equation 5.13)

$$\parallel \underline{d} \parallel := d(V_A, V_B) = \sqrt{\sum_{i=1}^{i=n} \mid V_{A_i} - V_{B_i} \mid^2} \tag{5.13}$$

between these two points lying on the hypersphere.

We now have a measure of the distinctiveness of the state vector solutions with which to construct parent sets. The construction starts with a null set, picking the first state vector solution - the 'principal state vector' - to go in the set, and adding to this set all other state vectors solutions that have a larger arc length l from state vector solutions already in the set than a certain arc length threshold $l_{threshold}$. Thus we create a maximally spanning set, or in graph theory terminology a maximally independent set. Indeed, there are efficient algorithms to generate such maximally independent sets of a graph [Bron, Kerbosch 1973, Christofides 1973], however our method to accomplish this task is simple to understand and particularised to our problem in question. We choose $l_{threshold}$ to be at a specific quantile of the arc length separations of the existing population of solutions $\mathbf{V}$, namely the median, although we could ensure greater diversity of the parent set by increasing $l_{threshold}$, however this would also have the effect of creating a greater number of smaller parent sets with which to create children, thus potentially restricting the variability

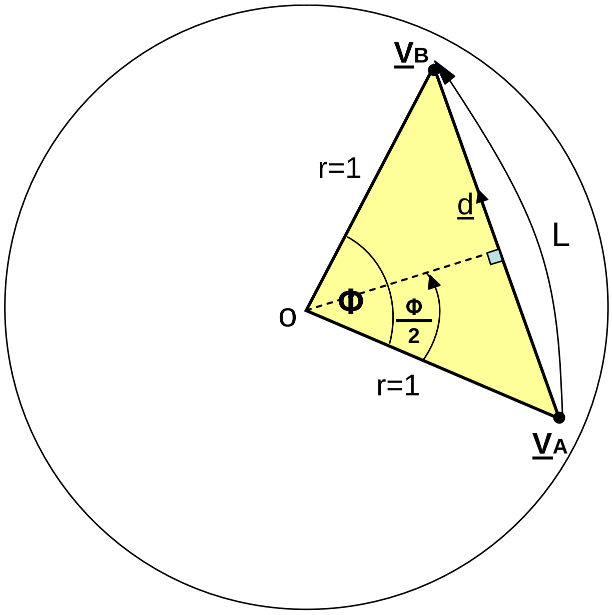


Figure 5.4: The arc length l between two state vector solutions V_A and V_B can be calculated using trigonometry on the plane defined by the two state vectors and the origin of the hypersphere. As the radius of the n-sphere is 1, $l = \phi$

of the children (as fewer parents means a smaller span over the hyper-surface of the n-sphere). By setting $l_{threshold}$ as the median arc length, we typically produce around twenty to thirty parent sets containing around five to ten state vector solutions from an initial parent population of 72. Figure 5.5 on the following page is a histogram showing the typical distribution of the arc lengths between an initial population of locally optimized state vector solutions (before generating children) - the minimum arc length is 1.13 radians, the median is 1.57 radians, and the maximum is 2.00 radians.

There are several ways to pick a principal state vector - the first solution to go in each parent set - such as picking with a probability of inclusion related to the solution's fitness (in this case its objective value, the hit-rate). However in this case we would also have to produce a mechanism for ensuring that if we picked the same principal state vector in two parent sets, we wouldn't repeat the same procedure of picking other state vectors and end up with two identical parent sets. This could be achieved by avoiding adding the same state vector solution to be the second parent in the set, and / or by adding some stochastic mechanism to the order in which we test the other state vectors to

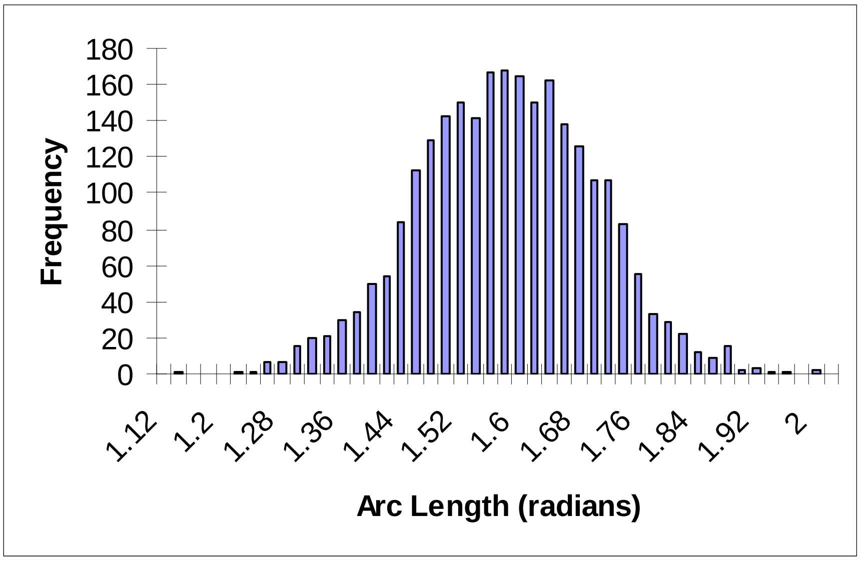
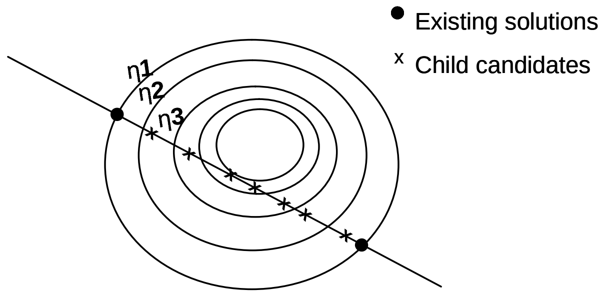
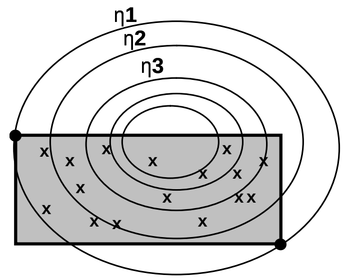


Figure 5.5: Histogram of arc lengths l in initial (locally optimized) population. The dispersion of state vector solutions follows an approximately normal distribution. As the radius of the n-sphere is 1, $l = \phi$.

see if they have a large enough arc length from the existing state vectors in the set. However, we go for a simple sequential procedure in which we pick the fittest state vector solution to be the principal state vector in the first parent set S_1, then move to the second fittest state vector and, on the condition that it was not included in the first parent set, use it as the principal state vector of the second parent set S_2. If it was included in the first parent set S_1, then we skip this solution and move to the next fittest solution that was not included in the first parent set, in order to prevent much overlap between the two parent sets. We extend this to further parent sets S_q, where we pick the fittest state vector solution that has not already been included in any of the previous parent sets, continuing this until a certain proportion of the population have been chosen as principal state vector solutions to start a parent set and we define $S := S_{max}$. In our studies, to explore as wide a search area as possible, we chose to use all of the population, so we also included the worst performing (lowest hit-rate) state vector as a principal state vector. This is reasonable, as the arc length threshold $l_{threshold}$ ensures that all other solutions in the parent set will be far enough away from this principal vector that we are likely to get some good solutions in this parent set, and thus are not restricting the search



(a) Linear line search



(b) Hypercube search

Figure 5.6: Two methods of combining parents to generate children in Euclidean space, with the contours $\eta_{1,2,3}$ signifying lines of constant objective value.

to a poor area of the solution space.

5.2.4.6 Generating a pool of children

As mentioned previously, Genetic Algorithms are generally restricted to parent sets with cardinality $k = 2$. This restricts the variability of child solutions, as children are often generated as points on a straight line (in Euclidean space) or arc length (in hypersphere space), however Karavias also highlights the method of using the two parents to define the corners of a hypercube volume within which new children are generated [Karavias 2006]. Figure 5.6 demonstrates the two methods in Euclidean space.

The latter hypercube method (b) doesn't easily generalise to the hypersphere space, however we can develop our own method benefiting from the increased cardinality $k_q \geq 2$ of the parent sets in the Bionomic Algorithm. Thus we define a hyperplane on which to generate new points from an initial state vector V_A^q (in order to move onto a point on the hyperplane) and the $k_q - 1$ difference vectors $\underline{d_{AB}^q}$ between the initial state vectors $\underline{V_A^q}$ and all other state vector solutions $\left\{\underline{V_B^q}\right\}$ in the parent set S_q. See Figure 5.7 for a $k = 3$ representation, that can be generalised to higher cardinality.

We use a method analogous to Scatter Search [Glover 1994, Glover 1995] [Fleurant et al. 1996, Glover 1997], where children are constructed from linear

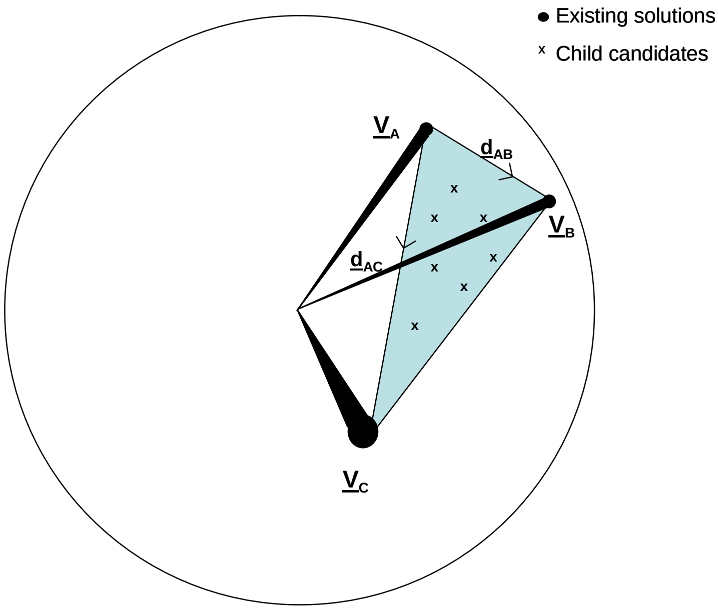


Figure 5.7: Hyperplane defined by difference vectors d_{AB} and d_{AC} between parent state vectors $\{V_A, V_B\}$ and $\{V_A, V_C\}$ respectively. The difference vectors also determine the boundary within which child candidate solutions are created.

combinations of members of the parent set, by making randomly scattered points on this hyperplane. We could weight the members of the parent set, such that parents with highest fitness have stronger weighting leading to a cluster of points generated around the fittest parents. However, in order to search the space within the area on the hyperplane as thoroughly as possible looking for optima, we generate $k_q - 1$ uniformly distributed random variables U_{AB} of interval $[0, 1]$, one for each difference vector d^q_{AB} between the initial state vector point V^q_A and the other state vectors V^q_B in the parent set, where $A < B \leq k_q$, to specify how far along each difference vector d^q_{AB} we have to move from the initial state vector to get to the point. We also multiply by an enlargement factor f_{en} to allow points to lie on the hyperplane outside the boundary of the parent points; we choose $f_{en} = 1.2$, so that points aren't too far away from the existing parent state vector points.

The newly generated points on the hyper-plane do not generally lie on the surface of the hypersphere of state vectors, lying within parent state vectors inside the sphere in the event that $f_{en} \cdot U_{AB} < 1$ or outside the sphere in the event that $f_{en} \cdot U_{AB} > 1$. These points must therefore be normalised to bring them back to the surface of the hypersphere. If we consider the vector $\underline{v}$ that connects the origin of the hypersphere to a newly generated point on the hyperplane, we can normalise them by simply dividing by the length (2-norm) $\| \underline{v} \|$,so that the newly scaled points are guaranteed to have length equal to 1. So for the w-th new child vector $\underline{C_w^q}$ generated by parent set q created:

$$\underline{C_w^q} = (v_1 + v_2 \ldots + v_{n-1} + v_n)^{\frac{-1}{2}} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{bmatrix} \tag{5.14}$$

We must choose how many child solutions to create in each parent set. Typically we decided to set the child-parent ratio to 5 as generating children and evaluating their objective function value (hit-rate) is a fairly fast procedure compared to, for example, the local optimization search procedure, and ensures a more thorough search on the hyperplane.

5.2.4.7 Maturation of the best child solutions

As dictated in steps 3b and 3c of the Bionomic Algorithm, for each parent set S_q, we rank the pool of child solutions generated in $\underline{C_w^q}$, picking only the fittest, i.e. the one with the highest objective value and perform the same local optimization search as in subsection 5.2.4.4 in order to improve the solution so that the child solution is perturbed to become $\underline{C_{w*}^q}$. The resulting set of locally optimized best child vectors $\left\{\underline{C_{w*}^q}\right\}$ are then put forward as candidates to replace the existing population of state vector solutions $\mathbf{V}$.

We note there is some flexibility to choose more than just the fittest child solution from each parent set as candidates to replace the existing population, such as a certain quantile the child population, however this could cause there to be many similar children in the candidate set $\left\{\underline{C_{w*}^q}\right\}$, which if accepted into the existing population could cause premature convergence of the search over the surface of the hypersphere.

5.2.4.8 Acceptance criteria for the best child solutions

Step 4 of the Bionomic Algorithm stipulates that we compare the objective values (hit-rates) of the child candidates with a sub-list of the population and replace the most dominated member, however as Karavias attests, we must be careful to avoid adding too many similar child candidates to the population **V** [Karavias 2006]. This can be achieved by calculating the arc lengths l^q_{Aw*} between the candidate new solution $\underline{C^q_{w*}}$ of parent set q and all the existing state vector solutions to find the nearest existing solution $\underline{V_{A*}}$. We then compare the objective solution between $\underline{V_{A*}}$ and $\underline{C^q_{w*}}$, and replace the existing solution in **V** if the objective value η (hit-rate) of C^q_{w*} fulfils the following criteria:

$$\eta\left(C^q_{w*}\right) > \eta\left(V_{A*}\right) \tag{5.15}$$

As seen in step 4b of the Bionomic Algorithm, the existing solution can also be replaced regardless of the objective value comparison if an age criterion is satisfied. The age Ω of V_{A*} measures the number of 'iterations' of the Bionomic Algorithm (the number of times we evolve the population by creating child candidates and offering them up to replace the existing population) that V_{A*} has existed in the population of state vector solutions **V**. If this age $\Omega\left(V_{A*}\right)$ is above a certain critical age threshold $\Omega_{critial}$:

$$\Omega\left(V_{A*}\right) \geq \Omega_{critical} \tag{5.16}$$

we also accept the child candidate to replace the existing state vector solution V_{A*} in **V**. As Christofides explains, the reason for replacing an 'old' solution with a child that does not necessarily dominate it is to avoid premature convergence of the Bionomic Algorithm, in the case that the same state vector solutions are repeatedly chosen to form parent sets, but their offspring are continuously rejected [Christofides et al. 2003].

If neither equation 5.15 or equation 5.16 are satisfied, we reject the candidate child solution. We note that to increase the frequency of child candidates being accepted, if the nearest solution dominates the child, we could test other solutions near to the child such as the second, third, fourth nearest existing solutions etc., until a certain proportion of the total population has been tested, stopping if the child dominates the existing solution. However, this would lead to more similar solutions and thus potentially causing the search space to converge prematurely.

We then repeat this procedure for all the other child candidates in the set $\left\{\underline{C^q_{w*}}\right\}$. This marks the end of one iteration of the Bionomic Algorithm.

5.2.4.9 Stopping criteria

Steps 2, 3 and 4 of the Bionomic Algorithm - the procedure of creating parent sets, locally optimizing the candidate child solutions and offering them up to replace the current population of state vector solutions - are repeated until either of the following stopping criteria have been met:

- The replacement rate RR, defined as the proportion of child candidates accepted to replace existing state vector solutions and thus calculated as the number of replacements divided by the number of parent sets q, falls below a certain threshold rate, $RR_{threshold}$. We typically choose $RR_{threshold} = 0.05$, so that at least a twentieth of the children need to be accepted. Our parameters and procedure generally creates around twenty to thirty parent sets, so this is equivalent to demanding that at least one child is accepted every iteration.
- A maximum number of iterations (creation of parent sets and population evolution etc.) has been reached. Typically given the value of our other parameters, we choose the maximum to be 100 iterations as almost all runs of the procedure have finished due to the replacement rate threshold $RR_{threshold}$ being breached before reaching this maximum iteration limit.

5.3 Method

We add to the parameter set $\{m,\ T,\ c,\ \tau\}$ another parameter to examine - the population size of the state vector solutions SP. For a specific set of parameters $\{m,\ T,\ c,\ \tau,\ SP\}$, we run the Bionomic MGFF over a sample of data (the in-sample data), usually around 500 to 900 time-steps in length in order to calculate the hit-rates and relative hit-rates of the population of vector solutions. Using the BA to evolve the population, the BA finishes when it reaches a stopping criteria, and returns the best state vector solution of the most recent iteration - the state vector that resulted in the highest in-sample hit-rate κ^{in}. The best solution is never replaced during the BA unless it is dominated by a child solution, so this best solution is guaranteed to be optimal state vector solution found during the whole analysis. Using this optimal state vector that has been calibrated on the in-sample data, we then test this solution on out-of-sample data of approximately equal length, to determine the out-of-sample hit-rate κ^{out}.

We then repeat the whole process for a different set of $\{m, T, c, \tau, SP\}$, looping through the parameters such that we cover a range of parameter space. Finally we report the solution and parameter set that achieved the highest hit-rate measure. To increase the speed of our computation, we chose not to extend the forecasting to multiple time-steps, so we keep $e = 1$ and only focus on the 1-step ahead hit-rate $\kappa_1 := \kappa$, although we note the possibility of relaxing this restriction in future work. Our measurement of the hit-rates is taken over the moving window of $M = 100$ time-steps.

In order to ensure consistency in our average out-of-sample hit-rate and standard deviation, we require all moving hit-rates used to estimate the sample average to be calculated over a constant sample size. We therefore ignore the first few moving hit-rate calculations up to a number of data points equivalent to the length of the sample size. The first few time-steps after the start of the out-of-sample hit-rate calculations give rise to particularly volatile hit-rate measures, being that they are calculated over only a few time-steps. Thus we only derive our average hit-rate measures presented in the following results tables after allowing there to be an equivalent number of data points for each moving hit-rate.

	κ	$\widehat{\sigma}$	$\frac{\kappa-\kappa_0}{\widehat{\sigma}}$	p-value	κ^{rel}	$\widehat{\sigma^{rel}}$
in-sample	0.524	0.047	0.51	0.32	1.05	0.14
out-of-sample	0.526	0.054	0.48	0.30	1.07	0.136

Table 5.3: Results of Bionomic MGFF with data from Stochastic Phase of a matched Minority Game. κ is the observed hit-rate, $\widehat{\sigma}$ is the observed standard error, $\kappa_0 = 0.5$ is the null hit-rate, $\frac{\kappa-\kappa_0}{\widehat{\sigma}}$ is the separation of means criterion, p-value is the probability of achieving the observed hit-rate or larger given the null-hypothesis is true, κ^{rel} is the observed relative hit-rate and $\widehat{\sigma^{rel}}$ is the observed standard error of the relative hit-rate.

5.4 Results

5.4.1 Applications to the Minority Game

Using data generated by a matched Minority Game ($m = 3$, $T = 100$, $c = 0.53$, $\tau = 1\}$ in the periodic phase, we obtain an in-sample hit-rate $\kappa^{in} = 1$ and an optimal out-of-sample hit-rate of $\kappa^{out} = 1$, with many of the population of solutions giving similarly high hit-rates. This suggests that our method is indeed able to recover the correct strategy pair weights, such that we can predict the Minority Game data to perfection. With the stochastic phase of a matched Minority Game, we achieve the results as displayed in Table 5.3, similar to that discussed in Section 4.4.1 on page 125, where there appears to be no statistically significant evidence of predictive power. Again, this is likely to the Minority Game system being in the phase where the two agent group populations are frequently equal in number, thus the system is dominated by stochastic 'coin-tosses' to resolve this.

5.4.2 Applications to Financial Data

For details of the assets, see Tables 4.3 on page 119 and 4.6 on page 128 in Chapter 4. Note that in comparison to the parameter set used in the Simple and Expanded MGFF using the Kalman Filter, we have the solution population size SP as an additional parameter, which we set as a fraction of the total number of dimensions of the hyperspace, usually from the lowest fraction of half the number of dimensions to the largest of twice the number of dimensions. This is to reflect the dependence of the area of the hypersphere on the number of dimensions - if we increase the value of m and thus the number of dimensions, we also increase the surface area of the hypersphere[5]. Therefore we need more solutions to ensure a good search across the surface area.

We initially choose to investigate Copper Futures (LMCADY <Cmdty>) due to the significance of the results we obtained by applying the Simple and Expanded MGFF to this asset in Chapter 4. This particular asset is also interesting in that it was found to give rise to consistency in the optimal parameter set as can be seen in Table 4.9 on page 135, with both the Simple and Expanded MGFF choosing $\{m = 4, T = 50, c = 0.52, \tau = 0.9\}$ as the optimal parameter set. We restrict our search of the agent memory parameter m in this investigation of the Bionomic MGFF to values $m < 4$ owing to the computational time constraints involved in such a large parameter search (there are 272 hyperspace dimensions when m=4 and this increases exponentially with m, as can be seen in Table 5.2 on page 157). It is hoped that using sub-optimal values of m will not make a significant difference on the outcomes of our analysis; indeed, we found evidence to support this assumption in the Kalman Filter case as discussed in Section 4.4.1 on page 125. We therefore choose to explore the parameter set $\{m = 3, T, c, \tau = 0.9\}$ over different values of T and c, in order to assess the impact of these parameters. We present the 1-step hit-rate and relative hit-rate achieved using the Bionomic MGFF applied to daily Copper Futures data from 02/12/92 to 19/09/00 ('Copper Fut a') in Table 5.4a, and with a later time-period from 20/09/00 to 22/08/08 ('Copper Fut b') in Table 5.4b on the following page. The data sample length consists of approximately 900 in-sample points and 900 out-of-sample points, though as we remove repeated prices, the exact number may be slightly lower.

We also apply the Bionomic MGFF to other assets. In doing so, we find that the parameter set $\{m, T, c, \tau, SP\}$ that produces the best in-sample hit-rate does not necessarily produce the best out-of-sample hit-rate. Therefore we choose to present the parameter set that results in the best out-of-sample hit-

[5]Using the convention of Geometers, where a circle is a 2-sphere and a normal sphere is a 3-sphere, the surface area of an n-sphere of unit radius is $\frac{2\pi^{n/2}}{\Gamma(n/2)}$

κ_{in}	$\widehat{\sigma_{in}}$	$\frac{\kappa_{in}-\kappa_0}{\widehat{\sigma_{in}}}$	p-val	κ_{out}	$\widehat{\sigma_{out}}$	$\frac{\kappa_{out}-\kappa_0}{\widehat{\sigma_{out}}}$	p-val	κ_{in}^{rel}	$\widehat{\sigma_{in}^{rel}}$	κ_{out}^{rel}	$\widehat{\sigma_{out}^{rel}}$	$\{m, T, c, \tau, SP\}$
0.535	0.046	0.76	0.24	0.509	0.060	0.15	0.43	1.20	0.16	1.00	0.16	{3, 50, 0.52, 0.9, 36}
0.547	0.036	1.31	0.17	0.506	0.054	0.11	0.45	1.22	0.14	1.00	0.16	{3, 50, 0.52, 0.9, 108}
0.542	0.044	0.95	0.20	0.497	0.047	-0.06	0.52	1.22	0.13	0.973	0.13	{3, 100, 0.51, 0.9. 36}
0.535	0.042	0.83	0.24	0.522	0.041	0.54	0.33	1.21	0.18	1.03	0.14	{3, 100, 0.51, 0.9. 72}
0.531	0.047	0.66	0.27	0.516	0.048	0.33	0.37	1.20	0.19	1.01	0.15	{3, 100, 0.51, 0.9. 108}
0.527	0.068	0.40	0.29	0.479	0.079	-0.27	0.66	1.18	0.15	0.93	0.13	{3, 100, 0.52, 0.9, 36}
0.529	0.076	0.38	0.28	0.468	0.077	-0.42	0.74	1.19	0.16	0.91	0.13	{3, 100, 0.52, 0.9, 72}
0.527	0.075	0.36	0.29	0.469	0.074	-0.42	0.73	1.19	0.17	0.91	0.13	{3, 100, 0.52, 0.9, 108}

(a) Copper Fut a: Period 02/12/92 to 19/09/00

κ_{in}	$\widehat{\sigma_{in}}$	$\frac{\kappa_{in}-\kappa_0}{\widehat{\sigma_{in}}}$	p-val	κ_{out}	$\widehat{\sigma_{out}}$	$\frac{\kappa_{out}-\kappa_0}{\widehat{\sigma_{out}}}$	p-val	κ_{in}^{rel}	$\widehat{\sigma_{in}^{rel}}$	κ_{out}^{rel}	$\widehat{\sigma_{out}^{rel}}$	$\{m, T, c, \tau, SP\}$
0.551	0.069	0.74	0.15	0.522	0.037	0.59	0.33	1.23	0.12	1.09	0.13	{3, 50, 0.51, 0.9, 72}
0.536	0.035	1.03	0.24	0.534	0.038	0.89	0.25	1.20	0.12	1.12	0.17	{3, 50, 0.51, 1, 72}
0.538	0.067	0.57	0.22	0.512	0.055	0.22	0.41	1.20	0.16	1.08	0.16	{3, 50, 0.52, 0.9, 36}
0.545	0.076	0.59	0.18	0.504	0.054	0.07	0.47	1.22	0.18	1.06	0.18	{3, 50, 0.52, 0.9, 72}
0.574	0.087	0.85	0.07	0.545	0.082	0.55	0.18	1.29	0.18	1.15	0.24	{3, 50, 0.52, 0.9, 108}
0.545	0.065	0.69	0.18	0.492	0.059	-0.14	0.56	1.22	0.16	1.03	0.16	{3, 50, 0.52, 0.9, 144}
0.548	0.054	0.89	0.17	0.513	0.027	0.48	0.4	1.23	0.11	1.07	0.11	{3, 50, 0.52, 1, 72}
0.571	0.089	0.80	0.08	0.547	0.083	0.57	0.17	1.28	0.18	1.15	0.24	{3, 100, 0.52, 0.9, 36}
0.573	0.088	0.83	0.07	0.536	0.086	0.42	0.24	1.29	0.19	1.13	0.25	{3, 100, 0.52, 0.9, 72}
0.569	0.086	0.80	0.08	0.542	0.084	0.5	0.2	1.28	0.19	1.14	0.24	{3, 100, 0.52, 0.9, 108}
0.573	0.088	0.83	0.07	0.536	0.083	0.43	0.24	1.29	0.19	1.13	0.24	{3, 100, 0.52, 0.9, 144}

(b) Copper Fut b: Period 20/09/00 to 22/08/08

Table 5.4: Hit-rates of Bionomic MGFF with daily Copper Futures (LMCADY) prices. κ is the observed hit-rate, $\widehat{\sigma}$ is the observed standard error where the superscript 'in' and 'out' refer to the in-sample and out-of-sample. $\kappa_0 = 0.5$ is the null hit-rate, $\frac{\kappa-\kappa_0}{\widehat{\sigma}}$ is the separation of means criterion, κ^{rel} is the observed relative hit-rate and $\widehat{\sigma^{rel}}$ is the observed standard error of the relative hit-rate. $\{m, T, c, \tau, SP\}$ is the parameter set used in the Bionomic MGFF.

rate over our small parameter search for each asset in Table 5.5b on the next page.

5.5 Observations

As can be seen from the data tables of 5.4 on the preceding page and 5.5b on the next page, despite most of our in-sample hit-rates and a majority of out-of-sample hit-rates being above the value of the null hit rate $\kappa_0 = 0.5$, almost none of the hit-rates we achieved were found to be statistically significant for either the in-sample or out-of-sample hit-rates, owing to the large values in the observed standard deviation of the hit-rates. The exceptions to this are the in-sample hit-rates of the Topix Real Estate(c) Index and the NASDAQ(d) Index, which are 1.72 and 2 standard deviations away from the null hit-rate respectively. For both cases however, the p-value is greater than 0.05 (0.16 for Real Estate(c) Index and 0.11 for NASDAQ(d) Index) meaning that there is only little evidence against the null hit-rate, such that we cannot rule out the possibility of these hit-rates being down to luck.

If we consider the null relative hit-rate as being the expected value of the naive prediction method, $\kappa^{rel} = 1$, then the average relative in-sample and out-of-sample hit-rates achieved with the USD-EUR(c) tick data does appear to be statistically significant, in that it is greater than 2 standard deviations away from the null relative hit-rate value. However this is most likely down to the relative weakness of the naive predictor in forecasting high-frequency financial data, most likely because of the bid-ask bounce effect, a result of impatient traders demand immediacy in their deals [Harris 2003]. This means the naive predictor's assumption that the next price increment will be in the same direction as the previous one is more often incorrect. However, we could also construct another 'mean-reverting naive predictor', that predicts the next price increment will be in the opposite direction to the previous one. This should not be so influenced by the bid-ask bounce, and it is likely that the new relative hit-rate that compares the success of our method with the mean-reverting naive predictor is unlikely to be statistically significant, in line with the normal values κ_{in} and κ_{out}.

Asset	Freq.	Time Period	Size	$\{m, T, c, \tau, SP\}$
Real Estate(c)	daily	20/08/98 - 27/01/03	450	{3, 100, 0.53, 1, 72}
Copper Fut(b)	daily	02/12/92 - 19/09/00	900	{3, 100, 0.51, 0.9, 72}
Copper Fut(c)	daily	20/09/00 - 22/08/08	900	{3, 100, 0.52, 0.9, 36}
Oil Fut(d)	daily	02/07/86 - 29/08/04	900	{2, 100, 0.52, 0.9, 20}
S&P Fut(f)	tick	13/08/07 - 14/08/07	700	{3, 100, 0.53, 1, 96}
NASDAQ(d)	daily	11/12/02 - 19/08/08	600	{3, 100, 0.53, 1, 20}
USD-EUR(b)	daily	01/01/99 - 26/10/04	700	{3, 100, 0.53, 1, 72}
USD-EUR(c)	tick	27/05/07 - 28/05/07	900	{3, 100, 0.53, 120}
USD-JPY(b)	daily	02/01/90 - 07/12/95	650	{3, 100, 0.51, 1, 72}

(a) Data sample details and optimal parameter set $\{m, T, c, \tau, SP\}$

Asset	κ_{in}	$\widehat{\sigma_{in}}$	$\frac{\kappa_{in}-\kappa_0}{\widehat{\sigma_{in}}}$	p-val	κ_{out}	$\widehat{\sigma_{out}}$	$\frac{\kappa_{out}-\kappa_0}{\widehat{\sigma_{out}}}$	p-val	κ_{in}^{rel}	$\widehat{\sigma_{in}^{rel}}$	κ_{out}^{rel}	$\widehat{\sigma_{out}^{rel}}$
Real Estate(c)	0.550	0.029	**1.72**	0.16	0.481	0.031	-0.61	0.65	1.10	0.11	0.91	0.04
Copper Fut(b)	0.535	0.042	0.83	0.24	0.522	0.041	0.54	0.33	1.21	0.18	1.03	0.14
Copper Fut(c)	0.571	0.089	0.80	0.08	0.547	0.083	0.57	0.17	1.28	0.18	1.15	0.24
Oil Fut(d)	0.552	0.096	0.54	0.15	0.550	0.107	0.47	0.16	1.08	0.10	1.13	0.21
S&P Fut(f)	0.510	0.038	0.26	0.42	0.503	0.042	0.07	0.48	1.14	0.19	1.05	0.16
NASDAQ(d)	0.562	0.031	**2.00**	0.11	0.488	0.026	-0.46	0.59	1.19	0.11	0.99	0.10
USD-EUR(b)	0.551	0.057	0.89	0.15	0.496	0.049	-0.08	0.53	1.10	0.14	1.12	0.11
USD-EUR(c)	0.511	0.045	0.24	0.41	0.482	0.063	-0.29	0.64	1.70	0.31	1.83	0.37
USD-JPY(b)	0.530	0.04	0.75	0.27	0.51	0.030	0.33	0.42	1.10	0.14	1.15	0.11

(b) Hit-rates, relative hit-rates and statistical significance values

Table 5.5: Bionomic MGFF applied to a range of assets. κ is the observed hit-rate, $\widehat{\sigma}$ is the observed standard error where the superscript 'in' and 'out' refer to the in-sample and out-of-sample. $\kappa_0 = 0.5$ is the null hit-rate, $\frac{\kappa-\kappa_0}{\widehat{\sigma}}$ is the separation of means criterion, κ^{rel} is the observed relative hit-rate and $\widehat{\sigma^{rel}}$ is the observed standard error of the relative hit-rate. $\{m, T, c, \tau, SP\}$ is the optimal parameter set used in the Bionomic MGFF that produced the highest (in-sample) hit-rate values. The number of in-sample and out-of-sample calculations are each approximately equal to the 'Size' column in (a)

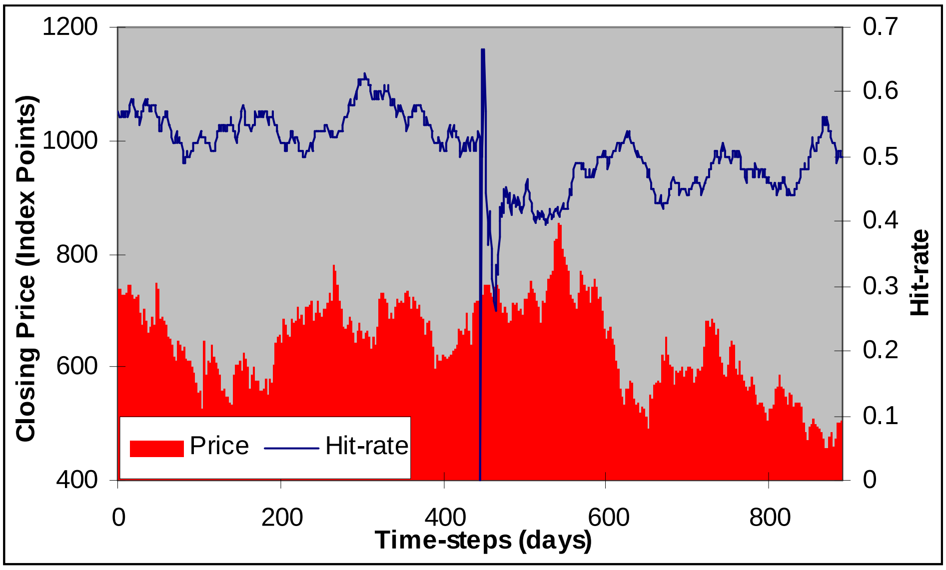


Figure 5.8: Hit-rate of daily Topix Real Estate(c) data

We also note that the in-sample hit-rates are usually higher than the out-of-sample hit-rates, sometimes by as much as 6-7% as in the Real Estate(c) and NASDAQ(d) cases, but also for the USD-EUR(b) daily rate. An example of the dramatic change in behaviour from the in-sample hit-rate to the out-of-sample hit-rate is illustrated in Figure 5.8 for Real Estate(c) data from 20/08/98 to 27/01/03, with the crossover marked by increased volatility in the hit-rate around time-step 440, owing to the sudden shortening of the sample window over which it is calculated, as we remove the data from the in-sample time period. Note that the correlation between price and hit-rate $\rho_{\kappa-price} = 0.22$ for the in-sample period and $\rho_{\kappa-price} = -0.05$ for the whole period (in-sample and out-of-sample).

5.6 Discussions and Future Work

In Section 4.5.3 on page 138, we discussed some studies by other authors using different techniques to make directional forecasts of various assets. Although one study found significant hit-rates by applying feed-forward neural network models to design non-parametric trading strategies to forecast the sign changes in daily Dow Jones Industrial Average data [Gencay 1998], and another by using logit, probit, probabilistic and feed-forward neural network models to monthly S&P500 data [Leung et al. 2000], there are plenty of counter-examples of techniques that did not produce statistically significant directional forecasts at the 95% level. Such examples in these studies include discriminant analysis, adaptive exponential smoothing, vector auto-regression with Kalman filter and multivariate transfer function techniques applied to S&P500 monthly data. In particular [Kuan, Liu 1995] finds little significance by applying different versions of feed-forward and recurrent neural networks to foreign exchange data. These statistically insignificant methods usually involve smaller numbers of observations compared to our Bionomic MGFF study, therefore although our results are comparable to theirs in this regard, it may be that the Bionomic MGFF is more suitable to use over shorter periods of time. Certainly, with the results we have achieved in this thesis, the Simple and Expanded MGFF using the Kalman Filter estimation procedure appear more promising.

The inability to make long-term predictions using the Bionomic MGFF could suggest the existence of time-varying unexplained variables in the system. In particular, the assumption that the strategy pair weights are constant over the period from in-sample to out-of-sample seems to be unsupported by these results. It may be that we should change our focus from analysing long term behaviour of the market, with an out-of-sample length equivalent to the in-sample length, to instead considering the hit-rate success immediately after the end of the in-sample period. Certainly, if one were hoping to use the Bionomic MGFF to predict short term market movements in order to execute trades frequently, the long term behaviour would not be important. In fact, our research may be useful in that it shows the limits of this method for long-term prediction. There is therefore the hope that by focusing on the short-term predictions, we may be able to improve the hit-rates of the most poorly performing asset forecasts. We leave the analysis of the short-term predictions of the Bionomic MGFF for future investigation.

In this chapter, we have created a new particularisation of the Bionomic Algorithm suitable for the Minority Game, and we could easily extend the concepts to cater for other forms of agent-based models such as the Mix Game, and indeed any situation in which the Bionomic Algorithm is to be applied to estimate an equality constrained multi-variate quantity (a vector of constant length in hyper-space). We have also demonstrated a technique to find the optimal parameters set $\{m, T, c, \tau, SP\}$, which could be used to find any parameter set specific to the problem in question. In particular, we apply it to data generated by a 'Black Box' MG, and retrieve the correct parameters, before applying it to financial time-series data. Whilst we acknowledge that we have performed a parameter search over only a small range of the possible parameter space of $\{m, T, c, \tau, SP\}$ and have thus far found only statistically insignificant out-of-sample results, the examination of the Bionomic Minority Game Forecasting Framework is still in the early stages. There are several ways in which we can improve our method. In particular, we would endeavour to establish a better understanding of the relationship between the optimal hit-rates achieved and the parameter set $\{m, T, c, \tau, SP\}$, thus allowing us to direct the search across parameter space, to increase the quality of the solutions obtained, and to shorten the time it takes to find solutions that are 'good-enough' for our requirements.

With further work, we hope to investigate the general dependency of our technique on the parameter space, and in particular to increase the size of the parameter search, with a wider range of parameters $\{m, T, c, \tau, SP\}$ used, and to explore this space in finer detail by using smaller increments over the parameter range. We would also like to expand the parameter search space to include other parameters of the BA, such as the arc length threshold $l_{threshold}$, the critical age $\Omega_{critical}$, the child-parent ratio etc. As suggested in Section 4.6 on page 141, we could simplify the parameter space by changing the strategy score updating of the finite Time-horizon T version of the Minority Game to an iterative technique (discussed in 2.5.1 on page 55), thus creating the infinite Time-horizon version and removing the need to search over T.

We would also like to study in detail the effects of the parameters on the convergence to the optimal solution. This can be achieved by setting the replacement rate threshold $RR_{threshold} = 0$ and evaluating how long it takes the Bionomic Algorithm to find solutions with a specific minimum hit-rate, whilst varying the arc length threshold $l_{threshold}$, the critical age $\Omega_{critical}$, the child-parent ratio etc. This assumes we have some input data that is well described by the Minority Game, or by whichever agent game forms the basis of the Bionomic Algorithm calibration - in particular we could use data generated from a Minority Game simulation.

Also, we would like to use a 'low-dispersion' algorithm to generate maximally avoiding points on the surface of the n-dimensional hypersphere to improve initial parent solution coverage of the search space, thus hopefully increasing the likelihood of finding the optimal hit-rate. We could also take several data samples of a single asset at various points in time, using this to assess the change in the optimal strategy pair weights over time, analogous to the work in [Gou 2006c]. Indeed, our method has focused on long term predictive behaviour of the Bionomic MGFF, with the embedded assumption that the strategy pair weights change only very slowly over time. Our research has not provided evidence to support this claim, so we would like to focus on the short-term predictive ability, assessing the behaviour of the out-of-sample data over much shorter time-scales.

As in Chapter 4, we have merely focused on the hit-rate as we have been primarily interested in using this measure to find the optimal parameters of our model. However, we note the importance of the monetary gain possible using the Bionomic MGFF technique if one were to use it to trade in the market place. It would be necessary to construct suitable trading strategies to exploit any directional predictability found by the the Bionomic MGFF in an efficient manner (for example, to reduce the transaction costs). We leave this endeavour for future work.

5.7 Concluding Remarks

In this chapter we have introduced the Bionomic Algorithm, and have developed a particularization of it that can be specifically applied to the Expanded MGFF estimation problem, or more generally to any equality constrained (or conserved quantity e.g. the length of a vector) optimization problem. We applied this particularization of the Bionomic Algorithm to the calibration of the Minority Game Forecasting Framework to create a Bionomic MGFF. We used this Bionomic MGFF to examine the forecasting ability of MG simulation data and also the long term forecasting ability of data from a range of financial assets. We find little evidence of long-term predictability in the financial data, suggesting that we should focus on using the Bionomic MGFF to generate short-term forecasts in future work.

Chapter 6

Conclusions

6.1 Summary

This thesis has addressed questions relating to the use of agent-based models, in particular the Minority Game and Mix Game, and their role in forecasting financial markets. Specifically we have considered:

1. *the building of agent-based models of financial markets.* We have tested and validated the Minority Game and Mix Game, comparing their behaviour with results in the literature. We notice a phenomenon in the finite time-horizon version of Minority Game that has not been discussed in the literature to our knowledge: that there is an initial phase of the Minority Game simulation that is dominated by stochasticity, which we show is not amenable to our forecasting methodology of later chapters. After the stochastic phase, there is a cyclic phase, where a seemingly stochastic period of length equal to the Time Horizon T is continuously repeated. We later show that this cycle can be tracked using the Forecasting methodology that we develop. We perform a new analysis of the S&P500 Composite Index daily closing prices from the perspective of the Minority Game, and find evidence to suggests a forecasting method using Minority Game principles could have success in forecasting short-term movements in the S&P500 Index. We also outline the conceptual

framework in which to build a new model of an Oil market, using the principles of the Minority Game and Mix Game.

2. *the use of Kalman Filtering in a Minority Game Forecasting Framework.* We build a forecasting framework based on existing literature, creating two new variants of the existing framework by using the Reduced Strategy Space. One creation is a simpler version of the existing models in that it only estimates single strategy weights (the Simple MGFF). This simple version is computationally faster and therefore has the potential to find the optimal parameter set in a shorter time, and to process longer data sets in a reasonable amount of time. It can also handle larger values of the agent memory parameter m than the existing models, and therefore can examine a higher degree of complexity of patterns in the input data. We test this Simple Minority Game Forecasting Framework with data generated from a 'black-box' Minority Game engine to demonstrate the estimation of the correct Minority Game parameters and then apply it to the prediction of short-term price increments of financial data. We develop a second version which is similar to existing forecasting models in that it estimates strategy pair weights (the Expanded MGFF), and compare the results with the the Simple MGFF, finding that they give comparable results. We apply both versions to the analysis of new financial data sets.

3. *the use of Population Heuristics in a Minority Game Forecasting Framework.* We develop a novel particularisation of the Bionomic Algorithm that is suitable for use with the Minority Game Forecasting Framework. This particularisation could also be applied to the estimation of any equality constrained multi-variate quantity (a constant length vector in hyper-space), and could be extended to the Mix Game model. We build the Bionomic Minority Game Forecasting Framework, test it with data from a matched Minority Game, and apply it to long-term forecasting of financial markets. We find little evidence to support its use in predicting long-term price movements in financial market data, although a logical next step would be to test its use in short term predictions.

6.2 Future Developments

Throughout this thesis, we have raised a number of different areas for further research. These include:

1. Apply ensemble forecasting techniques to the Minority Game Forecasting Framework by running multiple versions of the Forecasting Framework with different parameter sets in a bid to improve the reliability of the predictions.

2. Investigating the short-term out-of-sample forecasting results of the Bionomic Minority Game Forecasting Framework to ascertain whether there is significant predictive power at a much shorter scale than to the time-scales we examined.

3. Examining the behaviour of the Bionomic Minority Game Forecasting Framework over a large range of parameters to better understand how both the Minority Game parameters and Bionomic Algorithm parameters influence the speed of convergence to the optimal state vector solution, and also the value of the optimal hit-rate.

4. Constraining the Kalman Filter to improve the speed of convergence and stability, or replace it with the Square Root Filter in order to improve numerical stability and accuracy.

5. Performing an exhaustive survey of the three versions of Minority Game Forecasting Framework discussed in this thesis over a wide range of markets, time periods, data frequencies to establish how general these techniques can be applied. We would also be able to investigate how the prediction successes, and thus applicability of such techniques, may have changed over time with the development of Algorithmic based trading, and also determine if there is an optimal frequency of data that should be used. We suggest a wider and finer search of the parameter space in order to increase the likelihood of finding the optimal hit-rate. We would also be in a better position to determine whether the Simple version can improve upon the results from the existing Expanded version.

6. Investigating how to make optimal use of the MGFF as a model to trade in the market place. Design trading strategies that make use of any predictive information provided by the MGFF in an efficient manner, such that transaction costs and other considerations are taken into account, and we maximise the potential monetary gain.

7. The creation of an Oil market model based on Minority Game principles, examining its behaviour over a range of parameters, and building a forecasting framework around it in order to forecast Oil commodities and derivatives, and the stocks of Oil companies.

8. Applying the Minority Game to the valuation of derivative contracts. Once we have calibrated a Minority Game to a specific market, a logical extension would be to use it in the pricing of derivatives such as Option contracts - this might be particularly useful for pricing high-dimensional problems such as path-dependent derivatives. We can use the Minority Game as an engine to drive the simulation of asset prices, analogous to existing Monte Carlo techniques for pricing derivatives. In this case, the magnitude of fluctuations is important, so the MG engine would need to be calibrated to exhibit the correct volatility for the underlying asset in question (see Figure 3.3 on page 66 for an indication on how this could be done). It may also be possible to tune the parameters to achieve a realistic value for the kurtosis of the price increments of the underlying asset, thus providing an improvement over the basic wiener process Monte Carlo simulations.

9. Develop a Forecasting Framework for the Mix Game using Kalman Filters to estimate the strategy weights (Simple version) or strategy pair weights (Expanded version).

6.3 Closing Remarks

No model is perfect, and the lack of significance of our results may mean that the Minority Game is either unable to capture real market dynamics, or is overly simplified, missing some necessary features required to achieve significant forecasting power. Alternatively, perhaps our optimization techniques need further refinement. These results do not, however, prove the notion that the techniques introduced in this thesis cannot work given a better underlying model - an improved agent-based model or improved optimization techniques.

In fact, the usefulness of the Minority Game and agent-based models in general, lies in their adaptability - we can readily add more realistic features to the model if required and assess whether this leads to the desired improvements. Indeed, perhaps the Minority Game used in our investigations is too general, and improvement is possible by adapting it to better characterize the specific market in question. The Oil market proposed in this thesis is a prime example that is ripe for further development.

Bibliography

[Econophysics website, University of Fribourg] University of Fribourg, Department of Physics, *Econophysics Website* http://www.unifr.ch/econophysics/minority [accessed February 2009]

[Website, University of Fribourg] University of Fribourg, Department of Physics, *Interactive Minority Game* http://www.unifr.ch/econophysics/minority/game [accessed February 2009]

[Adelman 1980] Adelman, M.A. (1980) 'The Clumsy Cartel', *The Energy Journal*, 1(1)

[Al-Turki 1994] Al-Turki, S. (1994) Autocorrelation in Static Economic Models and their Dynamic Respecifications: An Application to OPEC Behaviour, *Journal of King Saud University*, Administrative Sciences

[Al-Yousef 1998] Al-Yousef, N.A. (1998) *The Role of Saudi Arabia in the World Oil Market 1974-1997*, PhD Thesis, Guildford, University of Surrey

[Andersen, Sornette 2005] Andersen, J.V. & Sornette, D. (2005) A mechanism for pockets of predictability in complex adaptive systems, *Europhysics Letters*, 70(5), 697-703

[Anderson 1972] Anderson, P.W. (1972) More Is Different, *Science*, New Series, Vol. 177, No. 4047, pp. 393-396

[Anderson 2001] Anderson, P.W. (2001) *More is Different: Fifty Years of Condensed Matter Physics*, Woodstock, UK, Princeton University Press

[Anghel et al. 2004] Anghel, M, Toroczkai, Z., Bassler, K. & Korniss, G. (2004) Competition-Driven Network Dynamics: Emergence of a Scale-Free Leadership Structure and Collective Efficiency, Physical Review Letters, 92, 058701-1

[Aristotle] Aristotle (1998) *The Metaphysics*, H.Lawson-Tancred [translator], Penguin Group

[Arthur 1994] Arthur, B.W. (1994) Inductive reasoning and bounded rationality: the El Farol problem. *Am. Econ. Rev.* 84, 406-411

[Bachelier 1900] Bachelier, L. (1900) Theorie de la speculation [PhD thesis], *Annales Scientifiques de l'Ecole Normale Superieure* III-17, 21-86

[Bak 1997] Bak, P. (1997) *How nature works: The Science of Self-Organized Criticality*, Oxford, UK, Oxford University Press

[Bak, Chen 1993] Bak P. & Chen, K. (1993) Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics, *Ricerche Economiche* 47, 3-30

[Bak, Paczuski 1995] Bak, P. & Paczuski, M. (1995) Complexity, contingency, and criticality , *Proc. Nat. Acad. Sci. USA* 92 (1995) 6689

[Bar-Shalom et al. 2001] Bar-Shalom, Y., Rong Li, X. & Kirubarajan, T. (2001) *Estimation with Applications to Tracking and Navigation*, John Wiley & Sons, Inc.

[Beasley 1999] Beasley, J. (1999) *Population heuristics*, Working Paper, The Management School, Imperial College London

[Bird, Casavecchia 2007] Bird, R. & Casavecchia, L. (2007) The Market Impact of Momentum Investors: a Simulation Study, paper presented at the Conference on Trading Strategies and Financial

Market Inefficiency, Tanaka Business School, Imperial College London

[Black 1986] Black, F. (1986) Noise, *Journal of Finance* 41, 529-543

[Black, Scholes 1973] Black, F. & Scholes, M. (1973) The Pricing of Options and Corporate Liabilities, *Journal of Political Economy.* 81, 637-654

[Bollerslev 1986] Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31:307-327

[Bollerslev et al. 1992] Bollerslev, T., Chou, R.C. & Kroner, K. (1992) ARCH modeling in finance, *Journal of Econometrics*, 52, 5-59

[Bouchaud, Potters 2003] Bouchaud, J.P. & Potters, M. (2003) *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, Cambridge, UK, Cambridge University Press

[Box, Muller 1958] Box, G.E.P. & Muller, M.E. (1958) A Note on the Generation of Random Normal Deviates, *The Annals of Mathematical Statistics*, 29, 2, 610–611

[Breen et al. 1988] Breen, W., Glosten, L. R. & Jagannathan, R. (1988) Economic significance of predictable variations in stock index returns, *Journal of Finance*, 44, 1177–1189

[Bron, Kerbosch 1973] Bron, C. & Kerbosch, J. (1973) Algorithm 457 - Finding all cliques of an undirected graph, *Communications of the ACM*, 16, 575-577

[Buffet 2001] Buffett, Warren (2001) Lawrence Cunningham [ed.] *The Essays of Warren Buffett*, The Cunningham Group

[Campbell, Shiller 1988a] Campbell, J.Y. & Shiller, R.J. (1988) Stock prices, earning, and expected dividends, *Journal of Finance*, 43, 661-676

[Campbell, Shiller 1988b] Campbell, J.Y. & Shiller, R.J. (1988) The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies*, 1(3), 195-228

[Campbell et al. 1997] Campbell, J., Lo, A.H. & McKinlay, C. (1997) *The Econometrics of Financial Markets*, New Jersey, USA, Princeton University Press

[Challet, Zhang 1997] Challet, D. & Zhang Y.C. (1997) Emergence of cooperation and organization in an evolutionary game, *Physica A,* 246, 407-418

[Challet, Zhang 1998] Challet, D. & Zhang Y.C. (1998) On the minority game: Analytical and numerical studies, *Physica A,* 256, 514

[Challet, Marsili 1999] Challet, D. & Marsili, M. (1999) Phase transition and symmetry breaking in the minority game, *Phys. Rev. E* 60, R6271

[Challet et al. 2000] Challet, D., Marsili, M. & Zhang, Y.C. (2000) Modeling market mechanism with minority game, *Physica A* 276, 284

[Challet et al. 2001] Challet, D., Chessa, A., Marsili, M. & Zhang, Y.C. (2001) From Minority Games to Real Markets, *Quantitative Finance,* 1, 168–176

[Challet, Marsili 2003] Challet, D. & Marsili, M. (2003) Criticality and market efficiency in a simple realistic model of the stock market, *Phys. Rev. E.* 68, 36132-36134

[Challet et al. 2005] Challet, D., Marsili, M. & Zhang, Y. (2005) *Minority Games: Interacting Agents in Financial Markets*, Oxford, UK, Oxford University Press

[Challet et al. 2006] Challet, D., de Martino, A., Marsili, M. & Perez Castillo, I. (2006) Minority Games with finite score memory, *Journal of Statistical Mechanics: Theory and Experiment*

[Chen 1996] Chen, L. (1996) *Stochastic Mean and Stochastic Volatility - A Three-Factor Model of the Term Structure of Interest Rates and Its Application to the Pricing of Interest Rate Derivatives*, Blackwell Publishers.

[Chen 2002] Chen, S. (2002) *Genetic algorithms and genetic programming in computational finance*, Springer

[Chen et al. 2008] Chen, F., Gou, C., Guo, X. & Gao, J. (2008) Prediction of stock markets by the evolutionary mix-game model, *Physica A*, 387, 3594-3604

[Choe et al. 2004] Choe, S.C., Johnson, N.F. & Hui, P.M. (2004) Error-driven global transition in a competitive population on a network, *Physical Review E*, 70, 055101(R)

[Christoffersen, Diebold 2006] Cristoffersen, P.F. & Diebold, F.X. (2006) Financial asset returns, direction-of-change forecasting, and volatility dynamics, *Management Science*, 52, 1273–1287

[Christofides 1973] Christofides, N. (1973) *Graph theory: an algorithmic approach*, (Academic Press)

[Christofides 1994] Christofides, N. (1994) The bionomic algorithm, *Paper presented at the AIRO'94 conference*, Savona, Italy

[Christofides 1996] Christofides, N. (1996) Optimal design of telecommunication networks, *Paper presented at the CO96 Conference*, Imperial College, London

[Christofides et al. 2003] Christofides, S., Christofides, A. & Christofides, N. (2003) The design of corporate tax structures, *Mathematical Programming* 98, 493-510

[Coolen 2005] Coolen, A.C.C. (2005) *The Mathematical Theory of Minority Games: Statistical mechanics of interacting agents*, Oxford, UK, Oxford University Press

[Cont, Bouchaud 2000] Cont, R. & Bouchaud, J.P. (2000) Herd Behavior and Aggregate Fluctuations in Financial Markets, *Marcoeconomic Dynamics*, 4, 170-196

[Cont 2001] Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance* 1, 223-236

[Cont, Tankov 2004] Cont, R. & Tankov, P. (2004) *Financial Modelling with Jump Processes*, Chapman & Hall/CRC Financial Mathematics Series

[Cremer, Isfahani-Salehi 1991] Cremer, J. & Isfahani-Salehi, D. (1991) *Models of the Oil Market*, Chur, Harwood Academic

[Cutler et al. 1989] Cutler, D.M., Poterba, J.M. & Summers, L. (1989) What moves stock prices? *Journal of Portfolio Management,* Spring, 4-12

[Dacorogna et al. 2001] Dacorogna, M.M., Genay, R., Muller, U.A., Olsen, R.B., & Pictet, O.V. (2001) *An Introduction to High-Frequency Finance* , London, Academic Press

[Dahl, Yucel 1991] Dahl, C.A. & Yucel, M. (1991) Testing Alternative Hypothesis of Oil Producer Behaviour, *Energy Journal*, 12(4)

[de Martino et al. 2003] de Martino, A.D., Giardina, I. & Mosetti G. (2003) Statistical mechanics of the mixed majority–minority game with random external information, *Journal of Physics A* 36, 8935

[de Martino et al. 2004a] de Martino, A., Giardina, I., Marsili, M. & Tedeschi, A. (2004) Generalized minority games with adaptive trend-followers and contrarians, *Physical Review E*, 70, Issue 2, id. 025104 (http://prola.aps.org/pdf/PRE/v70/i2/e025104)

[de Martino et al. 2004b] de Martino, A.D., Giardina, I., Marsili, M. & Tedeschi A. (2004) Generalized minority games with adaptive trend-followers and contrarians, *Physical Review E*, 70, 025104

[Eldredge, Gould 1972] Eldredge, N. & Gould. S.J. (1972) Punctuated Equilibria: An Alternative to Phyletic Gradualism. In Schopf, T.J.M. (ed.), *Models in Paleobiology*, pp. 82-115, San Francisco, USA, Freeman, Cooper and Co.

[Elliott, Ito 1999] Elliott, G. & Ito, T. (1999) Heterogeneous expectations and tests of efficiency in the yen/dollar forward foreign exchange market, *Journal of Monetary Economics*, 43 435–456

[Engle 1982] Engle, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, *Econometrica*, 50, 987

[Engle 2001] Engle R.F. (2001) GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics, *Journal of Economic Perspectives*, 15(4):157-168

[Erdos, Renyi 1960] Erdős, P. & Rényi, A. (1960) The Evolution of Random Graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5: 17–61

[Evans 2004] Evans, T.S., (2004) Complex networks, *Contemporary Physics*, 45, 6, pp.455-474

[Fama 1965a] Fama, E.F. (1965) The Behavior of Stock Market Prices, The Journal of Business, 38, 1 pp. 34-105 http://links.jstor.org/sici?sici=0021-9398%28196501%2938%3A1%3C34%3ATBOSP%3E2.0.CO%3B2-6

[Fama 1965b] Fama, E.F. (1965) Random Walks In Stock Market Prices, *Financial Analysts Journal*, 21 (5): 55–59 http://www.cfapubs.org/toc/faj/1965/21/5 [accessed 10th February 2010]

[Fama, French 1988] Fama, E.F. & French, K.F. (1988) Dividend yields and expected stock returns, *Journal of Financial Economics*, 22(1), 3-25

[Fama, French 1989] Fama, E.F. & French, K.F. (1989) Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics*, 25(1), 23-50

[Farmer 1999] Farmer, J.D. (1999) Physicists attempt to scale the ivory towers of finance, *Computing in Science and Engineering*, 1, 26-39

[Farmer 2002] Farmer, J.D. (2002) Market force, ecology and evolution, *Industrial and Corporate Change*, 11, 5, 895-953

[Flanders 2010] Flanders, S. (2010) *Stephanomics*, British Broadcasting Corporation, http://www.bbc.co.uk/blogs/thereporters/stephanieflanders [accessed 4th February 2010]

[Fleurant et al. 1996] Fleurant, C., Glover, F., Michelon, P. & Valii, Z. (1996) A scatter search approach for unconstrained continuous optimization, *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, 643-648

[Foldy et al. 2003] Foldy, C., Somogyvari, Z. & Erdi, P. (2003) Hierarchically organized minority games, *Physica A*, 323, 735-742

[Forrest 1993] Forrest, S. (1993) Genetic Algorithms: Principles of Natural Selection Applied to Computation, *Science*, 261, 872-78

[French, Roll 1986] French, K.R., & Roll, R. (1986) Stock Return Variances: The Arrival of Information and the Reaction of Traders, *Journal of Financial Economics*, 5-26

[Frisch 1933] Frisch, R. (1933) Propagation and Impulse Problems in Dynamic Economics, *Economic Essays in Honor of Gustav Cassel*, London, Allen

[Gately 1984] Gately, D. (1984) A Ten-year Retrospective: OPEC and the World Oil Market, *Journal of Economic Literature*, 22

[Gately 1995] Gately, D. (1995) Strategies for OPEC Pricing and Output Decisions, *Energy Journal*, 16(3)

[Gencay 1998] Gencay, R. (1998) Optimization of technical trading strategies and the profitability in security markets, *Economics Letters*, 59 249–254

[Gershenfeld 1999] Gershenfeld, N. (1999) *The Nature of Mathematical Modelling,* Cambridge, UK, Cambridge University Press

[Glover 1994] Glover, F. (1994) Genetic algorithms and scatter search: unsuspected potential, *Statistics and Computing*, 4, 131-140

[Glover 1995] Glover, F. (1995) Scatter search and star-paths: beyond the genetic metaphon, *OR Spektrum*, 17, 125-137

[Glover 1997] Glover, F. (1997) A template for scatter search and path relinking, In J.K. Hao et al. (editors), *Artificial evolution: Lecture Notes in Computer Science*, 1363, 3-51

[Gorman 1961] Gorman, W.M. (1961) On a class of preference fields, *Metroeconomica*, 13, August, 53-56.

[Gou 2005] Gou, C. (2005) Predictability of Shanghai Stock Market by Agent-based Mix-game Model, *Proceedings of the 2005 International Conference on Neural Networks and Brain*, Vol 3 pg:1651 -1655 http://arxiv.org/ftp/physics/papers/0505/0505180.pdf

[Gou 2006a] Gou, C. (2006) Dynamic behaviours of mix-game model and its application, *Chinese Physics*, IOP Publishing Ltd.

[Gou 2006 b] Gou, C. (2006) The Simulation of Financial Markets by an Agent-Based Mix-Game Model, *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 3. Available at http://jasss.soc.surrey.ac.uk/9/3/6.html

[Gou 2006c] Gou, C. (2006) Deduction of initial strategy distributions of agents in mix-game models, *Physica A*, 371, 633-640

[Gou 2007] Gou, C. (2007) The relations between agent performances and their intellective abilities in mix-games, *Physica A*, 378, 459-464

[Gou et al. 2008] Gou, C., Guo, X. & Chen, F. (2008) Study on system dynamics of evolutionary mix-game models, *Physica A*, 387, 6353-6359

[Gourley et al. 2004] Gourley, S., Choe, S.C., Hui, P.M. & Johnson, N.F. (2004) Effects of local connectivity in a competitive population with limited resources, *Europhysics Letters*, 67(6), 867-873

[Grewal 2008] Grewal, M.S. & Andrews, A.P. (2008) *Kalman filtering: theory and practice using MATLAB*, Wiley IEEE-Press

[Griffin 1985] Griffin, J.M. (1985) OPEC Behavior: A Test of Alternative Hypotheses, *American Economic Review*, 75(5)

[Gulen 1996] Gulen, G. (1996) Is OPEC a Cartel? Evidence from Cointegration and Causality Tests, *EnergyJournal* 17(2)

[Gupta 2005a] Gupta, N. (2005) *Predicting Behaviour in Agent-Based Models for Complex Systems*, MPhil Transfer Thesis, University of Oxford

[Gupta et al. 2005a] Gupta, N., Hauser, R. & Johnson, N.F. (2005) Using Artificial Market Models to Forecast Financial Time-Series, *Oxford University pre-print* http://arxiv.org/PS_cache/physics/pdf/0506/0506134.pdf

[Gupta et al. 2005b] Gupta, N., Hauser, R. & Johnson, N.F. (2005) Forecasting Financial Time-Series Using Artificial Market Models, *Oxford University Computing Laboratory* Report 05/09 ftp://ftp.comlab.ox.ac.uk/pub/Documents/techreports/NA-05-09.pdf

[Gupta et al. 2006] Gupta, N., Hauser, R. & Johnson, N.F. (2006) *Deducing the Multi-Trader Population Driving a Financial Market*, Proceedings of SPIE: Complex Systems, The International Society for Optical Engineering. Brisbane, Australia ftp://ftp.comlab.ox.ac.uk/pub/Documents/techreports/NA-06-09.pdf

[Gupta et al. 2007a] Gupta, N., Hauser, R. & Johnson, N.F. (2007) Inferring the Composition of a Trader Population in a Financial Market, *Econophysics of Markets and Business Networks*, Part 1, Milan, Springer

[Gupta et al. 2007b] Gupta, N., Hauser, R. & Johnson, N.F. (2007) Robust methods for tracking intelligent agents playing in an artificial financial market, *in Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, AAMAS 2007: 176 http://portal.acm.org/citation.cfm?id=1329338

[Gupta 2007] Gupta, N. (2007) Kalman Filtering in the Presence of State Space Equality Constraints, *Oxford University Computing Laboratory* Report 07/14 http://www.comlab.ox.ac.uk/files/732/NA-07-14.pdf

[Gupta, Hauser 2007] Gupta, N. & Hauser, R. (2007) Kalman Filtering with Equality and Inequality State Constraints, *Oxford University Computing Laboratory* Report 07/18 http://www.comlab.ox.ac.uk/files/728/NA-07-18.pdf

[Hansen, Snyder 1997] Hansen, C. & Snyder, S. (1997) *Active control of noise and vibration*, Taylor & Francis

[Harris 2003] Harris, L. (2003) *Trading and exchanges: market microstructure for practitioners*, Oxford, UK, Oxford University Press

[Hart et al. 2001] Hart, M.L., Jefferies, P., Hui, P. & Johnson, N.F. (2001) Crowd-anticrowd theory of multi-agent market games, *Eur. Phys. J.* B 20, 547-550

[Hart et al. 2002] Hart, M.L., Lamper, D. & Johnson, N.F. (2002) An investigation of crash avoidance in a complex system, *Physica A*

[Harvey 1989] Harvey, A. (1989) *Forecasting, structural time-series models and the Kalman Filter*, Cambridge, UK, Cambridge University Press

[Herrick 2001] Herrick, T. (2001) Opec seeks price stability for crude oil with latest planned cutback, *Wall Street Journal*, 9th January

[Heston 1993] Heston, S.L. (1993) A closed-form solution for options with stochastic volatility,

[Holland 1975] Holland, J.H. (1975) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, Michigan, USA, University of Michigan Press.

[Hong, Chung 2003] Hong, Y. & Chung, J. (2003) Are the directions of stock price changes predictable? Statistical theory and evidence, *Manuscript*, Cornell University, USA

[Hull 2003] Hull, J.C. (2003) *Options, Futures, and Other Derivatives, Fifth Edition*, New Delhi, India, Prentice-Hall of India

[Jefferies et al. 2001] Jefferies, P., Hart, M.L., Hui, P.M., & Johnson, N.F. (2001) From market games to real-world markets, *Eur. Phys. J. B* 20, 493-501

[Jefferies et al. 2002] Jefferies, P., Lamper, D. & Johnson, N.F. (2002) Anatomy of extreme events in a complex adaptive system, *Physica A*

[Johnson 2007] Johnson, N.F. (2007) *Two's Company, Three is Complexity*, Oxford, UK, Oneworld Publications

[Johnson et al. 1999] Johnson, N.F., Hui, P.M., Jonson, R. & Lo, T.S. (1999) Self-Organized Segregation within an Evolving Population, *Physical Review Letters*, 82, 3360

[Johnson et al. 2000] Johnson, N.F., Hart, M., Hui, P.K. & Zheng, D. (2000) Trader Dynamics in a Model Market, *International Journal of Theoretical and Applied Finance* Vol 3, No. 3, 443-450

[Johnson et al. 2003] Johnson, N.F., Jefferies, P. & Hui, P.M. (2003) *Financial Market Complexity*, Oxford, UK, Oxford University Press

[Johnson et al. 2001] Johnson, N.F., Lamper, D., Jefferies, P., Hart, M.L. & Howison, S. (2001) Application of multi-agent games to the prediction of financial time-series, *Physica A* 299, 222

[Kahneman et al. 1991] Kahneman, D., Knetsch, J.L. & Thaler, R.H. (1991) Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias, *Journal of Economic Perspectives*, 5, 1, 193-206

[Kalman 1960] Kalman, R.E. (1960) A new approach to linear filtering and prediction problems, *Transactions of the ASME. Series D, Journal of Basic Engineering*, 82, 35–45 http://www.elo.utfsm.cl/~ipd481/Papers%20varios/kalman1960.pdf

[Kalman, Bucy 1961] Kalman, R.E. & Bucy, R. S. (1961) New Results in Linear Filtering and Prediction Theory, *Transactions of the ASME. Series D, Journal of Basic Engineering*, 83, 95-108

[Karavias 2006] Karavias, A. (2006) *Optimisation Algorithms for Sampling and Estimation in Finance*, PhD Thesis, Imperial College London

[Kauffman 1993] Kauffman, S. (1993) *The origins of order: self-organization and selection in evolution*, New York, Oxford University Press

[Keim 1983] Keim, D.B. (1983) Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence, *Journal of Financial Economics*, 12

[Keynes 1936] Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money*, London, MacMillan

[Kirkpatrick et al. 1983] Kirkpatrick, S. Gelatt Jr., C.D. & Vecchi, M.P. (1983) Optimization by Simulated Annealing, *Science*, 220, 671-80

[Kirkpatrick 1984] Kirkpatrick, S. (1984) Optimization by Simulated Annealing: Quantitative Studies, *J. Stat. Phys.*, 34, 975-86

[Kuan, Liu 1995] Kuan, C.M. & Liu, T. (1995) Forecasting exchange rates using feedforward and recurrent neural networks, *Journal of Applied Econometrics*, 10 347–364.

[Lamper et al. 2001] Lamper, D., Howison, S. & Johnson, N.F. (2001) Predictability of Large Future Changes in a Competitive Evolving Population, *Phys. Rev. Lett.* 88, 01790

[Lamper 2002] Lamper, D. (2002) *Problems in Mathematical Finance : Market Modelling and Derivative Pricing*, PhD thesis, University of Oxford

[Laureti et al. 2004] Laureti, P., Ruch, P., Wakeling, J. & Zhang, Y.C. (2004) The Interactive Minority Game: a Web-based investigation of human market interactions, *Physica A*, 331, 651-659

[LeBaron et al. 1998] LeBaron, B., Arthur, W.B. & Palmer, R. (1998) Time Series Properties of an Artificial Stock Market, *Journal of Economic Dynamics & Control*

[Leitch, Tanner 1991] Leitch, G. & Tanner, J. (1991) Economic forecasts evaluation: Profits versus the conventional error measures, *American Economic Review*, 81, 580–590

[Lettau, Ludvigson 2001] Lettau, M. & Ludvigson, S. (2001) Consumption, aggregate wealth and expected stock returns, *Journal of Finance*, 56(3), 815-849

[Leung et al. 2000] Leung, M.T., Daouk, H. & Chen, A.S. (2000) Forecasting stock indices: A comparison of classification and level estimation models, *International Journal of Forecasting*, 16 173–190

[Levy et al. 2000] Levy, H., Solomon, S. & Levy, M. (2000) *Microscopic Simulation of Financial Markets: From Investor Behavior to Market Phenomena*, London, Academic Press

[Li et al. 2000a] Li, Y., Riolo, R. & Savit, R. (2000) Evolution in Minority Games. (I). Games with a Fixed Strategy Space, *Physica A* 276, 234-264

[Li et al. 2000b] Li, Y., Riolo, R. & Savit, R. (2000) Evolution in minority games. (II). Games with variable strategy spaces, *Physica A*, 276, 265

[Lioui, Rangvid 2009] Lioui, A. & Rangvid, J. (2009) Stock Return Predictability in a Monetary Economy, http://staff.cbs.dk/JRangvid/PredMonEcon.pdf

[Lowenstein 2001] Lowenstein, R. (2001) *When Genius Failed : The Rise and Fall of Long-Term Capital Management*, London, Fourth Estate

[Lucas 1978] Lucas, R.E. (1978) Asset Prices in an Exchange Economy, *Econometrica*, 46, 6, 1429-1445

[Lux 1996] Lux, T. (1996) The stable Paretian hypothesis and the freqency of large returns: an examination of major German stocks, *Applied Financial Economics*, 6, 463

[Lux, Marchesi 1999] Lux, T. & Marchesi M. (1999) Scaling and criticality in a stochastic multi-agent model of a financial market, *Nature* 397: 498-500

[Mabro 1992] Mabro, R. (1992) *OPEC and the Price of Oil*, Oxford, Oxford Institute for Energy Studies

[Malkiel 1996] Malkiel, B.G. (1996) *A Random Walk Down Wall Street*, W. W. Norton

[Mandelbrot 1963] Mandelbrot, B.B. (1963) The Variation of Certain Speculative Prices, *J. Business,* 36, 394-419

[Mandelbrot, Wallis 1968] Mandelbrot, B.B. & Wallis, J.R. (1968) Noah, Joseph, and Operational Hydrology, *Water Resources Research*, Vol. 4, No. 5, P. 909–918

[Mandelbrot 1997] Mandelbrot, B.B. (1997) *Fractals and Scaling in Finance*, New York, USA, Springer-Verlag

[Mantegna, Stanley 2000] Mantegna, R.N. & Stanley, H.E. (2000) *An introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge, UK, Cambridge University Press

[Markowitz 1959] Markowitz, H. (1959) *Portfolio Selection: Efficient Diversification of Investment*, New York, J. Wiley

[Marsaglia 1972] Marsaglia, G. (1972) Choosing a point from the surface of a sphere, The Annals of Mathematical Statistics, 43, 2, 645-646

[Marsili 2001] Marsili, M. (2001) Market mechanism and expectations in minority and majority games, *Physica A* 299, 93

[Marsili et al. 2000] Marsili, M., Challet, D. & Zecchina, R. (2000) Exact solution of a modified El Farol's bar problem: Efficiency and the role of market impact, *Physica A*, 280, 522

[Maybeck 1979] Maybeck, P.S. (1979) *Stochastic Models, Estimation and Control*, Academic Press Inc.

[Menzly, Santos, Veronesi 2004] Menzly, L., Santos, T. & Veronesi, P. (2004) Understanding predictability, *Journal of Political Economy*, 112(1), 1-47

[Merton 1973] Merton, R. (1973) Theory of Rational Option Pricing, *Bell J. Econ. Management Sci.* 4, 141-183

[Merton 1976] Merton, R. C. (1976) Option Pricing When Underlying Stock Returns Are Continuous, *Journal of Financial Economics*, 3, 125-144

[Met Office] Met Office, *Ensemble and Probability Forecasting* http://www.metoffice.gov.uk/science/creating/daysahead/ensembles [accessed 10th September 2009]

[Metropolis et al. 1953] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E. (1953) Equation Of State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, 21, 1087-92

[Mitman et al. 2005] Mitman, K.E., Choe, S.C., Johnson, N.F. (2005) Competitive advantage for multiple-memory strategies in an artificial market, *Proceedings of SPIE–the international society for optical engineering*, 5848, 225 -232 http://arxiv.org/abs/physics/0503031

[Muller 1959] Muller, M. (1959) A note on a method for generating points uniformly on N-dimensional spheres, *Comm. Assn. Comp. Mach.* 2, 4, 19-20

[Nagel 1995] Nagel, R. (1995) Unraveling in Guessing Games: An Experimental Study, *American Economic Review*, 85 (5),1313-1326

[Nash 1950] Nash, J. (1950) Equilibrium points in n-person games, *Proceedings of the National Academy of Sciences of the United States of America*, 36 (1): 48–49, doi:10.1073/pnas.36.1.48 http://www.pnas.org/content/36/1/48

[Niederreier 1992] Niederreier, H. (1992) Random Number Generation and Quasi-Monte Carlo Methods, *Society for Industrial and Applied Mathematics*, Philadelphia

[Nyberg 2008] Nyberg, H. (2008) Forecasting the Direction of the U.S. Stock Market with Dynamic Binary Probit Models, *Helsinki Centre of Economic Research*, Discussion Paper No. 227, Helsinki, Finland http://ethesis.helsinki.fi/julkaisut/eri/hecer/disc/227/forecast.pdf

[OPEC] Organization of the Petroleum Exporting Countries, http://www.opec.org/home

[Osborne 1959] Osborne, M.F.M., (1959) Brownian Motion in the Stock Market, *Operations Research*, March-April

[Orlean 2005] Orlean, A. (2005) The self-referential hypothesis in finance. In Touffut, J.P., editor, *The Stability of Finance in Europa*, Paris, France, Albin Michel http://www.pse.ens.fr/orlean/depot/publi/CHA2005tSELF.pdf

[Parra 2004] Parra, F. (2004) *Oil Politics: A Modern History of Petroleum*, London, I.B.Taurus

[Pinol 2006] Pinol, H., Beasley, J.E. (2006) Scatter Search and Bionomic Algorithms for the aircraft landing problem, *European Journal of Operational Research* 171, 439-462

[Post et al. 2008] Post, T., Van den Assem, M.J., Baltussen, G. & Thaler, R. (2008) Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show, *American Economic Review*, 98, 1

[Roehner 2002] Roehner, B.M. (2002) *Patterns of Speculation: A Study in Observational Econophysics*, Cambridge, UK, Cambridge University Press

[Salehi-Isfahani 1995] Salehi-Isfahani, D. (1995) Models of the Oil Market Revisited, *Journal of Energy Literature*, 1(1)

[Sardar, Abrams 1998] Sardar, Z. & Abrams, I. (1998) *Chaos for Beginners*, Cambridge, Icon Books Ltd.

[Savit et al. 1999] Savit, R., Manuca, R. & Riolo, R. (1999) Adaptive Competition, Market Efficiency, and Phase Transitions, *Phys. Rev. Lett.* 82(10): 2203-2206

[Savit et al. 2004] Savit, R., Koelle, K., Treynor, W. & Gonzalez, R. (2004) Man and Superman: Human Limitations, Innovation, and Emergence in Resource Competition, page 199 in *Collectives and the Design of Complex Systems*, Tumer, K., Wolpert, D., (ed.) New York, USA, Springer

[Seymour 1980] Seymour, I. (1980) *OPEC Instrument of Change*, London, MacMillan

[Shiller 1989] Shiller, R. (1989) *Market Volatility* , Cambridge, Massachusetts, USA, MIT Press

[Shleifer 2000] Shleifer, A. (2000) *Inefficient Markets: an Introduction to Behavioral Finance*, Oxford, UK, Oxford University Press

[Simon 1997] Simon, H. (1997) *Models of Bounded Rationality*, Cambridge, Massachusetts, USA, MIT Press

[Slanina, Zhang 1999] Slanina, F. & Zhang, Y.C. (1999) Capital flow in a two-component dynamical system, *Physica A* 272, 257

[Slutsky 1937] Slutsky, E. (1937) The Summation of Random Causes as the Source of Cyclical Processes, *Econometrica*

[Sornette 2003] Sornette, D. (2003) *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*, Springer

[Sornette 2004] Sornette, D. (2004) *Why Stock Markets Crash: Critical Events in Complex Financial Systems*, New Jersey, USA, Princeton University Press

[Soros 1994] Soros, G. (1994) *The Alchemy of Finance: Reading the Mind of the Market*, John Wiley & Sons

[Stevens 05] Stevens, P. (2005) Oil Markets, *Oxford Review of Economic Policy*, 21, 1, Oxford University Press

[Taleb 2008] Taleb, N.N. (2008) *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets*, Random House

[Tedeschi et al. 2005] Tedeschi, A., Martino, A.D. & Giardina, I. (2005) Coordination, intermittency and trends in generalized minority games, *Physica A*, 358, 529-544

[Turner 2009] Turner, J.A. (2009) *The Turner Review: A regulatory response to the global banking crisis*, London, UK, Financial Services Authority, p.8, http://www.fsa.gov.uk/pubs/other/turner_review.pdf

[UBS Global Outlook 2008] UBS Global Outlook (2008) *UBS Wealth Management Research* http://www.ubs.com [von Neumann, Morgenstern 1947] von Neumann, J. & Morgenstern, O. (1947) *Theory of games and economic behavior*, New Jersey, USA, Princeton University Press

[Waldrop 1994] Waldrop, M.M. (1994) *Complexity: The Emerging Science at the Edge of Order and Chaos*, UK, Penguin Science

[Watts, Strogatz 1998] Watts, D.J. & Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks, *Nature* 393: 440–442

[Zhang 1998] Zhang, Y.C. (1998) Evolving Models of Financial Markets, *Europhys. News*, March/April: 51-54