THE IMPLICATIONS OF INVESTORS CONSIDERING THEIR SOCIAL NETWORK

by

Matthew Oldham
A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Arts
Interdisciplinary Studies

Committee:

____ Director

____

____

____ Program Director

____ Dean, College of Humanities and Social Sciences

Date: ____________________________________ Spring Semester 2016
George Mason University
Fairfax, VA

The Implications Of Investors Considering Their Social Network.

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts Interdisciplinary Studies at George Mason University

By

Matthew Oldham
Bachelor of Economics (Hons)
University of Tasmania, 1995

Director: Rob Axtell, Professor
Department of Computational and Data Sciences, College of Science

Spring Semester 2016
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my loving wife Lora and two children; Téa and Jenson, for all their support and understanding during this journey.

ACKNOWLEDGEMENTS

Dr. Robert Axtell, thank you for your advice; being an exceptional sound blocking and helping me navigate my way through the production of this thesis.

Dr. Andrew Crooks, thank you for your timely advice and being there to answer my questions.

Dr. Carlos Ramirez, thank you for agreeing to be part of the committee for this thesis.

Steve Scott and Stefan McCabe for all their assistance in helping me ‘code it’ throughout my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

Agent Based Model .... ABM
Capital Asset Pricing Model .... CAPM
Complex Adaptive System .... CAS
Dividend Per Share .... DPS
Efficient Market Hypothesis .... EMH
Earnings Per Share .... EPS
EPS forecast .... EPSF
Global Financial Crisis .... GFC
International Monetary Fund .... IMF
National Association of Securities Dealers Automated Quotations .... NASDAQ
Probability Density Function .... PDF
Standard & Poor's 500 Index .... S&P 500

ABSTRACT

THE IMPLICATIONS OF INVESTORS CONSIDERING THEIR SOCIAL NETWORK.

Matthew Oldham MAIS

George Mason University, 2016

Thesis Director: Dr. Rob Axtell

The behavior of financial markets has, and continues, to frustrate investors and academics. With the advent of new approaches, including a complex systems framework, the search for an explanation as to how and why markets behavior as they do has greatly expanded. The use of agent-based models (ABMs) and the inclusion of network science has played an important role in increasing the relevance of the complex systems. Through the use of an artificial stock market utilizing an Ising model based agent-based model, this thesis has been able to provide significant insight into the mechanisms that drive the returns in financial markets, including periods of elevated prices and excess volatility. In particular, the thesis demonstrates the following: the network topology that investors form; along with the dividend payout ratio of a stock significantly impact the behavior of the market. The model also investigates the impact of introducing multiple risky assets, something that has been absent in any previous attempts. By successfully addressing these issues this thesis has helped refine and shape a variety of further research tasks.

1 INTRODUCTION

1.1 Background and Motivation

A common feature of financial markets since their advent has been periods where price movements and trading volumes have been much higher than what is commonly experienced, that is, the markets become more volatile. In particular, the semi regular appearance of bubbles[1] and their subsequent collapse has left academia and the general public searching for answers. The repercussions of these boom and bust cycles are severe, with over-investment and excessive trading occurring in the boom times, while the busts have on occasions led to devastating financial crises and depressed real economies. One of the most recent occurrences of such an event saw the Dow Jones Industrial Average close at a record level on October 9, 2007 yet one year later the Dow dropped 21% in the first nine days of October 2008 and the world plunged into the Global Financial Crisis (GFC), which according to the International Monetary Fund (IMF) cost the global economy $USD11.9 trillion (Conway, 2009).

There has been much debate about whether it is possible to predict the future movements of the financial markets and whether a bubble can even be detected in advance (Gupta, Hauser, & Johnson, 2005). The 'mainstream' doctrine and practices

[1] A bubble can be defined as a period during which the market value of assets vastly exceed reasonable assessments of their fundamental value. Alternatively, Kindleberger & Aliber (2011) state more generally that a bubble occurs when there is "an upward price movement over an extended range that then implodes."

have tended to follow Fama's (1970) Efficient Market Hypothesis (EMH) which states that market prices fully reflect all available information. The implication being that as news arrives in a random unpredictable fashion, prices follow a similar pattern, which is one of a random walk. An implication of the EMH is that asset bubbles and subsequent crashes will not and cannot occur. The view that asset prices do follow a random walk, a view popularized by Malkiel (1999), has found some empirical support. However, the reality of continued episodes of boom and bust, and mounting statistical evidence, provides stronger evidence that markets do not function in accordance with the EMH[2].

1.2 Overview of Approach

An alternate approach to the EMH is to consider financial markets as a complex system. Considering financial markets as a complex system is to accept that outcomes in financial markets are the result of an emergent process, based on the self-organized behavior of independently acting, self-motivated individuals (Farmer et al., 2012). A process is defined as emergent when larger entities, patterns, and regularities arise through interactions among smaller or simpler entities that themselves do not exhibit such properties (Wikipedia, 2014). The main attraction of utilizing a complex system framework to analyze financial markets is that they are able to generate extreme events and asset returns in line with what has been experienced in the real world. The complex system approach is consistent with the thoughts of Sornette (2014), who concluded that after 20 years of research, the key concepts required to understand stock market returns

[2] Bollerslev, Engle, and Nelson (1994) and Mandelbrot (1963) are classic papers that highlight that financial returns may not follow the statistical distribution as prescribed by the EMH.

are; imitation, herding, self-organized co-operativity and positive feedbacks. These factors, along with a viable explanation for the extreme returns experienced in real asset markets are lacking from the EMH.

Another advantage of utilizing a complex systems framework is that it allows researchers to introduce networks within their analytical framework. The importance of networks is highlighted by Newman (2010), who suggested that networks are a "powerful means of representing patterns of connections or interactions between the parts of a system". The rationale for utilizing a network is that the behavior of a system can vary greatly depending on which network structure (the topology) is formed within a system[3]. The relevance of networks to financial markets is their ability to explain investor trading decisions and portfolio performance (Ozsoylev & Walden, 2011).

Given the characteristics of a complex system, traditional analytical approaches are rendered ineffective, and researchers have been forced to turn to computer simulation to understand the dynamics of them. In particular, agent based models (ABMs) are extensively utilized in the study of complex systems. ABMs allow for the interaction between individual agents, investors in the case of financial markets, who act and undertake actions based on the context of their environment using basic rules. Importantly, the agents' behavior is not fixed and can evolve in response to the behavior of others and their environment. Therefore ABMs are not constrained to equilibrium conditions (Sornette, 2014). Axtell (2000) provides a comprehensive review supporting these arguments and the benefits of ABMs in analyzing complex systems.

[3] For classic papers demonstrating the implications of different network structures see; Albert, Jeong, & Barabási (2000), Santos and Pacheco (2005) and Callaway, Newman, Strogatz, & Watts (2000).

1.3 Research Question

The purpose of the research topic is to implement an ABM that is capable of understanding the impact that particular network structures have on the performance of financial markets. And in particular to understand whether certain network structures lead to greater volatility and the dynamics behind whether or not the population forms large common groups ("herds") in terms of their investment strategies. The relevance being that understanding the dynamics of herd formation should provide insights into how bubbles can form and then collapse in financial markets. In turn, these insights can potentially inform market participants and be used to reduce the risk of another crash.

While there is a large volume of work of utilizing ABMs to simulate financial market returns (LeBaron (2006) and Sorrnette (2014) provide extensive reviews of the application of ABMs to financial markets), the utilization of a network structure within the various frameworks has been limited. While attempts to implement a network structure have commenced, there is a rich field of research questions to consider with multiple avenues of investigation. Panchenko, Gerasymchuk and Pavlov (2013) were able to show that network structures are capable of influencing the stability of, and the fluctuations of an asset's price, while Harras and Sornette (2011) demonstrated how bubbles may emerge as a result of agents considering different information sources, including the expected actions of their neighbors.

The literature relating to ABMs and artificial stock markets, including a simulation of the NASDAQ exchange (Darley & Outkin, 2007), utilize a single risky asset, therefore reducing the problem to one of asset allocation rather than choosing amongst risky assets. In what is believed to be a first for an artificial stock market, the author introduces the ability for agents to consider multiple risky assets and analyzes the actions of the agents through the formation of a dynamic quasi-efficient frontier. The rationale for this inclusion is to see if, and under what circumstances, agents with basic rules are able to produce an outcome prescribed by traditional financial literature.

To achieve the aforementioned goals, the model used by Harras and Sornette (2011) formed the foundation for the implemented model. This framework enabled the key questions of: how investors undertake their decision-making process; to what extent they utilize their network in that process; what, if any, network structure are the agents linked by; is the network static or dynamic, and if it is dynamic in what regard is it dynamic; to be assessed. The results of these simulations provide an insight into how financial markets operate and in particular, identify the factors around the formation of large “herds” and whether particular network structures lead to excess volatility.

1.4 Purpose of Thesis

By investigating the role of how a network’s structure and dynamics contribute to the performance of financial markets, it provides important insights into the mechanism that generate inefficient behavior. Once a mechanism can be identified, the research effort can be linked to identifying real world investor networks and analyzing which network structure they take and how they change through time.

Current examples of this aim, include judging the impact of social networks - using Twitter as a proxy - on financial markets. Bollen, Mao and Zeng (2011) and Jaffe (2015) report on the advent of investment vehicles based on investor sentiment as expressed via Twitter. Ozsoylev and Walden (2011) were able to find empirical support for their theoretical model that, among other things, predicted that investors with higher connectedness earn higher profits from trading more aggressively, and that price volatility is higher in markets with intermediate connectedness.

The continued monitoring of investor networks may provide a vital early warning to the formation of an asset bubble – something that the EMH fails to deliver and something only very few investors are able to lay claim to.

1.5 Thesis Outline

To effectively answer the research question, this thesis has been divided into distinction sections. Greater detail with regards to the EMH and the case for and against it are provided in Section 2.2. The background and rationale for an alternative approach utilizing a complex system approach is described in Section 2.3 before Section 2.4 details the rise of ABMs as a viable framework to assess how and why financial markets operate in a manner inconsistent with that proposed by the EMH. The justification for the use of the base model and its various extensions, that were mentioned in Section 1.3, along with the specifications of the model are detailed in Section 3. The results of the various experiments, which were designed to meet the research question and ultimately provide significant insight into the dynamics of financial markets, are provided in Section 4.

2 LITERATURE REVIEW

2.1 Background and Introduction

Since their advent, financial markets have experienced episodes of extreme volatility interspersed with periods of relative calm[4]. One of the first attempts to document and explain the volatile nature of financial markets was Mackay in *Extraordinary Popular Delusions and the Madness of Crowds* (1841), who pointed to the behavioral traits of investors in creating this volatility. The argument that markets were driven by 'animal spirits' and hence difficult to understand and prone to periods of 'excitement' was proposed by Keynes (2007, pp. 161–162). This view has since gained further support from the likes of Shiller (2015) and Thaler (2015) and coupled with the principle of Simon's (1955) bounded rationality, can be seen as the genesis of behavioral economics/finance, a field that has become increasingly relevant in describing the behavior of financial markets.

However, the more accepted approach to modern finance has not followed these works. Rather it was the advanced mathematics of Bachelier (1995) that became the genesis of how financial markets were analyzed and forecasts developed. This approach ultimately culminated with the Efficient Market Hypothesis (EMH) - Fama (1970), becoming the center piece of modern finance. This point is reaffirmed by Summers

[4] Relative calm is defined as, price movements that are not dramatic.

(1986) who claimed that the assumption of market efficiency "forms the basis for most research in financial economics". Further, the zealot like faith in the EMH is demonstrated by Jensen (1978), when he stated his belief that "there is no other proposition in economics which has more solid empirical evidence supporting it than the Efficient Market Hypothesis". Subsequently, other predominate financial models such as Black Scholes option model - Black and Scholes (1973), and the Capital Asset Pricing Model (CAPM) – Sharpe (1964), have been underwritten by the EMH and another key theory - rational expectations, as proposed by Friedman (1953).

However, there have been numerous examples – the October 1987 stock market crash, the Japanese Bubble of the 1990s, the demise of Long Term Capital Management, the dot-com bubble and the 2008 collapse (De Bondt, Muradoglu, Shefrin, & Staikouras, 2008), where the EMH - and the models which have utilized it – have proven inadequate in explaining how and why the markets functioned as they had. These episodes have led to the search for alternative theories with behavioral finance and the study of financial markets as a complex system as two leading examples. However, the path to dislodging the incumbent models remains difficult, as witnessed by Fama and French (2004) who state that despite its empirical shortcomings, the CAPM remains a "theoretical tour de Force".[5] This point is further discussed by Thaler (2015), who makes the point that despite the shortcomings of the various models underwritten by the EMH, the true believers refuse to discard it and simply propose updated models that are more capable of

[5] Note it is not possible to empirically test the EMH and it must be done in combination with an equilibrium model (Sorropago, 2014).

predicting previously unexplained market movements.

2.2 The Efficient Market Hypothesis

2.2.1 In Greater Detail

The EMH reached its dominance in the 1970s (Shiller, 2003) but could not have done so without the advent of “positive economics”, by Friedman (1953). By successfully arguing that economics should be the study of “what is” rather than “what ought to be”, Friedman shifted the focus of economics (and finance as a subset) from basing models on realistic assumptions and observation to producing models that had superior predictive ability. Under this banner, a theory becomes better insofar as it made less plausible assumptions (Buchanan, 2014).

Friedman (1953) also presented the theory that a market would only contain rational investors. Rational investors being those who employ a decision-making process that is based on picking the option that results in the most optimal level of benefit or utility for the individual. By assuming that all investors were rational meant that the population would have homogeneous expectations and beliefs. This approach did not deny that investors may at times make mistakes but there would be a few smart people who would trade against them and correct prices quickly. A further implication of Friedman’s world is that irrational investors would be forced from the market due to losing their money at the hands of the rational investors. The acceptance of positive economics and the concept of the rational investor led to the development of the rational expectations model by Muth (1961), which in turn became the key plank for Fama’s EMH (Shiller 2003).

A major implication of rationality and agents having homogenous beliefs is that economists are able to create a sole representative agent who embodies the collective preferences of the population (Sornette, 2014). It is the creation of this agent that allows closed form mathematical solutions for various financial models underwritten by the EMH to be derived. Under this approach, agents employ deductive top-down reasoning, which is where conclusions are reached by reductively applying general rules that hold over the entirety of a closed system.

In essence, the EMH states that market prices fully reflect all available information and, as a consequence, asset prices are unpredictable. Two key implications coming from the EMH, are:

- The 'price is right'; and
- There is no 'free lunch' for investors.

The 'price is right' refers to the fact that prices should remain close to their fundamental value[6] (Buchanan, 2014). It is this assumption that precludes the existence of a bubble – a price greatly exceeding its fundamental value. Interestingly, as Thaler (2015) points out, this part of the EMH is extremely difficult, if not impossible to test, as it cannot be tested in isolation but rather in unison with an asset pricing model that has utilized the assumptions of the EMH; a point discussed later.

'No free lunch' implies that investors should not be able to outperform the market because all publicly available information is already factored into an asset price. Therefore investors have no additional information that can help them reliably predict

[6] Fundamental value as per Buchanan (2014) is a realistic value based on an analysis of a company's current and future profit prospects.

future prices. New news or information will arrive in a random fashion, meaning investors having no ability to accurately forecast it, prices will be updated instantaneously and follow the same random path as the news. In addition, no patterns in past price movements exist that investors can exploit (Johnson, Jefferies, & Hui, 2003). In the event that an opportunity did exist, the EMH predicts it would vanish quickly.

There are three accepted versions of the EMH, which relates to the information available to investors. As outlined by Jensen (1978) these are the:

- Weak Form - where the information available to investors is solely the past price history;
- Semi Strong Form – In addition, to the Weak Form, investors have access to all publicly available information; and
- Strong Form – This form includes all possible information.

As Jensen (1978) suggests, the most widely accepted form is the Semi Strong Form, with the implication being that by trading solely on publicly available information, including the past price history, it is impossible to make an economic profit. Efficiency therefore refers to the fact that all information is rapidly reflected in the price and there is no information capable of moving the share price that is not already incorporated in the price. For this to occur, it is assumed that investors are capable of collecting all relevant information and have the computational power to correctly form probabilistic assessments and calculate their expected utility.

Simon (1955) questioned whether human's possessed these capabilities, before offering the alternative of human's acting with bounded rationality. Further, in making these calculations, it is assumed that there is no cost associated with collecting or analyzing the information. Grossman and Stiglitz (1980) challenged this point by arguing that perfectly informationally efficient markets are not possible because if markets were perfectly efficient the returns from gathering information are non existent. The implication being, investors have no incentive to trade or collect the information.

An interesting implication of the EMH, if it is held to its upmost, is that no trading would occur. This results from the fact that a rational investor wishing to sell an asset would be unable to find a rational buyer willing to acquire the same asset because both would have the same value for the asset. As Thaler (2015) mentions, while no one expects this to hold, proponents of the EMH have been forced to concede that trading volumes are generally in excess of what the theory would predict.

A return series that follows a random walk, the returns implied by the EMH, is illustrated in Figure 1. The series has a Gaussian distribution, implying that returns fall within a well-defined probability distribution function (PDF), with 99.7% of all returns falling within three standard deviations of the mean. Under this function, the probability of an extreme event - a 'Black Swan' event, such as the 1987 crash, is extremely small (1 in 50 billion (Mandelbrot & Hudson, 2006).

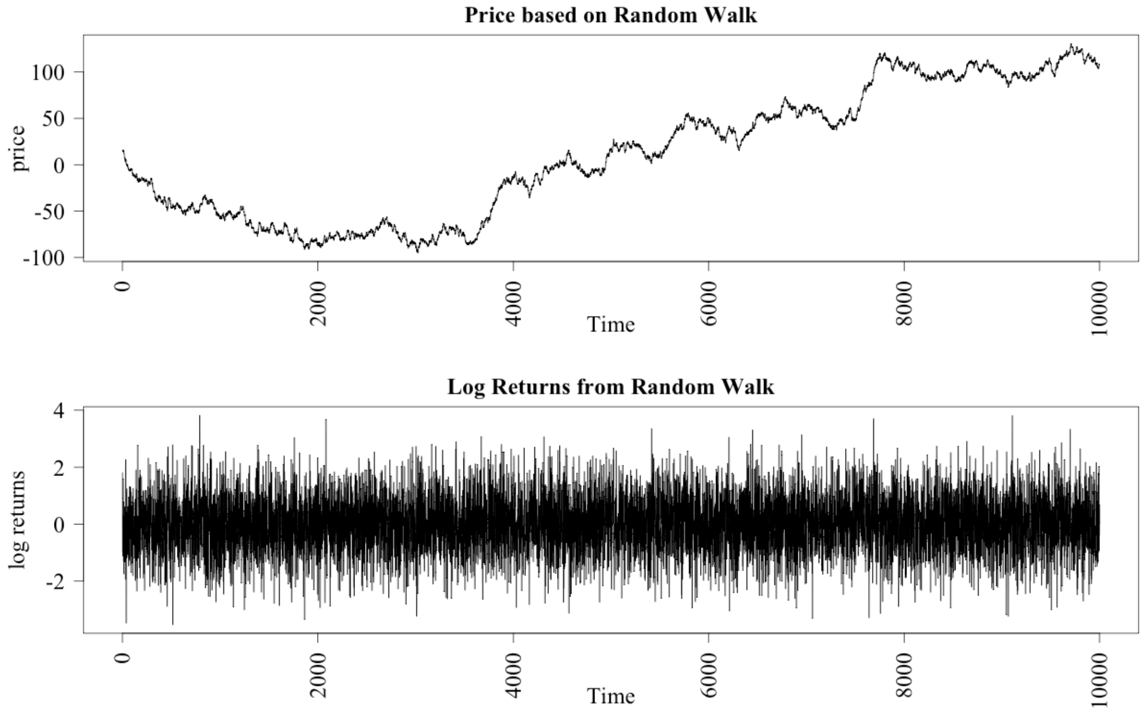


Figure 1: Returns for a random walk series over time

Kirou, Ruszczycki, Walser, & Johnson (2008) make the point that a Gaussian model provides a rough approximation of financial market returns, but fails to explain outlying events. It is the search for an understanding of what drives these outlying events that has driven a vast quantity of intellectual power, including this thesis.

An example of a real world asset market is seen in Figure 2 (data sourced from Yahoo Finance via a user defined query in R (2015)), which plots the log daily returns of the S&P500 index between 1985 and 2016. It illustrates that while returns for the S&P 500 have at times followed the prescribed distribution, there have been a number of periods where returns have deviated greatly, including the arrival of the 'Black Swan' in

1987 and 2008. Given the cost associated with the periods of extreme volatility, to this author a theory consistent with 'close enough is good enough' may not suffice.

Despite its apparent shortcomings, the EMH is not without empirical support in terms of the 'no free lunch' hypothesis. Malkiel (1999) has popularized the fact that investors have over the long term been unable to outperform the market. This argument is supported by further studies that have shown that money managers on average have failed to outperform the market (Rubinstein, 2001), a point that even the most stringent critics of the EMH have been forced to concede, see for example Thaler (2015). With regards to the 'price is right', the evidence is not as supportive; a point discussed later.

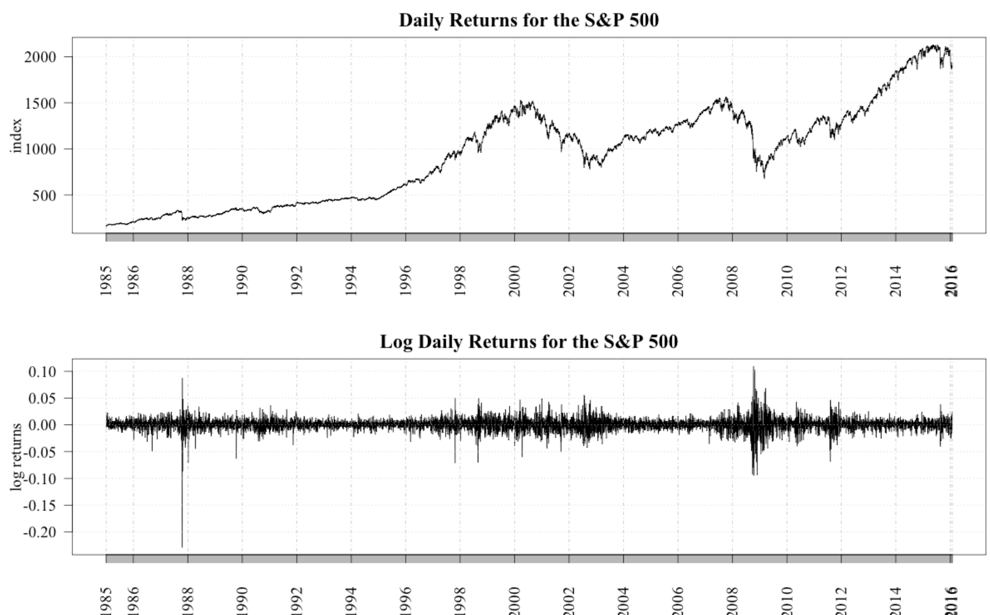


Figure 2: Daily returns for the S&P500 (1985-2016)

It should be noted that the EMH does not entirely preclude investors from ‘beating the market’. However, to achieve this, investors must be willing to accept greater risk than the market, the implication being that investors can not outperform the market on a risk adjusted basis. Therefore, any analysis of excess returns is required to determine if the excess returns were generated by mispricing (inconsistent with the EMH) or due to risk (consistent with the EMH), Thaler (2015).

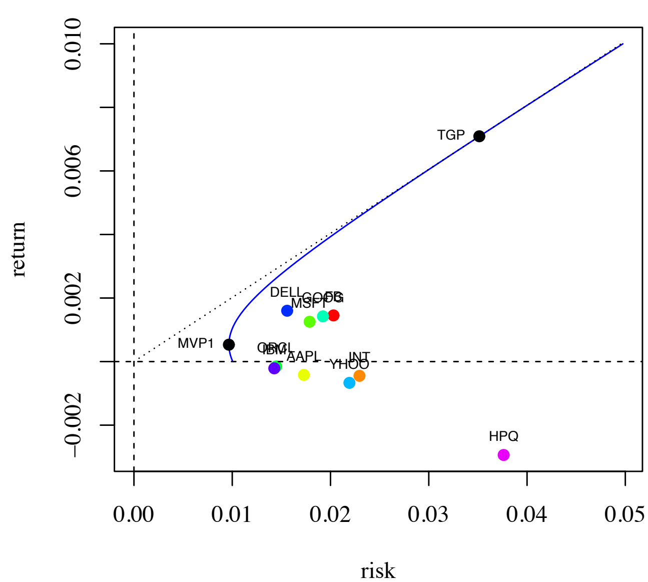
To assess this point, the EMH must be used in combination with financial models that makes use of its assumptions. The Capital Asset Pricing model (CAPM) – Sharpe (1964), which predicts that investors would only be compensated for the time value of money and the level of non-diversifiable risk (or systematic risk), is one such model. After initial promise and wide acceptance, the CAPM has subsequently been shown to be deficient (Fama & French, 2004). However, proponents of the EMH only see this outcome as reflecting poorly on the CAPM and not the EMH.

With regards to the existence of bubbles within the rational expectations approach (and therefore the EMH), an explanation was put forward by Blanchard (1979), who channeled the ideas of Keynes (2007). Investors can be seen to be acting rationally if they recognize a bubble and adjust their objective to “beating the gun” or in simpler terms, trying to sell out before a collapse. Under this assumption, speculative bubbles become consistent with rational investing. To counter this, following their study on financial returns, Diba and Grossman (1988) concluded that “stock prices do not contain explosive rational bubbles” as prices being no more explosive than dividends. Regardless of who is right, the process of detecting the presence of a bubble or rejecting their presence ex-ante remains difficult and as Kirman (1991) suggests, the debate remains open as to whether bubbles can or have been detected.

Another defense of the EMH comes from Rubunstein (2001), who in response to the criticism that markets show greater volatility than anticipated, points out that rational does not mean certain, and therefore returns can remain uncertain yet investors remain rational.

2.2.2 The Efficient Frontier

Markowitz (1952) first proposed the idea that agents would select a portfolio of risky assets on a frontier, the efficient one, where a portfolio would have the minimum risk for a given return or alternatively have the maximum return for a given risk. Figure 3 provides a demonstration of the efficient frontier as generated from the code provided in Andrecut (2013) which in turn sourced its data from Yahoo Finance.



Legend

Ticker	Company Name
AAPL	Apple Inc.
DELL	Dell Inc.
FB	Facebook Inc.
GOOG	Alphabet Inc.
HPQ	HP Inc.
IBM	International Business Machines Corp.
INT	World Fuel Services Corp.
MSFT	Microsoft Corporation
MVP1	Minimum Variance Portfolio
ORCL	Oracle Corporation
TGP	Tangency Portfolio
YHOO	Yahoo! Inc.

Figure 3: An example of the efficient frontier (Source: Andrecut (2013))

Figure 3 plots 10 stocks, as detailed in the legend, along with the minimum variance portfolio (MVP1) and the tangency portfolio (TGP) in the risk (x axis) and return (y axis) space. In addition, the code solves for the efficient frontier, which is illustrated by the blue line. From the chart it can be seen that individually the stocks sit below the frontier. However, investors can form portfolios of the stocks such that the risk and return of the portfolio is on the efficient frontier.

While the efficient frontier was released prior to the EMH, it utilizes many of the same principals, including, the existence of a representation rational investor. This investor, who seeks to maximize their utility given their budget constraint, will choose either on purpose or inadvertently, a portfolio that generates the largest possible return with the least amount of risk. Given this theory was developed utilizing a static equilibrium approach, the challenge now exists to see if an ABM, utilizing out of equilibrium dynamics, is capable of replicating the results. By achieving this, certain questions, as detailed by Arthur (2006), can be addressed, including how agents' actions, expectations and strategies change, and what impact this has on the system as a whole.

An attempt at such a process was provided by Steinbach et al. (2010), who successfully utilized a discrete time and state model with interacting agents connected via a network to demonstrate that agents were capable of selecting efficient portfolios without the optimization equation as set out by Markowitz (1952). However, the returns were exogenously provided to the agents, which contrasts to the proposed model, where asset and therefore portfolio returns are generated endogenously. The inclusion of endogenous returns provides a richer data source and allows the questions posed by Arthur (2006) to be explored in greater detail.

2.2.3 Evidence Against the Efficient Market Hypothesis

Evidence against the EMH remains somewhat of a mixed bag. While there is no disputing the fact that some investors have done very well at times, there are very few that have maintained this record over the longer term, hence the 'no free lunch' doctrine has remained difficult to dislodge. With regards to the 'price is right', the evidence is more damming. For example, Shiller (1980) was able to demonstrate that stock prices over an extended period did indeed deviate from their inferred intrinsic value.

Further evidence against the EMH is provided by the set of stylized facts financial markets have exhibited, which are inconsistent with those proposed by the EMH. As outlined by Cont (2005) and further supported by Johnson, Jefferies, & Hui (2003), market returns have demonstrated:

- Excess volatility – the existence of large movements which is not supported by the arrival of new news;
- Heavy Tails – returns exhibit "heavy tails" indicating returns deviate more than anticipated and do not follow a Gaussian distribution;
- Volatility Clustering – large changes are followed by further large changes; and
- Volume/volatility clustering – trading volumes and volatility show the same type of long memory.

Further to this, Mandelbrot (1963) first provided evidence that returns followed a very unique distribution – a power law distribution. Lux (2006) provides a detailed review of the empirical evidence supporting the existence of power laws in financial markets. A summary of the research into why power law distribution for stock market returns exist is provided by Sornette (2014), with the possibilities including:

- The interplay between the power law distribution of the sizes of large financial institutions and the trading in these firms;
- The close link with Pareto wealth distribution and market efficiency; and
- Sudden drops in liquidity rather than outsized orders.

For investors, the main implication of returns following a power law is that the risk of large losses is much higher than suggested by the EMH suggests, and markets are more volatile. While the original works of Mandelbrot (1963) fell out of favor (MacKenzie, 2008), the reality of continued episodes of boom and bust, and mounting statistical evidence, provides strong evidence that in fact markets do not function in accordance with the EMH. The existence of power law returns suggests the presence of a complex system, thus providing a clue into a mechanism that is capable of generating such returns.

2.3 Alternate Approaches

2.3.1 Introduction

The evidence against the EMH coupled with the questioning of its underlying assumptions has supported the rise of behavioral finance, which according to De Bondt et al. (2008) "studies investor decision processes, which in turn shed light on anomalies i.e. departures from neoclassical finance theory". Further, Sornette (2009) states that the intention of behavioral finance "is to join the objective approach with the interpretative approach, with the intention of understanding how markets reflect the actions of people acting with thoughts and emotion as opposed to the idealized investor" (an idealized investor as defined by EMH). Within the realm of behavioral economics/finance there are two important considerations relevant to the development of this thesis – bounded rationality and herding.

These two points challenge one of the key underlying assumptions of the EMH, and the one that raises the most questions is that investors are homogenous and share the same rational expectations regarding asset prices (Arthur, Holland, LeBaron, Palmer, & Tayler, 1997). These assumptions were made under the umbrella of "positive economics" and were required to produce the eloquent closed form mathematical solutions that cover the finance landscape. It is the development of modern simulation techniques (including ABMs) that allows these assumptions to be removed and alternative approaches taken.

2.3.2 Bounded Rationality

Simon (1955) raised the prospect that humans do not have the computational power to make the calculations required of them to act as fully rational agents. Instead, agents are forced to act as if they have bounded rationality, forcing them to search for solutions by trying solutions that seem appropriate – i.e. heuristics. This approach is consistent with the thoughts of Keynes (1937) - who indicated that when individuals are making a decision, their basic decision making framework should include the following rules:

- The present is more relevant than the past;
- They have to accept the current price until something new comes along; and
- Understand that individual judgments are worthless so it is best to fall back on the

view of the rest of the world.

The link between the theory of bounded rationality, complex systems and ABMs was made by Arthur (1994) when he solved his El Farol Bar problem[7]. Within his solution, Arthur posed and the answered the question of what occurs if humans cannot rely on other humans to act in a rational manner? The answer was that humans needed to employ inductive (bottom-up) reasoning, as opposed to the deductive top-down approach of the EMH and normative economics. The bottom up process involves agents forming a number of hypothesis/belief models, acting on the most credible, and then updating or discarding those that do not work. A vital aspect of this approach was that while a system is able to reach a solution close to an efficient outcome, it never settles into an exact equilibrium - a characteristic of today's financial markets.

The El Farol problem and its relevance to financial markets has been demonstrated through the Minority Game (Challet, Marsili, & Zhang, 2013) and the $ Game (Andersen & Sornette, 2003). Despite these successes, a shortcoming, according to Johnson et al. (2003) is that the model can not explain everything, as crowds form unintentionally, rather than by a defined process. Despite this, the principle of bounded rationality remains a core theme within the literature of behavioral economics and ABMs.

2.3.3 Herding

The aforementioned shortcoming of the El Farol model is addressed by the analysis of how and why herds form. Sornette (2009) defines herding as "many people taking the

[7] A problem in which if everyone reached the same conclusion via a deductive approach they will ultimately all be wrong.

same option because they are mimicking the action of others". This raises the key questions of why people wish to mimic another and is it relevant to financial markets?

One explanation is that if agents have bounded rationality or lack vital information, it may indeed be optimal for them to mimic, rather than to solve the problem for themselves. Evidence of herding behavior has been found in financial markets (Cont & Bouchaud, 2000) and includes investors switching their investment strategies together – changing from buying to selling simultaneously, resulting in a crash or the continued buying of an asset that is already priced in excess of its fundamental value. The latter is seen as being responsible for the formation of bubbles.

If investors mimic others, then their information network becomes an important consideration (Ozsoylev & Walden, 2011). Within this network, the social network of investors is a key consideration because within social networks it is often found that people copy or imitate what others do, or think in that network (Ormerod, 2012).

Keynes's (1937) beauty contest analogy presents another insight into why herding can occur. In it he proposes that professional investors become more interested in understanding what the market thinks the value of a particular asset (the perceived value) is, rather than understanding what the fundamental value of the asset is. The implication being that prices move not by changes in the fundamental value of stocks, but by changes in the perceived value of stocks. Importantly, under this process price movements can generate momentum due to a positive feedback process – something that is not considered in a rational expectation model.

Scharfstein and Stein (1990) provide further support for this argument by suggesting that herding can arise as a rational response of investment managers trying to enhance their reputation as they try and match the performance of their peer group. In addition, they suggest managers may herd in an attempt to protect their reputation by avoiding the blame from an investment decision outside of the herd. Importantly, as Banerjee (1992) points out, any equilibrium found under these conditions is inefficient. Another possible source of herding is word-of mouth. Among others, studies from Hong, Kubik, & Stein (2005) and Shiller and Pound (1989) support its presence amongst both professional and amateur investors. This point provides support to the theory that opinions of investors flow across a network.

In support of herding as a driver of asset prices, Cont and Bouchaud (2000) and Kirman (1991) pioneered the use of models that utilized a herding framework. Their models produced results that were quantitatively comparable to the existing empirical findings in relation to the distribution of stock market returns.

2.3.4 Networks

The relevance of networks in understanding how herding may occur amongst investors comes from trying to understand how an idea or opinion can spread between investors. The principle of contagion or the cascading of an idea across different network structures was demonstrated by Watts (2002) and extended by others including Santos and Pacheco (2005). In addition, the general importance of networks, and the analysis of them, is outlined by Newman (2010) who suggests that they are “powerful means of representing patterns of connections or interactions between the parts of a system”. In support of this research project, Schweitzer et al. (2009), make a compelling argument that the analysis of economic networks is essential for extending existing economic theory. The genesis of the argument being that the 2007 financial crisis highlighted that existing theories, and the policies associated with them, were inadequate in understanding and analyzing the growing interdependencies that had formed across global trade, supply chains and investments networks.

There is a growing body of literature from the likes of Ozsoylev (2005), Colla and Melle (2010), Ozsoylev and Walden (2011), Han and Young (2013), Ozsoylev, Walden, Yavuz, & Bildik (2014) and Walden (2014) that have taken the ideas from the likes of Hong, Kubik, & Stein (2005) and Shiller and Pound (1989), that networks may exist between investors and formalized them into closed form models utilizing network structures. The work has produced key insights, including some supported with empirical proof, including the following:

- The information of a socially influential agent, the agent whom many learn from, has a higher impact on the risky asset price compared to information of those with less influence (Ozsoylev 2005);
- The topology of a social of network impacts information efficiency; with one implication being price volatility decreases with the average number of information sources agents have. Therefore greater social communication improves efficiency (Ozsoylev and Walden 2011);
- Agents who are close in the network have positively correlated trades, while distant agents have negatively correlated trades (Colla and Melle 2010). Interestingly, these authors make the point that the delay in forming a closed form

model was due to the complexity of combining networks, rational agents and endogenous price formation; and

- Centrality is directly related to acting early on information (Ozsoylev, Walden, Yavuz, & Bildi 2014) and more specifical it is the eignvalue centrality measure which is important (Walden, 2014).

Within network science literature there are four general types of network; regular/lattice, random, small world, and scale free networks. In high levels terms, the differences relate to how each agent is assigned their neighbors and the number of neighbors they have. These small differences are capable of generating non-trivial differences in the outcome of the system and for the agents within the network – hence their inclusion in the proposed ABM (see sections 3.3.3 and 4.2.2).

Figure 4 illustrates a traditional lattice (left) and small world (right) network. A lattice network has an agent joined to a given number of agents that immediately surround them. While lattice networks are not common in the real world they may be relevant to in financial markets given the location of trading desks and the like.

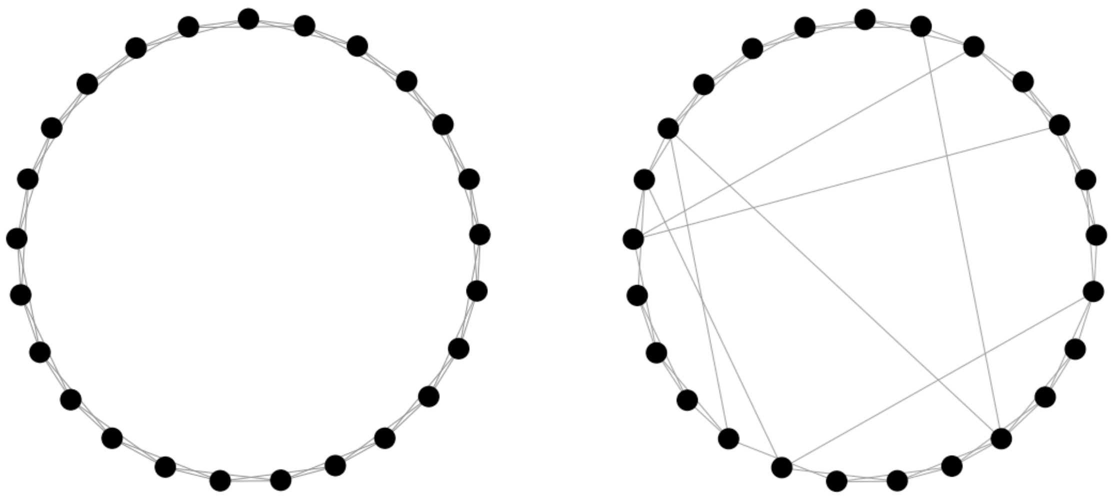


Figure 4: Lattice (left) and small world (right) networks

Small world networks were first proposed by Watts and Strogatz (1998) as a network structure that was consistent with social networks of people. The model is premised on the fact that geographical proximity plays an important role in the formation of social networks. A well-known feature of observed social networks is that they show a high degree of clustering, yet a relatively small diameter. The ramifications being that small-world networks tend to be more robust to perturbations than other network architectures, but are vulnerable to targeted attack.

A random network – one consisting of N nodes, where each node pair is connected with probability p, was first proposed by Erdös & Rényi (1960), is illustrated on the left of Figure 5. Given its abstract nature, it generally serves as a theoretical base line only. Using this baseline, Barabási & Albert (1999) were able to create a special form of a random network - the scale free network. This results in the degree distribution of the network following a power-law, with the consequence being that in contrast to a small world network, a scale free network tends to be more vulnerable to perturbations but is robust to targeted attack. A popular mechanism for generating a scale-free network is through preferential attachment - a process where an agent chooses their neighbors based on how many neighbors the potential neighbor already has. This process delivers a ‘the rich get richer’ outcome and is illustrated on the right of Figure 5.

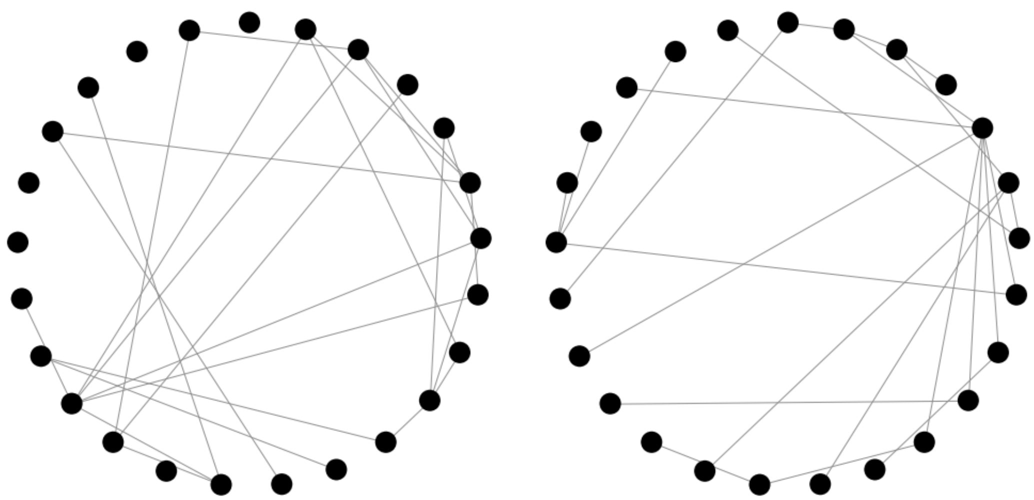


Figure 5: Erdos Renyi (left) and scale free (right) networks

In direct support of this research topic, Panchenko, Gerasymchuk and Pavlov (2013) have demonstrated the relevance of networks by implementing network frameworks within an ABM based artificial stock market. The model showed that the network structure did indeed influence the price dynamics of their artificial stock market. Their model utilized a variation of the Brock and Hommes (1998) evolutionary algorithm framework but the possibility of utilizing other market frameworks exist. The authors also left open numerous research questions including the effects of dynamic network formation and network structures other than small world networks. In another example, Alfarano and Milaković (2009) extended Kirman's (1993) model, which was based on the processes of social interaction among agents who displayed a tendency to follow the crowd, by overlaying a network structure on it. While the model was deemed to be quite robust with respect to behavioral heterogeneity, it showed that "the network structure describing the very feasibility of agent interaction turns out to have a crucial and non-trivial impact on the macroscopic properties of the model".

Harras and Sornette (2011) utilized an Ising model based ABM that saw investors consider the expected actions of their neighbors as one of three information sources used to determine their action. This model is discussed in greater detail in Section 3 as it formed the foundation of the model used in this thesis.

2.3.5 Complex Systems

A defining characteristic of a complex system is its ability to self-generate large endogenous changes (Jefferies, Lamper, & Johnson, 2003). Therefore, if one considers market bubbles and crashes as such a change, then utilizing a complex system framework is an appropriate research method for trying to understand the behavior of financial markets. Further support for this approach is provided by Simon (1982) who indicated that there is a need to understand why certain regularities (booms and crashes) occur despite no top down planning. It was proposed that this regularity comes from the local interaction of autonomous heterogeneous agents - a trademark of a complex system.

Johnson et al. (2003), see complex systems containing some if not all of the following characteristics:

- Feedback – the nature of and magnitude of this may change over time;
- Non-stationarity – the dynamics or properties of a system may not continue as is;
- Many interacting agents – the system contains heterogeneous agents that interact;
- Adaptation – an agent can adapt to the environment;
- Evolution – the population evolves as a result of interacting and adapting;
- Single realization – at any particular time the system is a single realization among many possible ones; and
- Open system – there are both endogenous and endogenous impacts.

Empirical support for employing a complex system framework, including ABMs (see LeBaron, 2006), has come from their ability to generate the extreme events and asset returns that are consistent with the stylized facts outlined earlier, as opposed to the theoretical solution put forward under the EMH framework.

2.4 Agent Based Models (ABMs)

Given the characteristics of a complex system, more traditional analytical approaches are rendered ineffective, and researchers have turned to computer simulation to understand their dynamics. In particular, ABMs are extensively utilized in the study of complex systems. ABMs allow for the interaction between individual agents (investors in this case), who act and undertake actions based on the context of their environment and basic rules. Important considerations of these models are that they are capable of addressing issues such as; heterogeneous expectations and investment approaches amongst agents, out of equilibrium dynamics, the impact of a changing external environment (including shocks) and the ability for the population to adapt and evolve. ABMs are able to achieve this because they dispense with optimization and hence are not constrained to equilibrium conditions (Sornette, 2014).

2.4.1 Artificial Stock Markets

The utilization of ABMs to create artificial stock markets, which was first mentioned in Section 1.3, can be traced back to the Santa Fe Institute in the late 90's, where Arthur, Holland, LeBaron, Palmer, & Tayler (1997) attempted to comprehend under what conditions agents would or would not behave in accordance with a rational expectations model. The previously discussed Minority Game (see Section 2.3.2) also appeared around this time. Sornette (2014), LeBaron (2000) and (2006) provide detailed surveys of the various analytical frameworks that have subsequently utilized ABMs in developing artificial stocks markets.

According to LeBaron (2006) the rationale for continuing to utilize ABMs to create artificial stock markets are:

- The question of whether markets are rational remains unanswered and many other questions regarding the behavior of markets remain unanswered and therefore alternate approaches are warranted; and
- Given the large amount of data in financial markets the output of ABMs can be validated.

To this point, Darley and Outkin (2007) utilized an ABM to successfully answer questions regarding the NASDAQ Stock Market's decimalization which in occurred 2001 and Cui, Wang, Ye and Yan (2012) produced an ABM that matched the macro characteristics of the Chinese market. In another approach, Duffy and Ünver (2006) were able to replicate the actual experimental results of the Smith, Suchanek and Williams (1988) paper, which demonstrated that asset bubbles can develop despite investors being fully informed, via an ABM.

The various approaches have the common theme of utilizing heterogeneous agents in terms of both expectations and investment strategies with the intention of; studying how agents act, how prices are set, reproducing the stylized facts of the markets and the understanding influence of the market's microstructure. The differentiating factors for each framework is how they handle agent preferences, the price setting mechanism, whether evolution is allowed, and how strategies were stored. Cont (2007) classified the various frameworks into four categories:

- Heterogeneous arrival of information;
- Evolutionary models;
- Behavioral switching; and
- Investor inertia.

One of the downsides of the early models, were that they tended to be complex, making it difficult to determine the influence of each of the inputs and what the key determining factors were (LeBaron, 2000). Sornette (2014) further suggests that the predictive power of a given model is constrained and it is unclear how to generalize a given result.

An important consideration in designing an ABM based artificial stock market is the mechanism under which the price in determined. Options include, establishing order books and fulfilling those orders through an auction process or making use of a market maker, who co-ordinates the market. The early models tended to favor the market marker approach while the later models have turned to a formal auction market. In what some may consider 'hand waving', the market marker is able to clear the market by providing liquidity and also standing on the other side of all trades at each tick. This means that investors are guaranteed to have their trades executed and allows for price determination to occur at each step. Therefore, the market will not become frozen due to a lack of liquidity or an inability to match orders, something that is a real world consideration. In addition, returns are not impacted by large gaps in a discrete order book. The downside to this approach is that the model losses the ability to assess what the investors were willing to buy and sell the asset for.

One shortcoming not explicitly mentioned in the review of the ABM based artificial stock markets is the fact that agents do not create a portfolio of stocks but rather just decide to allocate their funds between a risk free asset, such as cash, and a risky asset, which is seen as a proxy to investing in an index. I feel that this approach may gloss over important aspects of investing such as stock selection and the benefits that result from diversifying risk across risky assets. It is for this reason the quasi-efficient frontier framework and multiple assets were introduced into the research topic.

2.5 Section Summary

The justification for the utilization of ABMs in analyzing financial market was made in this section of the thesis. The argument for the need to explore alternative approached was made in Section 2.2, where the shortcomings of the EMH were presented. Section 2.3 then presented the emergence of one alternative approach, treating financial markets as a complex adaptive system (CAS). Finally, Section 2.4 provided the background of, and the justification of utilizing ABMs to model a CAS. More specifically it provided evidence and high-level considerations of how artificial stocks markets can be created through an ABM.

Section 3 provides justification for the model implemented in this thesis. The model is one that proved capable of having investors consider: the actions of their neighbors, multi assets and for those investors to adapt and evolve. The exact details of the various agent classes and how the model considers the various research questions is provided in Section 3.3 through 3.8.

3 APPROACH AND MODEL DESIGN

3.1 Introduction

This chapter provides a detailed description of the model that was implemented to address the research question as presented in Section 1.3. The justification for the selection of the implemented model is provided in Section 3.2, with Section 3.3 detailing the agent classes implemented in the model and their relevance to research question. Sections 3.4 through 3.8 provides a walk through of the actual implementation of the model, as well as providing initial insights into the expected behavior of the model. All extensions to the original model are detailed throughout this section. The final two sections of the chapter detail the verification steps undertaken to ensure the model performed as intended and the outputs the model created.

3.2 Model Background

The model utilized for the research topic utilized and extended the Ising based ABM of Harras and Sornette (2011) (H&S hereafter). Ising models had their genesis in mathematics before being adapted by physics, where they are seen as the "simplest representation of interacting elements with a finite number of possible tasks" Sornette (2014). They have subsequently been adapted to the fields of finance and economics because, as per McCoy and Wu (1973), they allow competition between the ordering forces of imitation or contagion and the disordering impact of varying news sources that result in heterogeneous decisions. The effects of imitation and contagion are also commonly analyzed with the aid of networks.

The basic premise of the H&S model is that boundedly rational investors[8] (for completeness, an investor(s) is defined as an agent who invests in the model) have access to three sources of information; the expected actions of their neighbors ($E_{ij}[a_{ik}(t)]$), public information ($pi_i(t)$) and private information $\epsilon_{ij}(t)$. The investors utilize these sources of information to determine their propensity to invest (ω_{ij}). Equation 1 details the exact calculation that the investors perform at each time step (defined as a tick), while Sections 3.5, 3.5 and 3.6 provide greater details on each coefficient, their relevance and the processes involved in updating them. The model then allows the investors to transact, with the new price endogenously determined for the asset(s) along with a variety of accompanying asset and portfolio statistics.

From Equation 1 it can be seen that the level of influence of each information source is weighted by two variables[9], with one of these being fixed and the other variable. The fixed values are given by c_{1ij}, c_{2ij} and c_{3ij}, while the variable coefficients are network trust (nt_{jk}) and public trust (pt_i). By altering the c_{1ij}, c_{2ij} and c_{3ij} coefficients, different dynamics are generated in the H&S model. In particular, when the upper limit for c_{1ij} is set at 4, bubbles in the risky asset's price appear. Given this, all parameter sweeps for the original and revised model will include values of c_{1ij} ranging from 1 to 4.

---

[8] In this instance they are only considering past information.
[9] The exception is private information, which has a single variable.

Equation 1: The decision equation

$$\boldsymbol{\omega_{ij} = c_{1ij}\left(\sum_{k=1}^{K} nt_{jk}\,(t-1)E_{ij}[a_{ik}(t)]\right) + c_{2ij}pt_i(t-1)pi_i(t) + c_{3ij}\epsilon_{ij}(t)}$$

Or in simpler terms;

$$Decision\ score\ = Network\ score\ + Public\ score + Private\ score$$

The model considers the processes of adaption and evolution through investors continually reassessing and adjusting the trust in each of their information sources (Step 9 and 10 of Figure 6) based on the ability of each source to predict the appropriate action. An appropriate action is when the information tells the investor to buy and the price subsequently increases (and vice versa for a sell signal).

The key findings of the H&S model, which form the justification for utilizing the model as a foundation for this thesis, were:

- Price movements were impacted by how strongly the agents are influenced by their neighbors;
- The asset returns, which were fat tailed, did not match the Gaussian distribution of the public and private information, thus indicating the model's ability to generate the stylized facts of financial markets: and
- The model is able to identify the conditions under which a bubble forms.

Prior to making the necessary extensions to the model to meet the specific research questions of this thesis, I attempted and satisfactorily succeeded in replicating the output of the original H&S model. This process formed a key part of the verification process for the model, a subject that is expanded upon in Section 3.9. The key metrics that were replicated were:

- Asset returns and their volatility (as measured by Equation 10) were consistent with Figure 1 and Figure 2 of H&S (see Appendix 1 for comparative charts); and
- The range of, and variability in the level of network and public trust.

Changes to the model included; introducing multiple risky assets, an alternate source of public information, as outlined in Section 3.5.2 and varying network structures, as outlined in Section 3.3.3. The justification for adding multiple risky assets is, as previously mentioned (see Section 2.4.1), one of the shortcomings of the previous artificial stock markets in that investors are faced effectively with an asset allocation problem – invest in a risk free asset or a risky asset, and not how to choose between risky assets. By way of definition, a risk free asset is assumed to have a guaranteed return with no deviation (in the H&S model it is assumed to be zero), while a risky asset has a variable return, with no guarantee that it is positive. The higher the variability in the return of an asset, as given by its standard deviation, the riskier it is.

In reality investors are faced with the decision of how to allocate their wealth across and within asset classes. For example, a typical investor will invest in domestic and international bonds, domestic, international and emerging market equities and various alternative assets. Within these classes they will form a portfolio, whether it be through direct investment or through multiple mutual funds. By investing in this manner an investor is attempting to diversify the risk of their investment portfolio. Complicating this process is the fact that some assets/asset class returns are correlated while others are not, yet all of these assets will tend to become correlated in times of financial distress. As previously stated in Section 2.2.2, the Markowitz's (1952) efficient frontier was one of the original mathematical based frameworks offering investors a solution as to how to diverify their risk. Hence, its inclusion within this research topic.

The author's model was implemented in NetLogo 5.3 (Wilensky, 1999). The NetLogo network extension was utilized to generate the Erdos Renyi network and the associate network statistics for all the network structures. The agents were initialized as per the NetLogo default of a random asynchronous order. The assumed time period per tick is a calendar quarter. Figure 6 provides an overview of how the model flows, with each step described in greater detail in the following section.

- **Initial Steps**
- Step 1: Assets initialised (Section 3.2.1)
- Step 2: Network formed (Section 3.2.3)
- Step 3: Network populated with investors (Section 3.2.2)

- **Repeated Steps**
- Step 4: Asset information updated (Section 3.4.2)
- Step 5: Agents assess information sources (Section 3.4)
- Step 6: Agents make investment decision (Section 3.5)
- Step 7: Order book formed (Section 3.6)
- Step 8: Market cleared and asset returns calculated (Section 3.6)
- Step 9: Agents reassess the value of public information (Section 3.7.1)
- Step 10: Agents reassess the value of their neighbors (Section 3.7.1)
- Step 11: Agent portfolios updated (Section 3.7.2)
- Step 12: Portfolio valuation performed

Figure 6: Overview of the model's implementation

3.3 Agent Classes

3.3.1 Assets

To address multiple assets, the implemented model has the capability to have between 1 and 10 risky assets (denoted as *I* assets with *i* being the i^{th} asset). With the introduction of multiple assets, assets effectively become agents and have the ability to be assigned, maintain and evolve heterogeneous characteristics. Table 1 details the key attributes of the asset class and the role they play in the model.

In a key difference from the H&S model, the extended model sees the asset(s) maintain an earnings per share (EPS) value for each period. EPS reflects the income generating ability of the asset and is a key component in determining the fundamental value of an asset. The model also makes use of the past values EPS values to generate a future earnings forecast for investors. The importance of including an earnings forecast is articulated by Sornette (2014) when he suggests "in a given financial bubble, it is expectation of future earnings rather than present economic reality that the average investor".

In a further difference to the H&S model, by combining the EPS of the asset with its payout ratio, each asset returns a dividend per share (DPS), as shown in Equation 2. The justification of introducing dividends comes from the fact that dividends are a key component in the total returns for most assets. For example for the S&P 500, dividends are responsible for 42% of total returns (Ro, 2013). Each investor's dividend is held in their dividend bank and investors do not have access to those funds for the purpose of further investing.

Equation 2: Dividend per share (DPS) per period

$$d_i(t) = eps_i(t) * pay\_out\_ratio_i$$

Table 1: Asset characteristics

Characteristic (model variable)	Function and Description
Price in $'s (price)	The price of each asset is initiated at $1 as per the H&S approach and is updated endogenously at each tick as per the process outlined in Section 3.7. In turn, a return series is determined given the changing prices.
EPS in cents (EPS)	To allow for an alternate process by which investors asses public information, each asset has an earnings per share (EPS) value. With the initial price per asset being $1, an acceptable P/E ratio for an equity being 15 and the model simulating quarterly updates, the initial EPS value (in cents) is given as (1/4)/15 * 100. The initial EPS is also the mean parameter for the probability distribution function (PDF) required for Step 4 of Figure 6 (nb. only for the extended model). The role that an asset's earnings have on the behavior of each investor is outlined in Section 3.5.2.
Asset Returns (ave_return and stddev_a_rt)	Through the market clearing process an asset's returns are generated. These returns are stored in a list and this allows the average and standard deviation of those returns to be calculated. The values of those calculations are stored in these variables.
Correlation (corr_r)	To allow for the varying effects of multiple assets, the public information of the assets are correlated to the first asset (Asset 0) by this variable. Greater detail of its use is provided in Section 3.5.2. The value is set by the `corr_r` parameter and is constant for all risky assets and ranges between 0 and 1.
Pay_out_ratio	The payout ratio determines the percentage of earnings that are returned to the investors. In reality, a low growth stock will generally have a higher payout ratio and a high growth stock a lower one. Given the H&S model did not include a dividend (or EPS), any benchmarking sets this value at 0. The variable will be utilized to assess the impact of an asset's payout ratio on price volatility utilizing the extended model.
EPS_dev H&S_dev	Regardless of which model is implemented, Step 4 of Figure 6 requires a standard deviation for the PDF. For the H&S model the

	first asset has a standard deviation (H&S_dev) of 1 and a mean of 0. For each subsequent asset added the standard deviation increases by 0.1. For the revised model, the EPS's standard deviation (EPS_dev) for the first asset is set by the `std_eps` parameter, which is a fraction (between 0 and 1) of the mean EPS value. It then increases by .1 for each incremental asset.
Consensus EPS	The extended model also requires each asset to maintain a consensus EPS forecast (EPSF). The forecast is homogenous for the population and its use is detailed in Section 3.5.2. The forecast is formed as the ensemble average of past EPS result as per the following: $$\langle epsf_i(t)\rangle = \alpha * \langle epsf_i(t-1)\rangle + (1-\alpha) * eps_i(t-1)$$ α relates to the `memory_weight` parameter which is detailed in Section 3.8. Given the EPS results are drawn from a normal distribution, the consensus forecast should mean revert over time but the use of the above equation does allow for the development of trends.
Consensus_Accuracy	The revised model replaces the public information process of the H&S model with investors assessing the accuracy of the consensus forecast. This attribute captures the outcome of the processes described in greater detail in Section 3.5.2.

3.3.2 Investors

The population size of investors is variable and is set by the user. However, to remain consistent with H&S, 2,500 investors (denoted by J with j referring to the j^{th} investor) are used in all experiments. At all times they hold a combination of the risk free asset (a proxy for cash) - which can be redeemed to purchase the risky asset(s), and the risky asset(s). The objective of the investors is to improve/increase their wealth by buying the risky asset(s) when they think the value will increase and selling it(them) if they think the value will decrease. At initiation the investors are provided one unit of the risk free asset and one unit of each of the risky assets. Table 2 summarizes the key variables that the investors own and how they are utilized in the model.

Table 2: Investor variables

Symbol	Name	Purpose
c_{1ij}	Network influence	Each investor is initiated with a fixed value (float) that is drawn randomly from a uniform distribution between 0 and a value up to 5 that the user decides. The variable is used to weight the information the investor generates from their network. As investors have a different value for c_{1ij} this introduces a level of heterogeneity within the population. While beyond the scope of this paper it is a worthier consideration that if the value of c_{1ij} were allowed to be less than 0, it would introduce contrarian investors. Analyzing the impact of different levels for this variable and c_{2ij} forms a key component of this thesis and the H&S paper. An acceptable interpretation of these variables is that a higher value (such as 4), indicates a higher initial bias to that information source.
c_{2ij}	Public information influence	Similar to the above with the exception of weighting the public information by a value between 0 and the value of the `public_influence` parameter.
c_{3ij}	Private information influence	Similar to the above, with the exception of weighting the private information being set by the user defined `private_influence` parameter. Note that the private information has no adaption variable.
nt_{jk}	Network information trust	While this variable is initiated at 0, investors update the value at each tick (see Section 3.8.1) to reflect an increasing or decreasing level of trust in the information coming from each of the investors in their neighborhood. c_{1ij} is then used to compound the information from the investor's network. It should be noted that an investor places trust in a neighbor's overall ability to make the right selection (their average ability) and they do not keep track of a neighbor's ability to pick individual assets.

pt_i	Public information trust	Similar to the above with the exception of being the trust an agent has in public information for each specific asset ($asset_i$). As detailed by Section 3.8.1, the level of public trust is homogenous across the population as all investors receive and assess the public information in the same manner. However, in contrast to the trust an investor has in their neighbors, an investor maintains public trust at an asset level. The alternative approach - taking the average from the assets - was assessed but it was felt that by aggregating the trust a level of freedom was lost in model. It is not unrealistic to expect investors to have varying levels of faith in a stock's ability to surprise based on past results.
$\overline{\omega}_j$	Transaction threshold	Each investor is initiated with a fixed value (float) that is drawn randomly from a uniform distribution between 0 and the value of the `threshold` parameter. The default level for the model is 2. The variable is used as the value by which an investor decides to either buy, hold or sell (see Table 4 and Section 3.6 for a detailed description). As investors have different values for $\overline{\omega}_j$, another level of heterogeneity exists within the population. H&S attribute this value to the risk aversion of the investor. Indeed an investor with a high $\overline{\omega}_j$ requires significant evidence before they commit to a transaction, while a low value will see the agent act on the slightest change in information.
tr	Transaction ratio	The value is set by the `transaction_ratio` parameter and represents the fixed fraction that an agent is willing to trade. The default value is 0.02. While the transaction ratio is fixed, future research may look to have this variable vary based on how confident an investor is.

3.3.3 Networks

Prior to the initiation of the investors, their network topology is generated (Step 2 in Figure 6). The use of alternative network structures is a variation on the H&S model, which only utilized a lattice in assigning neighbors. The authors did comment that their results held for both random and complete graphs but they made no mention of the other network structures outlined in Section 2.3.4. Once the network is initialized, the model then populates the nodes with the investors.

The network structures utilized in the model are those illustrated in and Figure 4 Figure 5 in Section 2.3.4, with the exact characteristics used for the formation of each network being detailed in Table 3.

Table 3: Network characteristics

Network	Key Characteristics
Lattice	Number of links per investor = four The `Ring_M` parameter is set to 2
Small world	Number of initial links per investor = four The probability of rewiring (`prob_of_rewire` parameter) = .10
Random network	Probability of connection (`prob_of_link` parameter) = .0016
Scale Free	Number of hubs (set by the `Ring_M` parameter) = 10 Probability of connection (`prob_of_link` parameter) = .20

In terms of generating the networks, with the exception of the Erdos Renyi network, where the network extension of NetLogo was utilized, the following user determined algorithms were used.

For the lattice network the first step in generating the network is for the user to set the number of neighbors they want each investor to have via the `Ring_M` parameter. Given undirected links are being formed by the investors, the parameter is set as half the number of required neighbors because an investor becomes a neighbor with another investor regardless of whether they create the link or the neighbor creates the link with them. Next the investors are placed into a list, which is sorted by the investor's numerical identification number. Each investor is then asked to form an undirected link with the next highest investor in the list. The process is repeated for the investor by value of the `Ring_M` parameter, with the following link being formed with the next highest neighbor.

The small world procedure firstly implements the lattice network procedure using the same `Ring_M` parameter. Next, each investor forms a list of their links, which they cycle through asking each link to rewire, with the probability provided by the `prob_of_rewire` parameter, to another investor within the population that they are not connected with. Rewiring involves creating an undirected link to the new investor and then removing the link to the existing neighbor.

To create a heavily skewed degree distribution with approximately 5,000 links for the scale free network, the user needs to decide on the number of hubs (set by the `Ring_M` parameter) they require and the probability (`prob_of_link` parameter) of an investor connecting to each of those hubs. The procedure operates by each investor identifying the hubs by finding and then forming a list of the investors with the most number of neighbors. The number of investors in the list is determined by the `Ring_M` parameter. Next, each investor with a probability determined by the `prob_of_link` parameter, links with each of the investors in the list. This process is sufficient to create a scale free network via a preferential attachment process.

The above approaches were taken to ensure that the number of edges (5,000) and the average number of neighbors (4) would be consistent across the different network structures. This approach ensures that any difference in the outcome is not influenced by the number of edges, but solely by the degree distribution of the networks.

In the H&S model the number of neighbors was consistent at 4 for all investors. However, with the introduction of the different network structures, the number of neighbors is no longer consistent across all investors. The ramification being that if an investor has only one neighbor they will have an initial bias towards public and private information, as they collect less opinions and if they have a lot of neighbors there will be an initial bias to the information coming from their network as they collect more opinions. Consideration was given to normalizing the network score but this would minimize the impact of the different network structures. In addition, as investors continually reassess their trust in each information source, it does not preclude a single neighbor becoming very persuasive. Conversely, an investor with a large number of neighbors may end up attributing very little trust in them.

The links between neighbors are undirected and not specifically weighted nor are the links dynamic. However, given investors vary the level of trust they have in each neighbor, the network does become quasi dynamic i.e. as the trust in a neighbor increases, the weight of a directed link between the two effectively increases. Future iterations of this model could look to have investors jettison untrustworthy neighbors and search for investors that have superior performance. Also, directed links may be a worthwhile investigation - just because you listen to a neighbor there is no guarantee they listen to you.

One of the key differences in the network structures is their centrality. The concept of centrality captures the implications of knowing not only who your direct neighbors are is important, but also who your neighbors' neighbors are, who your neighbors' neighbors' neighbors are, etc. (Walden, 2014). The implication for financial markets is that investors who are centrally placed tend to receive information signals earlier than peripheral agents, and therefore tend to perform better (Walden, 2014).

One centrality measure is betweenness centrality, which was introduced by Freeman (1978), with the intent of capturing the idea that an agent (node) can be more or less important based on the number of chains that pass through it (Caldarelli & Catanzaro, 2012). The process involves finding the frequency with which a node lies along the shortest path between two other nodes. The agent with the highest betweenness is the one with the most 'best paths'[10]. Its relevance as a measure is that it is seen as an index of potential gate keeping and indicates the power and access to the diversity of what information is flowing through the network.

Another relevant measure is closeness centrality, which aims to find the agent that is closest to all other agents. It is calculated as the sum of distances[11] of an agent to all other nodes and is seen as an inverse measure of centrality. Its relevance is that it is seen as the index of the expected time until arrival, for a given node, of whatever is flowing across the network. Therefore, an agent with high closeness is seen as a key player because they hear things first.

Walden (2014) showed that having an information advantage (i.e., the advantage an investor has because of his position in the network) allows an investor to earn excess returns. This advantage was closely related to their eigenvector centrality, but less so the

[10] The shortest path is defined as the route with the least number of edges between two agents.
[11] Distance is defined as the length of the shortest path between two agents.

other centrality measures. The eigenvector score for an agent assesses how connected the neighbors for an agent are. The investor with the highest eigenvector will be the one with the best overall connection and is seen as the leader of leaders. Its popularity comes from providing a good balance between connections at all distances compared with other measures (Walden, 2014).

3.4 Market

Section 2.4.1 outlined the two alternatives for the market clearing process: the existence of a market marker or a formal auction market. To remain consistent with the H&S model, a market marker model is employed for this thesis. In addition, the market remains closed with regards to the number of investors and their ability to access additional cash or raise debt to acquire risky assets, which is again consistent with the H&S model. According to H&S, this enables a greater amplification of the bubble. The author concedes that this is a possible weakness to the model because as Xiong (2013) points out, among other things, assets bubbles see new investors enter the market plus markets become vulnerable to increases in asset supplies.

In addition, to the market being closed, the dividends ($d_i(t)$) which are declared at each step, are not available to be re-invested into the market. This dampens the absolute movement of the price and should see the price mean revert to 1 in the absence of any other dynamics within the model. While dividends are not available for re-investment, their value is tracked because the value is needed for the portfolio statistics for the investors.

3.5 Decision/Opinion Sources

3.5.1 Private Information

At each time step investors update their private information/opinion with regard to each asset. Each opinion ($\epsilon_{ij}(t)$) is drawn randomly from a normally distributed population (N(0,1)) as part of Step 5 in Figure 6. This process should ensure that the information is uncorrelated across and between assets and investors. The justification for the inclusion of the private information variable is that it allows for investors to form independent decisions on how an asset will perform, and covers the possibility that investors may have access to private information sources, something, as discussed in Section 2.2.1, that is precluded by the generally accepted Semi Strong Form of the EMH.

Investor trust in their private information remains constant at the initiated value of c_{3ij}. While this approach follows the H&S model, the field of behavioral finance can further justify the position, as DeBondt and Thaler (1985) comment “perhaps the most robust finding in the psychology of judgment is that people are overconfident". This overconfidence is manifested in both overconfidence and self-attribution bias. Daniel, Hirshleifer and Subrahamanyam (1998) define an overconfident investor “as one who over estimates the precision of his private information signal, but not of information signals publicly received by all”. Self attribution refers to how agents will attribute a correct prediction to their skill while dismissing an incorrect prediction to noise (Daniel et al., 1998).

3.5.2 Public Information

Regardless of which model is used, the updating of the public information occurs at Step 6 in Figure 6. To allow for the implications resulting from multiple assets, certain changes to the H&S model were required. In the H&S model, public information is determined in a similar manner to the private information, i.e. a Gaussian white noise process with a mean of 0 and variance of 1 (N(0,1))[12]. With multiple assets, a key question is how correlated are those assets to each other? Therefore, the model needed to allow for the possibility that the public information of the assets is fully, partially, or not at all correlated. The impact of varying the correlation will be discussed in Section 4.2.4, but early findings suggest that having no correlation between the assets hampered the ability of the investors to generate any sufficient trust in their information sources.

The process for updating the public information for each asset is given by Equation 3. The process works by the news for the first asset (asset 0) being randomly chosen from the Gaussian distribution and if there are multiple assets, the process is repeated for each of the assets. The final value is then determined by weighting the values by β or 1 - β. While the equation only addresses the correlation between the first asset and each asset, it can be seen that if β = 1 then the process is the equivalent to the original H&S model because all risky assets will have the same public information.

Equation 3: Public information update

$$\boldsymbol{pi_i(t)} = \beta * AssetNews\_0(t) + (1 - \beta)AssetNews_i(t)$$

[12] The exception is in the multiple situation where the standard deviation increments by .1 per asset.

As discussion in Section 3.3.1, the introduction of EPS in the extended model required further changes. In effect, required further changes. In effect, $eps_i(t)$ for the asset replaces 'news' in Equation 3 as the earnings for the the earnings for the asset for each period are determined in a similar manner to the original model, with the original model, with the exception that the earnings are drawn from a PDF based on the parameters as outlined parameters as outlined in Table 1, that is, one with a non zero mean. This factor also sees the consensus forecast the consensus forecast always being strictly greater than 0. However, in an important difference under the difference under the extended implementation, investors assess the actual EPS ($eps_i(t)$) delivered at each step delivered at each step against the consensus forecast, as per Equation 4. In terms of the consensus forecast ($epsf_i(t)$), each agent holds the same forecast, which is determined as the ensemble average of past EPS results as provided by

Equation 5.

Equation 4: Public information

$$pi_{ti} = \frac{eps_i(t) - epsf_i(t)}{epsf_i(t)}$$

Equation 5: Consensus forecast

$$\langle epsf_i(t)\rangle = \alpha * \langle epsf_i(t-1)\rangle + (1-\alpha) * eps_i(t-1)$$

If the actual earnings for an asset exceed the consensus estimate, this is considered an earnings surprise and this would be reflected positively with a score greater than zero. Alternatively, if earnings miss to the downside this is a negative and agents will read it as a signal to sell down their holding in the asset. If earnings meet expectations then the information adds no value in the decision-making process because the investor assumes the information is already reflected in the price (they are assuming that the EMH actually holds true!) and they will hold their current position. As investors have the same public information the level of trust they have in it for each asset is the same.

The primary support for the above deviation comes from the existing volume of work that has analyzed the impact of earnings announcements (see Kothari (2001) for a extensive review of the literature). In summary, the findings show that stock prices react positively to positive earnings news, yet it takes time for this information to be fully reflected in the price of the asset (Kothari, Lewellen, & Warner, 2006). Further support comes from Barberis, Shleifer and Vishny (1998), who produced a model of investor sentiment that was successful in explaining and replicating an asset's price movement following an earning surprise.

3.5.3 Network Information

To capture the information from their network, each investor polls their neighbors in terms of their intentions (buy, hold or sell) for each asset in Step 5 of Figure 6. The justification for this variable is that if an investor utilizes a bounded rationality framework, it becomes more efficient for them to imitate their neighbors (see section 2.3.2). It is also known that investors consider their peers (see section 2.3.3) in their investment process.

The results of the process are captured by the $E_{ij}[a_{ik}(t)]$ term in Equation 1, with the $a_{ik}(t)$ reflecting the action of the neighbor, as per Table 4. The investor will then weigh the action by the amount of trust they have in the particular neighbor $nt_{jk}(\boldsymbol{t-1})$. The investor then sums the results from each neighbor before finally multiplying the value by their fixed c_{1ij} term. Given the updating process of the model, an investor may poll some neighbors before those neighbors have processed their new private and network information. This is not considered a problem because agent initiation in the model is consistent with the process described by H&S, thereby accepting the justification of varying reaction times between investors.

3.6 Investment Decision

After investors assess their new information and before they update their level of trust, they must make a decision at each time step (Step 6 in Figure 6). As detailed in Section 3.3, investors are provided with a threshold value $\bar{\boldsymbol{\omega}}_{j,}$ and it is this value that they compare with their score ($\boldsymbol{\omega}_{ij}(t)$) when deciding their actions ($a_{ij}(t)$) for each asset at time *t*. Table 4 details the conditions by which they make their decisions.

Table 4: Agent decision thresholds

Scenario	Action	Variable	Trading Volume
$\boldsymbol{\omega_{ij}(t) > \bar{\omega}_j}$	Buy	$a_{ij}(t) = +1$	$v_{ij}(t) = tr * \frac{rf_j(t)}{p_i(t-1)}$
$\boldsymbol{\omega_{ij}(t) < \bar{\omega}_j * -1}$	Sell	$a_{ij}(t) = -1$	$v_{ij}(t) = tr * holding_{ij}(t)$
Otherwise	Hold	$a_{ji}(t) = 0$	

Having decided to buy, hold or sell, the investors must decide how much they are willing to buy or sell (Step 6 in Figure 6). Table 4 again provides the formula by which they determine the transaction value ($v_{ij}(t)$). In determining the value of each transaction, the agents apply the coefficient *tr.* H&S indicate that as long as this coefficient does not approach 1 the general findings of the model are not significantly affected[13]. With regard to the calculation, investors face the following constraints within the model:

- No leverage - Agents must have a positive holding of the risk free asset ($rf_j(t)$) to enable them to trade; and
- No short selling – Agents must have a positive holding of the asset ($holding_{ij}(t)$) they wish to sell.

These constraints raise the possibility that an investor may wish to undertake an action but are unable to. It is for this reason that it is the intention of an investor's neighbor (given by $a_{ij}(t)$) that is polled rather than the transaction value.

It should be noted that when deciding how much to invest ($v_{ij}(t) > 0$), investors do not attempt to forecast what the price will be at the completion of the trade (p(t)), rather they use the existing price, which is provided by p(t-1). H&S again suggest that the alternative approach does not impact the results.

3.7 Market Clearing Process

Following the accumulation of the investors' orders for each asset, the market is cleared and the returns for the asset(s) are determined via Equation 6 (see Steps 7 and 8 in Figure 6). The first term of Equation 6 remains consistent with the H&S model, where the λ term is used to weight the investor's actions by the market depth. A detailed analysis and justification for the use of this term and the general market pricing mechanism is provided by Farmer (2002).

[13] This was tested with regards the author's model and a similar conclusion was reached.

The $\sum_{j=1}^{N_j} a_{ij}(t) * v_{ij}(t)$ term, which is the accumulation of the investor's individual orders, provides the surplus/deficit demand for the asset at each step. If there is net buying (a surplus in buy orders) then the return for the period will be positive and the price will increase, the opposite occurring if there is a net selling (deficit in demand). Excess prices movements for any period will be generated by large surplus/deficits and indicates that the population has become a herd. A bubble and crash will occur if the population remains in that herd for an extended period, that is, there are multiple periods of net buying or selling.

Equation 6: Return determinant for asset i

$$r_i(t) = \frac{1}{\lambda * J} \sum_{j=1}^{N_j} a_{ij}(t) * v_{ij}(t) + \log\left((d_i(t) + p_i(t))/p_i(t)\right)$$

In a deviation from the approach of H&S, an asset's dividend ($d_i(t)$) is included when an asset's return is determined as per the last component of Equation 6. Section 3.5.2 detailed the process by which the dividend is determined for each asset at each step.

The value of $r_i(t)$ is placed into a list so that the average and standard deviation of the asset's returns can be tracked and used to plot the quasi-efficient frontier. However, before this occurs, the inverse log is taken to ensure the values are in the same scale as the investor's returns as determined in Section 3.8.2.

Within step 8 the price of each asset is updated as per Equation 7 and then the model takes the inverse of the $\log[p_i(t)]$ term to finalize all book keeping for the investors (Section 3.8.2). Technically, an asset's price should fall by its dividend value once the dividend is paid. However, in the interest of preserving the intention of Equation 7, the price is not impacted by the dividend of the stock.

Equation 7: Log price update

$$\log[p_i(t)] = \log[\, p_i(t-1)] + \frac{1}{\lambda * J} \sum_{j=1}^{N_j} a_{ij}\,(t) * \, v_{ij}(t)$$

3.8 Agent Updating

3.8.1 Trust Updating

After the investors become aware of their returns, they reassess the level of trust they have in the information provided by their network and public sources, as used in Equation 1, and update (nt_{jk}) and (pt_j) respectively (Step 9 and 10 in Figure 6). The rationale for these variables, which are initiated with a value of 0, is that the investor will place greater trust in a source if it provides the correct advice. That is, if the agent receives a buy (sell) signal from the information source, and the price subsequently increases (decreases), then the weight (trust) increases. The weight decreases if the signal is erroneous. As H&S detail, the investors are "looking for persistent sources of information, which impact on returns". There is an important difference in updating the two trust variables. For public trust, the population maintains trust at an individual asset level to preserve the heterogeneous nature of the asset's performance. In contrast, for an investor's trust in their neighbor, they will assess the individual recommendations of their neighbors before updating the trust based on the average performance of the recommendations.

Investors follow the process as illustrated in Figure 7 when assessing the usefulness of the public information ($pi_i(t-1)$) or network information ($\sum_{i=1}^{I} E_{ij}[a_{ik}(t)]$).

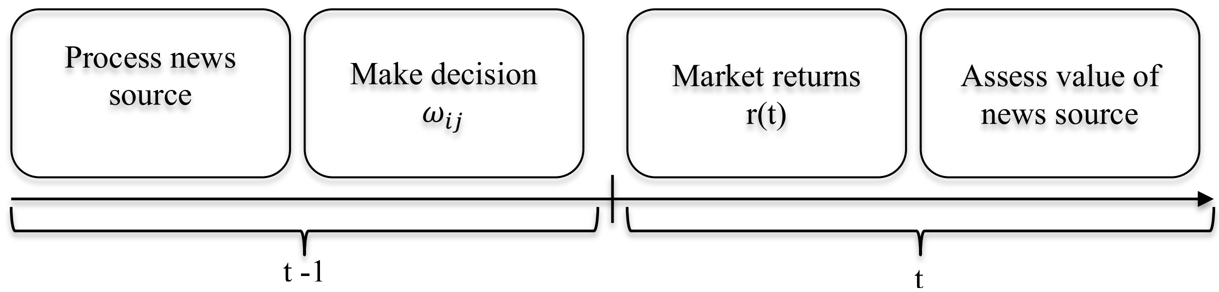


Figure 7: The simplistic flow of agents assessing the usefulness of information

The key implication of the process, as detailed by H&S is "that for any information source to have any predictive power it must have some persistence". For this statement to hold, investors process their information sources and make their decision in t-1. The market will subsequently move in t-1 (see Figure 7) based on the combined decisions of the investors in t-1. However, it is not until period t that investors assess the value of the information that they processed in t-1 and update their trusts levels according to the standard autoregressive update as detailed in Equation 8 and Equation 9.

Equation 8: Public trust update

$$pt_i(t) = \alpha pt_i(t-1) + (1-\alpha)pi_i\,(t-1) * \frac{r_i(t)}{\sigma_{ir}(t)}$$

Equation 9: Network trust update

$$nt_{jk}(t) = \frac{\sum_{i=1}^{I} \alpha nt_{jk}(t) + (1-\alpha)E_{ij}[a_{ik}(t)] * \frac{r_i(t)}{\sigma_{ir}(t)}}{I}$$

The first part of the above equations discounts the previous trust value variable by the variable (α), which is set by the `memory_weight` parameter. The significance of the α variable, which is also used in the consensus forecasting process, is that it sets the time scale over which past performance impacts a variable's value. The time scale as per H&S is given as $\frac{1}{|\ln(\alpha)|}$. The second part of the equation adds the assessment of the immediately preceding information, which has been discounted by (1- α) after it is multiplied by the $\frac{r_i(t)}{\sigma_{ir(t)}}$ term. The $\frac{r_i(t)}{\sigma_{ir(t)}}$ term is used to normalize the past return of an asset by the standard deviation of its past returns. The rationale, as articulated by H&S, is that a larger return scaled by its volatility ($\sigma_{ir(t)}$) will enhance the trust to a greater degree. This process provides the potential mechanism for the development of a bubble or crash, as investors place more and more trust in an information source when it provides greater profits (or saves losses), hence the potential to create a positive feedback loop.

Equation 10 details how $\sigma_{ir}(t)^2$ is computed as a moving standard deviation with an exponentially decreasing kernel. The $\langle r_i(t)\rangle$ and $\langle r_i(t-1)\rangle$ terms represent the ensemble average of the return series, where the ensemble average is defined as the expected object of the stochastic process.

Equation 10: The variance and ensemble average for an asset's return

$$\sigma_{ir}(t)^2 = \alpha * \sigma_{ir}(t-1)^2 + (1-\alpha) * (r_i(t) - \langle r_i(t)\rangle)^2$$

Where $\langle r_i(t)\rangle$ is given by:

$$\langle r_i(t)\rangle = \alpha * \langle r_i(t-1)\rangle + (1-\alpha) * r_i(t)$$

3.8.2 Portfolio Updating

The final step of the process is to update the investors' portfolio and calculate their portfolio statistics (Steps 11 and 12 in Figure 6). It should be noted that the dividend payments are not added to $rf_j(t)$, the balance of the risk free asset, as the investors cannot reinvest dividends. The implication being that the model remains a closed system and therefore does not deviate too far from the original H&S model. The alternate approach would introduce a wealth effect, adding a layer of complexity, thus making it harder to interpret the implications of introducing the other new factors to the market.

Equation 11 provides the book keeping formulas used to update the portfolio. The investor's risk free asset (rf_j) is updated by the proceeds of any sales and the cost of any purchases, noting that the proceeds are impacted by the price realized in the period. In a similar manner, the holding for each of the investor's risky assets are updated.

Equation 11: Portfolio updates

$$rf_j(t) = \; rf_j(t-1) - \sum_{i=1}^{I} s_{ij}\,(t) * \; v_{ij}\,(t) * p_j(t)$$

$$holding_{ij}(t) = \; holding_{ij}(t-1) + \; s_{ij}(t) * v_{ij}t$$

Given the model's design, each investor's average portfolio return and its standard deviation are calculated at each tick. The results of these calculations allow for the construction of the quasi-efficient frontier, as illustrated in Figure 8, where the '+' symbols represent the average return and standard deviation of the investors, and the $\otimes$ symbol represents the risky asset. The quasi-efficient frontier plots the average return for each investor's portfolio against its standard deviation, which is the proxy for risk. Equation 12 provides the formula for the value of an investor's portfolio value. The inclusion of the dividend bank should be noted as it does form part of the investor's returns.

Equation 12: Portfolio value

$$Portfolio\ Value_{ij}(t) = rf_j(t) + \sum_{i=1}^{I} holding_{ij}(t) + dividendbank_{ij}$$

While further discussion will be undertaken in the results section, the reader should note that the top of Figure 8 represents a scenario where a bubble is growing and the bottom represents the situation as the bubble implodes. The implication being that while the market is going up, investors move towards the risky asset but cannot catch the asset. In this situation, investors with a lower threshold move first. In an exploding market, investors move away from the risky asset and the investors in the top left corner may well have timed the market perfectly and sold at the top as they have outperformed the market with considerably less risk than the risky asset.

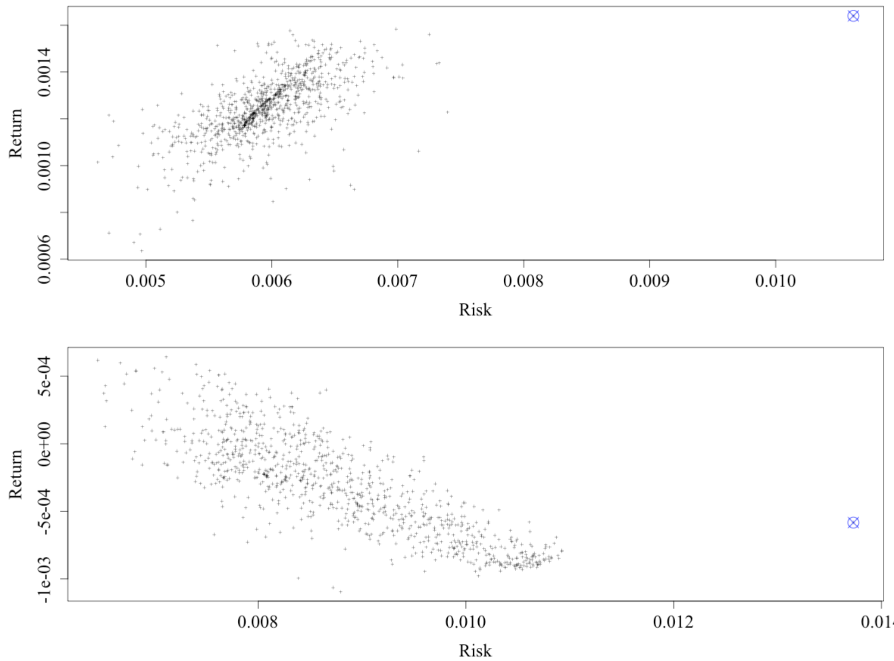


Figure 8: Examples of the quasi-efficient frontier pre and post bubble

3.9 Verification

An important consideration in the ABM building process is to verify that the model performs as it is was designed. While the actual results cannot (and should not) be assessed ex ante, certain steps can be taken. For the implementation of this model the following verification steps were successfully undertaken:

- Matching the baseline output of the H&S model;
- The implementation of a journal to allow manual calculations to be undertaken to ensure calculations within the model were correct;
- Visual inspection of various charts plotting the behavior of the variables;

- A code walk through was undertaken to firstly ensure no coding errors were made and secondly, to produce flow charts to ensure the code implemented the intended model as outlined in Sections 3.2 through 3.8; and
- Parameter sweeps of the extreme values.

3.10 Model Outputs

An inherent value of ABMs is being able to collect data at both the agent and system level (the data from this model extensively analyzed in Section 4). At the system level the model generates a new price at each step of the model as described in Section 3.7. Associated with this new price are the returns of the asset plus the mean and the standard deviation of those returns. The model also collects the average network and public trust of the investors and the cumulative actions of the investors, which in turn allows for the analysis of the order book at each step.

In terms of the data collection at the agent level, the focus of this thesis was to collect the average return and the standard deviation of the returns of the investors. This was necessary to create the quasi-efficient frontier. Given the infrastructure of the model the potential exists to extract far more information at the agent level. Obvious candidates include the centrality of the individual agents and their investment thresholds. The extraction of these variables will be helpful to uncover the determinant of what makes an investor successful or not.

The Netlogo network extension package is also utilized to calculate betweenness and closeness centrality measures for whichever network topology is implemented. The results are maintained at the agent level, which again allows for agent and system analysis to be undertaken.

3.11 Section Summary

The detail of the implemented model, which has made multiple extensions to the H&S model, was provided in this section. The extensions all relate to providing greater clarity to how financial markets operate. At a high level the extensions were; the introduction of different network topologies, an alternate source of public information and the introduction of a dividend and multiple assets. Sections 4.2.2 through 4.2.4 provide the results of various experiments that test the implications of these extensions. While Section 4.2.1 details the creation and collapse of bubble in greater detail and finally Section 4.2.5 outlines the quasi-frontier.

4 RESULTS AND FINDINGS

4.1 Introduction

The successful replication of the H&S model and the implementation of the various extensions provided the opportunity to experiment with a vast array of scenarios. However, the scope of the experiments for this thesis have been restricted to those outlined in Table 5. The key components and findings of the various experiments are also provided, with greater detail provided in the appropriate subsections. In addition to these experiments insights were gained with regards to the trading behavior of the investors and the dynamics of the quasi-efficient frontier.

Table 5: Summary of results

Model Design	Key Components	Summary of Findings
H&S with different networks (see Section 4.2.2)	All four networks topology were tested with the level of c1 varied from 1 to 4 in increments of 1.	The scale free network has very different characteristics to the other three networks. When c1 is increased to 4, all network experience much higher volatility.
Revised model with a single asset (see Section 4.2.3)	Utilized a lattice network with the revised source of public information and the dividend payout ratio varied between: 0, 0.33, 0.66 and 1.	The revised public information process generates similar outcomes when the payout ratio is 0. Increasing the payout ratio sees asset prices remain elevated.
H&S with multi assets (see Section 4.2.4)	Utilized a lattice network with 3 risky assets and the H&S model for the source of public information.	When the correlation in public information increases, price peaks appear but they are not as high or as consistent as under the H&S model.

To ensure consistency across the experiments, the parameter sweeps for the various scenarios were performed with the following characteristics, which were the default settings for the H&S model:

- 2,500 steps per run;
- 30 runs per setting;
- The number of investors (J) was 2,500;
- The upper bound for the conviction threshold was set at 2;
- The market depth (λ) was set at .25;
- The transaction ratio (`tr`) was set at .02; and
- Memory length (α) was set at 0.95.

All charts and statistical tests presented in this section were generated in R (2015), unless stated otherwise. The presentation of the results is also consistent with each experiment containing the following items; fan plots[14] illustrating the combined runs of a particular setting, boxplots of the mean and standard deviation of the prices for each of the combinations and where necessary boxplots of other variables of interest. With regards to the fan plots, the reader should note that the x-axis in these charts is time as determined by tick/step number of the experiment. The combination of these plots and appropriate statistical tests provides sufficient information to support meaningful inferences about the behavior of the market.

The most relevant finding from the H&S paper was the existence of a phase transition when the level of the c_{1ij} variable (the fixed coefficient with regards to an

[14] The fan plots are generated via the fanplot library (Abel, 2015).

investors trust in their neighbors, or in other terms, their initial bias to that source) was set at 3[15]. The authors suggest that at this setting a positive feedback loop with regards to investors adapting the actions of their neighbors becomes material, that is, the investors will ‘herd’. Further, at values greater than 3, the system becomes “excitable” and the existence of a large bubble occurs, with the authors classifying these events as outliers. Based on these findings and in the interest of time and space, this author designed the various experiments around contrasting the behavior of the system when c_{1ij} is set at 1 versus when it is set at 4.

4.2 Detailed Results

4.2.1 Trading Behavior in a Bubble

The H&S paper provides a detailed explanation of how and why bubbles are created in their model. However, one issue they do not directly address is what the intended actions of investors are during the formation and deflation of a bubble. This is an important consideration because the question of what pops a bubble remains open. H&S mention that the rise in the asset price slows because investors do not have the infinite resources to drive the price higher. Closer inspection of Figure 9, which is a representation of the system when a bubble forms, provides evidence that the there may be more to it than investors not having infinite funds. The series titled ‘buying’, which measures the number of investors with an intention to buy ($a_{ij}(t)$ = 1), reaches its peak before the top of the cycle and remains high for a considerable period. It should be noted that both the

[15] With c_{2ij} and c_{3ij} set at 1 and the remaining variables set at their default settings as per the base experiments for this thesis.

‘buying’ and ‘selling’ series can be used to judge the size of a herd as they measure the cumulative actions of the population. An interesting point is that at no time are all investors buying, that is the number of buyers does not reach J, the number of investors, as some super cautious investors sit on their cash.

The question of what is driving the buying behavior appears to be given by the factor variable series. The factor series is the $\frac{r_i(t)}{\sigma_{ir}(t)}$ term utilized in Section 3.8.1 through Equation 8 and Equation 9. The term and relates to by what magnitude investors update their trusts values. The peak in the factor value is aligned with the peak in the agents buying behavior. The factor variable starts to decline as the returns starts to slow, which as H&S indicate is due to a lack of funds by the investors. At the same time, the influence of the stronger past returns start to diminish. It is once that the factor variable turns negative that the buying behavior ceases abruptly and the price starts to fall. Of particular interest is the fact that while selling does jump, the population does not immediately form a single herd and all head for the exit. That is there is a significant period where investors hold and do not make any trades because while the trust in the neighbors has been lost their information score is not sufficient to have them sell. Eventually, as the trust increases after the extended downward period the entire population wants to sell.

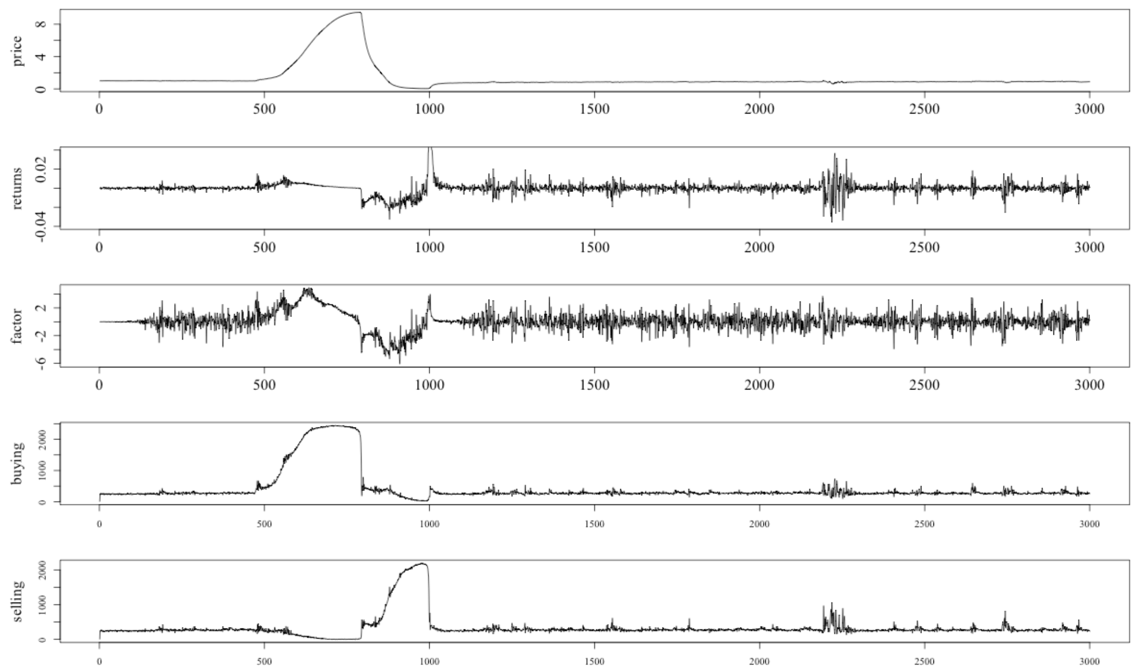


Figure 9: Trading within a bubble over time

The implication of the factor variable declining is that investors do not adjust the trust in their information sources with the same power and the trust levels start to decline, something that is illustrated in Figure 10. The peak in the factor variable is aligned with the peak in the investor's trust in the network information and it is the decline in this trust level that can be seen as being responsible for the bubble collapsing. It should also be noted that the trust in the public information is extremely volatile during the bubble period, adding further noise.

A very important implication of the above description is that some investors have jumped early and would have generated significant outperformance. Their exact return will be depend on when they bought in and whether or not they jump on and off during the inflation of the bubble. Also some investor will have jumped late and sustained heavy losses. The determining factor for the performance of the investors is likely to be a mix of their position in the network and their threshold leave. Intuition suggests that investors with a low threshold may well join the herd early and leave early while higher threshold will join and leave late. As discussed in Section 5.3, the confirmation of this will have to wait for another day.

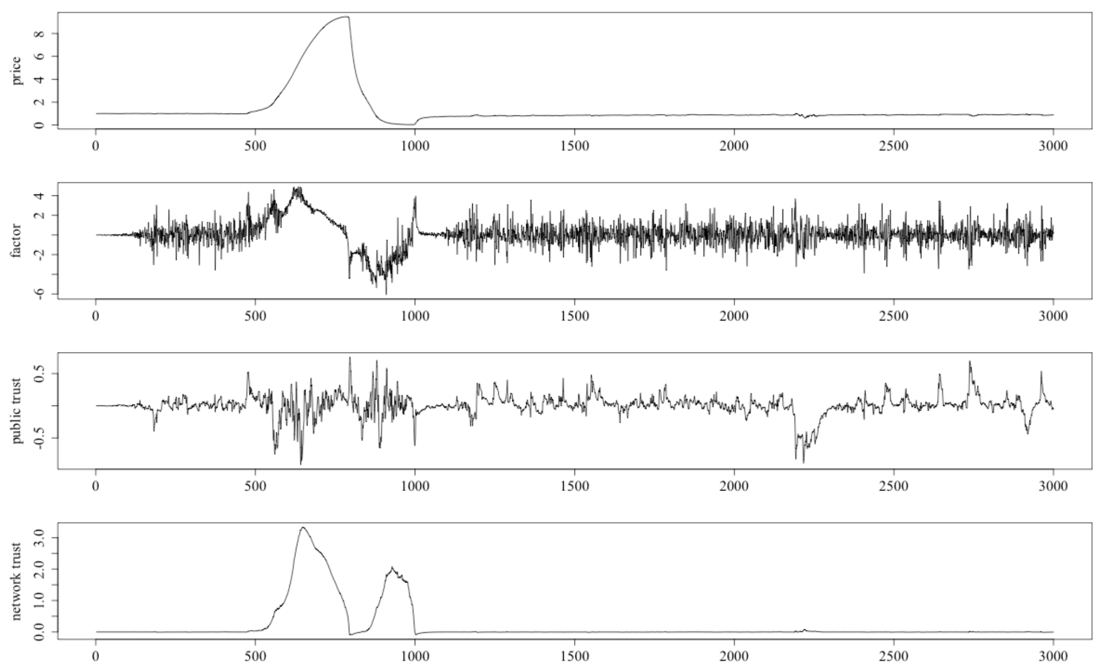


Figure 10: Trust levels through a bubble over time

The influence and power of the investor's trust in their network on the price behavior will become more evident throughout the analysis of the various models in the remainder of the thesis.

4.2.2 The H&S Model With Different Networks

As detailed in Section 3.3.3, the implemented model has the ability to create different network structures (topologies). The aim and objectives of this section are:

- To test whether the underlying network topology impacts the behavior of the system; and
- Test whether the claims of Ozsoylev and Walden (2011) with regards to the impact of centrality are confirmed by the model.

The topology of the network is potentially important because the manner in which the investors are linked varies. These diffences are capture by various metrics which measure and define the key characteristics of each topology. Table 6 provides a summary of the representative statistics relating to each of the networks.

Table 6: Network metrics[16]

	Topology	Neighbors	Closeness	Clustering[17]	Betweenness
Max.	Lattice	4	(c) 0.0032	N/A	389,688
	Small world	7	0.1272	N/A	126,773
	Scale free	562	(b) 0.5268	N/A	(b) 441,143
	Erdos-Renyi	13	0.2124	N/A	(c) 38,291
Average	Lattice	(a) 4	0.0032	(e) 0.500	389,688
	Small world	(a) 4	0.0976	0.380	11,646
	Scale Free	(a) 4.0208	(d) 0.3361	(e) 0.141	(d) 1,694
	Erdos-Renyi	(a) 4.0446	0.1723	(e) 0.002	5,631
Std. Dev.	Lattice	N/A	N/A	N/A	N/A
	Small world	0.6204	0.0083	N/A	12,130
	Scale Free	31.774	0.1199	N/A	24,965
	Erdos-Renyi	2.01	0.0292	N/A	5,422

[16] Metrics calculated using the Netlogo network extension.
[17] This metric is the clustering coefficient and was calculated in Gelphi.

There are a number of key points and possible explanations for the varying behavior provided by Table 6. The first being that while the average number of neighbors for each of the networks is 4 (refer to (a) in Table 6), the distribution is varied, with the lattice having no variance, while the scale free network has the largest variation. This fact relates to the degree distribution of networks as discussed in Section 2.3.4. Therefore the average degree for each of the networks is 4 but again the distribution is very different. At one end of the spectrum the lattice has a uniform degree distribution, while the scale free network has a power law like degree distribution.

In terms of the centrality measures, the scale free network has the highest absolute betweenness and closeness centrality measure (refer to (b) in Table 6), while the lattice and Erdos Renyi networks share the lowest absolute measures (refer to (c) in Table 6). In addition, the scale free network has the lowest average betweenness and yet the highest average closeness centrality measure (refer to (d) in Table 6). The lattice network records the lowest average closeness and the highest average betweenness measures. The scale free network's clustering coefficient is in the middle of the sample, while lattice and Erdos Renyi network are at either end of the spectrum (refer to (e) in Table 6). Therefore, we would expect the scale free network to be the most volatile and the others to be less volatile if the results of Ozsoylev and Walden (2011) hold. With the Erdos Renyi and scale free networks being incomplete graphs, the Eigen vector values cannot be determined. Therefore the claims made by Walden (2014) can not be tested.

Figure 11 through Figure 14[18] provide a graphical representation of the networks. Figure 11 illustrates the lattice network, which shows that each investor (node) is linked to exactly four neighbors.

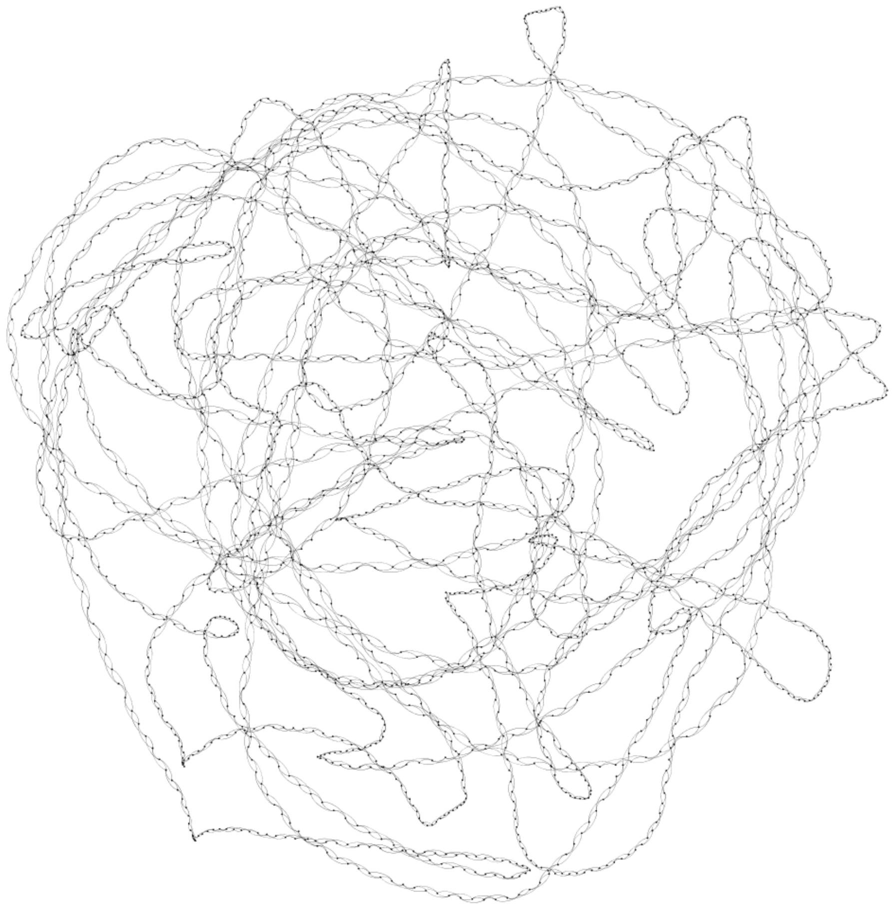


Figure 11: Representative investor network – lattice

A small world network is illustrated in Figure 12. The network was generated by first having each investor initially linked to their 4 nearest neighbors (as per a lattice network). Then with a probability of 10%, an investor will cut a link with their closest

[18] The figures were produced in Gelphi using data generated by Netlogo.

neighbor and rewire with another investor further away. The rationale behind this process is detailed in Watts (1999). In contrast to the lattice network, the creation of clusters, nodes with greater than four neighbors, is evident in the diagram.

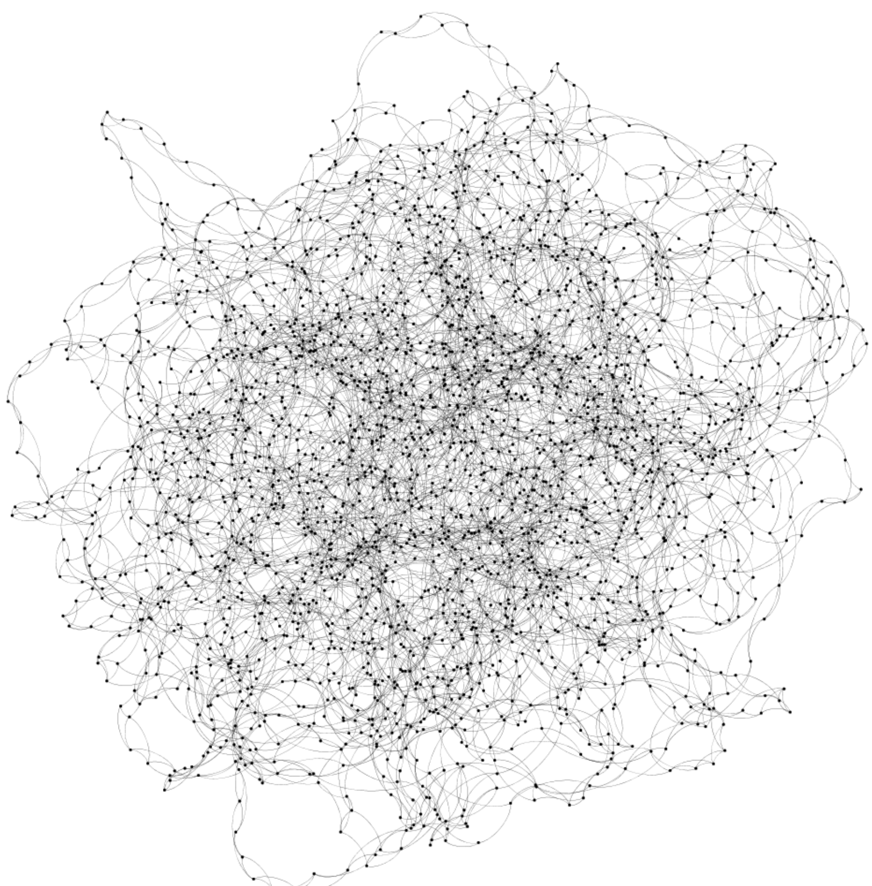


Figure 12: Representative investor network – small world

Figure 13 provides an illustration of a scale free network. The significant features are: the existence of isolates (nodes with no links) and the appearance of nodes that are heavily connected. The heavily connected nodes are known as hubs and contrast to the nodes with very few connections, which are known as spokes.

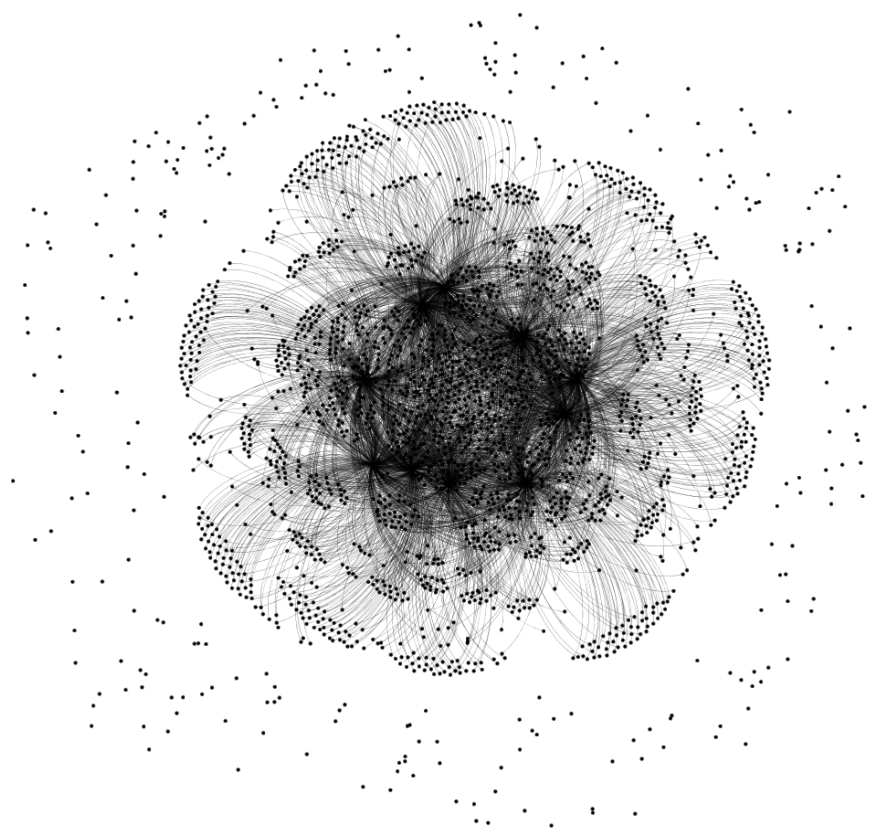


Figure 13: Representative investor network – scale free

The final network is the Erdos Renyi network, which is shown in Figure 14. While the network also has isolates, like the scale free network, there are not any distinguishable hubs and the number of connections per node shows no visible pattern – an implication of the nodes being connected in a random normally distributed fashion.

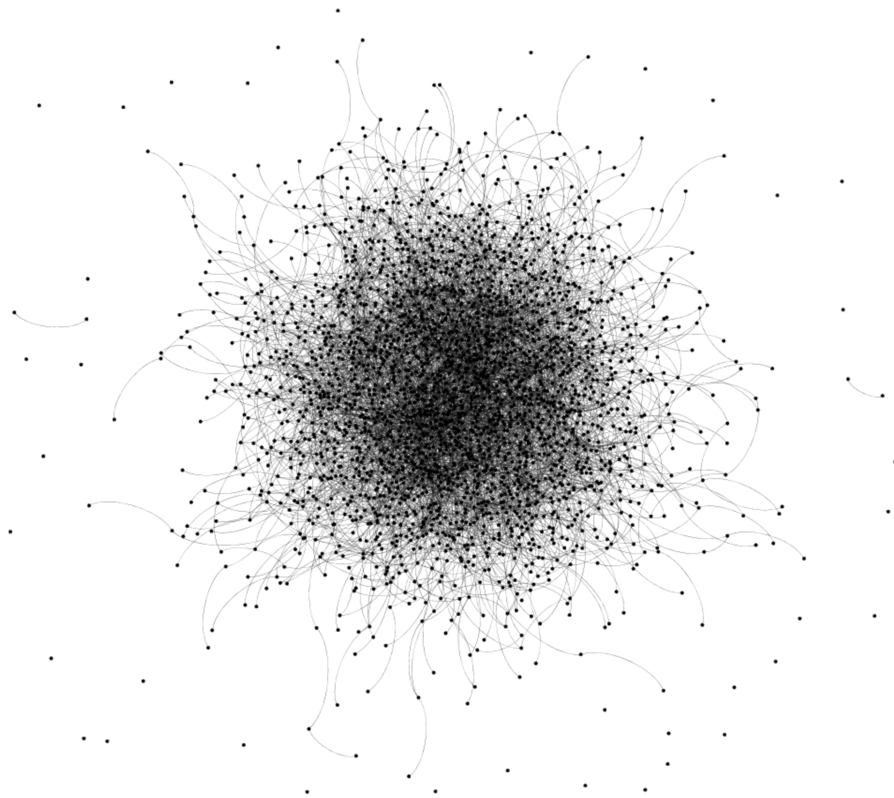


Figure 14: Representative investor network - Erdos Renyi

The impact on the price of the asset from varying levels of c_{1ij} (detailed in the heading by the c1 value[19]) across the varying network structures is seen in Figure 15 through Figure 18. The plots have the median price for the sample marked with the line marked with 50%. The differences (and indifferences) for the various levels of c1 and networks are detailed in Table 7 and Table 8. The reader should also note that the axes for c1 = 1 to 3 are same before increasing for c1 to 4 (the one exception being the scale free network).

---

[19] For convenience the term c1 and c2 will be used to describe the upper limit of the c_{1ij} and c_{2ij} terms.

Starting with the lattice network, the graph is consistent with the H&S model as it is not until the level of c1 is greater than 2 that the system starts to deviate in any meaningful manner from a price level of 1. The existence of bubbles and crashes, demonstrated by the price approaching 8 and 0, is seen when c1 is equal to 4. While the appearance of, or the exact timing of a bubble is not guaranteed, there are multiple occurrences within the sample and the price shows greater volatility when c1 is increased.

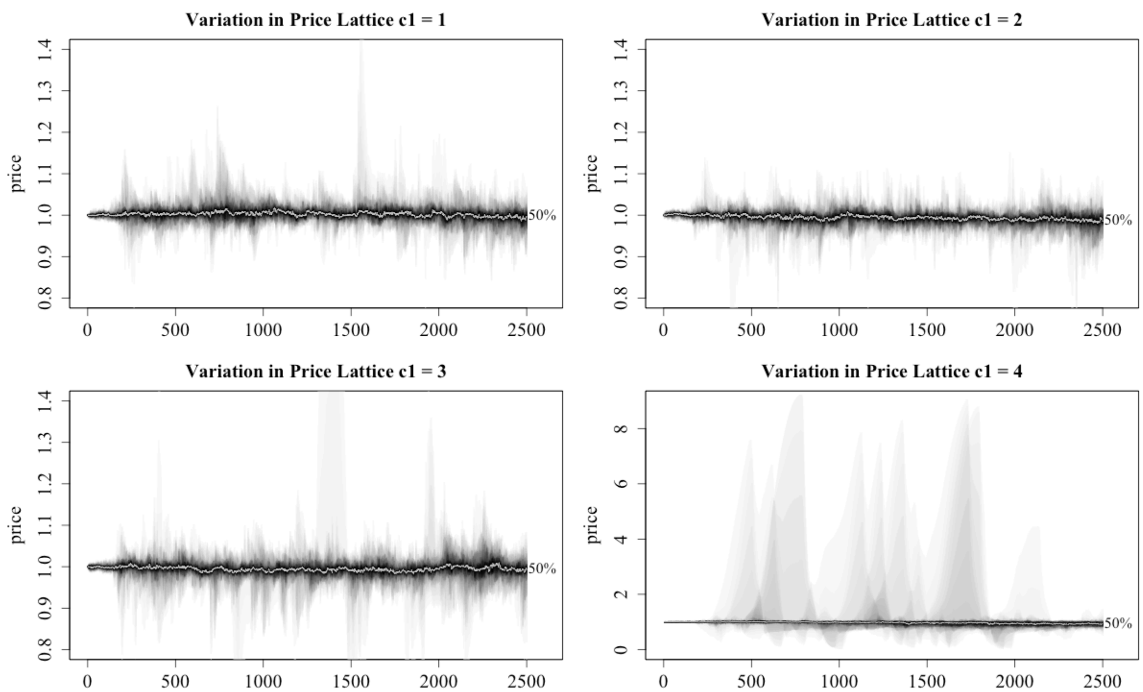


Figure 15: Lattice network with varying c1 over time

Figure 16 illustrates the outcome from using the small world topology with the probability of rewiring set at 10%. Visually there appears to be minor differences with lattice network, including, some bubble like periods when c1 = 3. One interesting difference is the median price when c1 = 4, which is more volatile and spends some time below 1 following the implosion of the bubble.

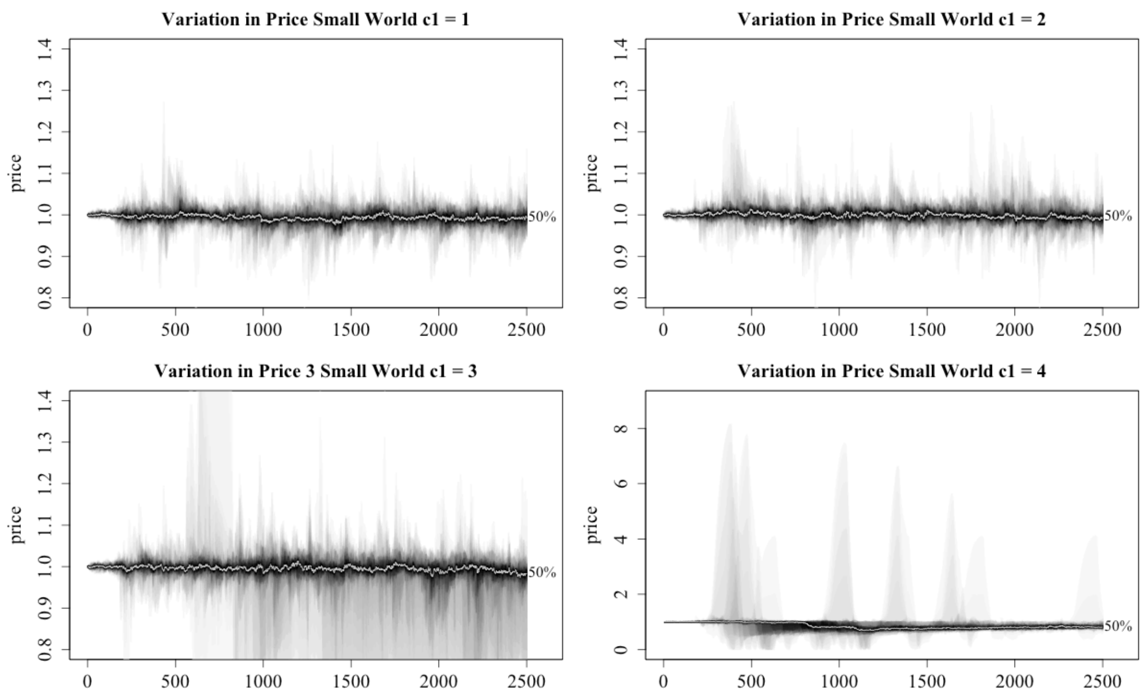


Figure 16: Small world network with varying c1 over time

The graphical representation of the scale free network is provided in Figure 17. Various settings in the formation of the scale free network were tested with similar results achieved. It is clearly evident that the behavior of this network is materially different from the previous two networks. The most notable feature is that there are dramatic price movements regardless of the starting level of c1. However, the level of c1 does appear to impact the peak of the initial bubble. There also appears to be a trend where the price oscillates before finally settling down to a median price, which is around 1. Another point is that despite all the movement, once the initial bubble implodes, the median price dips below 1 before gradually increasing throughout the remainder of the run.

Returning to Table 6, it is seen that the scale free network had the highest average closeness and lowest average betweenness and its clustering was in the middle of the sample. Therefore the findings appear in line with the model of Ozsoylev and Walden (2011), who suggested that centrality plays and important part in determining the volatility of a financial market.

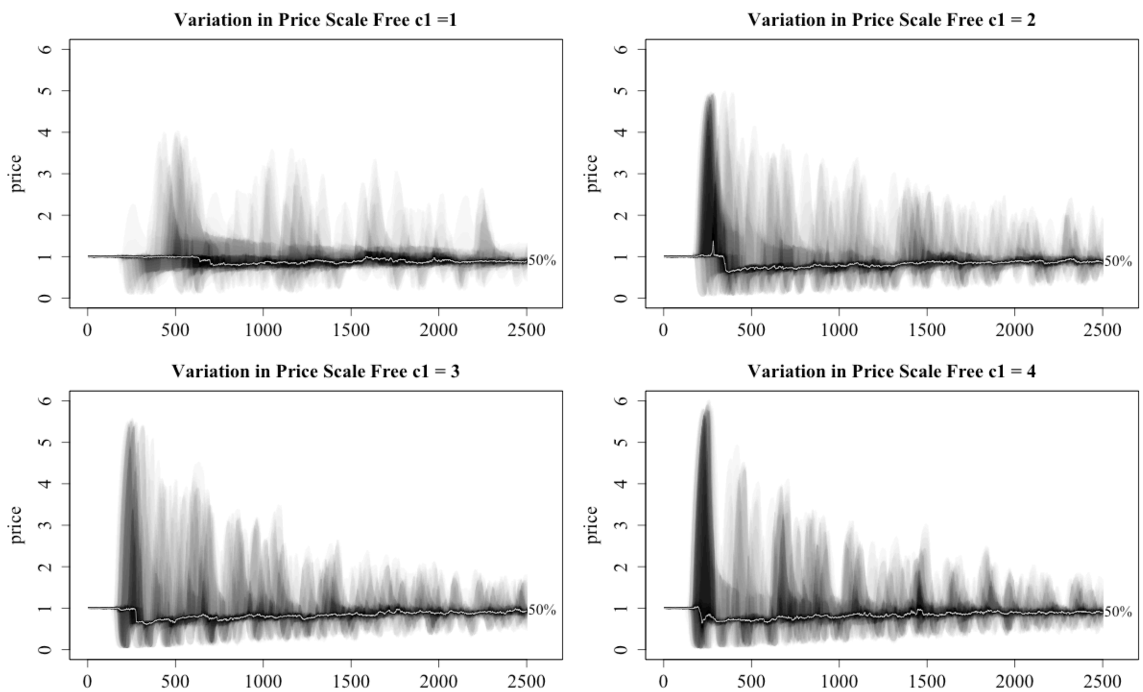


Figure 17: Scale free network with varying c1 over time

The fan plots for the Erdos Renyi network are illustrated in Figure 18. The initial impression is that the results show some similarity with the lattice network, which was to be expected. Possible exceptions are that the prevalence of bubbles and crashes is higher with the random network, yet the peak price is not as high when c1 is equal to 4.

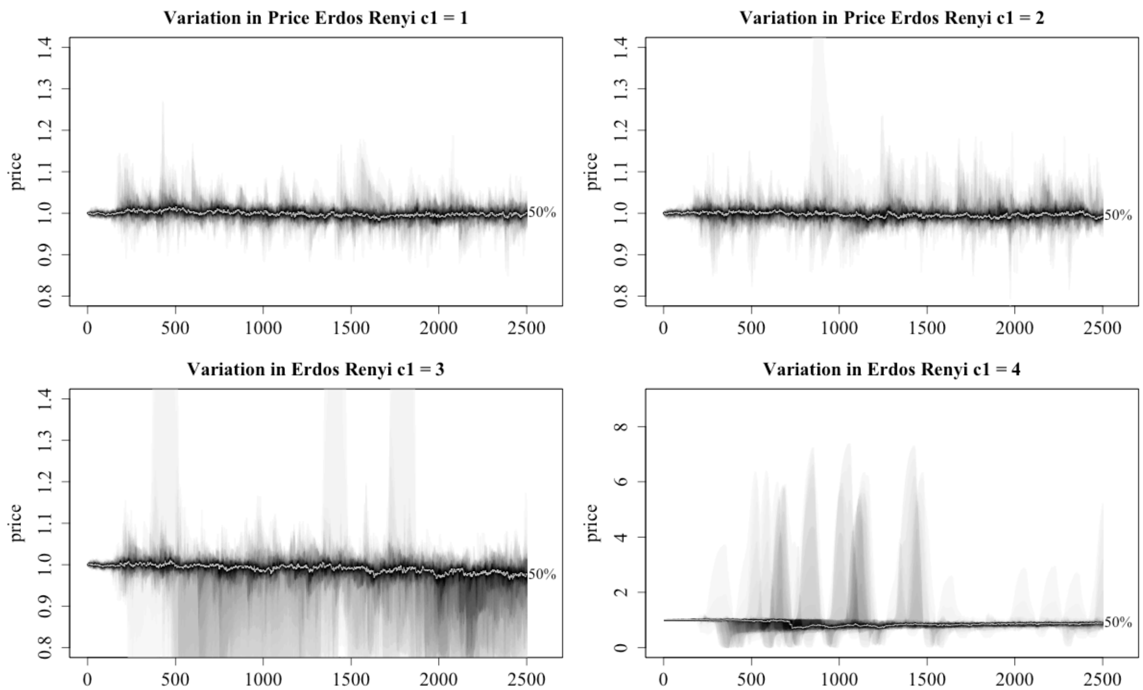


Figure 18: Erdos Renyi network with varying c1 over time

A more conclusive view of how the prices (and therefore returns) vary for the various networks is provided by the boxplots in Figure 19 (the mean price per series) and Figure 20 (the standard deviation of each series). The data for the plots comes from finding the mean price and the standard deviation from the 30 runs of 2,500 ticks for each of the particular settings in the parameter sweep. For example one box is for the 30 runs of a lattice network with a c1 value of 4.

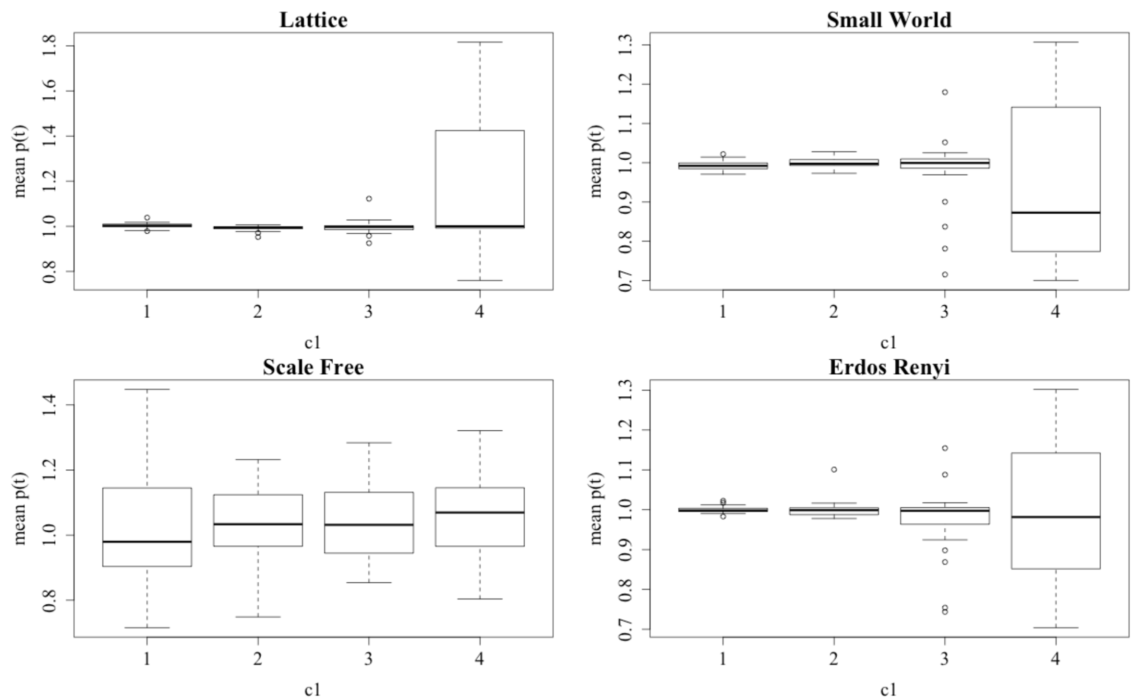


Figure 19: Boxplots showing the mean prices

From Figure 19 it can be seen that the median price is close to 1 for a majority of the scenarios. The one exception appears to be when c1 = 4 for the small world network. Given the previous fan plots, this result is neither surprising, nor that interesting, given the H&S model has a strong reversion characteristic built in. The more interesting result comes from the spread of the prices. Generally it can be seen that as c1 increases, the volatility in all the networks starts to increase. This is consistent with the H&S model. However, the level of volatility does not appear consistent across the networks, and it is this that produces the second finding of significance.

The difference in the volatility of the networks is seen more clearly in Figure 20, which displays boxplots of the standard deviation of the prices within the different scenarios. The rationale for using the standard deviation of a price series is that it is generally accepted as the best measure of volatility. Therefore the terms are interchangeable and have been used as such in describing the results.

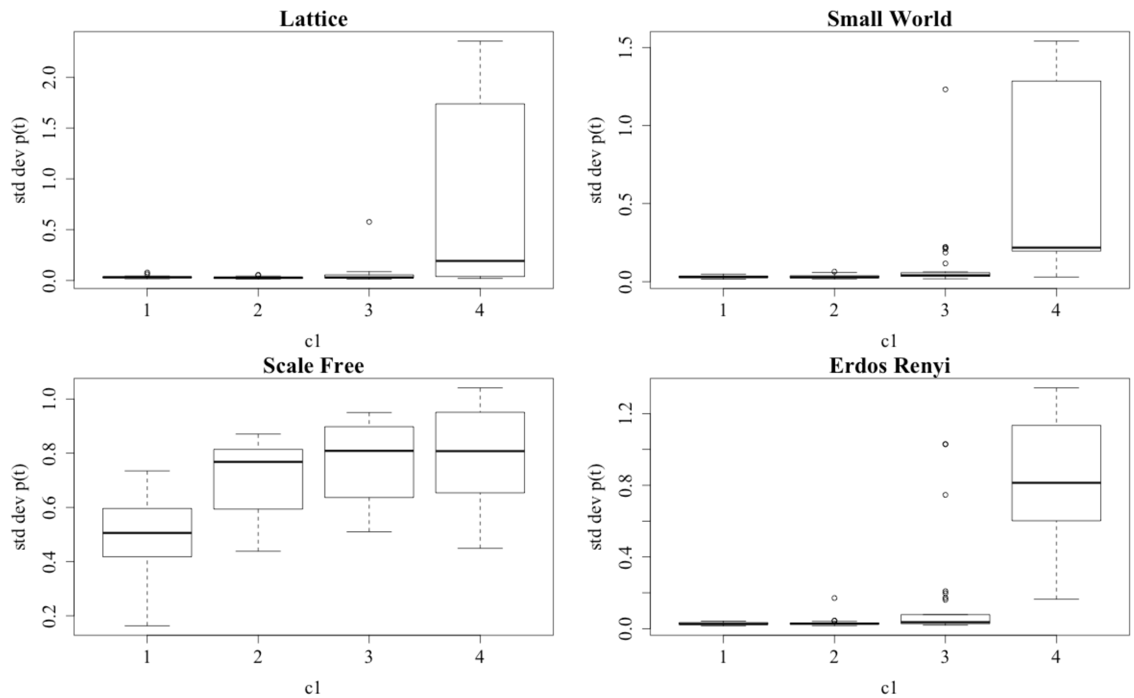


Figure 20: Boxplots showing the standard deviation in prices

From the above, it appears that increasing c1 impacts the volatility of the system, regardless of the network topology. The one special case is the scale free network because it has higher volatility at the initial level of c1 = 1, yet the levels do not increase materially across the spectrum of c1. The volatility of the small world and Erdos Renyi networks are impacted once c1 = 3, as seen by the appearance of outliers, which contrasts with the lattice that is not overly affected until c1 = 4. Another point of note is that the distributions are heavily skewed for the lattice and small world networks when c1 = 4.

To test the inferences put forward previously, a Kruskal Wallis rank sums test[20] was performed instead of a one-way ANOVA. The rationale being that the returns of the mean and standard deviations were not normally distributed, thus violating the assumptions of the one-way ANOVA. The two hypothesis tests that were carried out were:

- For a given level of c1, there are no differences in the mean and/or the standard deviation of the price for the various network types. The p-values for these tests are provided in the columns of Tables 7 and 8; and
- For a given network, there are no differences in the mean and/or the standard deviation of the price for various levels of c1. The p-values for these tests are provided in Tables 7 and 8.

Table 7 provides the mean prices for the various combinations of c1 and the network type, along with the overall average and statistical significance of any differences. The results from the null hypothesises, that there is no difference within the networks and across the networks at a significance level of 5% are:

- The null is rejected as the mean price of the various networks are statistically different (see (a) in Table 7) but varying the level of c1 does not statistically alter the mean price when all networks are considered (see (b) in Table 7);
- The null is rejected when c1 is set at 1, 2 or 4 as the mean price for the individual networks are statistically different (see (c) in Table 7); and

[20] The Kruskal-Wallis is a rank-based nonparametric test that is utilized to determine if there are statistically significant differences between two or more groups and is considered the nonparametric alternative to the one-way ANOVA (Laerd statistics, 2016)

- For the different network types, it is only the lattice network that has a statistically different mean price after varying the level of c1 (see (d) in Table 7).

Table 7: Mean prices under the various regimes

	Network Influence c_{1ij}				Average	p-value
Network	1	2	3	4		
Lattice	1.004	0.992	0.996	1.165	1.039	**(d) 0.003**
Small World	0.993	1.000	0.982	0.943	0.980	**0.159**
Scale Free	1.010	1.023	1.042	1.055	1.032	**0.507**
Erdos Renyi	1.000	1.000	0.976	1.003	0.995	**0.592**
Average	**1.002**	**1.004**	**0.999**	**1.041**	**1.012**	**(b) 0.808**
p-value	**(c) 0.035**	**(c) 0.025**	**0.095**	**(c) 0.007**	**(a) 0.0082**	

Table 8 provides the average standard deviation of the prices for the various combinations of c1 and network type, along with the overall average and statistical significance of any differences. The results from the null hypothesises that there is no difference within the networks, and across the networks, at a significance level of 5% are:

- The standard deviations of the prices across the various networks are statistically different (see (a) in Table 8). Additionally, the standard deviation of the prices within the networks are statically different (see (b) in Table 8);
- When c1 is set at 1, 2 or 3 the standard deviation for the various networks are statistically different (see (c) in Table 8). However, once c1 is set at 4 the level of volatility within the system are not statistically different (see (d) in Table 8); and
- For the different network types, all have statistically different standard deviations with varying levels of c1 (see (e) in Table 8).

Table 8: Standard deiviation under the various regimes

	Network Influence c_{1ij}				Average	p-value
Network	1	2	3	**4**		
Lattice	0.033	0.029	0.056	0.805	0.2308	**(e)<0.01**
Small World	0.031	0.033	0.104	0.604	0.1930	**(e)<0.01**
Scale Free	0.489	0.718	0.761	0.789	0.6893	**(e)<0.01**
Erdos Renyi	0.028	0.033	0.146	0.812	0.2548	**(e)<0.01**
Average	0.1453	0.2031	0.2666	0.7524	0.3419	**(b)<0.01**
p-value	**(c)<0.01**	**(c)<0.01**	**(c)<0.01**	**(d)0.347**	**(a) <0.01**	

Having established that there are significant differences within and across the various network types, the question turns to what is driving the difference. The answer is provided by Figure 21, which provides a boxplot of the mean level of trust that the investors have in the information provided by their neighbors/network. It should be remembered that the level of trust increases when the information from the particular source accurately predicts the correct investment decision. We have already seen from Figure 43 in the appendix that the creation of a bubble is triggered when the level of network trust increases, thus creating the positive feedback loop in buying behavior as the investors herd [21]. Alternatively, a collapse occurs when the investors switch camps/herds, almost instantaneously, and begin to sell. However, under the base H&S model it required the initial level of c1 to be at least 3 for the herding activity to become evident.

[21] The formation of a herd is defined by periods where the average network trust is significantly greater or less than 0 and investors share the same intentions.

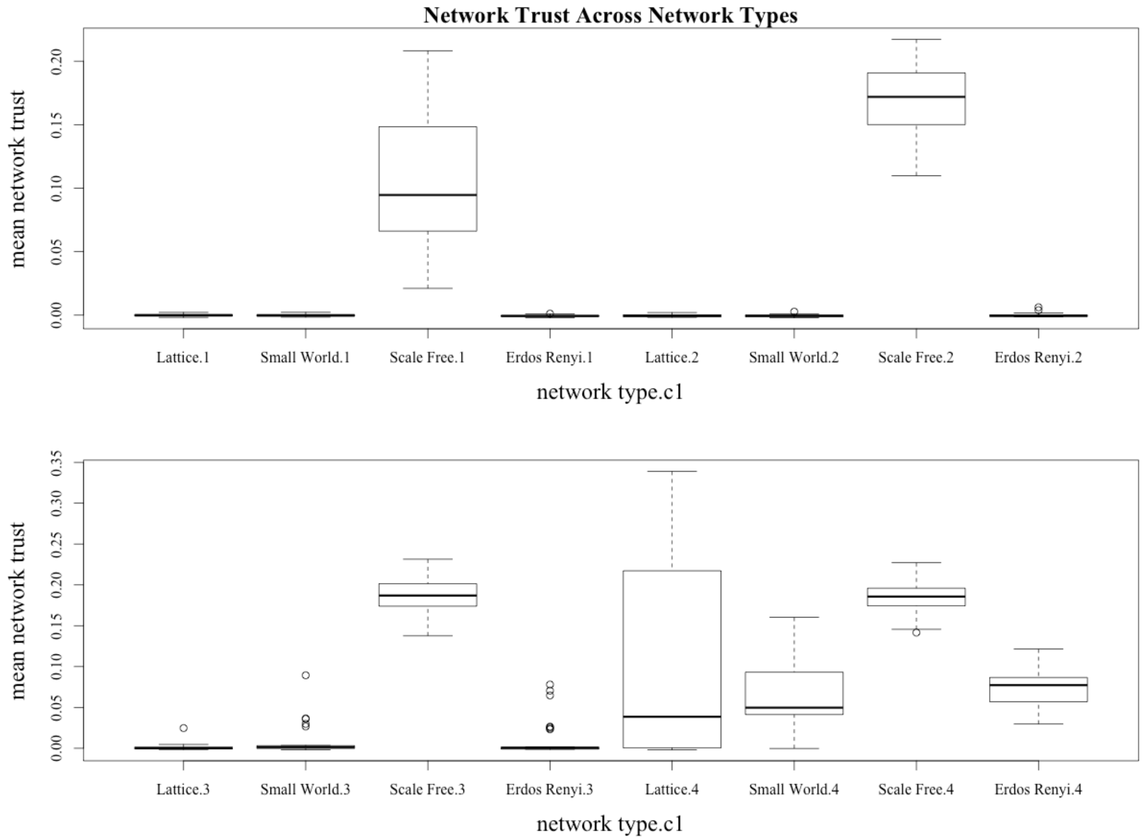


Figure 21: Differences in network trust when varying c1

What can be seen from Figure 21 is that the different network structures have very different behavior in terms of the level of network trust that they generate. Consistent with the results that the scale free network is the outlying structure, is the fact the median of network trust for the scale free network is both higher and requires a lower c1 for it to move away from 0. It can also be seen that when c1 is set at 4, the median of the network trust is greater than 0 for all the network types, which is consistent with the presence of bubbles under all network types once c1 is increased to 4.

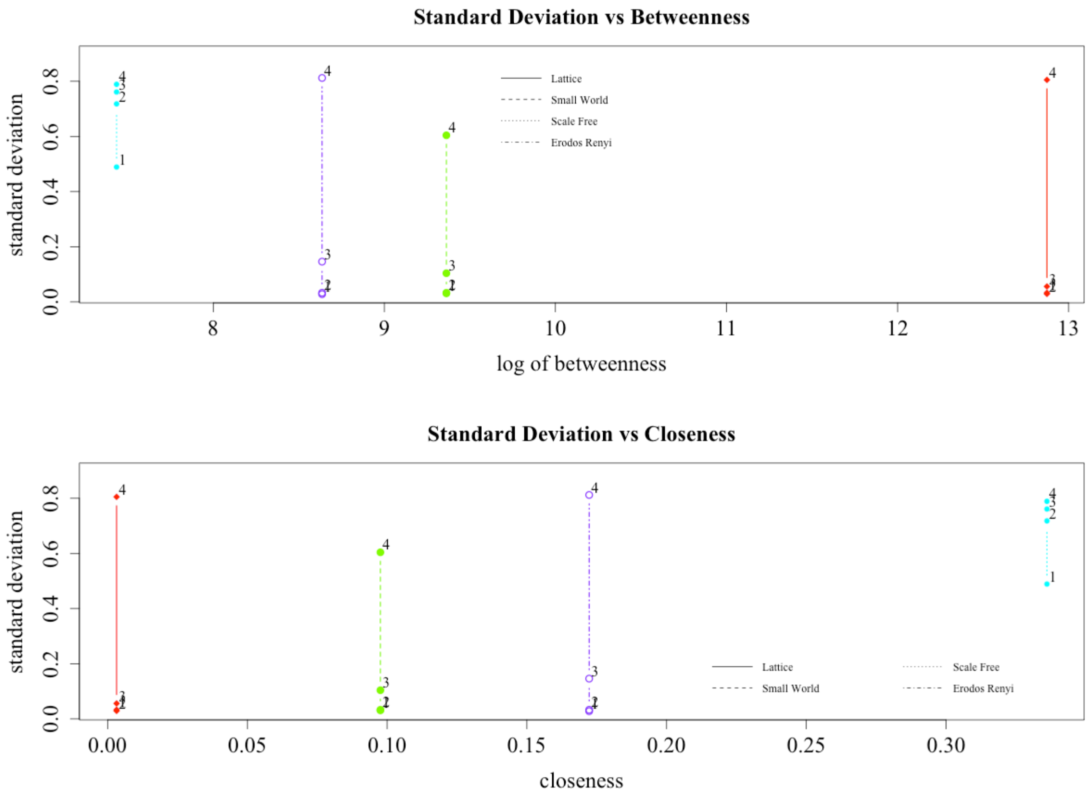


Figure 22: The relationship between risk and network centrality

Presented in Figure 22 is the relationship between the level of volatility, as given by the standard deviation, the centrality measures, the level of c1 (given by the numbers within the plot) and the various network structures. Noting from Table 8 that the level of volatility between the networks is statistically different, the key points are:

- When c1 = 1, the network with the lowest average betweenness measure (the scale free network) is the most volatile, while the other three are indistinguishable. However, when c1 is increased to 4, it is the Erdos Renyi network that becomes the most volatile. The Erdos Renyi network also has the greatest range in the volatility; and

- Given the relationship between closeness and betweenness it is to be expected that the previous observations are mirrored (and not replicated) when looking at closeness.

The first result is consistent with the finding of Ozsoylev and Walden (2011), who suggested that price volatility would be highest in markets with an intermediate level of connectedness yet lower in markets with higher or lower connectedness (the connectedness figures were provided in Table 6). The mechanism that drives the level of volatility higher yet closer for all network types is unclear. However, it appears that by increasing the initial level of trust in the information coming from the network (the c1 variable) it diminishes the impact of the network structure. This manifested itself with a bubble forming under each regime. Therefore, if investors are highly susceptible to listening to their network rather than other information sources, markets are likely to become more volatile regardless of the network structure.

The impact of investors being more susceptible to following their neighbor is illustrated by Thaler (2015) when he highlights the observation of Keynes (1937), who had suggested that markets had tended to be more efficient when professional investors using fundamental analysis controlled them. It had been the result of "uneducated" investors, who tended to follow the crowd, entering the market that created the greater volatility. This issue is further compounded when an asset bubble begins to inflate because, as Xiong (2013) points out, more and more less educated investors are attracted to the market.

4.2.3 Alternate Public Information

As detailed in Section 3.5.2, the revised model implements an alternate process where the investors compare the consensus forecast with the actual EPS result in assessing the value of the public information and the asset is capable of providing a dividend. The following section has the following aim and objectives:

- Assess the impact of changing the source of public information by comparing the revised model to the H&S model using only a lattice network. The rationale for the restricted testing is that it gives the best base to compare the results;
- Introduce a dividend payment by varying the payout ratio; and
- As part of introducing the new public trust functionality, adjusting the initial level of public information (c_{2ij} or c2 for short). The rationale for this change is to test the scenario that investors have greater initial faith in the value of public information and whether this can prevent the positive feedback loop with regards to adapting your neighbor's behavior. It is this feedback loop that is responsible for the creation of the bubble.

The top left corner of Figure 23 and Figure 24 provides the fan plots that show the results of the revised model with comparable settings to those in the previous section. The revisions do not materially change the model as it can be seen that with a setting of c1 = c2 = 1 and no dividend, no bubble forms and the price series is confined to a narrow bound around 1. While the band is narrower than the H&S model, test confirmed that the distribution of returns did not fit a Gaussian distribution. When the settings are changed to c1 = 4 and c2 =1 and no dividend, bubbles and crashes in comparable size to the original model appear, albeit they appear with greater regularity, something that may simply be the result of the random samples.

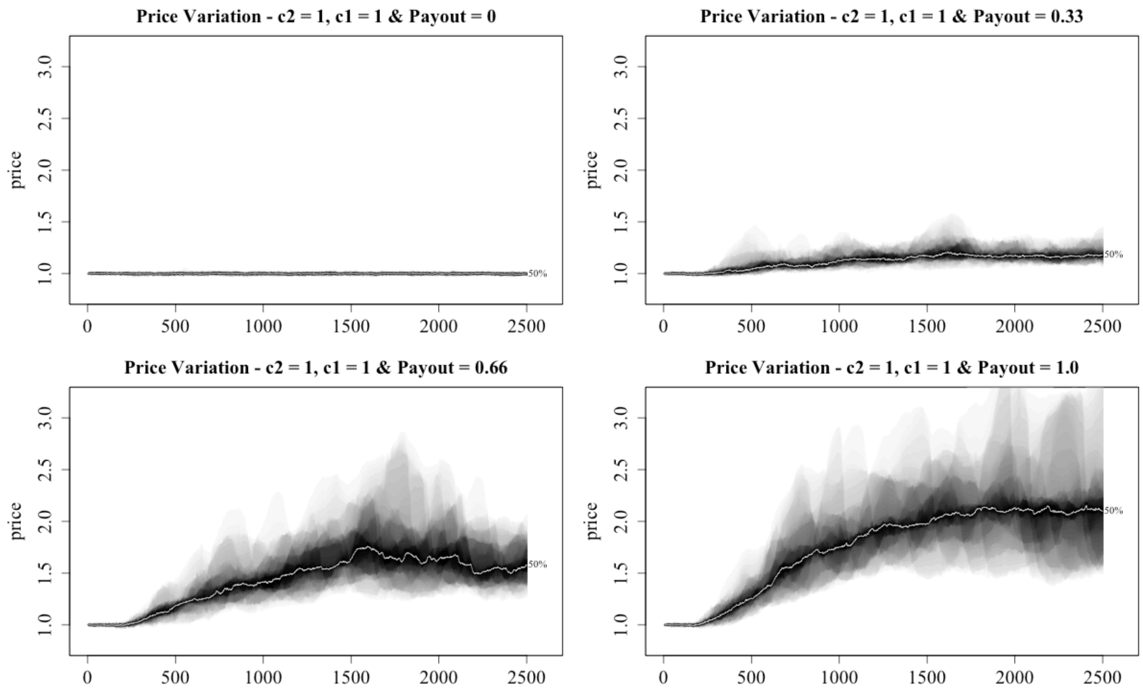


Figure 23: Revised model with varying pay-out ratios (c1 = 1 and c2 = 1) over time

From the figures it can be seen that the introduction of the dividend does impact the price series in numerous ways. Firstly, looking at Figure 23, it can be seen that as the payout ratio is increased from 0 to 1, the median and the volatility of the price series increases. At this point it is worth remembering that a dividend is only paid when the EPS for a period is greater than 0 and the investors cannot reinvest the proceeds. Also a sell signal is generated when the EPS result for the asset is less than the consensus forecast for the asset. In the instance that EPS < 0 and the result is below the consensus forecast,[22] the sell signal will not be diluted by the payment of the dividend. However, if the EPS result is positive, a dividend is paid and this will boosts the returns, thus reducing the power of the sell signal, which in turn will limit the growth in trust for the public information. The likely impact being that the trust that investors generate in their network is likely to go unchecked, resulting in the formation of more herds and therefore more volatile behavior in the market.

The introduction of a dividend above demonstrates a mild impact but the results in Figure 24 are far more explosive. The figures are prepared setting c1 to 4, a setting that is responsible for the creation of a bubble in the H&S model. Once the dividend is introduced the behavior post the inflation of the bubble is very different, to the point where the bubble does not deflate once the payout ratio is greater than 33%.

Even with a payout ratio of 33%, the median price remains in bubble territory, but the investors experience a wild ride. However, it is not enough to move the median price. When the payout ratio is 66% or greater, the median does not move once the bubble is formed. The significance of this being, that if there is a high initial bias to listening to your neighbors, the introduction of a dividend and the additional returns they provide sees investors form a buying herd and they can not be persuaded to join the selling herd regardless of what their other information sources are telling them. This is despite the fact that the EPS of the asset will miss consensus on average 50% of the time, thus creating a negative score for public information and providing a sell signal.

---

[22] It should be noted that the consensus forecast is very unlikely to be less than 0 given the distribution on the EPS and the forecasting mechanism. See Section 3.5.2 for more detail.

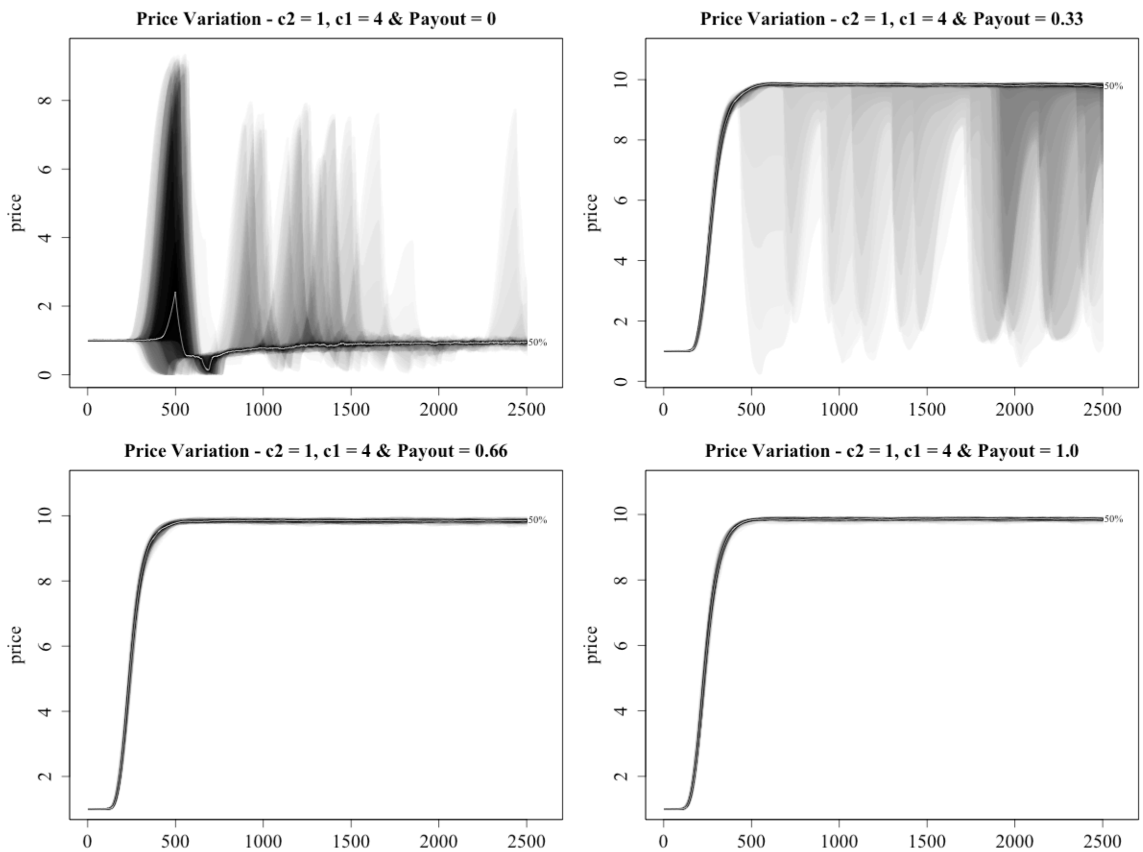


Figure 24: Revised model with varying pay-out ratios (c1 = 4 and c2 = 1) over time

Having seen the introduction of new dynamics resulting from the introduction of a dividend, the question moves to investigating if and how the bubble can be popped, and whether the market returns to the fundamental value of the asset (1 in this case). To try and achieve this, c2 was increased to 2, with results seen in Figure 25 and Figure 26.

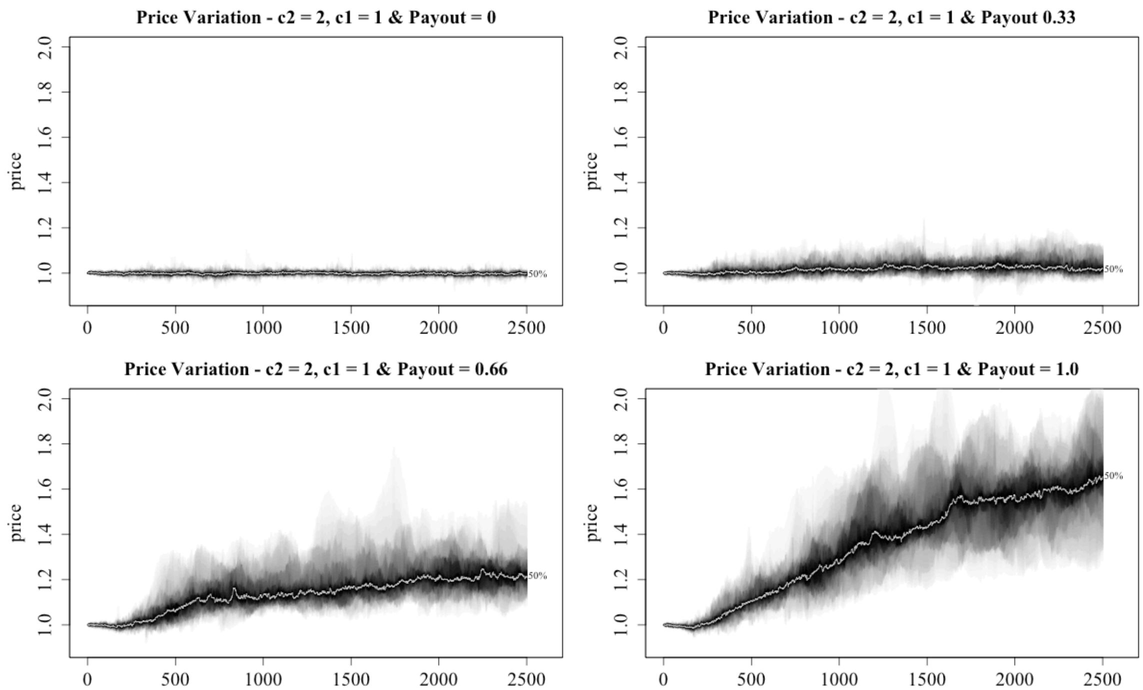


Figure 25: Revised model with varying pay-out ratios (c1 = 1 and c2 = 2) over time

The initial impression from setting c2 to 2 is that while the movement of the median price away from 1 still occurs, it is more gradual and does not reach the same level achieved by c1 = c2 = 1. One possible explanation, that is explored latter, is that an initial bias to public information slows the growth in the trust among neighbors and diminishes the probability of a herd forming. Under this regime investors will still be including a fundamental component in their decision making process.

In terms of the settings of c1 = 4 and c2 = 2, as per Figure 26, the main result is that when the payout ratio is 33%, the bubble sees a level deflation as the median price moves away from the upper limits. This contrasts with the higher payout ratios, where this behavior is not evident. In these instances the higher initial bias to public information is insufficient to prevent the herding of the investors.

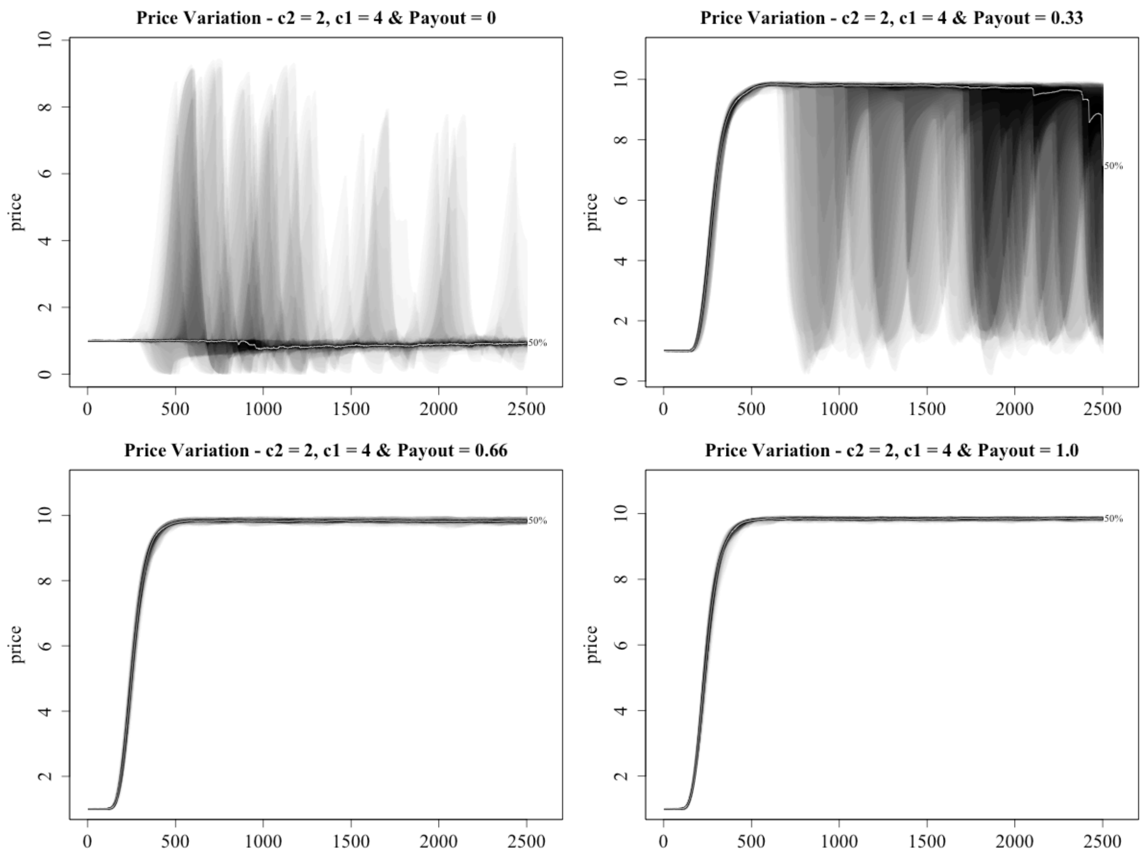


Figure 26: Revised model with varying pay-out ratios (c1 = 4 and c2 = 2) over time

A clearer contrast of the behavior of the various price series is provided in Figure 27, a boxplot of the mean prices with the various combinations of c1, c2 and the payout ratio. From the bottom chart, which illustrates the outcome when c1 is set to 4, the previously point of a higher initial bias to public influence having some influence when the payout ratio is 33% is seen. Not only is the median lower, but there is also a greater deviation in the mean. The influence that the higher c2 level has is seen in the top chart where median prices are lower. The ramification being that by increasing c2, the system has the ability to disrupt the herding behavior of the investors. This is an important insight into a mechanism that can prevent the inflation of a bubble, namely that if there is also a stronger initial consideration given to public information, the fundamental information provided by the public information is sufficient to dampen the herding instinct. The price series then becomes dependent on which source of information can gain the upper hand, assuming they are not correlated.

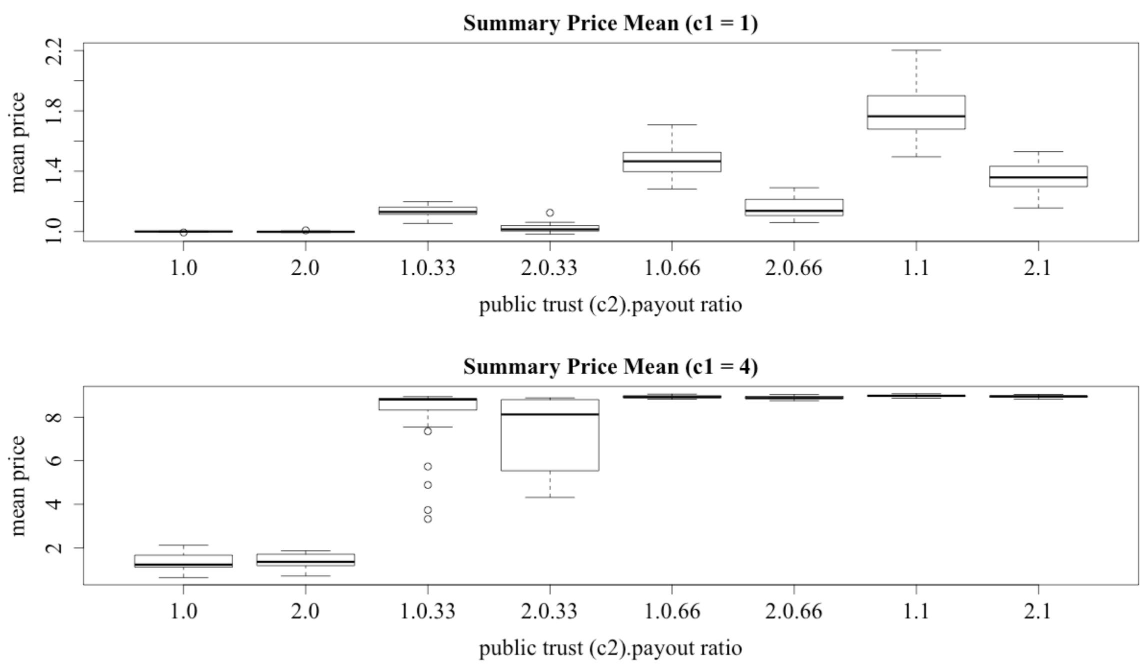


Figure 27: Boxplot with mean price variations

An interesting outcome, which is inferred in Figure 27, is illustrated in Figure 28, and shows the standard deviation of the price series. From the top chart, which is where c1 = 1, the price series where c2 = 1 show more range in terms of volatility yet the overall level of volatility is lower than the bottom chart where c1 = 4. This is no doubt a consequence of the bubble remaining inflated under the c1 = 4 regime. One exception is where there is no dividend and the volatility is higher under the c1 = 4 regime, a result that is in line with Section 4.2.2.

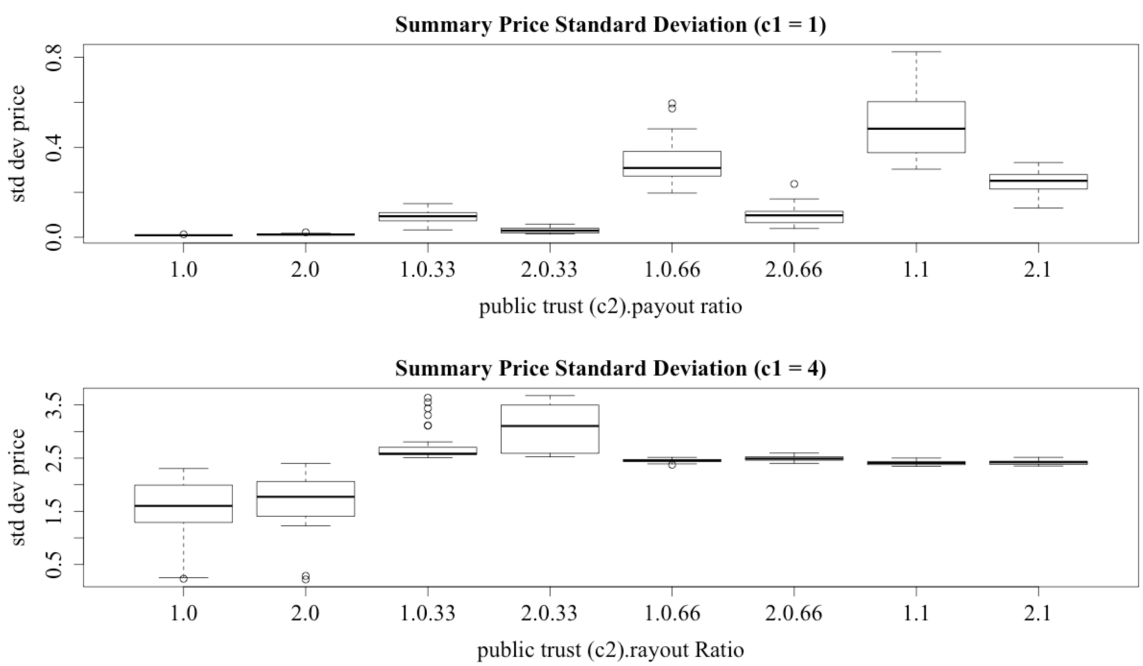


Figure 28: Boxplot with the standard deviation in price

Table 9 presents the mean price under the various combinations and the p-values for testing whether they are statistical significant. A Kruskal-Wallis rank sum test was used to generate the p-values because the distributions of the prices again violated the assumptions for a one-way ANOVA test.

Table 9: Mean prices under the various regimes

		Payout Ratio					
c1	c2	0	0.33	0.66	1	Ave.	P-value
1	1	1.000	1.133	1.474	1.785	1.348	**(a) <.01**
	2	0.999	1.024	1.155	1.360	1.134	**(a) <.01**
	Ave.	0.999	1.078	1.315	1.573	1.241	
	P-value	**0.060**	**(b)<.01**	**(b)<.01**	**(b)<.01**		
4	1	1.350	8.111	8.937	8.979	6.844	**(a) <.01**
	2	1.393	7.287	8.892	8.953	6.631	**(a) <.01**
	Ave.	1.371	7.699	8.915	8.966	6.738	
	P-value	**0.204**	**(b)0.013**	**(b)0.012**	**(b)0.035**		

The null hypothesis that dividends have no impact on price can be rejected for all combinations of initial public (c2) and network trust (c1) – see (a) in Table 9. The inference from this result is that increasing the payout ratio can have a positive impact on the price for an asset despite them having a similar earnings profile. To determine the exact relationship and to test its significance, more data points would need to be generated. The table also provides statistical support that in most cases if the initial bias to public information is increased, then the mean price is lower for a given payout ratio – see (b) in Table 9.

An explanation of the varying volatility is offered in Figure 29, boxplots of the average level for the network trust coefficient. From the bottom chart, where c1 was set to 4, the level of network trust is always very high. It is known that at this level investors join and stay in the herd. However, in the top chart when c1 = c2 = 1 the level of trust in the information coming from neighbors, despite being higher, is more variable, with the ramification being that the investors do not favor any particular information source and are therefore making use of all the information rather than just trusting their neighbors. Therefore, one can conclude that the variability in the network trust is driving the volatility in the price when c1 = c2 = 1.

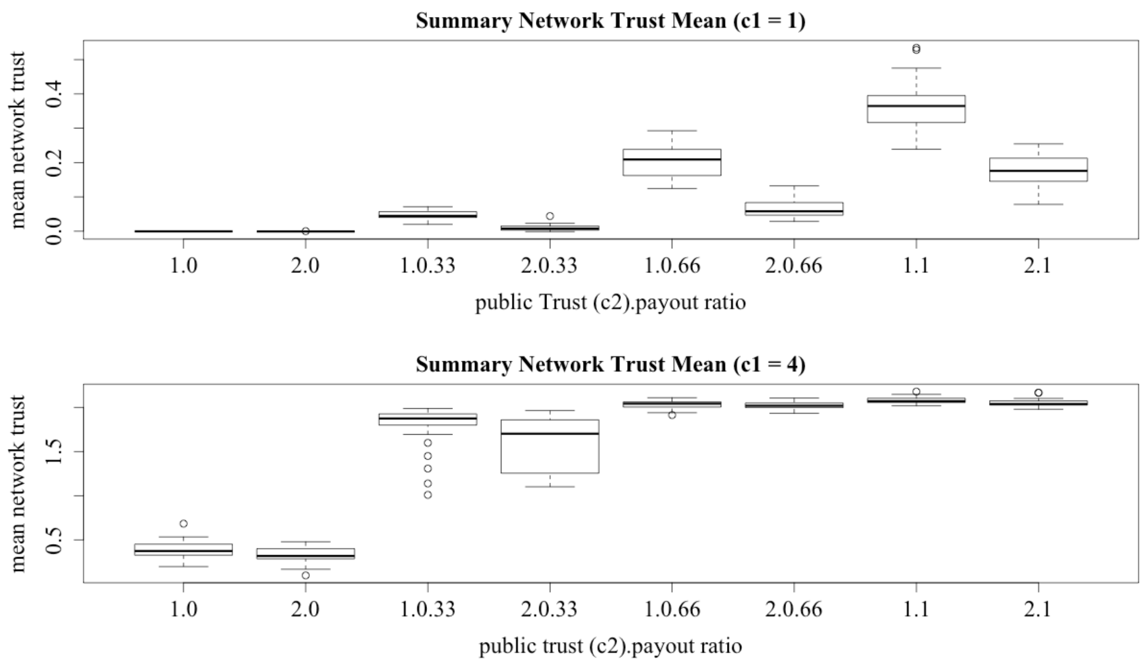


Figure 29: Boxplot of the mean network trust

To explore the previous conclusion further, an analysis of the state of the order book was conducted. The results can be seen in Figure 30, where the average of the net order books are displayed. The net order book is the number of bid (buying) orders minus the number of ask (selling) orders. This approach is preferable to the net value of the order book because it allows the reader to see the intention of the investors and is not impacted by the investors running out of funds. This is important because, given the closed system of this model; it is the lack of new funds that places a ceiling on the price.

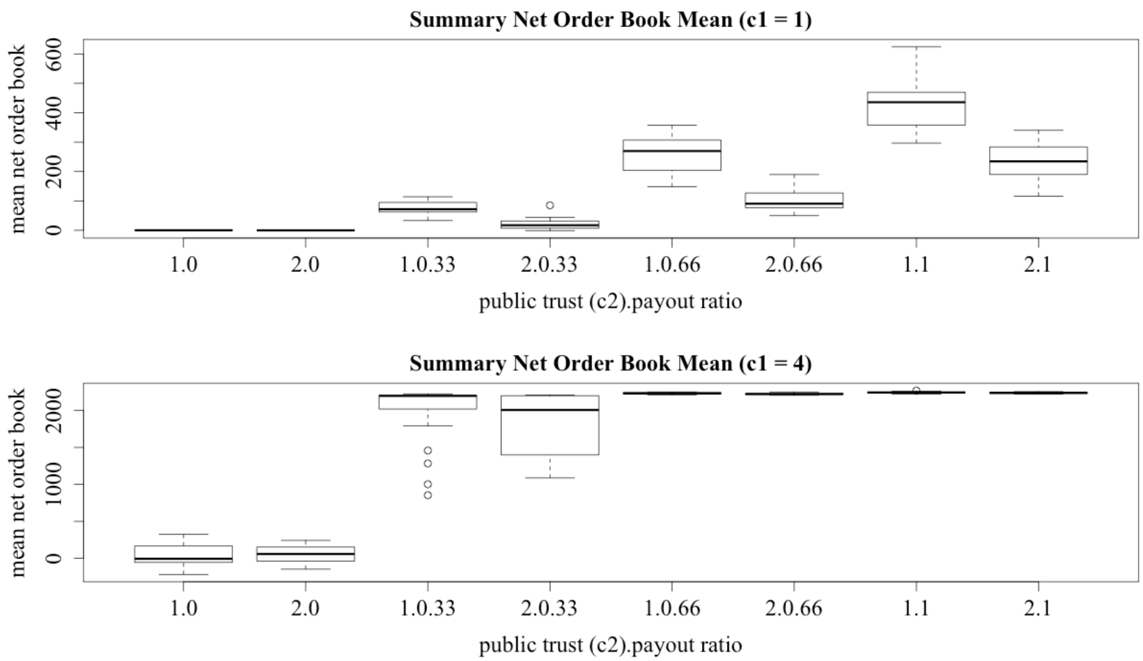


Figure 30: Boxplot for the net order book

An analysis of the net order book produces several important points. The first is that the bubble price is maintained under the c1 = 4 regime (the bottom chart) because investors still intend to buy, despite having no funds, as illustrated by the net order book having a median level of buyers over 2,000 once the payout ratio is at least 33%. Next, the variability in network trust flows through to order book when the payout ratio is 33%. Finally, when c1 = 1 (the top chart), it is seen the buying intentions of the investors do not match the bullishness of investors under the c1 = 4 regime, hence the lower median price. Consistent with the earlier analysis, is the fact that as the payout ratio increases, the buying intention of the investors increases. It is also seen that for the higher payout ratios, when c1 = c2 = 1, there is more volatility in the order book.

The conclusion that can be drawn from this section is that dividends play an important part in supporting a bullish sentiment amongst investors as they underwrite returns. This result contrasts with the capital structure irrelevance principle of Miller and Modigliani (1961). This theorem states that the market value of an asset is a combination of the earnings power and the underlying risk of a firm's assets, leaving the dividend policy as irrelevant. Unsurprisingly, this theorem was underwritten by the EMH.

4.2.4 The H&S Model With Multi Assets

As detailed in Section 3.3.1, the revised model provides the opportunity for the investors to consider multiple risky assets. To explore this space, the following section has the following experiments and output:

- The implemented model utilized the H&S framework; a lattice network, no dividend and white noise public information for three risky assets;
- Results were generated for varying the correlation for the public information of assets between 0, 0.33, 0.66 and 1 along with levels of c1 (initial network information bias) of 1 and 4 and c2 (initial public information bias) of 1 and 2. A decision based on keeping a consistent framework with the previous section; and
- Fan plots are used to illustrate the variation in prices for the three assets for the extreme settings correlation settings (0 and 1). Boxplots for the mean and

standard deviation of the multiple price series, network trust and the public trust for each asset are provided.

At this point it is timely to remember that the investors maintain only an overall trust level for each neighbor rather than a trust level for each neighbor for each asset. In contrast they do maintain a level of public trust for each specific asset. The significance being that for a neighbor to generate a high level of trust they, must consistently provide the correct action for all assets, as opposed to being a specialist in a particular asset. However, an investor may grow to trust the public information of one asset more than another.

The first price series as shown in Figure 31, provides the results of having 0 correlation between the asset's public information and a level of $c1 = 4$. These conditions within the single asset model were sufficient to see an asset bubble created. However, it is clear that no such event occurs for any of the assets under the multiple asset model.

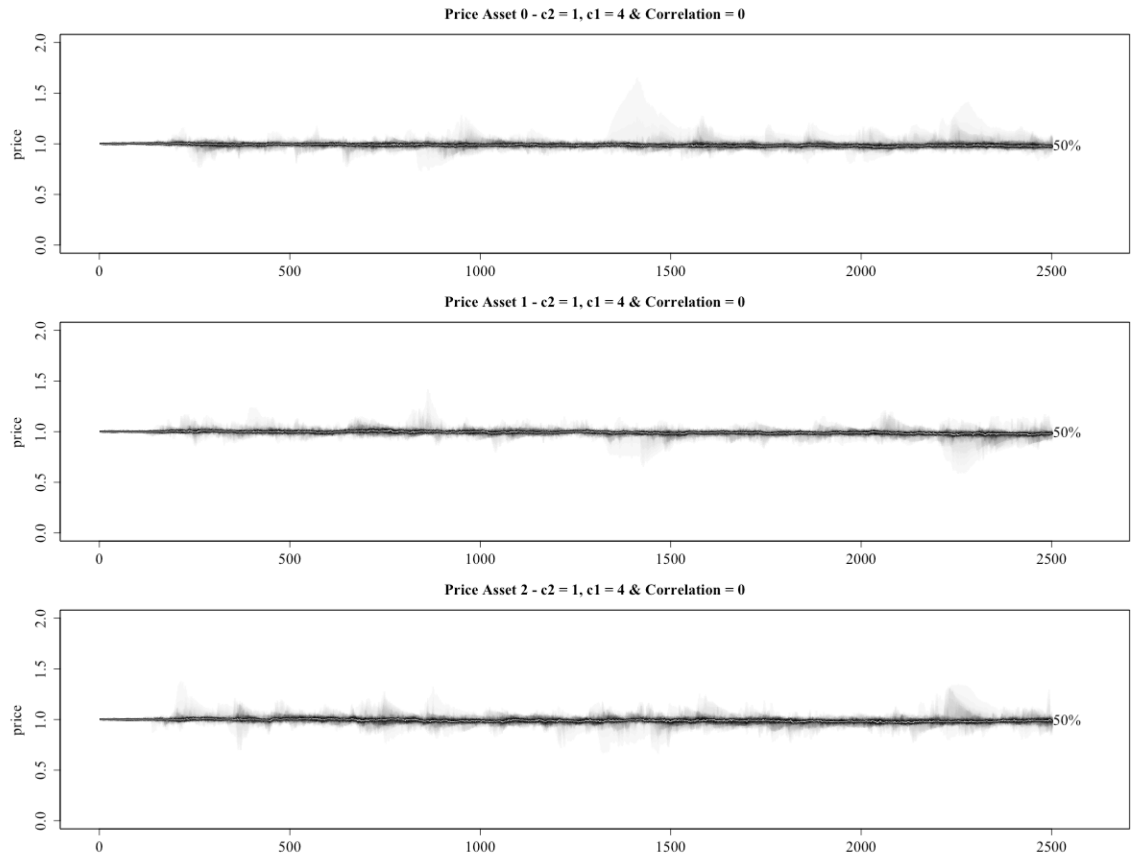


Figure 31: Price for 3 assets with c1 = 4, c2 = 1 and correlation = 0 over time

Having established the fact that the introduction of multiple assets with zero correlation delivers a more stable market, Figure 32 illustrates what occurs when the correlation in public information is increased to 1. The result is the appearance of some periods of elevated prices. However, neither the peaks nor the consistency of the appearance of them, matches the single asset case. Another point of note is that while the public information is the same for the three assets, the prices do not move in lock step as one might expect from a correlated series. This infers that the private information and network information must be providing contrary information to the investors. Another possible mechanism is that the trust investors have in public and/or network information may become negative, which will force them into taking a contrarian action to what the information suggests.

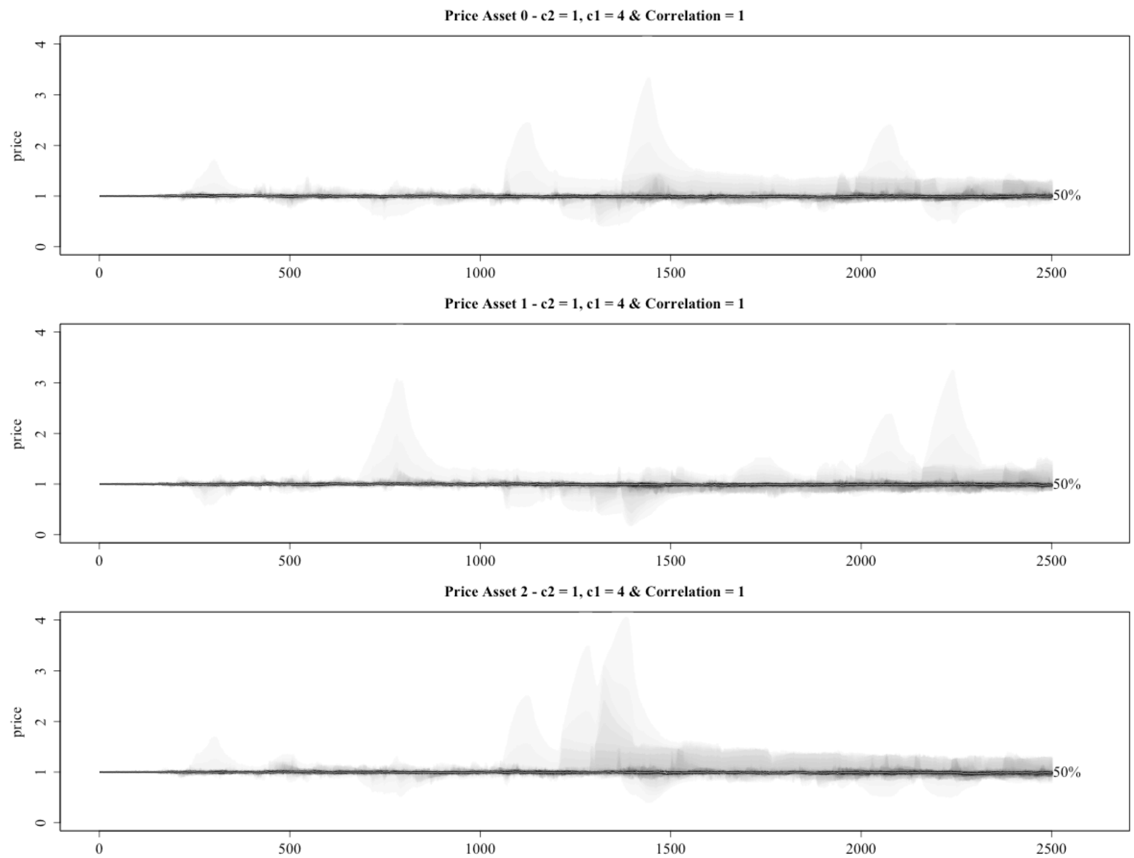


Figure 32: Price for 3 assets with c1 = 4, c2 = 1 and correlation = 1 over time

Under the revised model in the previous experiment, increasing the initial bias to public information (c2) was able to alter the behavior of the model, and this was attempted again. Figure 33 illustrates the results of zero correlation with c2 = 2 and c1 = 4. Again there is no bubble across any of the series, but any definitive comment regarding the difference is left to later in the section.

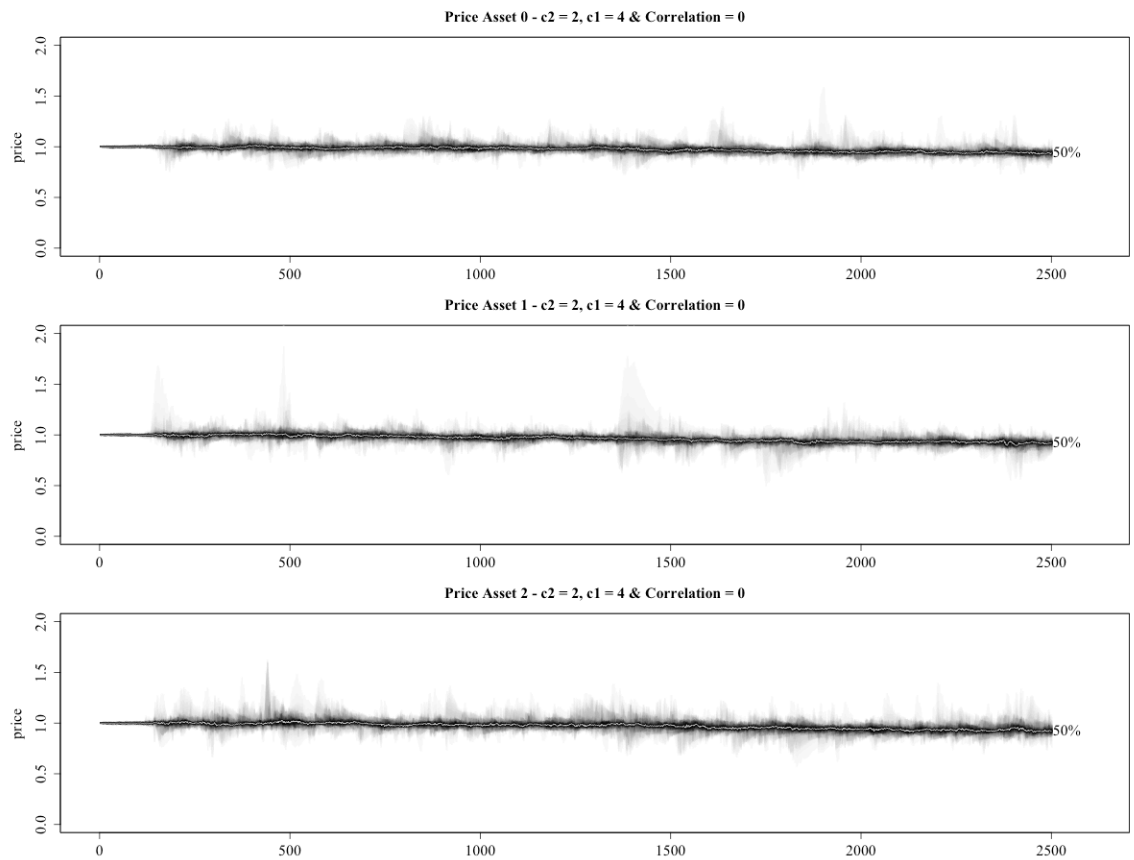


Figure 33: Price for 3 assets with c1 = 4, c2 = 2 and correlation = 0 over time

The final fan plot, as seen in Figure 34, illustrates the outcome of increasing the correlation between the assets. Again, there are periods of elevated prices and interestingly the price series of asset 0 and asset 2 appear to move in unison, while asset 1 appears to have more independence, suggesting that private and network information has greater influence than the public information for that asset.

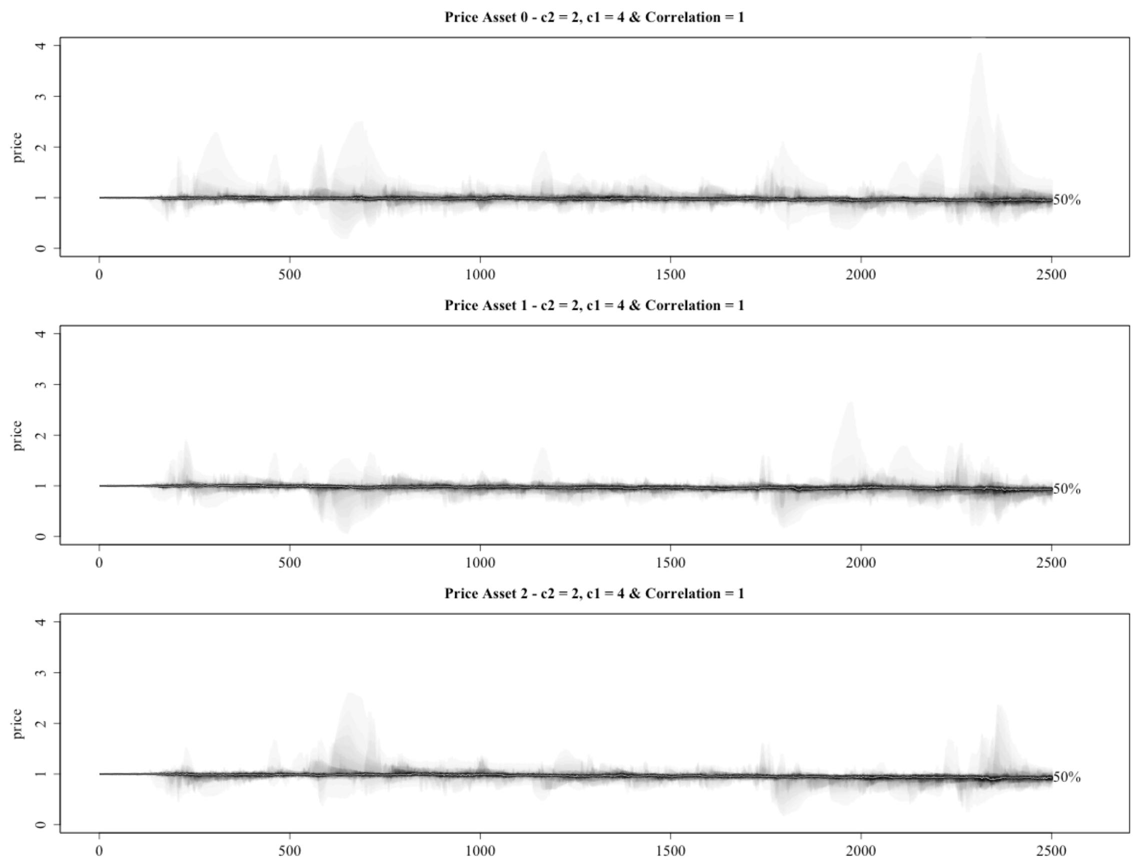


Figure 34: Price for 3 assets with c1 = 4, c2 = 2 and correlation = 1 over time

A clearer picture of the dynamics of the pricing is delivered via the boxplots of the mean and standard deviations of the prices for the assets in Figure 35 and Figure 36. In contrast to the single asset model, where the median price was generally above 1, is the fact that the median price for each of the assets is generally below 1 across the various scenarios. A price below 1 results from there being more selling than buying across the run. Given the information is generated in the same normally distributed manner, this would have resulted from investors having periods of negative trust in their information sources.

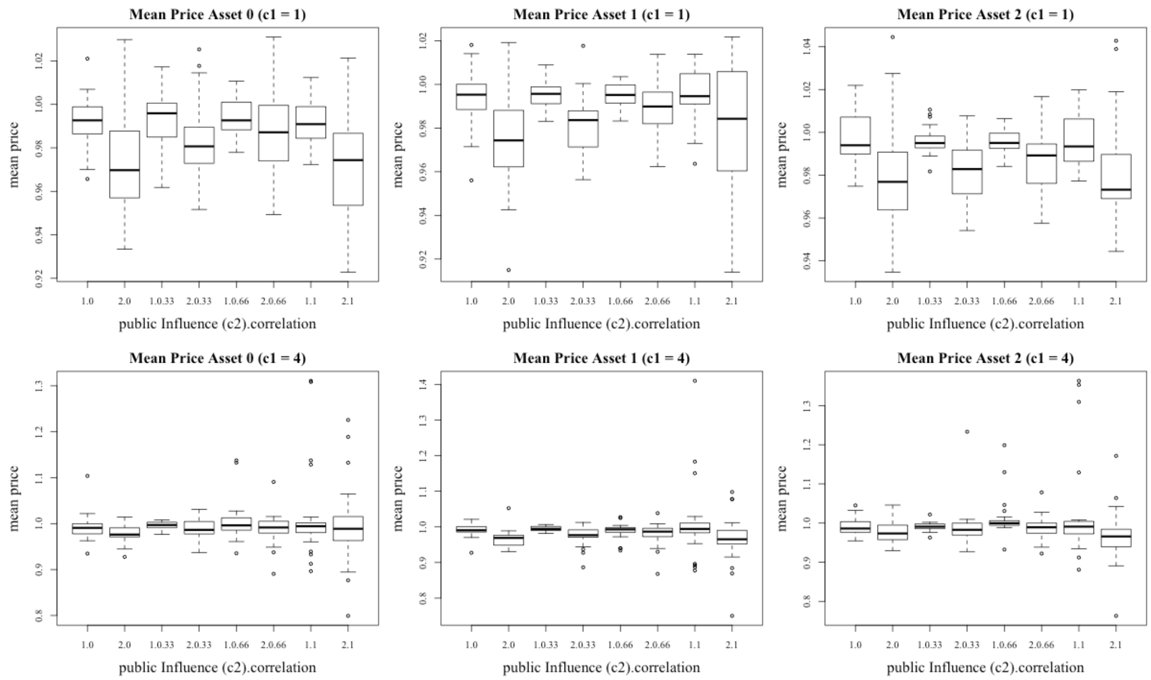


Figure 35: Boxplots for the mean prices for the 3 assets

Other results of note are; the volatility appears to increase when the level of c2 is increased to 2 when c1 = 1 ceteris paribus. Alternatively, when c1 is set to 4, the relationship is not as clear. However, as previously seen, overall the price series becomes more volatile when c1 is increased to 4. The level of volatility also increases as the level of correlation in the public information increases under the c1 = 4 scenario. This is unsurprising, given the previously noted periods of elevated asset prices, which are reminiscent of the single asset model. However, when comparing the level of volatility as seen in Figure 20, the introduction of multiple assets clearly suppresses the volatility.

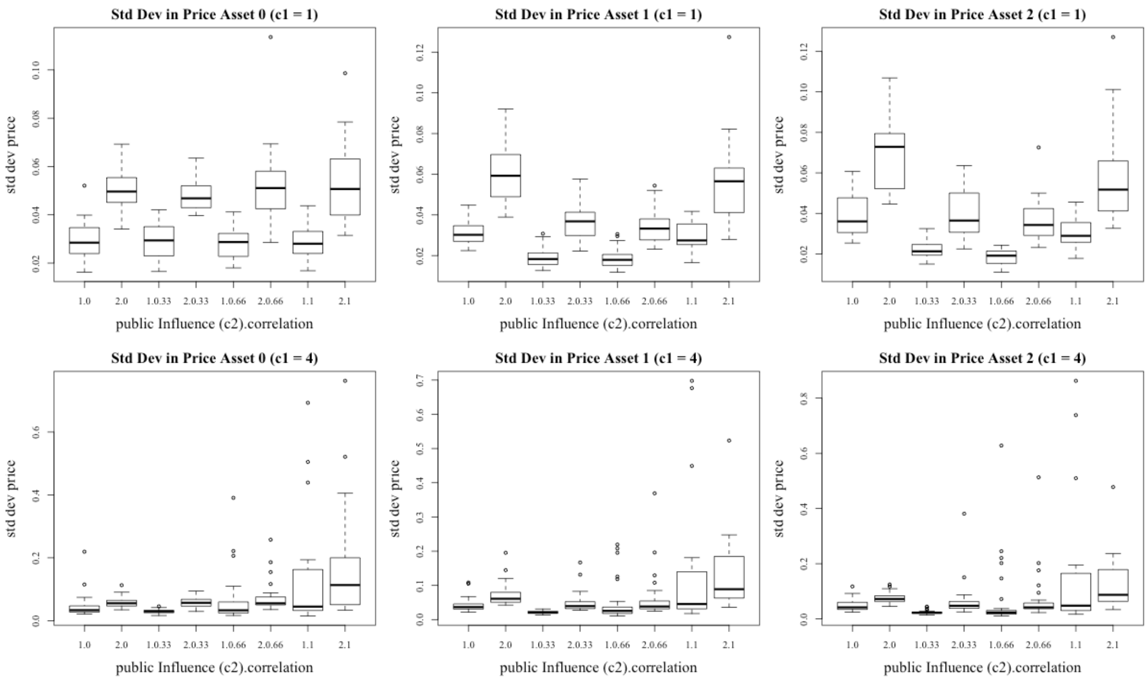


Figure 36: Boxplots for the standard deviation in price for the three assets

Having established the differences between the single and multiple asset model, and also within the multiple asset model (depending on the level of c1), the question turns to why these differences occur. From previous analysis, it is known that the level of trust for the various information sources will be an important source of variation. Boxplots of the average trust and standard deviations in trust for both network and public information are shown in Figure 37 through Figure 39.

Consistent with the other models, it can be seen that the level and the variability of network trust increases when c1 is increased to 4. Again it is this increased network trust that results in the formation of herds, which drives prices above and below what would be deemed the fundamental price of the asset.

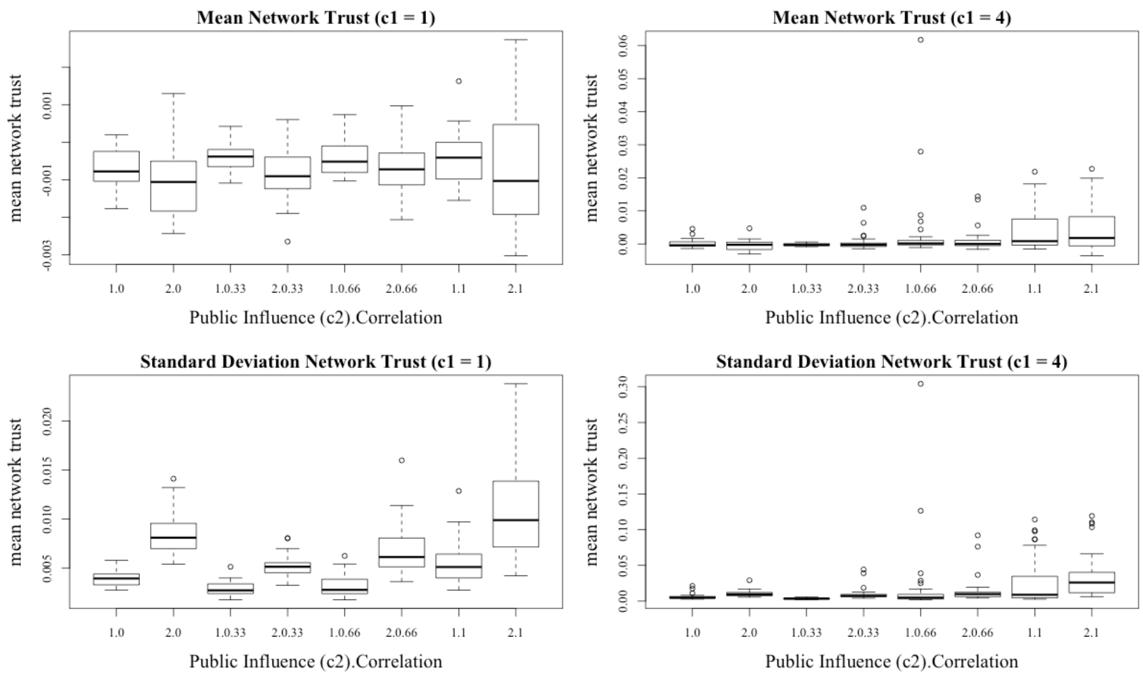


Figure 37: Mean and standard deviations for network trust levels

It was previously seen that the price series for the three assets were not comparable despite them having the same public information when the correlation is set to 1. Therefore, it is worth investigating how much trust the investors had in that source for each of the assets. From Figure 38 and Figure 39 it can be seen that indeed the level of public trust varies across the assets. The level of public trust is highest for asset 1 (careful attention should be given to the axis values) while the median public trust levels for asset 0 and 2 are around 0, indicating little overall trust. This outcome is capable of explaining why asset 1 performed differently to the other assets. That is the investors trusted the public information of Asset 1 to a greater degree and used that information to a greater degree in their decision making process.

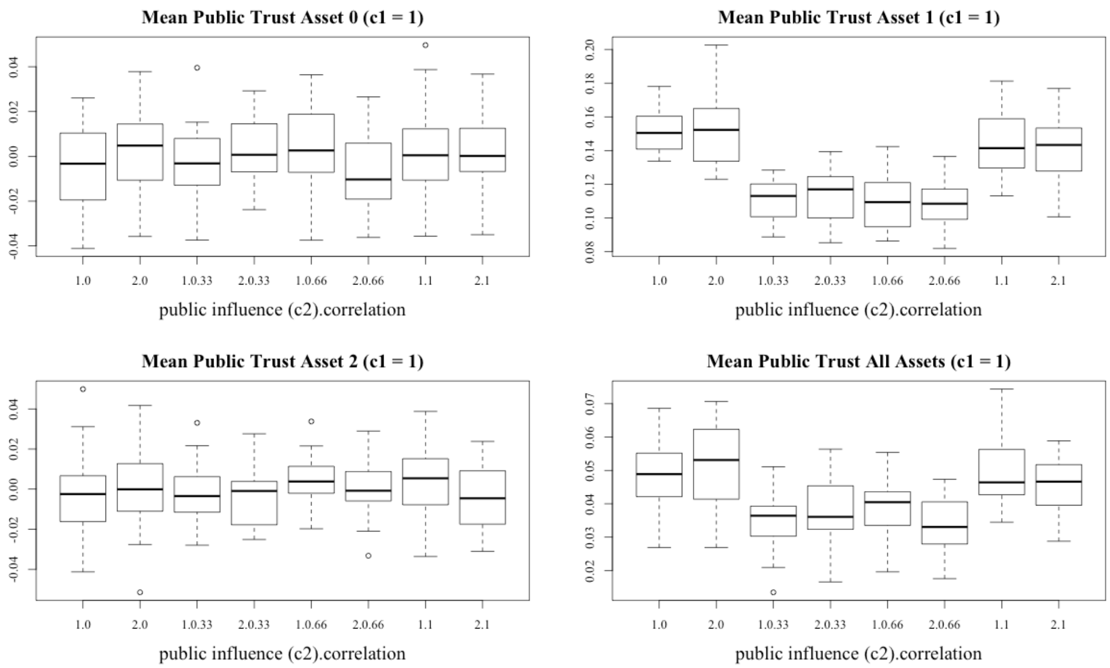


Figure 38: Mean public trust levels for the three assets (c1 = 1)

A curious outcome is how the median of the public trust for asset 1 decreases when the correlation with asset 0 is increased before increasing again when the correlation is set to 1. The mechanism that produces this result is worth further consideration but it is beyond the scope of this thesis.

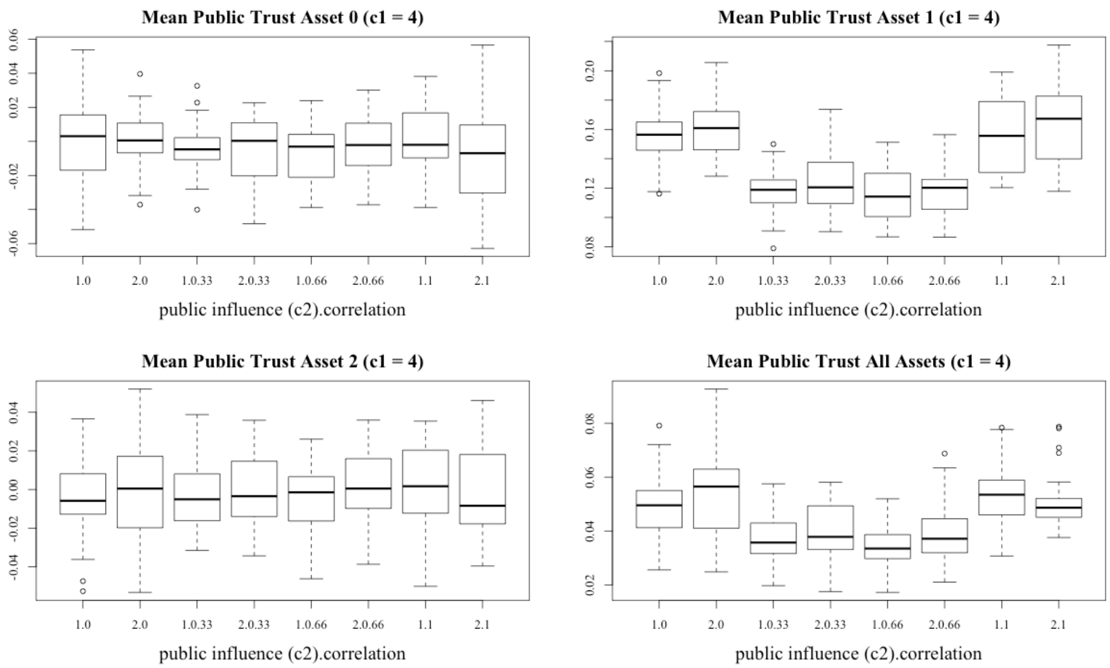


Figure 39: Mean public trust levels for the three assets (c1 = 4)

Comparing Figure 39 with Figure 38 suggests that increasing c1 to 4 has little impact on the level of public trust as both the magnitudes and patterns appear the same. Therefore, the increased price volatility that was seen when c1 was increased has been driven by the increased network trust. However, to confirm this, more work will be required, including generating data for a greater number of combinations.

4.2.5 The Quasi-Efficient Frontier

The final output for this thesis is to investigate the performance of the investors and assets in a risk return space. The intention of this is to establish whether the investors are capable of forming portfolios that are efficient, that is, there is no combination of assets that will deliver the same return with less risk. Before viewing the outcome, some background as to how the assets performed is illustrated by Figure 40. The series was generated with 5 assets with a correlation of 0.5, a seed of 100 and 1,500 investors. There are sufficient variations in the prices of the assets to create some interest but the prices generally revert to 1, which is a common consequence of the H&S model.

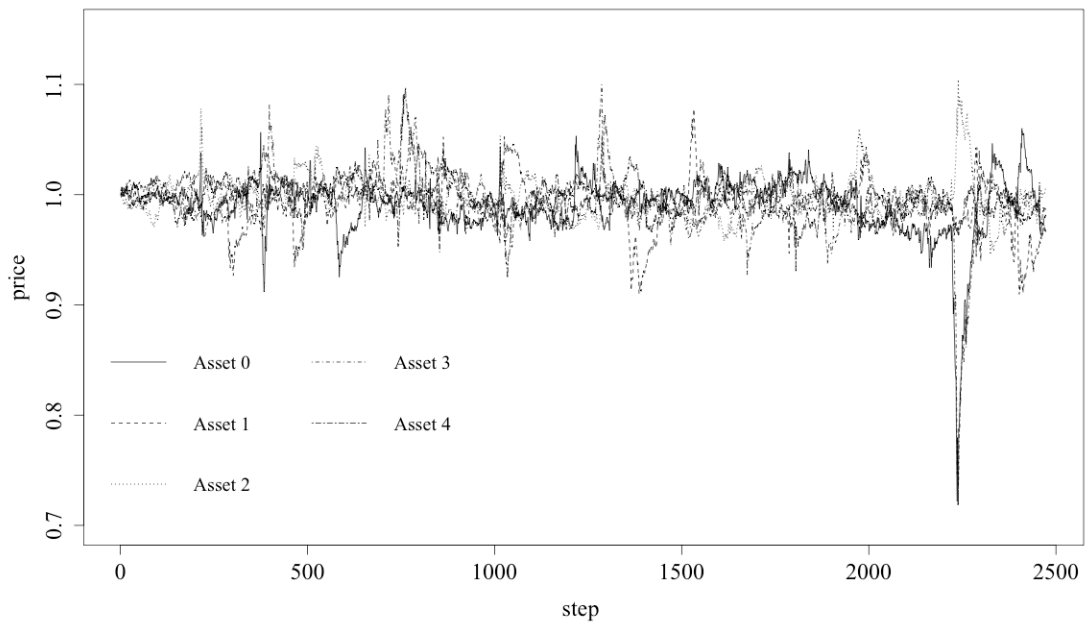


Figure 40: The price series for the 5 assets over time

The evolution of the quasi-efficient frontier can be seen in Figure 41, which plots the investors' (given by the + symbols) and assets' (given by the ⊗ symbols) positions at four time intervals in the return/risk space. To develop these charts the cash holding of the investors was removed, a step that is inconsistent with Figure 8, but was taken to condense the scale. The time intervals were specifically chosen and relate to volatile behavior in the times series plotted above. Given the axes are the same for each of the four plots, an evolutionary process can be identified. Initially the assets and the investors are spread ou,t but over time a thin band forms. This band could be construed as a frontier.

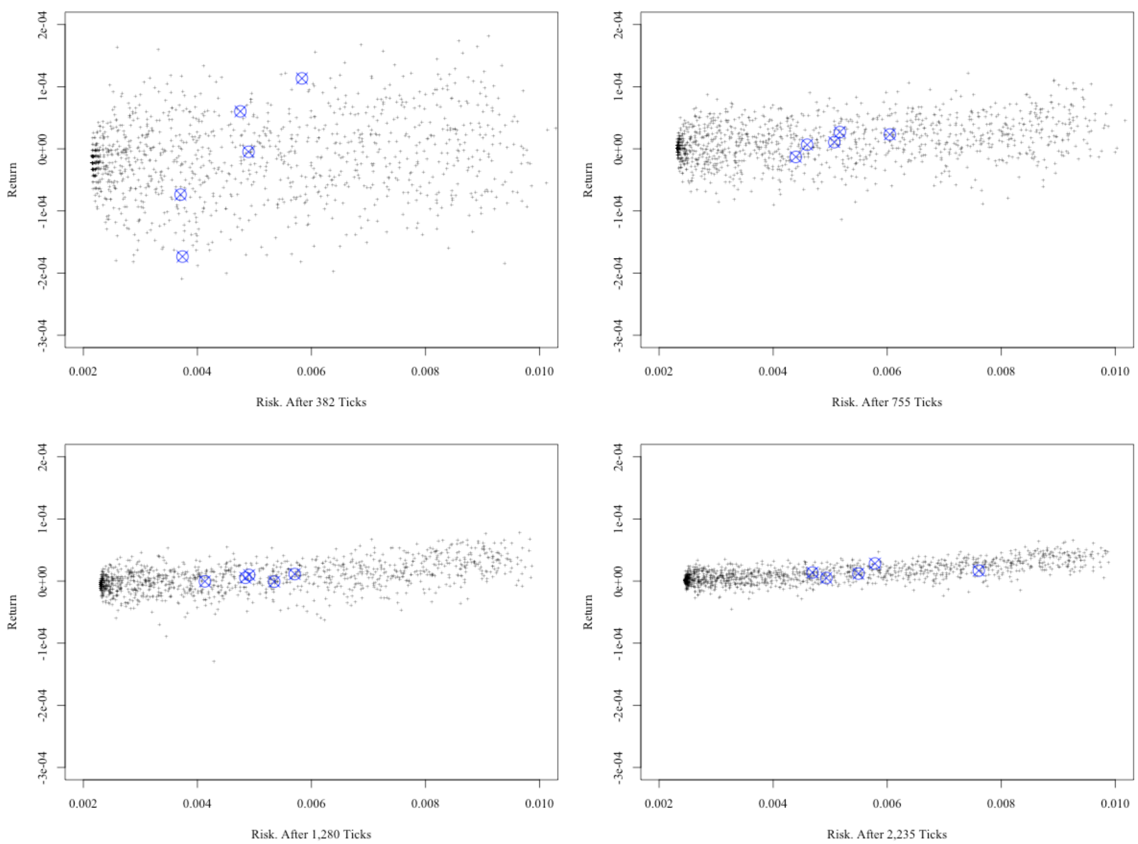


Figure 41: The evolution of the efficient frontier over time

Other observations of note are:

- There is a large group of investors who do not trade and they are grouped at the left hand end of the band. Given the returns for the assets are around zero, these investors have given up little in portfolio performance but have greatly reduced their risk. However, this result is quite abstract given the nature of the model and more realistic asset returns are required to test the utility of this finding; and

- In the final plot (bottom right), there does appear to be a trend that suggests that for agents to generate higher returns they had to be willing to tolerate higher risk[23]. This finding is consistent with finance theory, which dictates that for investors to tolerate higher risk they need to be compensated with higher returns.

In summary, the development of the quasi- efficient frontier is a pleasing result and validates the model. However, its presentation in this thesis was for the purpose of opening another field of examination rather than claiming any major breakthrough.

4.3 Section Summary

The research questions, that were first introduced in Section 1.3 and further expanded upon throughout this thesis, asked whether an ABM based artificial stock market could provide useful insight into the impact of different investor network topologies, dividend payout ratios and the introduction of multiple risky assets. Section 4.2.2 provided evidence that the different network topologies did indeed produce very different results in the situation that there was no initial bias to the information coming from your neighbors. Once there was a material initial bias to network information it was seen that the network structure became less relevant as bubbles appeared in all the regimes. The introduction of an alternative source of public information, as seen is Section 4.2.3, that also included a dividend, produced the insight that by varying the payout ratio very different returns were

[23] A basic OLS equation estimates a statistical significant relationship as return = 0.0047*risk - 1E-05.

generated. This is despite the assets having a similar earnings profile. The introduction of multiple assets, as seen in Section 4.2.4, and a first for an ABM based artificial stock market, saw the volatility of the market greatly reduced. However, increased correlation between the public information of the assets did see the volatility increase and outlying event still occurred. Finally, interesting insights with regards to the behavior of the investors were reported in Sections 4.2.1, trading behavior in a bubble, and 4.2.5, the quasi-efficient.

While the outcomes of this thesis are pleasing they are of little use unless further extensions can be generated and these extensions relate to actual financial markets. The case for both these arguments is made in Section 5.

5 DISCUSSION AND CONCLUSION

5.1 Introduction

This thesis has produced much to think about with regards to the possible mechanisms driving financial markets. This, in turn, forms the justification for multiple extensions. These extensions can be used to better inform regulators and investors alike in an attempt to avoid the inefficient behavior that has been experienced throughout the history of financial markets.

From the various experiments it is clear that the network investors form is important along with the dividend that an asset pays. However, it is also apparent that the outcome of an artificial stock market is very different once investors have to consider multiple assets. The implications of these findings are discussed in Section 5.2, while the further extensions are outlined in Section 5.3.

5.2 Implications of the Findings

The implications of the various experiments, which were first introduced in their relevant sections (Section 4.2.2 – 4.2.4), were varied and far-reaching. With regards to the different network topologies (see Section 4.2.2), it was clear that the different networks had very different characteristics when the c1 variable was set at 1. The first question to come from this is, what is the network structure that the actual market takes? This is not an easy question to answer, given there are over 10 million investors in the US stock markets (Ozsoylev & Walden, 2011). In addition, will this network remain static or does it change over time? Again this is a very difficult question to answer. However, from the results of the model used in this thesis, all is not lost. The model provided evidence that a bubble, which can be classified as the most inefficient market scenario, will form when investors have formed a scale free network, regardless of any initial bias to any of the information sources. This outcome is important because researchers can focus on understanding the hubs that exist in financial markets. Examples of hubs include: rating agencies, brokers, large pension funds and renowned stock pickers such as Warren Buffet. The case can easily be made that if these hubs became correlated then the rest of the investing universe would have little choice but to follow. For those who may doubt such a scenario, the sub prime meltdown in 2008 provides anecdotal evidence of such an outcome.

With regards to the implications of investors having a higher initial bias towards the actions of their neighbors, Harras and Sornettee (2011) cover this extensively. However, this thesis made the finding that the network topology becomes a redundant issue at a certain point. The implication stemming from this is the importance of the mindset of investors entering the market. While the implemented model assumed a fixed number of investors, it is not a large leap to see that if new investors are attracted to the market because of a period of abnormally higher returns, they will look to join the herd rather than taking the time to undertake fundamental analysis. From a quick reading of *The First Crash: Lessons from the South Sea Bubble* (Dale, 2004), one can see that this behavior is not beyond investors.

The findings with the regards to the impact of the payout ratio, see Section 4.2.3, presents various implications. Firstly, from the model it can be seen that a company has the ability to increase its share price by simply increasing its payout ratio, regardless of the underlying earnings. However, in reality this may be a short-term view because increasing the payout ratio comes at a cost, namely a lack of investment in further growth. Therefore, the earnings profile of the company will quite possibly be unsustainable. This point will be further discussed in Section 5.3.

Section 4.2.3 also presented the scenario where an asset bubble can be sustained indefinitely[24]. While this was interesting point, and provides some support to the argument of Blanchard (1979), in terms of the existence of bubbles within a rational expectations framework, the more important implication is the warning the model provides in the instance that a payout ratio cannot be sustained. Again while this was not tested, it is evident that a change in the payout ratio or a decline in the dividend payment would trigger a collapse of the bubble. This fact provides a clear warning sign, that if the market is being supported by abnormally high dividends the chance of a major correction appears to be higher. Parallels to current market conditions, where due to bond rates being at historically low forcing investors to seek yield elsewhere, can clearly be drawn.

In terms of the multi asset model, see Section 4.2.4, the implications are not quite as clear. The first implication is that for the general ABM community the bar should be moved from producing single risky asset markets to multi asset models. The rationale being that it provides a closer fit to the real world decisions facing investors and the

---

[24] Again this assumes that the earnings profile is sustainable regardless of the payout ratio

impact of multi assets is sufficient to alter the dynamics of the market. The second implication is that by adjusting, by a certain degree, the correlation between the assets, the behavior is again different. Therefore, further work can investigate how and why the correlation between assets may move.

5.3 Further Extensions

From the discussion in Section 5.2, the avenues for further extensions are plentiful but fall under two broad brackets. The first is to undertake calibration of the model so it better fits real world data. This step in the ABM building process is known as validation and obtaining a close fit to real world data is considered the Holy Grail and will go along way to removing the abstract nature of the model. In addition, understanding the network structure and dynamics of actual financial markets is an important task.

The second bracket is to make further extensions to the existing model. As raised in Section 4.2.2, there is the potential to expand the network dynamics. In particular, investors may look to disconnect with existing neighbors that they lose trust in, before looking for better performing investors in the population. Intuition suggests that this process has the potential to see any network structure transform into a scale fee network, as investors gravitate towards the better performing investor, with a herd and bubble resulting. However, will an investor be able to sustain a prolonged periods of outperformance, such that the population wants to replicate their decisions? The question also arises as to how the investors will search for the outperforming investors, is the search only within the neighbors of their neighbors or across the entire population?

The other obvious network related extension is to consider the implication of directed links. The current model sees information flow in both directions between neighbors. However, in the real world the flow of information may only flow in one direction for a number of reasons, including the nature of the relationship between neighbors and the fact that you do not always listen to someone who listens to you.

The possibility of extending the dynamics of both the investors and the assets exists. For the assets, the variables such as the payout ratio and the distribution of earnings are fixed rather than adjusting to environment. The key extension is developing an environment where the consequence of maintaining a high payout ratio is reflected in the earnings potential of the asset. Additionally, a non-stationary earnings series should be included to better reflect the reality that companies grow and decline, with investors attempting to understand which way the earnings are moving.

In terms of the investors, both their transaction ratio and investing threshold are obvious examples to add a dynamic element to. This extension would account for much of what behavioral finance has but forward with regards to how investors behave as their confidence and returns increase and decrease.

While the potential exists in the current model, further work in regards to understand the heterogeneity of investor performance across the population is an important step. As mentioned in Section 2.3.4, there are numerous theories with regards to the impact of an investor's position in a network and their investment performance. Future analysis should either confirm the existing work or provide insight as to why it does not hold. The other factor to consider is the decision-making threshold, or as Harras and Sornette (2011) refer to it, risk aversion. The intent will be to understand whether constant trading outperforms a buy and hold strategy. The other factor to consider, as mentioned in Section 4.2.1, is whether having a low decision threshold allows an investor to get in early when a bubble forms and then jump early, thus maximizing the investor's returns or whether these investors will be in and out and lose their early returns. Such analysis may have been of benefit to Isaac Newton, who after losing his initial gains when he reentered the market during the South Sea bubble, was forced to concede that "I can calculate the motion of heavenly bodies, but not the madness of people" (O'Hara, 2008).

5.4 Final Word (For Now)

It has become clear that the networks that investors form, and how they use information from that network, and information in general, are significant factors in unraveling the mysteries of financial markets. It is also apparent, that since their arrival, over 20 years ago, ABMs of artificial stock markets are ideally suited to helping unravel these mysteries. This thesis has made a small but important contribution in the development of these models, but more importantly, has provided a road map for any willing researcher to pickup and follow.

APPENDIX 1

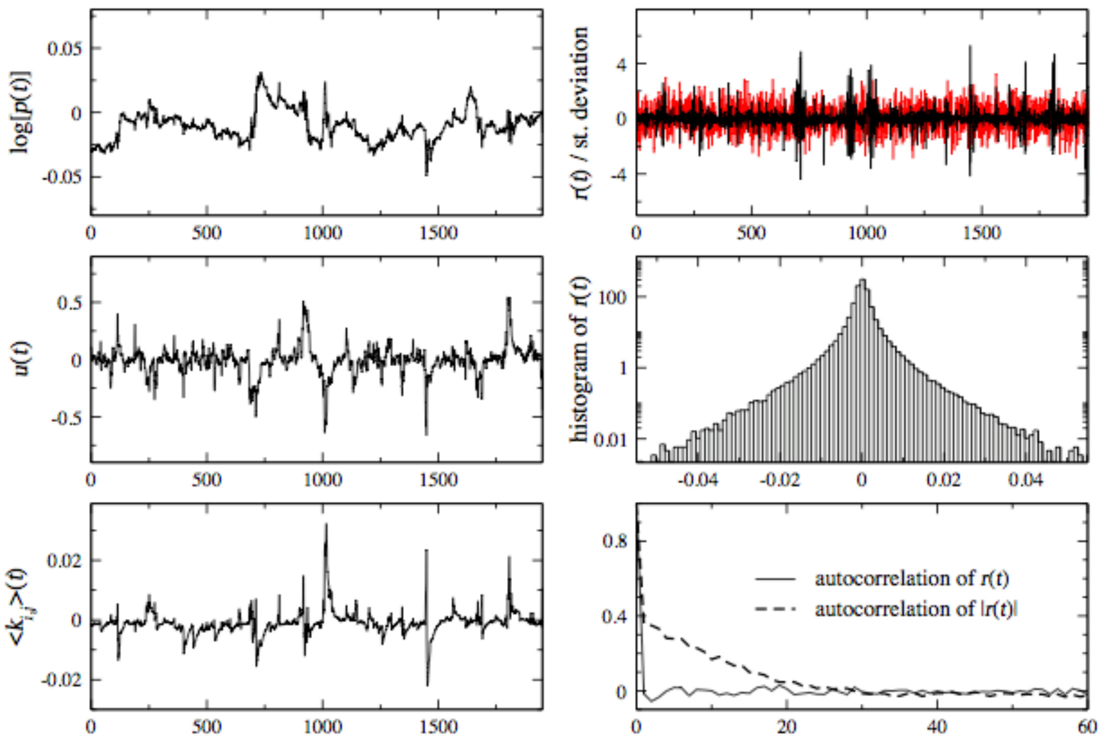


Figure 42: Typical runs from the H&S model (default settings c1 = c2 = c3 = 1)

The above chart is taken from the H&S paper, and provides the output for a 'typical' run.

These charts were used to ensure the correct calibration of the author's model.

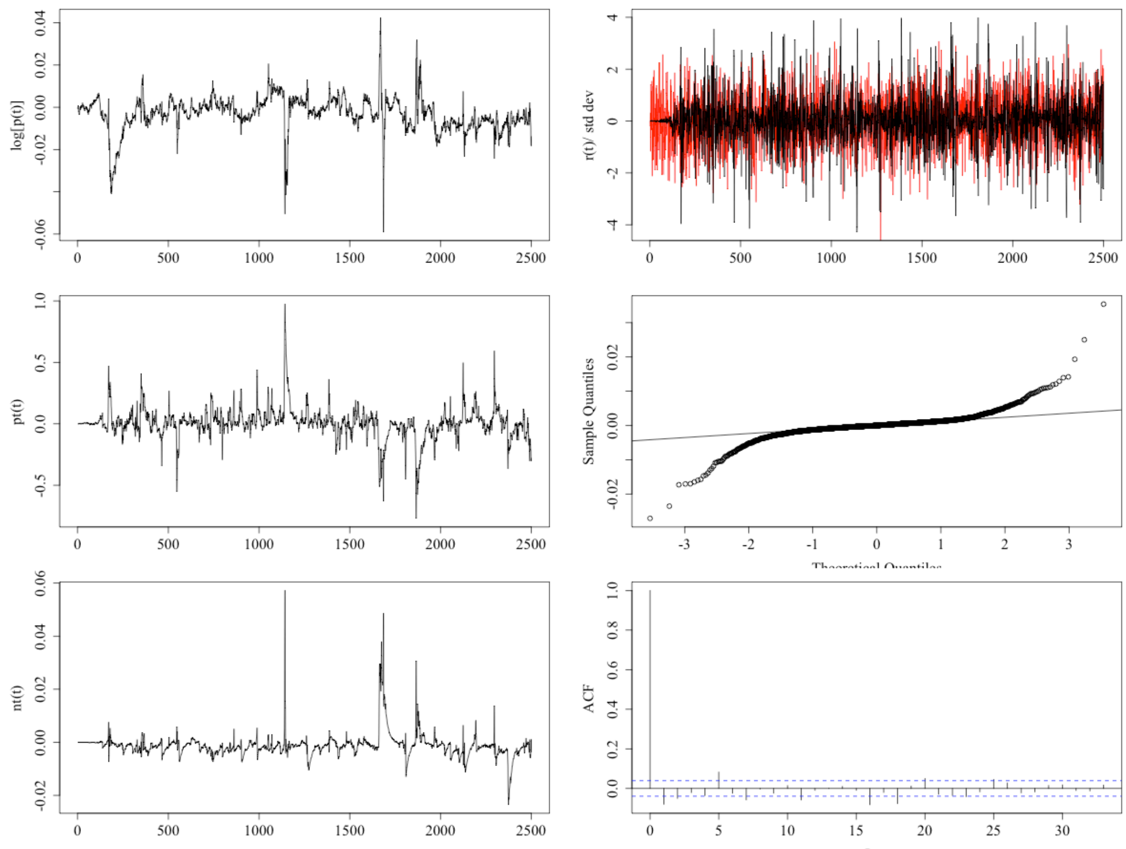


Figure 43: Typical run with default settings

The above figures provide an illustration of a typical run generated from the author’s model. The results can be reproduced using a seed of 100. While not an exact match with the H&S model, it does provide support that the model was able to produce results comparable to the H&S model.

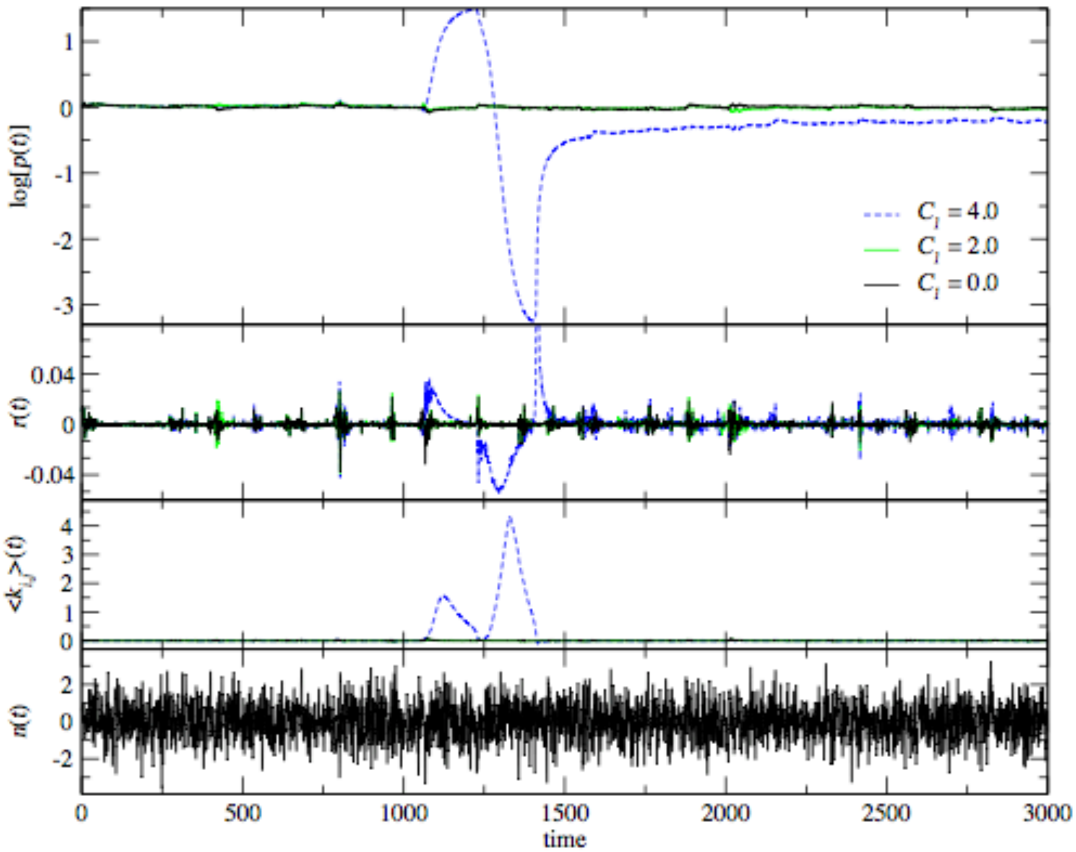


Figure 44: Examples of typical runs from H&S

This chart, from H&S, was provided to demonstrate the resulting dynamics from changing the c1 level. It can be seen that when c1 = 4 a bubble is generated.

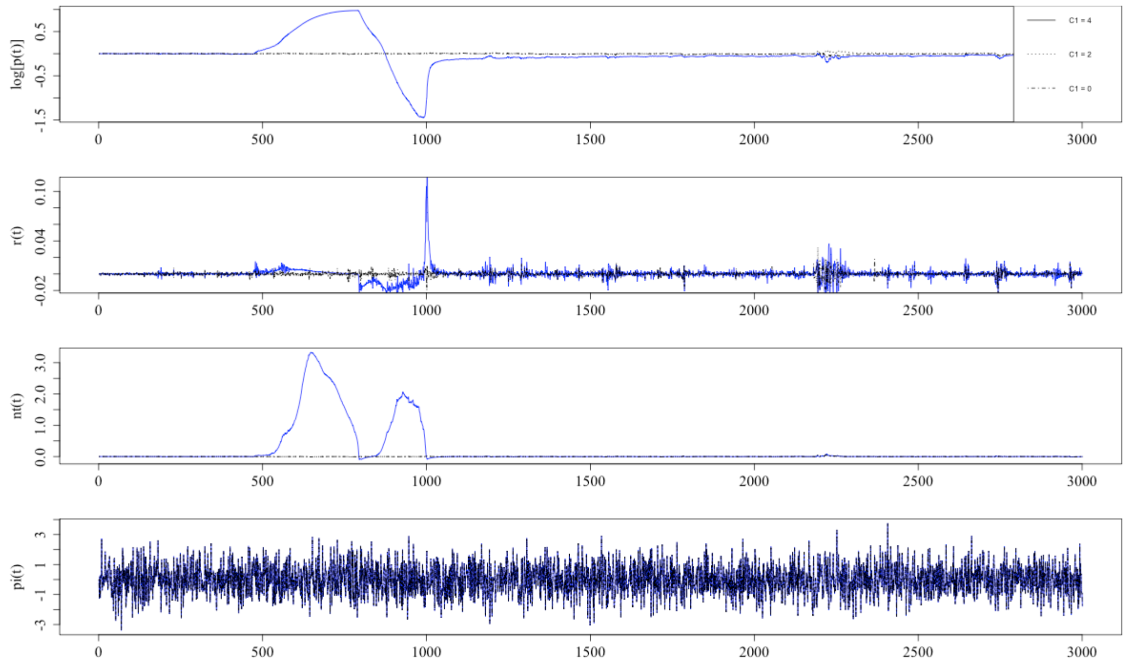


Figure 45: Replicated runs with different c1 values over time

The above figures provide an illustration of a typical run generated with the author's model when c1 is varied. Again the charts are sufficiently similar for the author to be satisfied that the implemented model performed as expected.

REFERENCES

Abel, G. (2015). fanplot: An R Package for Visualising Sequential Distributions. *The R Journal*, *7*(2), 15–23.

Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and Attack Tolerance of Complex Networks. *Nature*, *406*(6794), 378–382. http://doi.org/10.1038/35019019

Alfarano, S., & Milaković, M. (2009). Network Structure and N-dependence in Agent-Based Herding Models. *Journal of Economic Dynamics and Control*, *33*(1), 78–92. http://doi.org/10.1016/j.jedc.2008.05.003

Andersen, J., & Sornette, D. (2003). The $-Game. *The European Physical Journal B - Condensed Matter*, *31*(1), 141–145. http://doi.org/10.1140/epjb/e2003-00017-7

Andrecut, M. (2013). Portfolio Optimization in R. *Eprint arXiv:1307.0450*. Retrieved from http://arxiv.org/abs/1307.0450

Arthur, W. B. (1994). Inductive Reasoning and Bounded Rationality. *The American Economic Review*, *84*(2), 406–411.

Arthur, W. B. (2006). Chapter 32 Out-of-Equilibrium Economics and Agent-Based Modeling. In *Handbook of Computational Economics* (Vol. 2, pp. 1551–1564). Elsevier.

Arthur, W. B., Holland, J. H., LeBaron, B. D., Palmer, R. G., & Tayler, P. (1997). Asset Pricing Under Endogenous Expectations in an Artificial Stock Market. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.2252

Axtell, R. (2000). Why Agents?: On the Varied Motivations for Agent Computing in the Social Sciences. *The Brookings Institution Center on Social and Economic Dynamics Working Paper 17*. Retrieved from http://www.brook.edu/~/media/Files/rc/reports/2000/11technology_axtell/agents.pdf

Bachelier, L. (1995). *Théorie de la spéculation* (Réimpr). Paris: Gabay.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, *107*(3), 797–817. http://doi.org/10.2307/2118364

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509–512. http://doi.org/10.1126/science.286.5439.509

Barberis, N., Shleifer, A., & Vishny, R. (1998). A Model of Investor Sentiment. *Journal of Financial Economics*, *49*(3), 307–343. http://doi.org/10.1016/S0304-405X(98)00027-0

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*(3), 637–654. http://doi.org/10.1086/260062

Blanchard, O. J. (1979). Speculative Bubbles, Crashes and Rational Expectations. *Economics Letters*, *3*(4), 387–389. http://doi.org/10.1016/0165-1765(79)90017-X

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter Mood Predicts the Stock Market. *Journal of Computational Science*, *2*(1), 1–8. http://doi.org/10.1016/j.jocs.2010.12.007

Brock, W. A., & Hommes, C. H. (1998). Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model. *Journal of Economic Dynamics and Control*, *22*(8-9), 1235–1274. http://doi.org/10.1016/S0165-1889(98)00011-6

Buchanan, M. (2014). *Forecast: What Physics, Meteorology, and the Natural Sciences Can Teach Us About Economics*. New York: Bloomsbury.

Caldarelli, G., & Catanzaro, M. (2012). *Networks: A Very Short Introduction* (1st ed). Oxford: Oxford University Press.

Callaway, D. S., Newman, M. E. J., Strogatz, S. H., & Watts, D. J. (2000). Network Robustness and Fragility: Percolation on Random Graphs. *Physical Review Letters*, *85*(25), 5468–5471. http://doi.org/10.1103/PhysRevLett.85.5468

Challet, D., Marsili, M., & Zhang, Y.-C. (2013). *Minority Games*. Oxford; New York: Oxford University Press.

Colla, P., & Mele, A. (2010). Information Linkages and Correlated Trading. *Review of Financial Studies*, *23*(1), 203–246. http://doi.org/10.1093/rfs/hhp021

Cont, R. (2005). Volatility Clustering in Financial Markets: Empirical Facts and Agent-Based Models. *Long Memory in Economics*. Retrieved from http://dx.doi.org/10.2139/ssrn.1411462

Cont, R., & Bouchaud, J.-P. (2000). Herd Behavior and Aggregate Fluctuations in Financial Markets. *Macroeconomic Dynamics*, *4*(02). http://doi.org/10.1017/S1365100500015029

Conway, E. (2009). IMF Puts Total Cost of Crisis at £7.1 trillion. *The Telegraph*. London. Retrieved from IMF puts total cost of crisis at £7.1 trillion

Cui, B., Wang, H., Ye, K., & Yan, J. (2012). Intelligent Agent-Assisted Adaptive Order Simulation System in the Artificial Stock Market. *Expert Systems with Applications*, *39*(10), 8890–8898. http://doi.org/10.1016/j.eswa.2012.02.018

Dale, R. (2004). *The First Crash: Lessons from the South Sea Bubble*. Princeton, NJ: Princeton University Press.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance*, *53*(6), 1839–1885. http://doi.org/10.1111/0022-1082.00077

Darley, V., & Outkin, A. V. (2007). *A NASDAQ Market Simulation Insights on a Major Market from the Science of Complex Adaptive Systems*. Hackensack, NJ: World Scientific.

De Bondt, W., Muradoglu, G., Shefrin, H., & Staikouras, S. K. (2008). Behavioral Finance: Quo vadis? *Journal of Applied Finance*, *19*(2), 7–21.

De Bondt, W., & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, *40*(3), 793–805. http://doi.org/10.1111/j.1540-6261.1985.tb05004.x

Diba, B. T., & Grossman, H. I. (1988). Explosive Rational Bubbles in Stock Prices? *The American Economic Review*, *78*(3), 520–530. http://doi.org/10.3386/w1779

Duffy, J., & Ünver, M. U. (2006). Asset Price Bubbles and Crashes with Near-Zero-Intelligence Traders. *Economic Theory*, *27*(3), 537–563. http://doi.org/10.1007/s00199-004-0570-9

Erdös, P., & Rényi, A. (1960). On the Evolution of Random Graphs. *Publication of The Mathematical Institute of the Hungarian Academy of Sciences*, *5*, 17–61.

Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, *18*(3), 25–46. http://doi.org/10.1257/0895330042162430

Farmer, J. D. (2002). Market Force, Ecology and Evolution. *Industrial and Corporate Change*, *11*(5), 895–953. http://doi.org/10.1093/icc/11.5.895

Farmer, J. D., Gallegati, M., Hommes, C., Kirman, A., Ormerod, P., Cincotti, S., … Helbing, D. (2012). A Complex Systems Approach to Constructing Better Models for Managing Financial Markets and the Economy. *The European Physical Journal Special Topics*, *214*(1), 295–324. http://doi.org/10.1140/epjst/e2012-01696-9

Freeman, L. C. (1978). Centrality in Social Networks Conceptual Clarification. *Social Networks*, *1*(3), 215–239. http://doi.org/10.1016/0378-8733(78)90021-7

Friedman, M. (1953). The Methodology of Positive Economics. In *Essays in positive economics* (Vol. 3). Chicago, Ill.: University of Chicago Press.

Grossman, S. J., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, *70*(3), 393–408.

Gupta, N., Hauser, R., & Johnson, N. F. (2005). Using Artificial Market Models to Forecast Financial Time-Series. *arXiv Preprint physics/0506134*.

Han, B., & Yang, L. (2013). Social Networks, Information Acquisition, and Asset Prices. *Management Science*, *59*(6), 1444–1457. http://doi.org/10.1287/mnsc.1120.1678

Harras, G., & Sornette, D. (2011). How to Grow a Bubble: A Model of Myopic Adapting Agents. *Journal of Economic Behavior & Organization*, *80*(1), 137–152. http://doi.org/10.1016/j.jebo.2011.03.003

Hong, H., Kubik, J. D., & Stein, J. C. (2005). Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers. *The Journal of Finance*, *60*(6), 2801–2824. http://doi.org/10.1111/j.1540-6261.2005.00817.x

Jaffe, C. (2015). New "Twitter ETF" Sounds Crazy, but it's Actually Dangerous. Retrieved October 20, 2015, from http://www.marketwatch.com/story/new-twitter-etf-sounds-crazy-but-its-actually-dangerous-2015-10-19

Jefferies, P., Lamper, D., & Johnson, N. F. (2003). Anatomy of Extreme Events in a Complex Adaptive System. *Physica A: Statistical Mechanics and Its Applications*, *318*(3-4), 592–600. http://doi.org/10.1016/S0378-4371(02)01743-0

Jensen, M. C. (1978). Some Anomalous Evidence Regarding Market Efficiency. *Journal of Financial Economics*, *6*(2-3), 95–101. http://doi.org/10.1016/0304-405X(78)90025-9

Johnson, N. F., Jefferies, P., & Hui, P. M. (2003). *Financial Market Complexity*. Oxford ; New York: Oxford University Press.

Keynes, J. M. (1937). The General Theory of Employment. *The Quarterly Journal of Economics*, *51*(2), 209. http://doi.org/10.2307/1882087

Keynes, J. M. (2007). *General Theory of Employment, Interest and Money*. London: Macmillian.

Kindleberger, C. P., & Aliber, R. Z. (2011). *Manias, Panics and Crashes: A History of Financial Crises*. Palgrave Macmillan.

Kirman, A. (1991). Epidemics of Opinion and Speculative Bubbles in Financial Markets. *Money and Financial Markets*, 354–368.

Kirman, A. (1993). Ants, Rationality, and Recruitment. *The Quarterly Journal of Economics*, *108*(1), 137–156. http://doi.org/10.2307/2118498

Kirou, A., Ruszczycki, B., Walser, M., & Johnson, N. F. (2008). Computational Modeling of Collective Human Behavior: The Example of Financial Markets. In *Computational Science – ICCS 2008* (Vol. 5101, pp. 33–41). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kothari, S. . (2001). Capital Markets Research in Accounting. *Journal of Accounting and Economics*, *31*(1-3), 105–231. http://doi.org/10.1016/S0165-4101(01)00030-1

Kothari, S. P., Lewellen, J., & Warner, J. B. (2006). Stock Returns, Aggregate Earnings Surprises, and Behavioral Finance. *Journal of Financial Economics*, *79*(3), 537–568. http://doi.org/10.1016/j.jfineco.2004.06.016

Laerd statistics. (2016). Kruskal-Wallis H Test using SPSS Statistics. Retrieved from https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php

LeBaron, B. (2000). Agent-based Computational Finance: Suggested Readings and Early Research. *Journal of Economic Dynamics and Control*, *24*, 679–702.

LeBaron, B. (2006). Agent-based Computational Finance. In *Handbook of Computational Economics* (Vol. 2, pp. 1187–1233). Amsterdam: North Holland.

Lux, T. (2006). *Financial Power Laws: Empirical Evidence, Models, and Mechanism*. Economics working paper/Christian-Albrechts-Universität Kiel, Department of Economics. Retrieved from http://hdl.handle.net/10419/3923

Mackay, C. (1841). *Extraordinary Popular Delusions and the Madness of Crowds*. Start Publishing LLC.

MacKenzie, D. (2008). *An Engine, not a Camera: How Financial Models Shape Markets*. Cambridge, Mass.: MIT.

Malkiel, B. G. (1999). *A Random Walk Down Wall Street: Including a Life-cycle Guide to Personal Investing*. New York: WW Norton & Company.

Malkiel, B. G., & Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383–417. http://doi.org/10.1111/j.1540-6261.1970.tb00518.x

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, *36*(4), 394. http://doi.org/10.1086/294632

Mandelbrot, B., & Hudson, R. L. (2006). *The (Mis)behavior of Markets: A Fractal View of Risk, Ruin, and Reward* (1. publ. paperback). New York: Basic Books.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91.

McCoy, B. M., & Wu, T. T. (1973). *The Two-Dimensional Ising Model*. Cambridge, Mass: Harvard University Press.

Miller, M. H., & Modigliani, F. (1961). Dividend Policy, Growth, and the Valuation of Shares. *The Journal of Business*, *34*(4), 411. http://doi.org/10.1086/294442

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, *29*(3), 315. http://doi.org/10.2307/1909635

Newman, M. E. J. (2010). *Networks: An Introduction*. Oxford, New York: Oxford University Press.

O'Hara, M. (2008). Bubbles: Some Perspectives (and Loose Talk) from History. *Review of Financial Studies*, *21*(1), 11–17. http://doi.org/10.1093/rfs/hhn001

Ormerod, P. (2012). *Positive Linking: How Networks Can Revolutionise the World*. London, UK: Faber and Faber.

Ozsoylev, H. N. (2005). Asset Pricing Implications of Social Networks. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.686485

Ozsoylev, H. N., & Walden, J. (2011). Asset Pricing in Large Information Networks. *Journal of Economic Theory*, *146*(6), 2252–2280. http://doi.org/10.1016/j.jet.2011.10.003

Ozsoylev, H. N., Walden, J., Yavuz, M. D., & Bildik, R. (2014). Investor Networks in the Stock Market. *Review of Financial Studies*, *27*(5), 1323–1366. http://doi.org/10.1093/rfs/hht065

Panchenko, V., Gerasymchuk, S., & Pavlov, O. V. (2013). Asset Price Dynamics with Heterogeneous Beliefs and Local Network Interactions. *Journal of Economic Dynamics and Control*, *37*(12), 2623–2642. http://doi.org/10.1016/j.jedc.2013.06.015

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Ro, S. (2013). Dividends Were Responsible For 42% Of Stock Market Returns Since 1930. Retrieved from http://www.businessinsider.com/stock-returns-price-dividend-contribution-2013-1

Rubinstein, M. (2001). Rational Markets: Yes or No? The Affirmative Case. *Financial Analysts Journal*, *57*(3), 15–29. http://doi.org/10.2469/faj.v57.n3.2447

Santos, F. C., & Pacheco, J. M. (2005). Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation. *Physical Review Letters*, *95*(9). http://doi.org/10.1103/PhysRevLett.95.098104

Scharfstein, D. S., & Stein, J. C. (1990). Herd Behavior and Investment. *The American Economic Review*, *80*, 465–479.

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., & White, D. R. (2009). Economic Networks: The New Challenges. *Science*, *325*(5939), 422. http://doi.org/10.1126/science.1173644

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, *19*(3), 425–442. http://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Shiller, R. J. (1980). *Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?* (No. w0456). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w0456.pdf

Shiller, R. J. (2003). From Efficient Markets Theory to Behavioral Finance. *Journal of Economic Perspectives*, *17*(1), 83–104. http://doi.org/10.1257/089533003321164967

Shiller, R. J. (2015). *Irrational Exuberance* (Revised and Expanded Third Edition). Princeton, NJ: Princeton University Press.

Shiller, R. J., & Pound, J. (1989). Survey Evidence on Diffusion of Interest and Information Among Investors. *Journal of Economic Behavior & Organization*, *12*(1), 47–66. http://doi.org/10.1016/0167-2681(89)90076-0

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99. http://doi.org/10.2307/1884852

Simon, H. A. (1982). *Models of Bounded Rationality: Empirically Grounded Economic Reason* (Vol. 3). Cambridge, Mass: MIT Press.

Smith, V. L., Suchanek, G. L., & Williams, A. W. (1988). Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. *Econometrica*, *56*(5), 1119. http://doi.org/10.2307/1911361

Sornette, D. (2009). *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (5th ed.). Princeton, NJ: Princeton University Press.

Sornette, D. (2014). Physics and Financial Economics (1776–2014): Puzzles, Ising and Agent-Based Models. *Reports on Progress in Physics*, *77*(6), 062001. http://doi.org/10.1088/0034-4885/77/6/062001

Sorropago, C. (2014). *Behavioral Finance and Agent Based Model: The New Evolving Discipline of Quantitative Behavioral Finance?* (DIAG Technical Reports No.

2014-13). Department of Computer, Control and Management Engineering, Universita'degli Studi di Roma" La Sapienza." Retrieved from https://ideas.repec.org/p/aeg/report/2014-13.html

Steinbacher, M., Steinbacher, M., & Steinbacher, M. (2010). Challenging the Efficient Frontier Portfolio. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.1603926

Summers, L. H. (1986). Does the Stock Market Rationally Reflect Fundamental Values? *The Journal of Finance*, *41*(3), 591–601. http://doi.org/10.1111/j.1540-6261.1986.tb04519.x

Thaler, R. (2015, July 10). Keynes's "Beauty Contest." *Financial Times*. Retrieved from http://www.ft.com/intl/cms/s/0/6149527a-25b8-11e5-bd83-71cb60e8f08c.html

Thaler, R. H. (2015). *Misbehaving: The Making of Behavioral Economics* (First edition). New York: W.W. Norton & Company.

Walden, J. (2014). Trading, Profits, and Volatility in a Dynamic Information Network Model. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.2561055

Watts, D. J. (1999). *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton, NJ: Princeton University Press.

Watts, D. J. (2002). A Simple Model of Global Cascades on Random Networks. *Proceedings of the National Academy of Sciences*, *99*(9), 5766–5771. http://doi.org/10.1073/pnas.082090499

Watts, D. J., & Strogatz, S. H. (1998). Collective Dynamics of "Small-World" Networks. *Nature*, *393*(6684), 440–442. http://doi.org/10.1038/30918

Wikipedia. (2014). Emergence. Retrieved from https://en.wikipedia.org/wiki/Emergence

Wilensky, U. (1999). *Netlogo*. Northernwestern University. Evanston, IL: Center for Connected Learning and Computer-Based Modeling. Retrieved from http://ccl.northernwestern.edu/netlogo

Xiong, W. (2013). *Bubbles, Crises, and Heterogeneous Beliefs* (No. w18905). Cambridge, MA: National Bureau of Economic Research.

BIOGRAPHY

Matthew Oldham graduated from Launceston Church Grammar School, Launceston, Tasmania, in 1991. He received his Bachelor of Economics with honors from the University of Tasmania in 1995 and obtained his Chartered Financial Analyst charter in 2012.