ARTIFICIAL FINANCIAL MARKET: AN AGENT-BASED APPROACH TO MODELING THE INFLUENCE OF TRADERS CHARACTERISTICS ON EMERGENT MARKET PHENOMENA

by

Yasaman Kamyab Hessary

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Computing and Information Systems

Charlotte

2018

Approved by:

____

Dr. Mirsad Hadzikadic

____

Dr. Mohamed Shehab

____

Dr. Moutaz Khouja

____

Dr. Samira Shaikh

ABSTRACT

YASAMAN KAMYAB HESSARY. Artificial financial market: an agent-based approach to modeling the influence of traders characteristics on emergent market phenomena. (Under the direction of DR. MIRSAD HADZIKADIC)

Financial markets play a critical role in today's societies and are extremely important to the general health and efficiency of an economy. Solid understanding of the behavior of financial markets can benefit investors, governing organizations and in general, the whole economic growth. In financial markets, psychology and sociology of the traders have significant effects in giving rise to unique and unexpected (emergent) macroscopic properties, causing the classical economic approaches hardly effective to explain or reproduce them. Agent-based modeling (ABM) is a flexible methodology for simulating such complex systems and their behaviors. The literature using agent-based approach for analyzing patterns and phenomena observed in complex systems such as financial markets has grown into an important field of research in the recent years.

This research proposes an agent-based model of stock market to study the behavior of market and traders. The model is created using a bottom-up approach, taking into consideration the effects of cognitive processes and behaviors of the traders (e.g. decision-making, interpretation of public information and learning) on the emergent phenomena of financial markets. We use Genetic Algorithm (GA) to better explore the parameters space and find the best parameter set with respect to our optimization function, which is replicating some of the important statistical properties present in actual financial time series (stylized facts). This study suggests that local interactions, rational and irrational decision-making approaches and heterogeneity, which has been incorporated into different aspects of agent design, are among the key elements in modeling financial markets. In this model, all agents fall under three general trading systems; fundamentalist, optimist, and pessimist, while having different beliefs and reaction intensities within each category.

To evaluate the effectiveness and validity of our approach, a series of statistical analysis was conducted to test the artificial data with respect to a benchmark provided by the Bank of America (BAC) stock over a sufficiently long period of time. The results revealed that the model was able to reproduce and explain some of the most important stylized facts observed in actual financial time series and was consistent with empirical observations. Using this validated model, we estimate how and in what direction different market phenomena and states such as herding, volume, market volatility, number of traders in different categories and bullish/bearish markets, influence each other. Herding is an emergent property of financial markets, often leading to the creation of speculative bubbles. Bubbles inevitably make markets unstable and prone to major crashes, hence it is crucial to understand the origins and driving forces of herd behavior. The results from model generated time series suggest that herd behavior may cause or intensify volatility in the market, but not the other way around. Further, there is strong evidence of a bi-directional causal relationship between market volatility and trading volume in both model generated and BAC time series, implying that past values of trading volume in the market can help predict the current level of volatility, and vice versa. These investigations can help enhance the understanding of dynamics and structure of financial markets and paves the way for further explorations on causes and effects of the market and traders characteristics and behavior.

Using our proposed validated framework, we design and develop the necessary infrastructure that allows us to better explore and understand the relationships between traders behavioral factors and global market properties.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my mentor, Professor Mirsad Hadzikadic, for his continues support, motivation and intellectual contribution to this dissertation. Dr. Hadzikadic provided me with a scientific environment and the freedom to try new ideas and follow new paths, which helped me grow academically and personally over these five years. I thank my committee members, Dr. Mohamed Shehab, Dr. Moutaz Khouja, and Dr. Samira Shaikh for their invaluable comments and critiques on my research and dissertation. I am also very grateful to Sandy Krause, for her endless patience and support through my PhD journey.

I am enormously thankful for my wonderful family; my parents, Faezeh Ghadimi and Manouchehr Kamyab, for their unconditional love, support and encouragement; my sister, Shokoufeh, who has been my lifelong friend and confidant; fluffy Moochi and Bameshi; and my husband and best friend, Reza Yousefian, for his endless love and support, who has always believed in me, even when I didn't.

I would like to appreciate my great friends and colleagues: Narges Tabari, Farah Tokmic, Malak Abdullah, Elizabeth von Briesen, Mona Kashiha, Kaveh Daneshvar and Mojgan Ghadimi for their encouragement, support, humor, and friendship.

Finally, I am thankful for the financial support from the Graduate Assistant Support Plan (GASP) of UNC Charlotte, which provided me the financial assistance to accomplish this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: OVERVIEW

1.1 Introduction

Solid understanding of the behavior of financial markets and its participants is extremely important to the general health and efficiency of an economy, and plays a critical role in today's societies. The classical approach to the process of analyzing such systems requires the use of experimental and theoretical models. Much of these approaches are based on the Efficient Markets Hypothesis (EMH), the notion that assumes all market participants are perfectly rational and homogeneous. EMH claims that market price movements incorporate all information rationally and instantaneously, and that past information has no impact on the future market [3]. Under this theory, markets should be efficient and stable because traders always act perfectly rational and correct the mispricing in the market almost immediately. Therefore, crashes and other market inefficiencies can only be triggered by exogenous shocks such as earthquakes or evolutions. However, careful investigation of some of the financial crises, such as the stock market crash of 1987 and the 2008 global financial crisis, have shown otherwise [4]. The dramatic price volatility during these periods cannot be explained by the arrival of significant new information. This reveals that the price changes do not always reflect rational adjustments to the news in the market, and that other factors besides information-based trading can affect stock price volatility. In addition, these theoretical models fail to explain many of the important stylized facts that are present in financial time series, such as excess volatility, positive correlations of returns at short horizons, and negative correlation of returns at long horizons [5]. These empirical observations give rise to doubts in the overall efficiency of financial markets, and lead to the greater acceptance of the importance of psychological and behavioral aspects of the economy.

1.2 Financial Markets

1.2.1 Market Efficiency

Financial markets are characterized by their high complexity. They consist of large numbers of adaptive agents, interacting locally and giving rise to emergent phenomena and macroeconomic dynamics. Traditional theoretical and experimental economic models with their top-down approach have lacked the means to properly model financial markets in its complete dynamic complexity [6]. These conventional models are based on EMH that normally makes the following assumptions:

1. *Full Rationality:* All market participants are perfectly rational and always act and make decisions in an optimum fashion.

2. *Perfect Information:* Market price reflects all information rationally and instantaneously and all market participants have full knowledge of the situation.

3. *Common Expectations:* All agents assume that everyone is exposed to same information; therefore they are all expected to react to each situation with the same approach [7].

In the complex setting of financial markets, none of these three assumptions above can be fully satisfied. As for the first assumption, real traders cannot always be fully rational. They sometimes chose to do the wrong thing, either because they are not smart enough, or because they hope to speculate in the future. The second assumption cannot stand for many cases as well, such as financial crisis of 2008 and IT bubble burst in spring 2000. These are just two examples of many that reveal price dynamic do not always reflect rational adjustments to complete information in the market. The third assumption would also fail to hold by definition since it is dependent on the first two assumptions.

The fact that EMH has a powerful theoretical and empirical basis cannot be ignored [8]. However, as it was pointed out above, various empirical observations in financial markets such as volatility clustering, speculative bubbles and even financial crisis can not be explained in terms of this theory [6].

1.2.2 Behavioral Finance

The field of behavioral finance, which views finance from a broader social science perspective, is rapidly expanding. Previous works presented in this field have provided evidence that, in addition to information, emotions play a significant role in decision-making process [9]. As [10] presented in their survey, behavioral finance argues that some financial phenomena can plausibly be understood using models in which some agents are not fully rational. The goal of behavioral finance is to improve financial modeling by considering the psychology and sociology of agents. This field considers non-standard models, in which the price movements are influenced by the actions of heterogeneous boundedly rational traders, whose financial decisions are significantly driven by emotions such as fear or greed. This type of models can explain the aforementioned stylized facts and market inefficiencies, which could not be explained by EMH.

Viewing financial market participants as mutually interacting boundedly rational adaptive agents lead to a world of complexity. Such systems often create internal structure and dynamics that give rise to unique and unexpected (emergent) macroscopic properties. Consequently, it is difficult, if not impossible, for the experimental and theoretical approaches to replicate or describe many of the financial market characteristics. This has led to alternative approaches such as Agent-based Modeling (ABM), which is the methodology followed in this project for estimating models and testing hypotheses [11]. An agent based approach models the financial markets in an incremental bottom-up fashion, as an evolving adaptive system of autonomous interacting traders. This approach makes it possible to start from very simple setting and add different variables and components as the system goes on. Therefore, the agent's behavioral rules can be set along the way according to the requirements of each given experiment. This makes it possible to analyze the influence of different characteristics on various phenomena, as Lebaron [12] says: "It is important in agent based models not just to replicate features of real markets, but also to show which aspects of the model may have lead to them."

1.2.3 Agent-based Modeling

The state-of-the-art illustrates that agent-based approach is a well-suited modeling methodology for economical and financial analysis [6, 11, 13]. ABM is not just a tool, but also a mindset that describes a complex system from its constituent entities viewpoint, stressing interaction and learning among heterogeneous agents. With ABM, market or economy can be constructed in bottom-up fashion, populated with boundedly rational agents, each having certain characteristics, information and modes of behavior like communication and learning. Through the course of simulation the agents interact and evolve along with the market, leading to emergent phenomena that can then be further analyzed [14].

ABM provides the possibility of settings the agent's rules of behavior according to the experiment at hand, making it plausible to investigate macro phenomena with respect to micro aspects of market and agent's characteristics. Beyond this, financial markets have other features that make them further appealing for agent-based modelers including,

(i) Financial data are available plentifully and accurately in different frequencies, creating a detailed image of trade dynamics and market unfolding, which researchers can take advantage of.

(ii) The availability of experimental data can be used towards calibrating agent's characteristics and behavioral rules.

(iii) There are many puzzles in financial time-series that are yet to be understood which can be possibly solved with ABM approaches [15].

1.3 Motivation and Contribution

This research focuses mainly on developing an agent based model of financial market that can reproduce and explain the most important stylized facts observed in actual financial time series. The developed validated model is then used as a platform to study the effect of different market and trader parameters on the emergent properties of the market, such as herd and volatility.

As described in Section 1.2, traditional approaches to analyze financial markets are hardly adequate to explain or reproduce many of the stylized facts observed in real markets. Moreover, the trader's behavior is not always backed up by correct information and fundamentals, an some fluctuations in stock price can be due to the impact of certain inefficient social phenomena. The motivation of the present study is to develop a framework to fulfill the following research objectives:

- Develop an agent based model of financial market that can reproduce the key stylized facts observed in real financial markets.
- Investigate that under what conditions, the endogenously produced price of the model will resemble the statistical properties of real financial price time series.
- Investigate market and trader behavioral parameters and their contribution to the emergence of stylized facts.
- Investigate the relationships between different market phenomena and parameters such as herding, investment return, volatility, volume, number of different trader types and bullish/bearish markets.

In this thesis, an agent-based model of stock market is proposed to study the effects of trader's psychological and behavioral characteristics on the overall properties of financial markets. Two key features of this model are local interactions and heterogeneity, which allow for global market protocols and behavioral norms to arise from the bottom up. The model stresses interaction and learning between two general types of agents: fundamentalist traders and chartist traders. The behavior of the market as a whole, such as the dynamics of asset prices, is an emergent property of the agents' behavior. This study suggests that local interactions, rational and irrational decision-making approaches and heterogeneity, which has been incorporated into different aspects of agent design, are among the key elements in modeling financial markets. We use this validated model as a platform to investigate and understand the relationships between traders behavioral factors and global market properties, such as herding and volatility.

This research contributes to agent-base modeling of financial markets by designing a very simple market mechanism with non-simple agents, to study the co-evolution of agents and their trading behavior, and their consequences on the changes of the market as a whole. The proposed model and its mechanisms are simple enough to avoid the complications that would prevent us from understanding the critical origins of the reproduced statistical properties of real financial markets. Yet, the agents settings posses enough heterogeneity and sophistication to allow for co-evolution of the trading rules and reflect agents' strategic behaviours on the price changes of the market. As the number of parameters increases, the size of the search space rises enormously, since one needs to explore the space of parameters as well as thresholds to investigate their effects on market macro properties. To overcome this limitation, we tuned and calibrated the model using a standard Genetic Algorithm. By using such technique, we not only explore the search space more efficiently, but also improve aggregate dynamics of the market and the micro behavior of agents by means of increasing the degree of autonomy.

This research also contributes to the body of literature that studies herding behavior, as it uniquely investigates the relationships of different herding measures with other market parameters and properties, and advances the understanding of the phenomenon.

The market mechanism, the parameters of agent's behaviour and the analysis of the results will be described in detail in the following chapters.

1.4 Dissertation Organization

The dissertation is organized as follows:

Chapter 2 is devoted to the literature review of research in agent-based modeling of financial markets. First, it reviews literature on market mechanisms and agent design. In the second part, the herd related studies are investigates to understand the nature of the phenomenon. Finally, it explains the means of validation and provides a comprehensive list of stylized facts used in validation process.

Chapter 3 provides the description of the computational model and its specifications.

Chapter 4 focuses on calibration and validation of the model by the means of Genetic Algorithm, and provides the results on reproducing the stylized facts.

Chapter 5 presents the statistical analysis of the model's results and BAC time series, along with the investigation of relationships between different market phenomena.

Chapter 6 summarizes and concludes the thesis.

CHAPTER 2: LITERATURE REVIEW

2.1 Agent-Based Models of Financial Markets

In order to start modeling the market and economy, one should start with answering the design questions such as: What are the types of securities to be traded? What would be the fundamental value of the asset and how would it move? How and to what degree heterogeneity, learning and interactions are considered in the model? These are some of the large number of design questions that we try to attend in the following sections.

2.1.1 Agent Design

In try to categorize different agent-based models, it is helpful to classify them based on their agent's complexity in terms of three basic elements: heterogeneity, interaction and learning. Of course there are other taxonomies, which can be found in the works of [11], [5], and [15]. The complexity of agents can go from simplest settings with only two types of agents learning with imitation and reinforcement algorithms all the way to many type agents with complex learning algorithms such as genetic programming and neural networks. Similarly, as for interactions, it can range from simply interacting with neighbors to complex network typologies [2].

In actual financial markets, financial actors and traders can differ in so many aspects. They can have different levels of sophistication, exposure to public and private information, strategies, beliefs and so on. To decide on the level of heterogeneity that would serve the goal of artificial financial market is one of the important steps of such modeling. Obviously, keeping the model as simple as possible would make the validation process less complicated and allows for tractability of the model's parameters. On the other hand, it is not an easy choice to decide on keeping, cutting or simplifying different aspects of agents. A good practice can start from a simple setting and incrementally add necessary parameters and components to the model until reaching a certain goal.

2.1.1.1 Two-Type Design

This type of agent design is the simplest kind with respect to heterogeneity. Vast empirical investigation was done in 1990s by [16] and [17] on the behavior of financial agents. The data was gathered through different surveys and the result of these studies showed that in general, there are two different types of trading strategies that financial agents follow, fundamental and technical. *Fundamentalists* and *chartists* (technical, trend-follower, or noisy traders) have very different views on price dynamics. The former make decisions with the belief that the price of an asset return to their fundamental value in the long run, while the latter are mostly concerned about the trends and patterns observed in the past prices. Fundamentalists make the market stable by forcing the price towards the fundamental value while chartists are the destabilizing force in the market by causing positive feedback by extrapolating [5].

Evidence of the existence of these two types of traders in real markets and models can also be traced back to the fifties [18]. Later on, a number of literatures including [19] argued that chartist, due to their irrational beliefs, would eventually lose all their money and vanish from the market. Despite the popularity of this idea, some authors like [20] nearly immediately presented counterexamples to it. It is safe to say that the dollar bubble from the eighties and technology bubble which burst in 2000 have proved that not all financial agents are fully rational in real markets [11]. It is also noted that the market fraction, which is the proportion of fundamentalists and chartists, changes over time, indicating the adaptive features of financial agents [2]. These empirical discoveries created a stepping-stone for design and development of artificial financial agents.

The model with only these two types of traders is the simplest type of Heterogeneous Agent Models (HAMs). Zeeman's model [21] is one of the first HAMs of stock market with fundamentalists and chartists traders. He proposed a qualitative description of the stylized facts observed in short-term bull and bear markets. In the model, fundamentalists are assumed to know the intrinsic value of the stock and they only buy (sell) when the observed price is below (above) that value. Chartists, on the other hand, believe that prices move in trends so they buy when price rises and sell when it falls. He, based on behavioral assumptions of chartists and fundamentalists, explains the switching between bull and bear market [5]. Zeeman's model was a pioneer in financial market models with fundamentalists and chartists, and a large number of the literature was developed following the lead of him, e.g. [22], [23], and [24].

In the two-typed models, all agents pick among two behavioral rules at the start of each trading period and later on, in the next period they review and possibly switch rules based on a pre-defined fitness function. Switching mechanism, which is the *binary choice model*, captures all the interaction and learning in the model. At time t, the forecasting mechanism that each type uses to decide on the next action is based on its expectation of the price at time $t+1$, which can be written as [2]:

$$E_t^f[p_{t+1}] = p_t + \alpha^f(p_t^f - p_t) \tag{2.1}$$

$$E_t^c[p_{t+1}] = p_t + \alpha^c(p_t - p_{t-1}) \tag{2.2}$$

where, (2.1) and (2.2) are the forecast rule of fundamentalists and chartists, respectively, p_t is the current price, $\alpha^f \in [0,1]$, $\alpha^c \in [0,1]$, and p_t^f is the fundamental value of the asset. The second parts of the equations are formalization of orders generated by each trading rules:

$$D_t^f = \alpha^f(p_t^f - p_t) \tag{2.3}$$

$$D_t^c = \alpha^c(p_t - p_{t-1}) \tag{2.4}$$

where, α^f is a positive coefficient that describes fundamentalist mean-reverting belief and its basically specifying the speed with which fundamentalists expect the price to return to fundamental value. Fundamentalists would change their minds about an asset if the price departs-more than some threshold value- from the fundamental value. The threshold can be randomly generated for each fundamentalist in order to add further heterogeneity to the model and make each fundamentalist unique. In the special case $\alpha^f = 1$ they expect the price to return to its fundamental value immediately, while in the case of $\alpha^f = 0$ they naively expect the price to follow a random walk.

Furthermore, α^c is a reaction coefficient, conveying the sensitivity of chartists to price change and $p_t - p_{t-1}$ indicates the trend. Chartists extrapolate and predict the price change rate to be proportional to the latest observed change. Some literature (see, e.g., [25]) include a random term (a normal, Independent and Identically Distributed (IID) noise process with zero mean and constant standard deviation) in the equations to capture the diversity and uncontrollable elements in technical and fundamental analysis.

In the models discussed of fundamentalists and chartists, none of the two trader types can be entirely rational, as the Friedman hypothesis suggests, fully rational traders would eventually out-perform all other types of traders and make them vanish from the market. In this case, where all agents are rational and aware that all the others are rational as well, no trade will take place (see [26] and [27]). In this market, traders cannot benefit from their information, take a case for example that a traders wants to buy an asset, all the other traders anticipate that he has the positive information and would not sell the asset to him. As it can be seen in real markets, high volume of daily trading is in contrast with this no trade theory.

2.1.1.2 Three-Type Design

Three-type design is an extension of two-type design, adding to the heterogeneity of the agents. With the risen capability of observing the behavior of financial agents due to large amount of proprietary trading data, it can be confirmed that chartist, either individuals or institutions, differ in the way they react to past price and trends. Empirical investigations suggest that chartist can be further divided into optimists and pessimists, or as in most literature, momentum and contrarian traders. Momentum traders are the same as previously defined chartist, whereas contrarians act completely opposite to that. They still look at the past movements of the price to predict the future, but act against the trend, anticipating that the trend will soon be finish and reversed. Here is the formulation of their expectation of next period's price [2]:

$$E_t^{co}(p_{t+1}) = p_t + \alpha^{co}(p_t - p_{t-1}) \tag{2.5}$$

where $\alpha^{co} \leq 0$.

This class of models is proven to be very successful at reproducing many of the stylized facts observed in financial markets. [28] and [1] developed models with these three types of agents, either an optimistic or a pessimistic chartist or fundamentalist. Traders can switch between chartist and fundamentalist and within chartist, between optimistic and pessimistic strategies and the price is determined based on aggregated excess demand. With their approach, the appeal of chartist rule depends on achieved profits, while for fundamentalist it depends on the expected future profit opportunities.

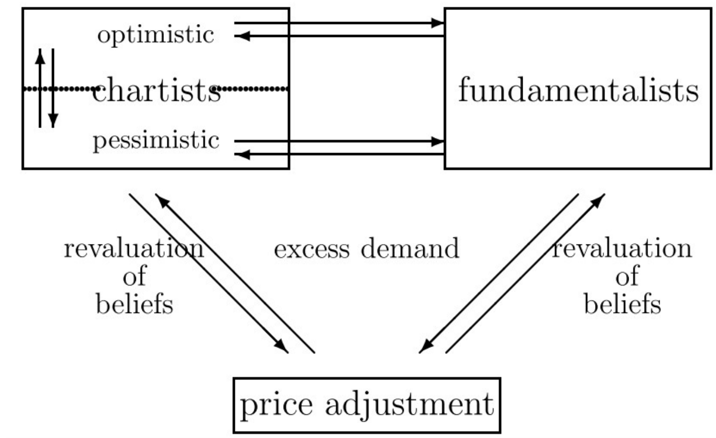


Figure 2.1: Switching dynamics of Lux-Marchesi model [1].

2.1.1.3 Fitness Function and Switching

Switching or adaptive belief dynamics that enables agents to shift between different strategy alternatives is considered to be a necessary source for creation of many stylized facts observed in the real markets ([29], [30], and [1]). In the N-type models, learning is encapsulated in the switching method. The fitness measure that drives agents to change their opinion and switch to a new trading strategy can be temporal realized profits from each rule or relative accuracy or pricing error following each rule [31]. Although, in the case of realized profit, [32] has critiqued that there maybe a bias that pushes traders towards chartist rule since chartists are driven by realized profit, while fundamentalists are driven by expected arbitrage profits that will not be recognized until the price has returned to its fundamental value.

In the two-type fundamentalist/chartist model, agents can only choose between two rules, therefore modeling the switch gets quite easy and is typically done using a the binary choice model, particularly, the logit model ([33]) which is a well-known behavioral optimization model. Consider having two alternative options (f (fundamentalist), c (chartist)), each having some random profit, to choose from. In this case the logit model assumes that the probability of choosing one of the options is equal to the probability that the gained profit from the chosen option is larger than those gained from choosing the other one. The following binary logit model can be derived under the assumption of random profit [2] as,

$$P(X = f, t) = \frac{\exp(\lambda V_{t-1}^{f})}{\exp(\lambda V_{t-1}^{f}) + \exp(\lambda V_{t-1}^{f})} \tag{2.6}$$

where, V_t^f and V_t^c are the deterministic components of the profits from the options f and c at time t, respectively, and the parameter λ is the intensity of choice, measuring how fast agents switch between the two strategies ad it is inversely correlated to the level of random utility. If $\lambda = 0$ no switching is going to take place between strategies, while all agents instantaneously switch to the best strategy if $\lambda = 1$.

In the model of [1], as shown in Fig. 1, agents are able to switch between chartist and fundamentalist strategy as well as switching between optimist and pessimist within the chartists. In the case of chartist, majority and current price trend plays the key role in the switching process. For example the probability that pessimistic turn into optimistic rise if the majority of chartists are optimistic and prices are going up. The number of agents in each category determines excess demand (the difference between demand and supply) and lead to price changes, which in turn affect agents' choices of strategies [11].

2.1.1.4 N-type Design

The two and three type models can be generalized to more heterogeneous agent setting by, for one, introducing a memory parameter to chartist equation. Therefore, instead of just looking at the last time step price, they have the option to look further into the past in order to choose their next move. This way the psychological bias, law of small numbers, is abided. It basically says that people put much more weight on recent events as oppose to long-term averages. For example if returns on an asset stays high for many years, traders having this bias would consider it as "normal". The expectation for momentum and contrarian traders can be formulated as follow [2]:

$$E_t^c(p_{t+1}) = p_t + \alpha^c(1-\beta^c)\sum_{k=0}^{T}\frac{(\beta^c)^k(p_{t-k}-p_{t-k-1})}{\sum_{k=0}^{T}(\beta^c)^k} \quad (2.7)$$

$$E_t^{co}(p_{t+1}) = p_t + \alpha^{co}(1-\beta^{co})\sum_{k=0}^{T}\frac{(\beta^{co})^k(p_{t-k}-p_{t-k-1})}{\sum_{k=0}^{T}(\beta^{co})^k} \quad (2.8)$$

where, β^c and $\beta^{co} \in [0,1]$. These two are the memory parameters of pessimist and optimist chartists and as they increase, the weight they put on the price changes farther in the past increases.

The memory parameter along with different thresholds and reverting and extrapolating coefficient for each of fundamentalists and chartists introduce further heterogeneity into the model and makes it an N-type design. However, although two chartists or fundamentalist can have different views and actions, but eventually the behavior of the agents would fall into one of the few vitally different groups.

2.1.1.5 Many-Type Design

As popular and useful N-type agent designs are in many cases, it still holds the restrictions of exogenously given rules. They can only choose their actions based on what is already in the system and are unable to follow new rules unless it is hard-codded into the model. The many type-design introduces new approach that lets artificial agents become more real by learning and creating new strategies on their own. One of the interesting characteristics of many of these models is that agents start with very homogeneous features and evolve in their features and rules endogenously over time [34]. The Santa Fe Artificial Stock Market (SFI-ASM) [35] is a pioneer and most influential work in this area. The may-type design approach uses Artificial Intelligence techniques to model evolution and learning, which are outside the scope of this research.

2.1.2 Market Maker

Many different methods are used in literature for determining prices in artificial financial markets. in this research, the category that uses a slow price adjustment process is investigated. In this category the market is mostly never in equilibrium and the price of an asset basically depends on excess demand, which is the difference between demand and supply. It rises if more agents are willing to buy instead of sell and falls if the supply exceeds the demand. An early example of such approach is [22], where there is a market maker who broadcasts the price and agents submit their buy and sell orders for this price. The orders are then summed and in the case of existence of excess demand (supply), the price is increased (decreased):

$$p_{t+1} = p_t + \alpha(D(p_t) - S(p_t)) \tag{2.9}$$

where, α is a positive coefficient, which can be explained as speed of adjustment of price, $D(p_t)$ is the number of buy and $S(p_t)$ is the number of sell orders. A random term (a normal, IID noise process with zero mean and constant standard deviation) can also be added to the equation to account for more variety and unknown facts contributing to price change, since this model only provides a simple representation of real financial markets.

2.2 Herding

Herd happens when traders observe the choices of others and, regardless of their private information and consequences of their choices; start to imitate them [36]. Suppose that there are 10 traders in the market, 2 optimists and 8 pessimists. The group of 2 believes it is in their best interest to buy an asset and the other 8 think that it is not. If these 2 optimists act first and buy, the other 80% will observe their action and may change their opinions and also decide to buy. This could eventually lead to all 10 traders to invest. As it can be seen in the above example, what will majority do can be based on the actions of the early traders and they may be most crucial in determining the direction of herd, which may well be incorrect.

The incentive behind an investor's decision to discard his own beliefs and information after observing others vary by different factors. Everyone in the market is trying to maximize their profit or utility and they may anticipate that others know something about the return on an investment that they do not. An investor may think that the action of others reveals true information about the value of an asset, so they start imitating them. There could be less sophisticated reasons behind herding such as the intrinsic need of individuals to match their beliefs and behaviors to group norms. Another reason which only relates to money managers who invest on behalf of others is that sometimes the compensation system or terms of employment may encourage and implicitly reward imitation [37].

It is worth injecting a note of caution that one should not confuse the "spurious herding" with true herd behavior that is, the decision to disregard one's private information to follow the behavior of others (see, [38] and [39]). Spurious herding happens when groups of traders face the same problems and information sets and therefore make similar decisions [37]. In this case even though everyone acts the same, their choice is based on fundamentals and will have an efficient outcome, as opposed to true herding that can cause information inefficiency in the market due to misalignment between the price we observe and the price we would have observed in the absence of herding [40]. In this research, the spurious herding is not going to be investigated, even though it is very difficult, if not impossible, to differentiate between true and spurious herding due to unknown motive behind a trade and lack of data on the private information available to traders. However, the reactions to public information can be distinguished by explicitly allowing for changes in fundamentals and then factoring out the effects. If collective behavior can still be seen in the data, then true herding may well be the cause [37].

2.2.1 Agent Based Models of Herding

2.2.1.1 Sequential vs. Non-Sequential Herding

There are various approaches towards modeling herd behavior in the literature. Some models like [41] and [42] attempt to model herding with a sequential characteristic, in which traders make decisions one at a time and consider the actions of the traders before them in their decision making process. On the other hand, [43] introduced a non-sequential model of herding, where agents interact simultaneously while modifying their decisions at each time step. He studies the Bayesian equilibria resulting from identical agents with the same imitation tendency who make binary decisions. By setting the imitation rate low, the model leads to a Gaussian distribution whereas strong imitation setting leads to a bimodal distribution with nonzero modes that he interprets as equivalent to crashes in market.

The decision structure of non-sequential models seems to be more realistic since traders participate in the market simultaneously and cumulative market variables like asset prices is determined by the aggregation of different orders. Although non-sequential setting is better suited for financial markets, neither of these two approaches can reproduce heavy tail of returns as such observed in empirical properties of stock returns. [44] have proposed a different approach that is non-sequential but avoids Orlean's impractical results by revising his assumption of all agents having the same imitation rate. In this methodology agents interact on a basis of random communication network, which leads to formation of groups of agent such that each group makes independent decisions while agents within each group mimic the behavior of each other.

2.3 Bubbles and Crashes

An overview of the relevant literature reveals that herding can cause information inefficiency and lead to speculative bubbles, which makes the market unstable and prone to major crashes. As Sornette argues in "Why Stock Markets Crash" [45], crashes are not solely the result of rational adjustments to news in the market, nor are always triggered by exogenous shocks such as earthquakes or evolutions alone. Sometimes, financial crisis happens due to the emergent instability in the market that is a result of escalation in herd behavior. Take, for example, the crash of 1929. For months preceding the crash, there was an increasing interest for commodities such as stocks, coins and diamonds and a bubble was forming throughout the society. Also, the election of Herbert Hoover as the president of the United States in that time period produced an overall optimistic mood and as a result, the stock purchase had the greatest increase up till then. After the boom of 1929, the investigations revealed the existence of bubble and the role of herd behavior in the crash. The results have shown that the actions of a large number of investors, who blindly followed the popular stocks, caused the prices for those favorite stocks to go way higher than their fundamental values. This had led to the creation of a bubble, subsequent volatility in the market and finally the great crash [46].

[47] presented a model of stock market in which traders have incomplete information about the price and they are aware that others may have different, incomplete information as well. Knowing that their information set is incomplete, they are reluctant to completely rely on the current value which, differs from the fundamental. Therefore, they might try and adjust their expectations to the opinion of some other investor who may have information which is not available to them. This creates the "fad" characteristic in the model. The results of the simulation explain that if agents make strong-form rational expectations, there will be no herd behavior observed in the market and the market price resembles the fundamental value. On the other hand, in the case of weak-form rational expectations, contagion could form in the market and price diverges from its fundamental value. The contagious behavior then leads to creation of speculative bubble and excess volatility.

There are other related literature that try to explain the creation of opinion clusters and bubbles in terms of herding parameter, in particular the paper of [48]. They propose a self-organized model for stochastic formation of opinion clusters. There is only one parameter in this model which is h, the rate of information diffusion per trade, and they call it the measure of herding behavior. The model consists of a network in which agents are vertices, and links between agents indicate they share the same information. Each agent can have three states, $\Phi = \{-1, 0, +1\}$. $\Phi = 0$ represents inactive, waiting state and $\Phi = 1$ and $\Phi = -1$ corresponds to active buying and selling states, respectively. At the beginning all the agents are inactive and at each time step, an agent is randomly selected. With probability a the state of the chosen agent becomes active to either 1 or -1 randomly. After this step, all the agents that belong to the same cluster follow the same action and then the cumulative state of the system and total size of the clusters are computed. Afterwards, all links inside the cluster are removed and all states change to inactive and, with the probability of $1 - a$, the chosen agent remains inactive and a new link between it and another random agent gets created and this process repeats. The connectivity of the network grows as agents remain inactive and decays whenever an activation happens. The herding parameter is defined as $h = 1/a - 1$ and $a \in [0, 1]$ controls the rate of trading activity versus information dispersion. As a grows closer to 1, the agents get more isolated and no herding behavior can be observed in the market. On the other hand, as a gets closer to 0, h increases and subsequently, the connectivity of the network rises until eventually, all the agents merge into a super-cluster close to the system size. This could lead to a relative increase in the probability of extremely high returns, which may end with a crash. Regarding these results, [48] conclude that herding can be held responsible for the occurrence of crashes.

2.4 Validation and Stylized Facts of Financial Time Series

Time series of financial markets exhibit many statistical features that are mostly common to a wide range of markets and time periods [49]. These properties are known as "stylize facts" and modelers of financial markets attempt to match the statistical properties of the artificial data produced from their models with those properties in order to make the model verifiable. The conditions under which stylized facts arise in the agent base models of financial markets provide an strong way to validate the model and confirm which factors matter in the simulation. For this purpose, econometric is used to examine the data generated from the models and investigate that if they are able to exhibit, grow and replicate a number of frequently observed empirical features of financial markets (stylized facts).

Some of the most universally accepted stylized facts are as follows:

- Fat tail of the returns: The most commonly used stylized facts in the validation process of financial market models by claiming that the distribution of high frequency returns exhibits a heavy tail with positive excess kurtosis. It has been first discovered by Mandelbrot [50] that there is higher density on the tails of the distribution in comparison to the tails' density under the normal distribution. In a normal distribution, the excess kurtosis is zero while in financial data it is usually greater than zero, which is an indication of fat tails. In order to determine the density of the tail and the tail index, one can also use the Hill tail index [51], which is a simple and efficient tool to the study the tail behavior of financial time series.

- Non normality: The stock returns on different frequencies are not normally distributed.

- Volatility clustering: Which describes the fact that high-volatile events tend to cluster in time. Empirical analysis on financial markets reveals that different measures of volatility display a positive autocorrelation over several days [49]. So if the market had been very volatile today, the chances are higher than average that it will have large positive or negative changes tomorrow.

- The absence of autocorrelation in raw returns
- The presence of autocorrelation in absolute returns (volatility clustering).

A comprehensive list of stylized facts is provided in table 1. From the total of 30 facts listed in the table, the highlighted codes represent the stylized facts of low-frequency financial time series of return and trading volume. The other codes refer to high-frequency time series of return, trading duration, transaction size, and bid-ask spread.

Table 2.1: Stylized Facts [2]

No.	Code	Stylized Facts	Reference
16	EPP	Equity Premium Puzzle	[52]
9	BC	Bubbles and Crashes	[53]
29	TLS	Thinness and Large Spread	[54]
30	TD	Turn-of-the-year Declining	
1	AA	Absence of Autocorrelations	[49]
2	AG	Aggregational Gaussianity	
3	CHT	Conditional Heavy Tails	
4	FT	Fat Tails	
5	LE	Leverage Effect	
6	LM	Long Memory	
7	VC	Volatility Clustering	
8	GLA	Gain/Loss Asymmetry	
19	PLBR	Power Law Behavior of Return	[55]
20	PLBTV	Power Law Behavior of Trading Volume	
21	PLBT	Power Law Behavior of Trades	
10	CE	Calendar Effect	[56]
11	AA-H	Absence of Autocorrelations	
12	FT-H	Fat Tails of Return Distribution	
13	LM-H	Long Memory	
14	PE	Periodic Effect	
15	BU	Bursts	
27	US	U Shape	[57]
28	SCPC	Spread Correlated with Price Change	
17	EV	Excess Volatility	[58]
18	VVC	Volatility Volume Correlations	
22	PLBV	Power Law Behavior of Volatility	[59]
24	CTD	Clustering of Trade Duration	[60]
25	DLM	Long Memory	
26	DO	Overdispersed	
23	VLM	Long Memory of Volume	[61]

CHAPTER 3: THE ARTIFICIAL FINANCIAL MARKET

In this chapter, the developed dynamic, heterogeneous stock market model and its main characteristics are presented and explained.

3.1 Introduction

In this work, the two-typed design methodology to model the stock market is followed. The two typed design is chosen to allow for a less complicated validation process and better tractability of the model's parameters. Zeeman's model [21] was one of the first agent-based models of stock market with fundamentalist and chartist traders, and a large number of literature was developed following the lead of him. He proposed a qualitative description of the stylized facts observed in short-term bull and bear markets. Although the model included several behavioral elements that are used as base for other financial market modelings, it lacked micro details. The study done by Gilli and Winker [62] showed that to have a better characterization of financial markets models, traders should be able to switch between different trading strategies than to simply assuming that the ratio of fundamentalists over chartists remains constant over time. The approach of stressing evolution and switching between fundamentalists and chartists, taken by Hommes [23], Lebaron [15] and Westerhoff [63], has proven to be quite successful. Although being able to reproduce some of the most important stylized facts of financial markets, they do not consider local interactions and heterogeneity within the two general belief systems.

The aim of this study is to investigate the effects of rational and irrational decision-making process and social interaction on overall market dynamics and the emergent of certain key stylized facts. In order to achieve this goal, a simple yet rich and flexible agent-based model of stock market is developed. In a first attempt to model this complex system, the methodology of two-type design suggested by Westerhoff [63] is followed. This method, being the simplest kind of heterogeneity, allows for a better tractability of the model's parameters and a less complicated validation process. However, in order to gain additional explanation power and a closer representation of real financial markets, further heterogeneity is incorporated as the model proceeds. The objective is to satisfy the two essential agent design principals, simplicity and heterogeneity by a well-defined scheme and parameter space.

The proposed model differs from the original model and other related works in the approach that is taken to model heterogeneity, interaction and learning behavior of the agents. Related research limits heterogeneity through only a few elements of two general agent types, fundamentalists and chartists. They also capture learning and interaction through switching mechanism, in that the overall population of fundamentalists and chartists is set due to the realized profit associated with their forecasting rules. On the other hand, the novelty of the approach used in this work is that, while consistent with other two-type design frameworks, it builds heterogeneity into different aspects of agent design, making each fundamentalist and chartist unique within the two general types. Also, the method to model learning and evolution is through local interactions. At each time step, financial agents exchange information with their surrounding neighbors. They either compare strategy profits and change their tactic accordingly, or blindly choose to follow the most popular trading decision in their neighborhood, regardless of its utility and their own private information. This method helps gaining a broader global insight to trading strategies and evolution of the agents.

In this work, heterogeneity is modeled through several key parameters:

- Agents with different reaction intensities to price and fundamentals changes.

- The initial wealth distribution among the agents.

- Agents with different levels of sophistication in their decision-making process.
- Agents with different memories who use different trading time windows.

3.2 Overview of the Model

The model describes daily stock trading with only one risky asset of the price p_t at time t. There is a fixed number of N traders in the market, scattered randomly across a pseudo-landscape grid of square cells. Each trader can interact with the traders on its eight neighboring patches. At the beginning of each simulation, all traders are endowed with different amount of cash drawn from a power law distribution. At each time step t, traders can choose between three actions, buy or sell one unit of the stock, or remain inactive. Each Trader is assigned one of the three general belief systems, optimist, pessimist or fundamental trading rules randomly at the beginning of simulation. Afterwards, at each time step traders can interact with the traders on its eight neighboring patches and change their strategy to maximize their profit or simply follow the crowd. This switching mechanism captures the interaction and learning in the model. The agent's action space is:

$$Action = \begin{cases} 1 & \text{Buy one unit of stock} \\ -1 & \text{Sell one unit of stock} \\ 0 & \text{No trade} \end{cases}$$

The overall flowchart of the model is presented in Figure 3.1, which will be discussed in details in this section.

3.3 Model Mechanism

3.3.1 Wealth Distribution

It was first observed by Pareto [64] that the income distribution across several countries follows a power law. Later, it was discovered that wealth is also distributed

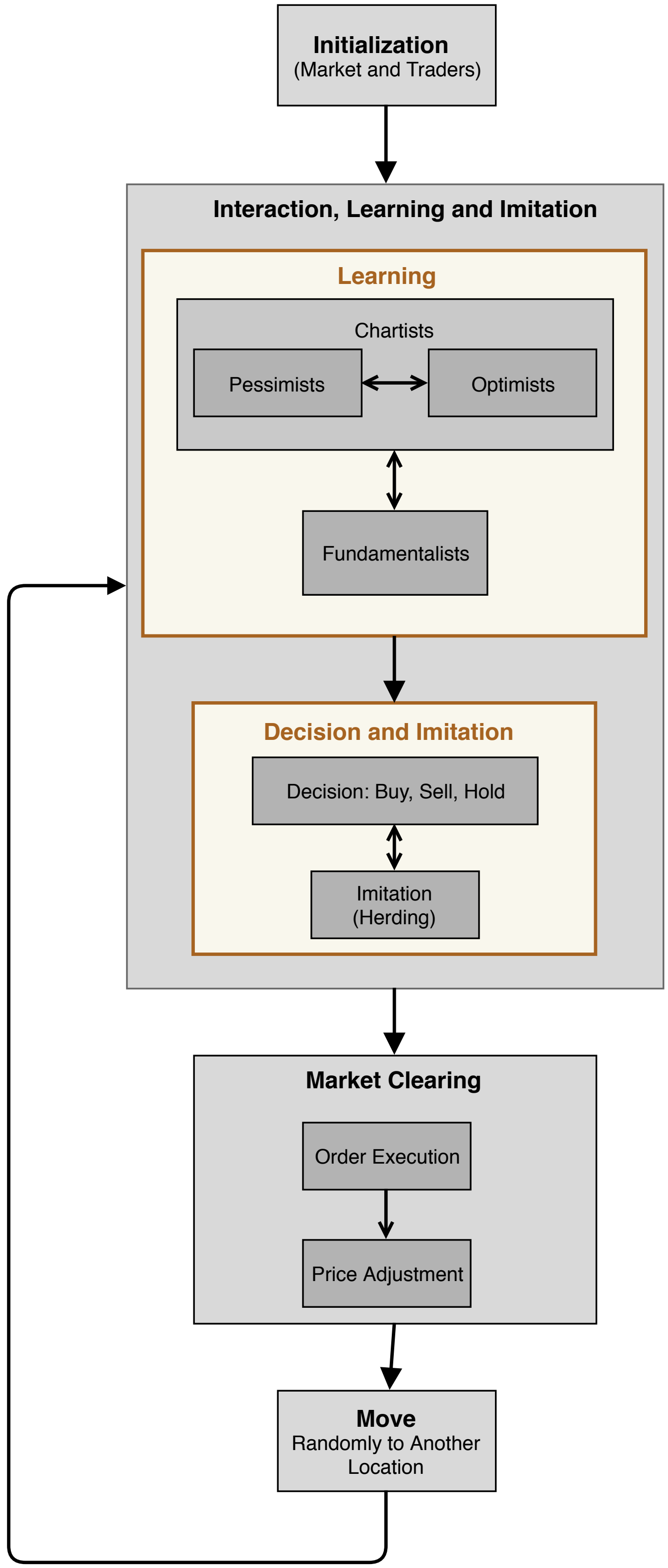


Figure 3.1: Model evolution flowchart.

according to a power law [65], which implies rather extreme wealth inequality. The probability density function that describes the Pareto distribution is of the form:

$$W(x) = x^{-(1+\alpha)} \tag{3.1}$$

where, α is the Pareto exponent, which measures the level of wealth inequality and was estimated by him as $\alpha = 1.5$. In this model, the same power law distribution is used to initially endow each agent with a different amount of money, in agreement with human population.

3.3.2 Traders

3.3.2.1 Fundamentalist

Fundamentalists make decisions with the assumption that in the long run, price of an asset returns to its fundamental value. In this model, fundamentalists are assumed to know the intrinsic value of the stock and they only buy (sell) when the observed price is below (above) that value.

The fundamental value of the asset F_t is publicly available to all traders. It is assumed that the fundamental price of the risky asset is given exogenously as a random walk process:

$$F_t = F_{t-1} + \eta \tag{3.2}$$

where, F_{t-1} suggests that yesterday's value of stock is carried into the future and η is a IID noise process with zero mean and constant standard deviation σ_η, added to take the fundamental shocks into account.

At the beginning of the simulation, a unique coefficient $(0 < f < 1)$ drawn from a Gaussian distribution is assigned to every fundamentalist. This element characterizes the belief of each fundamentalist and specifies the speed with which each agent expects the price to return to its fundamental value. By providing traders with different coefficients, every fundamentalist becomes unique in their beliefs and decision-making process, as practiced in real world markets. In the special case of $f = 1$, they expect the price to return to its fundamental value immediately, while $f = 0$ means that they naively anticipate that the price follows a random walk. With this being said, their expectation of the next period price can be written as:

$$E_t^f[p_{t+1}] = p_t + f * (F_t - p_t) + \tau \tag{3.3}$$

where, p_t is the current price. A random variable τ (a normal, IID noise process with zero mean and constant standard deviation σ_τ) is included in the equation to account for diversity and uncontrollable elements.

Furthermore, granted that fundamentalists trade less frequently than chartists, a threshold K is introduced. Fundamentalists become active only if the difference between the current price and the price they expect for the next period is above such threshold.

3.3.2.2 Chartist

Chartists, on the other hand, base their expectations on the past price changes, patterns and trends. An empirical study done by [66] suggests that chartist traders have varying investment horizons and thus use different time windows to analyze the trend and patterns in the price time series. Here, each agent is assigned a different memory length T, drawn randomly from a uniform distribution. The estimated maximum investment horizon used here is 50 days. By having different investment horizons in the market, the model not only conforms to real-world market traders, but also has an increased autonomy degree which will improve the macro dynamics of the artificial market.

In this model, following the methodology used by [1], two types of chartist are introduced; optimists and pessimists. Optimists extrapolate and predict the price change rate to be proportional to the latest observed pattern in the price history. Whereas pessimists, while still looking at the past movements of the price, act against the trend, anticipating that the trend will soon be finished and reversed. The formulation of their expectation of next periods price is as follow:

$$E_t^{co}[p_{t+1}] = p_t + c_o * (p_t - MA_t) + \beta_1 \tag{3.4}$$

$$E_t^{cp}[p_{t+1}] = p_t + c_p * (p_t - MA_t) + \beta_2 \tag{3.5}$$

where, $0 < c_o < 1$ and $c_p \leq 0$ are the reaction coefficients of optimists and pessimists respectively, conveying the sensitivity of chartists to price changes and trends. These two parameters are different for each chartist and are drawn from a Gaussian distribution at the beginning of simulation. β_1 and β_2 are random normal IID noise processes (with zero means and constant standard deviations σ_{β_1} and σ_{β_2}), added to capture the diversity and uncontrollable elements in chartist analysis. $p_t - MA_t$ indicates the trend, where MA_t is the moving average of the past prices, based on an an Exponentially Weighted Moving Average (EWMA) process:

$$MA_t = \phi \sum_{i=1}^{T} (1 - \phi)^{i-1} p_{t-i} \tag{3.6}$$

where, $0 < \phi < 1$ is a smoothing parameter and T is the memory length of the agent. Each chartist is assigned a unique memory length (T) drawn randomly from a uniform distribution between one and fifty days. Using this formulation, the psychological bias of law of small numbers is abided, which implies that people put much more weight on recent events as oppose to long-term averages [2]. As ϕ increases, the weight given to the recent prices further increases. It should be noted that, the weight $\phi * (1 - \phi)^{i-1}$ in (3.6) controls the MA_t and normalizes the average value. The weight of past prices fades by $(1 - \phi)$ at each time step, making the total area of the weight sequence upper bounded by 1. Meanwhile, the parameter ϕ is chosen with respect to T in order to make the area bellow the weights close to 1. Hence, the sign of trend value $(p_t - MA_t)$ is merely based on the current and previous values of p and the weights chartists put on past prices.

3.3.3 Price Adjustment

At the end of each trading day, the market maker sets the price according to observed excess demand. Following Day and Huang [22], a simple price adjustment scheme based on the aggregated excess demand is used as,

$$p_{t+1} = p_t * (1 + a * (D_t - S_t)) + \delta \tag{3.7}$$

where, a is a positive coefficient, which can be explained as the speed of price adjustment. D_t and S_t are the number of buy and sell orders at time t, respectively. Since this model only provides a simple representation of real financial markets, a random term, δ (a normal, IID noise process with zero mean and constant standard deviation σ_δ) is also added to the equation to account for unknown facts, contributing to price change.

3.4 Social Interaction, Adaptation and Learning

Local interactions and learning is a key factor in the proposed model and the simulations. The macro dynamics of the market as a whole is the result of endogenous switching between there groups of agents defined in the previous section. Switching or adaptive belief dynamics that enables agents to shift between different strategies is considered to be a necessary source for creation of many stylized facts observed in the real financial markets [1].

3.4.1 Logit Switching Models

In the N-type models, learning is encapsulated in the switching method. The fitness measure that drives agents to change their opinion and switch to a new trading strategy is considered to be the temporal realized profits from each trader's actions, which is formulated as follow:

$$S_t = \begin{cases} (1-m)*(p_{t+1}-p_t)+m*S_{t-1}, & \text{if agent was buyer} \\ (1-m)*(p_t-p_{t+1})+m*S_{t-1}, & \text{if agent was seller} \end{cases}$$

where, $m \in [0,1]$ is the normalization of memory parameter T, and is used to account for the past performance of each rule.

In this model, agents are able to switch between chartist and fundamentalist strategy as well as switching between optimist and pessimist within the chartist group. Since in the latter case agents can only choose between two rules, the switching process is modeled using the Logit binary choice model [33]. At the beginning of each trading day, agents meet and exchange information with their neighbors. Given the uniqueness of each fundamentalist,optimist and pessimist, the average utility associated with each type in the neighborhood is calculated and compared. The probability that an agent chooses optimist over pessimist trading rule is:

$$P(X=o) = \frac{\exp[\lambda \bar{S}^o]}{\exp[\lambda \bar{S}^o] + \exp[\lambda \bar{S}^p]} \tag{3.8}$$

where, $P(X=o)+P(X=p)=1$.

In the first case, where there is switching between fundamentalists, optimists and pessimists, the adaptation part of the switching mechanism is extended from the original Logit model (3.8) into the multinomial Logit model. Following is the probability that an agent chooses fundamental over optimist and pessimist trading rules:

$$P(X=f) = \frac{\exp[\lambda \bar{S}^f]}{\exp[\lambda \bar{S}^o] + \exp[\lambda \bar{S}^p] + \exp[\lambda \bar{S}^f]} \tag{3.9}$$

where, $P(X=f)+P(X=o)+P(X=p)=1$. λ is the intensity of choice, which measures how quickly agents switch if there are additional profits gained from choosing fundamental over chartist trading rule. If $\lambda = 0$, there is no switching between strategies, while for $\lambda = +\infty$ all agents immediately switch to the most profitable strategy.

3.4.2 Herding

In the proposed model, herding is investigated from two different aspects. The first side lies in the agent based design portion of the work. An imitation component is introduced to account for the herding behavior observed in real markets and documented by a number of studies (e.g. [67]). Such behavior arises when traders observe the choices of others and, regardless of their private information and knowledge, start to imitate them. To model this process, an imitation threshold, γ is defined. The trader will change his action (buy/sell) if less than γ percent of its neighbors chose the action similar to him. The incentive behind an investors decision to discard his own beliefs and information to follow the crowd varies by different factors. For example, while trying to maximize their profit, traders may anticipate that others know something about the return on investment that they do not. An investor may think that the action of others reveals true information about the value of an asset, so they start imitating them. On the other hand, there are less sophisticated reasons behind herding, such as the intrinsic need of individuals to match their beliefs and behaviors to group norms.

In the second phase of studying herding, a simple function is defined to capture herding phenomena and its intensity at each time-step:

$$Herd_t = \frac{|D_t - S_t|}{N} \tag{3.10}$$

When $Herd_t$ is close to zero, the number of buy and sell orders are close and no herding is taking place in the market, and vice versa for when $Herd_t$ gets closer to one. The relationship of herding parameter with other market parameters and phenomena will be investigated in Section 5.

CHAPTER 4: CALIBRATION AND VALIDATION

4.1 Introduction

A crucial stage in building an agent-base model of a complex system, such as stock market, is validation, and the first step in validation is to calibrate the model. In this work, the approach to solve the calibration and validation processes is done by decreasing the distance between the statistical properties of empirical and model generated time series. The proposed model is characterized by a large number of traders, each with different behavioral parameters, whose interactions and developments give rise to the global dynamics of the system. On the other hand, even the smallest change in model variables may cause a substantial shift in emergent macro properties of the market, hence finding the right solution space can be a very long and tedious process. To address these difficulties, the approach suggested in this research is to regard the calibration procedure as an optimization problem. Hence, Genetic Algorithm (GA) is used as optimization technique to traverse the search space and find the best combination of parameter values, based on the specified objective function.

4.2 Stylized Facts

As it was explained in Section 2.4, the statistical properties that are observed in a wide range of markets and time periods, also known as "stylize facts", are used to calibrate and validate the financial market models. In this research, the three important stylized facts used to validate the proposed model are: *heavy-tails in distribution of stock returns*, *non normality of stock returns*, *absence of autocorrelation in raw returns* and *volatility clustering.*

In order to detect the aforementioned properties in model generated time series and evaluate the goodness of the model with respect to empirical observations, the following four criteria are used:

- Excess kurtosis: The fat-tail property of stock returns in the model and BAC time series is explored by excess kurtosis. It should be larger than zero, which is the excess kurtosis of a normal distribution.

- Hill estimator: The kurtosis measure is somewhat ambiguous for measuring the fat-tail in a distribution. Hence, an estimator suggested by Hill [51] is used, which ranges approximately between 2 to 5 for empirical return distributions, and lower values express heavier tails [68].

- Return autocorrelation: The autocorrelation coefficients of raw returns are basically zero for all lags in real world financial time series.

- Absolute return autocorrelation: This feature describes the fact that high-volatile events tend to cluster in time. Different measures of volatility display positive autocorrelation over several days [49]. A common proxy for volatility is the absolute return.

4.2.0.1 Hill Estimate

The estimator proposed by Hill [51] has become an standard tool to study the exponent of the tails of distributions, mostly due to its simplicity and efficiency. The Hill tail index in most of the financial time series takes values from 5 to 2, and lower values usually point to a heavier tails [68]. The Hill estimator is calculated as follows:

$$\alpha_H = \frac{1}{\sum_{i=0}^{nk-1} (log(r_{n-i}) - log(r_{n-nk}))/nk} \tag{4.1}$$

where, n is the number of observations, k is the tail fraction, which makes nk the number of observations located in the tail, and r is the return. To calculate this estimator, the data points are required to be sorted descendingly, and ordered as:

$$r_n > r_{n-1} > ... > r_{n-nk} > ... > r_1$$

This formula calculates the right tail exponent, and can be reversed to calculate the left tail. Although the implementation of the Hill estimator is easy and straightforward, it is very sensitive to the choice of k, which also depends on the sample size. If k is chosen too high, then the extreme value region of the distribution might be missed. It is a common practice in applied economics for k to be set to 10%, 5%, 2.5% and 0.5% of the tail.

4.2.1 Volatility Clustering

As it was described in previous sections, measures of volatility in financial markets have positive autocorrelation and decay slowly as a function of the time lag. In other words, high-volatility events of either signs tend to cluster in time. A common proxy for volatility in financial data analysis is given by the absolute return, which is also used in this research granted that it has the advantage of considering the extreme negative and positive returns at the same time. Autocorrelation function or ACF is the correlation between a variable and lagged versions of it, and for the j^{th} order (j lags) of absolute returns, it is defined as:

$$ACF(j) = \frac{Cov(|r_t|, |r_{t-j}|)}{Var(|r_t|)}, \tag{4.2}$$

where

$$Cov(|r_t|, |r_{t-j}|) = \frac{1}{n-1} \sum\nolimits_{t=j+1}^{n} (|r_t| - |\bar{r}|)(|r_{t-j}| - |\bar{r}|), \tag{4.3}$$

and

$$Var(|r_t|) = \frac{1}{n-1} \sum\nolimits_{t=1}^{n} (|r_t| - |\bar{r}|)^2 \tag{4.4}$$

where, $t = 1, ..., n$, n is the number of observations, r and $|\bar{r}|$ are the returns and the mean of absolute returns (volatility).

4.3 Genetic Algorithm

As it was mentioned earlier in this Chapter, agent-based models are characterized by their unpredictable, emergent behaviors that are the product of many parameters interacting and evolving over the course of simulation. In this work, to better explore the parameters space towards the simulation goal, validation of the model is considered as an optimization problem solved by Genetic Algorithm.

John H. Hollands invented Genetic Algorithm (GA), describing how mathematical and optimization problems can be solved by applying the concept of evolutionary processes [69]. GA attempts to find a good (or the best) solution and improve the model by breeding individuals (chromosomes) in a population using biologically inspired operators (selection, mutation and recombination) over a series of generations. One of the crucial aspects of GA is choosing a proper fitness criterion, one to evaluate the quality of each candidate solution with regards to a certain goal. The GA then selects, combines and mutate individuals based on this fitness in a cycle called generation, to breed new children. The fittest new individuals then replace the older ones and are passed to the next generation. The flow of GA is shown in Figure 4.1.

During the selection process, individuals are chosen to be recombined, and the likelihood of each individual to be chosen as a parent depends on how fit they are. Here, recombination (crossover) is done by choosing random crossover point in two parents chromosomes (intuitively the best parts, in order obtain a better chromosome), and join them to create new off-springs. Mutation is changing a part of chromosome with a random probability, and avoids the prematurely converge to a local solution. As the generations go on, the population keeps getting improved, until a sufficiently fit individual is discovered or the time is exhausted.

GA, as an optimization technique, offers several key advantages. First, it is capable

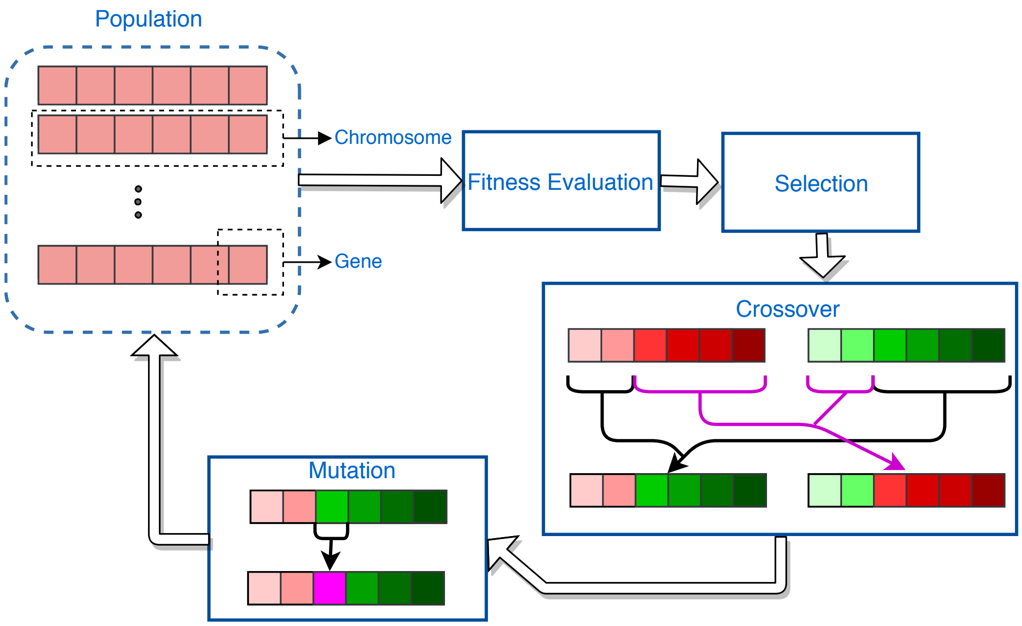


Figure 4.1: Flowchart of GA.

of parallelism from a population of points at once during each search process, hence avoids premature coverage to a local solution space. Second, GAs outperform traditional optimization techniques with their probabilistic rules, allowing them to act better on large, discontinuous problems and offer a reduced sensibility to noise. In addition, GAs use fitness score obtained from objective functions, and do not require other derivative or auxiliary information, which may not be available [70].

4.3.1 Fitness Function, GA Parameters and Optimization Results

The fitness function is a measure that rates the quality of a chromosome as a solution with respect a particular problem. As it was mentioned above, the choice of fitness function is a delicate and crucial aspect in tuning models with Genetic Algorithms. This task becomes even more challenging in the case of agent based models of financial markets, due to their dynamic nature and emergent properties that more often than not arise in their simulations. Provided that, one can end up with model characteristics and simulation results other than the initial goal if not careful with the choice and criteria of the fitness function.

In the case of proposed model, there are two approaches that can be taken to develop fitness function with the intention of decreasing the distance between model and real financial markets, the quantitative and the qualitative approaches [71].

In quantitative approach, the model generated time series is compared with the empirical and benchmark time series, and the Euclidean distance between two data vectors are computed and minimized.

In the second qualitative approach, which is adopted in this research, the stylized facts observed in real financial time series, that are desired to be reproduced in the model, are translated into a mathematical function (fitness function). In some cases, this approach may not be as simple as the first one, considering that its takes more effort, and some times its impossible, to characterize and formulate the observed emergent phenomena.

Earlier in this section, the stylized facts and their criteria that would help assess the model's validity and closeness to real financial market were explained. The same criteria are used to form a fitness function, which is a weighted sum of square distances measured between statistical moments of BAC and model generated time-series data:

$$Fitness(\theta) = \sum_{i=1}^{J} w_i (m^{model} - m^{bac})^2 \tag{4.5}$$

where, J is the number of criteria used in validation, i.e. number of model generated return moments (m^{model}) compared to empirical ones (m^{bac}), and w is the weight of each criteria, which is computed and set experimentally. θ is the parameter vector of statistics used to validate the model with respect to the empirical statistics of stock returns.

Most of the criteria considered here directly target the stylized facts that are desired to be observed it the model. The fat tail of the return distribution is measured by calculating the excess kurtosis of the returns, and the Hill estimate for 0.5% of the upper quantiles of the right tail. The volatility clustering is considered by calculating the autocorrelation of absolute returns for 1, 5, 10 and 20 lags. The other two moments considered in the fitness function are the standard deviation and mean of returns.

Finally, a simple Genetic Algorithm was used to solve the optimization problem. One of the crucial factors in running a successful GA is to have the right balance between exploration and exploitation, which includes the choice of population size, selection, and the rates of cross-over and mutation [72]. Many studies have explored and came up with different optimum parameter setting for GAs, since the task is hugely dependent on the problem at hand. Based on different experiments with different parameter settings, the GA is applied with generating an initial population of size 50, evaluating the mean fitness of each individual from 10 independent replications of the simulation, which reduces the fluctuations arise from stochastic nature of the model. At this point, the tournament selection schema with tournament size of 2 is adopted, and mutation and crossover are performed with probabilities of 0.05 and 0.7, respectively. Figure 4.2 illustrates two search progresses of GA minimizing the fitness function.

Table 4.1 provides the list of fitness function statistics and their empirical and model generated values for 3000 observations, covering the period of 03/22/2006 to 02/22/2018. Table 4.2 details the optimal parameter values suggested by GA along with the range of values allowed for each parameter, and Table 4.3 presents the values attributed to rest of the parameters used in the simulations. More details on why and how to calculate return from price time series and the process of obtaining and comparing the mentioned statistics are explained in Chapter 5.

As it can be seen from the first row of the table, Hill estimates over 0.5% of the right tail for both BAC and model return distributions are fairly close, and both support the existence of heavy tail. Also, the model's excess kurtosis of 8.2, though not as high as the BAC benchmark, is still a solid sign of fat tails in return distribution.

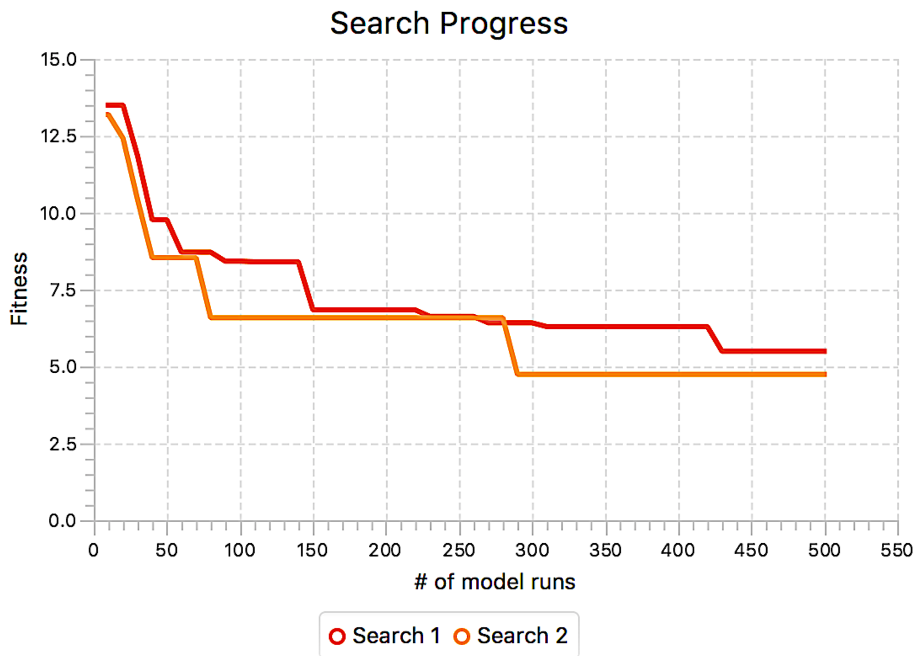


Figure 4.2: GA search process.

Table 4.1: Statistic criteria used in GA fitness function.

Statistic	Description	BAC	Model
Hill tail index	Hill estimate on 0.5% of right tail of return distribution	3.657255	3.461575
Kurtosis	Excess kurtosis of returns	22.1655	8.202247
Mean	Mean of returns	-0.000128233	-0.000486729
Standard-deviation	Standard deviation of returns	0.03422454	0.03243146
1_{st} Autocorrelation	ACF of absolute returns at lags 1	0.462	0.450
2_{nd} Autocorrelation	ACF of absolute returns at lags 5	0.420	0.460
3_{rd} Autocorrelation	ACF of absolute returns at lags 10	0.415	0.390
4_{th} Autocorrelation	ACF of absolute returns at lags 20	0.363	0.349

In fact, as it was explained earlier in the section, anything larger than zero point to departure of distribution from normal. The fat tail of the returns can also be confirmed from the quantile-quantile or q-q plots in Figures 4.3a and 4.3b. They compare

Table 4.2: Model parameter settings suggested by GA.

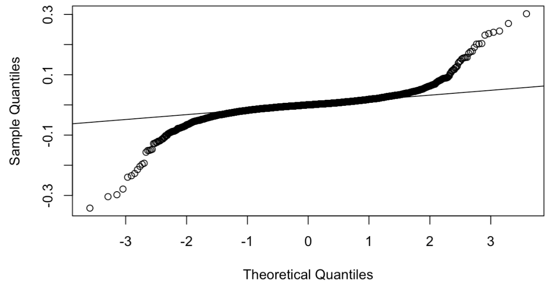
Parameter	Range	Optimized value
Fundamentalists trade threshold	[0,30]	14%
Optimist trade threshold	[0,30]	4%
Pessimist trade threshold	[0,30]	6%
Mean of fundamentalist reverting coefficient (f)	[0.01,0.8]	0.11
Mean of optimist reaction coefficient (c_o)	[0.01,0.8]	0.11
Mean of pessimist reaction coefficient (c_p)	[-0.8,-0.01]	-0.15
Imitation switch threshold (γ)	[0,100]	55%
Intensity of choice (λ)	[0,300]	272

the distributions of the returns against a standard normal distribution, and as it can be seen, both returns distributions are non-normal at the tails of the distribution. The non-Gaussian character of both distributions is clear from density plots, and the proposed model was successful in replicating the fat tail phenomena.

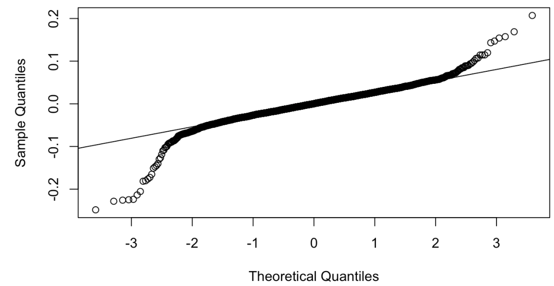
Table 4.3: Parameters of the stock market model.

Parameter	Value
Number of traders (N)	1000
Price adjustment-Positive excess demand (a)	0.41×10^{-4}
Price adjustment-Negative excess demand (a)	0.56×10^{-4}
Pareto exponent for wealth (power law) distribution (α)	1.5
Initial cash distribution	[100, 1000]
EWMA smoothing parameter (ϕ)	[0.3,0.99]
Standard deviation of random factor in price process (σ_δ)	0.025
Standard deviation of random factor in fundamental price process (σ_η)	0.026
Standard deviation of random factor in fundamental trading (σ_τ)	0.01
Reverting coefficient (f)	$N(\mu_f : 0.11, \sigma_f : 0.02)$
Reaction coefficient (c_o)	$N(\mu_{c_o} : 0.11, \sigma_{c_o} : 0.03)$
Reaction coefficient (c_p)	$N(\mu_{c_p} : -0.15, \sigma_{c_p} : 0.04)$
Standard deviation of random factor in optimistic trading (σ_{β_1})	0.05
Standard deviation of random factor in pessimistic trading (σ_{β_2})	0.05
Memory of traders (T)	U (1,50)
Normalized memory (m)	[0,1]

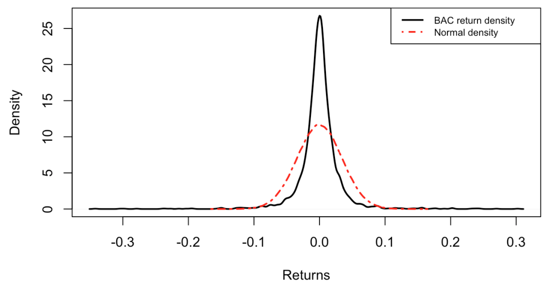
Figures 4.3e and 4.3f, as well as the last four rows of Table 4.1, examine the presence



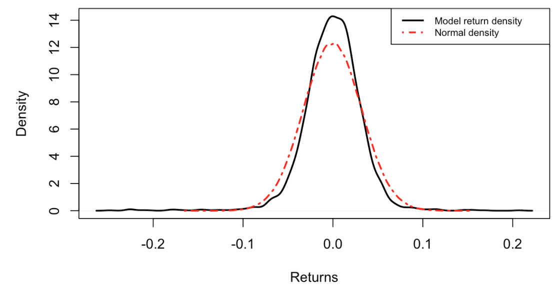
(a) BAC return q-q plot for normality comparison.



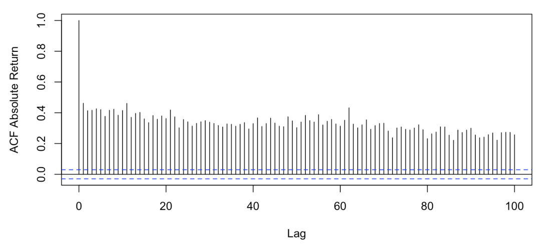
(b) Model return q-q plot for normality comparison.



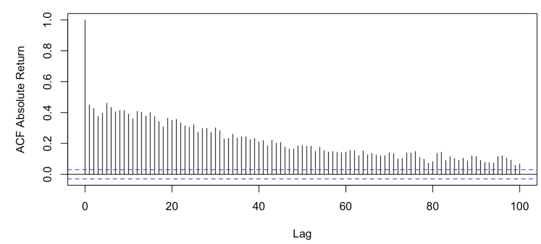
(c) BAC return density plot.



(d) Model return density plot.



(e) BAC autocorrelation plot for volatility measure(absolute return)



(f) Model autocorrelation plot for volatility measure(absolute return)

Figure 4.3: Normality and volatility clustering test results for BAC and model returns.

of volatility clustering in the market and demonstrates the autocorrelation of absolute returns in BAC and model. In Figures 4.3e and 4.3f, the dotted blue lines indicate the 95 percent confidence intervals of no autocorrelation. It should be noted that the absolute returns exhibit a slow decay of the autocorrelation in both BAC and model, and even stay positive for more than 100 lags. This is a clear sign of volatility clustering and can be observed in many real world financial markets.

CHAPTER 5: RESULTS AND ANALYSIS

5.1 Statistical Analysis

Having obtained a validated model, the extent of effects of different market phenomena on each other can now be investigated. These investigations can help enhance the understanding of dynamics and structure of financial markets and paves the way for further explorations on causes and effects of the traders characteristics on market phenomena.

The Bank of America (BAC) daily stock price is used as benchmark in this research. Data covers the period between March 2006 to February 2018 and is collected from the [73]. Therefore, for each model simulation run, 3000 observations are generated, which corresponds to the same time span of about 12 years.

5.1.1 Stationarity

Before starting any analysis on empirical and model generated time series, it is essential to check the stationarity of them. Non-stationary data are unpredictable; therefore they cannot give meaningful sample statistics and correlations with other variables. In non-stationary time series, the mean and variance change over time and series do not have tendency to come back towards their average. In this work, following the standard econometric practice to examine for stationarity, Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) unit root tests are conducted on all the variables of interest [74].

The DF tests for the most specific model of zero mean and no trend:

$$X_t = \alpha X_{t-1} + \epsilon_t \tag{5.1}$$

with the null hypothesis of $\alpha = 1$. Failing to reject the null hypothesis indicates that the series is characterized by a "unit root" and is non-stationary. However, DF test is not adequate in the cases where the data has "drift" or a time trend, in which case the ADF tests are used to allow for such variations.

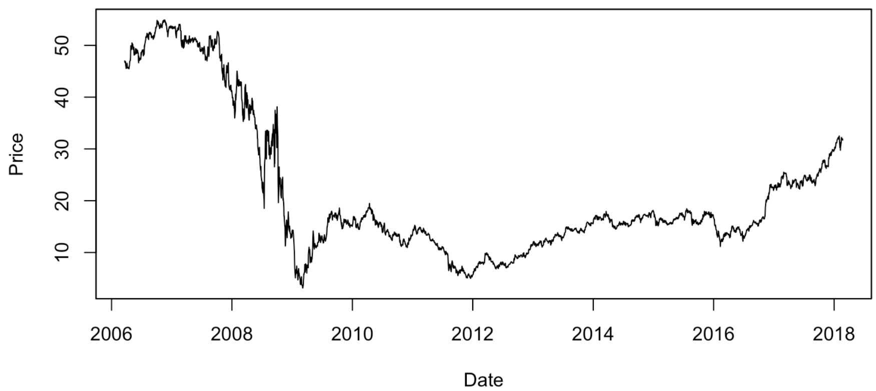
In the case of non stationary time series, the autocorrelation function (ACF), which is the correlation between a variable and the lagged version of it, decreases very slowly, and its coefficients remain significance for large number of lags.

In the case of non-stationary time-series, the difference transformation can be taken to try and make the time series stationary. If a variable proves to be stationary after the first difference, then it is integrated of order one or $I(1)$, and in the case where stationarity is gained in their Kth differences, they are said to be integrated of order K.

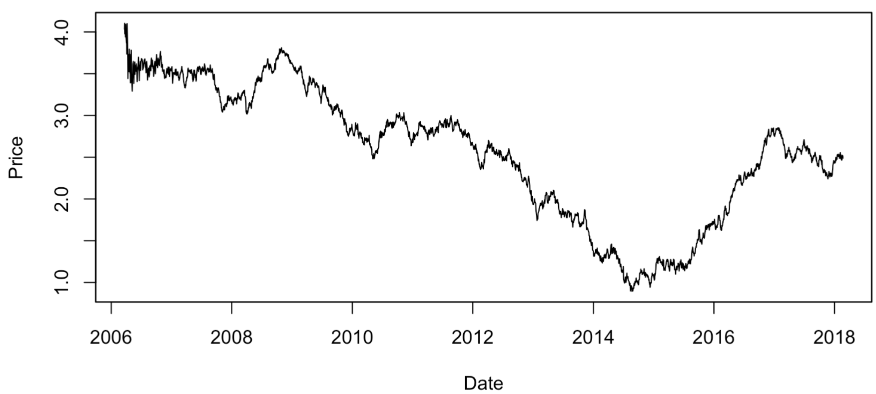
5.1.2 Data Description and Preparation

As the first step to analyze the empirical and model generated time series, a series of tests are conducted to check the stationarity of *price*, which is the core variable to study the market and validate the model. Figure 5.1 shows the daily development of price for BAC and model stock. Both series appear to have obvious trends. In non-stationary time series, the mean and variance change over time and prices do not have tendency to come back towards their average, which appears to be the case here.

Following the instructions in Section 5.1.1, ADF unit root tests are conducted on both price time series. Tables 5.1 and 5.2 provides the results, which tests the null hypothesis of non-stationarity. Zero Mean, Single Mean and Trend tests correspond to checking for no underlying time component, constant increase over time and acceleration over time respectively. The test statistics along with corresponding p-values indicate that the null hypothesis cannot be rejected, which confirms that BAC and model generated prices are in fact non-stationary. In addition, ACF plots for both



(a) BAC price.



(b) Model price.

Figure 5.1: Daily development of BAC and model generated price.

model and BAC prices in Figures 5.3a and 5.4a demonstrate that the correlation between the time series and its lags fall outside the 95% of confidence interval of no auto correlation, even for more that 100 lags. This is another proof of non stationarity of the price time series in both case, and also demonstrates that the model price is in agreement with empirical price.

In order to induce stationarity, first differences of logs of the price series, also known as the stock returns, are derived as:

$$return_t = \ln[p_t] - \ln[p_{t-1}] \tag{5.2}$$

Figure 5.2 depicts the plots of stock returns for BAC and model time series. They both appear to have constant mean and variance change over time. The ADF test is

Table 5.1: Augmented Dickey-Fuller unit root test for BAC price

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1.741	0.3628	-1.34	0.1673		
	1	-1.7203	0.3654	-1.35	0.1638		
Single Mean	0	-4.2468	0.5126	-1.77	0.3937	1.68	0.6414
	1	-4.1559	0.5225	-1.77	0.3944	1.68	0.6402
Trend	0	-2.6785	0.9497	-0.96	0.948	2.16	0.744
	1	-2.528	0.955	-0.92	0.952	2.23	0.7303

Table 5.2: Augmented Dickey-Fuller unit root test for model price

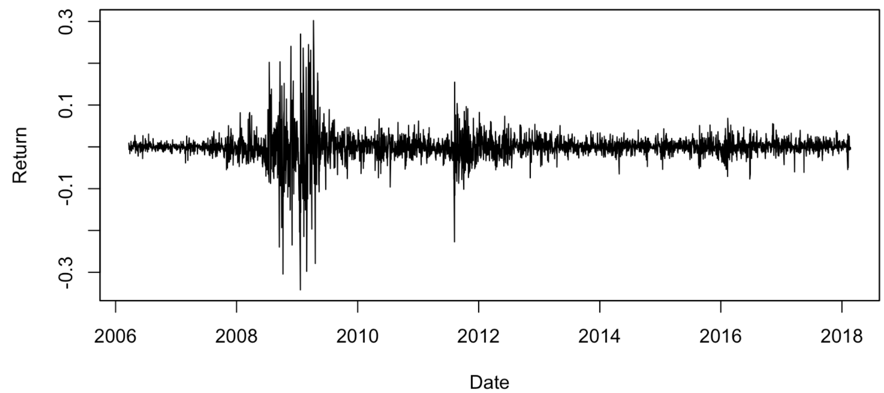
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-0.9263	0.486	-1.39	0.1523		
	1	-0.9259	0.486	-1.53	0.1195		
Single Mean	0	-4.2753	0.5096	-1.9	0.3302	2.19	0.5084
	1	-4.0173	0.5378	-1.96	0.305	2.41	0.4522
Trend	0	-4.4101	0.8614	-1.31	0.885	1.81	0.8142
	1	-3.5759	0.9102	-1.16	0.9167	1.94	0.7894

repeated on both return time series and the results are presented in Tables 5.3 and 5.4. The test statistics along with their p-value confirm that the null hypothesis of series being non-stationary can be rejected. Again, ACF plots for both model and BAC returns in Figures 5.5d and 5.6j show no sign of significant autocorrelation for return time series of both BAC and model, meaning that the price series became stationarity after first differences and also point to the fact that the model returns is in agreement with empirical observations.

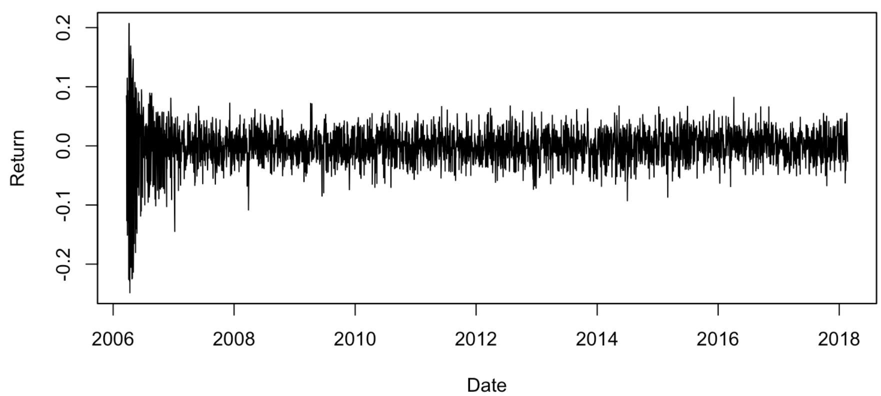
5.1.3 Other Time-Series Variables

In order to investigate the relationship between market states and herding, volatility and other market phenomena, a dummy variable, *bull*, is created. One of the common descriptions of bull and bear market is that they correspond to the periods of increasing and decreasing stock prices, respectively [75]. In this analysis, $bull_t$ takes the value of 1 if the return on day t is positive and market is bullish, and 0 when the return is negative, meaning that the market is bearish.

The parameters $fundamentalist_t$, $optimist_t$ and $pessimist_t$ are the number of



(a) BAC return.



(b) Model return.

Figure 5.2: Daily development of BAC and model generated return.

fundamentalists, optimist and pessimist present in the modeled market at each time step, t. As it was mentioned before, $volatility_t$ is the absolute return and $volume_t$ is the number of shares traded in the market at each time step t. Finally, $buyers_t$ and $sellers_t$ are the number of buyers and sellers that are present in the market at each time step, t.

Again, the standard econometric practice of ADF unit root test is conducted on all variables of interest. The *price* and *return* stationary test was covered earlier in this Section. In Figures 5.3 and 5.4, a series of ACF plots of BAC and model variables are shown. These plots clearly illustrate that except for *bull* variable of both BAC and model, the estimated autocorrelations for all the other variables fall

Table 5.3: Augmented Dickey-Fuller unit root test for BAC return.

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-3028.52	0.0001	-55.29	<.0001		
	1	-2787.72	0.0001	-37.32	<.0001		
Single Mean	0	-3028.56	0.0001	-55.28	<.0001	1527.82	
	1	-2787.84	0.0001	-37.32	<.0001	696.26	0.001
Trend	0	-3030.58	0.0001	-55.31	<.0001	1529.35	0.001
	1	-2793.36	0.0001	-37.35	<.0001	697.4	0.001

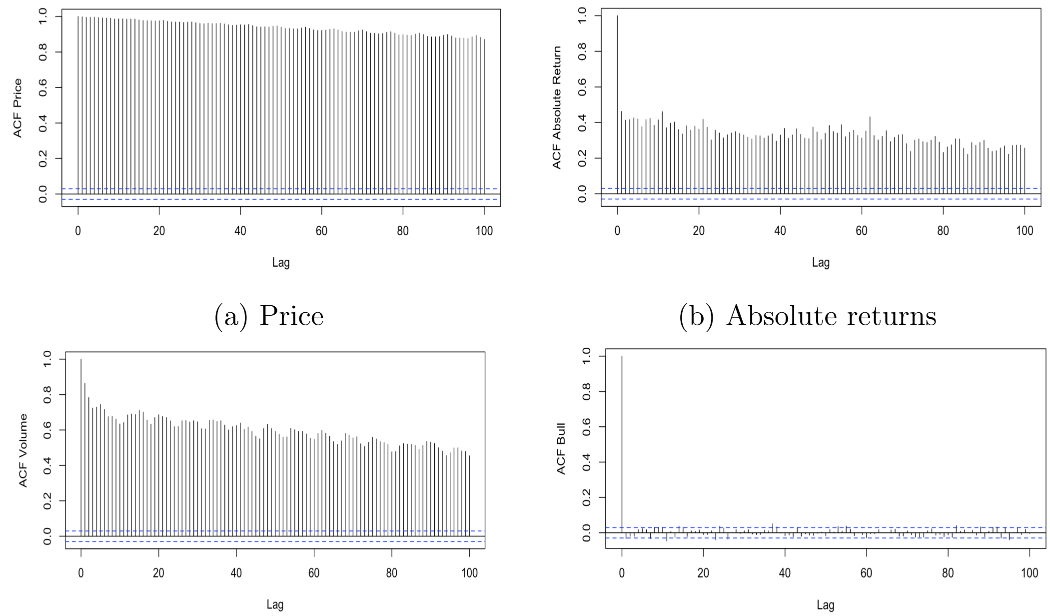
Table 5.4: Augmented Dickey-Fuller unit root test for model return.

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-3263.32	0.0001	-59.86	<.0001		
	1	-3638.68	0.0001	-42.74	<.0001		
Single Mean	0	-3264.08	0.0001	-59.87	<.0001	1792.15	0.001
	1	-3641.84	0.0001	-42.75	<.0001	913.75	0.001
Trend	0	-3266.06	0.0001	-59.9	<.0001	1794.04	0.001
	1	-3650.25	0.0001	-42.79	<.0001	915.69	0.001

outside the 95% confidence interval for at lease 20 lags, which is a sign of long-memory (non-stationary) time series.

The result of ADF test of the *bull* variable of BAC can be rejected, at 1% level significance, pointing to stationarity of it. This result can be backed up by its ACF plot in Figure 5.3d that shows no autocorrelation for the variable passed first lag. As for the other variables, it is a little more complicated than that. The ADF test can be rejected at 1% level significance, for both *volume* and *volatility* that the series are non-stationary, but the ACF plots in Figures 5.3b and 5.3c show that autocorrelation is decaying very slowly, and remains well above the significance range (dotted blue lines) for more than 100 lags, which is indicative of a non-stationary series.

Finally, the results of ADF tests on model variables is made it clear that the non-stationarity of the null hypothesis for the variables *price*, $fundamentalist$, *optimist* and *pessimist* cannot be rejected. Together with their significant, slow decaying ACF plots in Figures 5.4a, 5.4h, 5.4i and 5.4j the non-stationarity of these variables is confirmed. As for the *bull* variable, the null hypotheses can be rejected and the



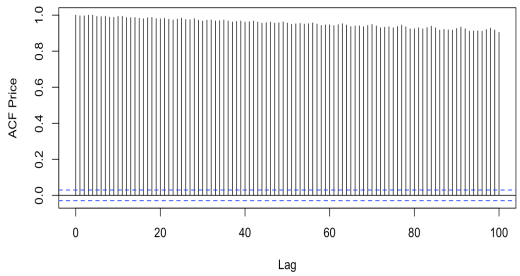
(a) Price (b) Absolute returns

(c) Volume (d) Bull market

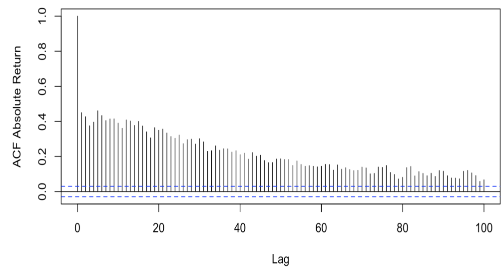
Figure 5.3: The autocorrelation plots for BAC variables in their levels.

ACF plots shows no significant autocrrelation passed the first lag, indication the stationarity of the time series. The ADF test for remaining variables (*volatility*, *volume*, *herd*, *buyers* and *sellers*) can be rejected at 1% level significance, however, as it was also the case for some BAC variables, the ACF function decreases slowly, and falls outside the 95% confidence interval for at lease 20 lags, suggesting non-stationarity. Appendix A provides the details of the ADF test results.

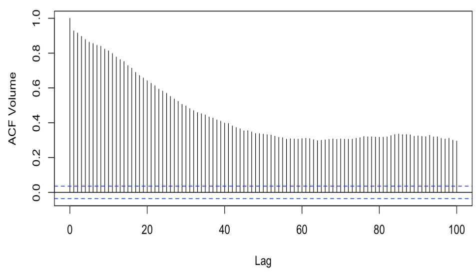
As explained earlier in this Section, to induce stationarity, the first differences the non-stationary time series are derived for both Model and BAC, and the tests are repeated. As it is illustrated in Figures 5.5 and 5.6, ACF plots demonstrate that there is no significant correlation between the time series and their lags after the first few lags. In addition, the results of ADF test statistic for all the time series reject the null hypothesis that the series are non-stationary, showing that all of the BAC and model non-stationary variables appear to be I(1) variables.



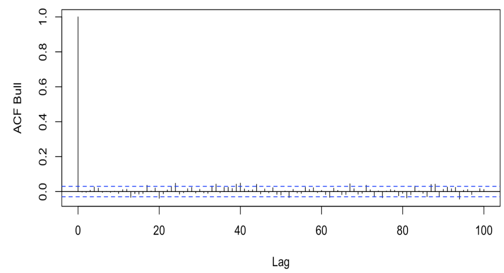
(a) Price



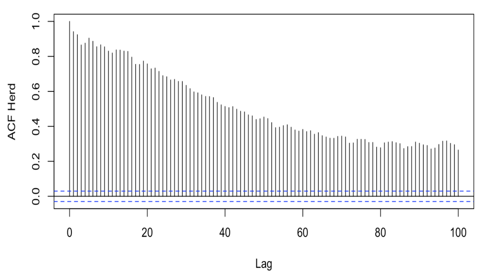
(b) Absolute returns



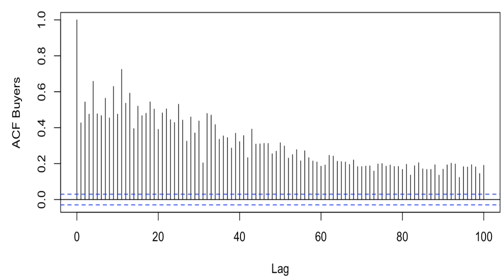
(c) Volume



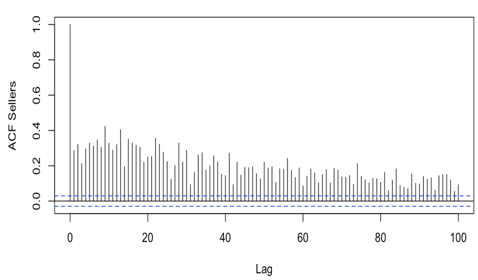
(d) Bull market



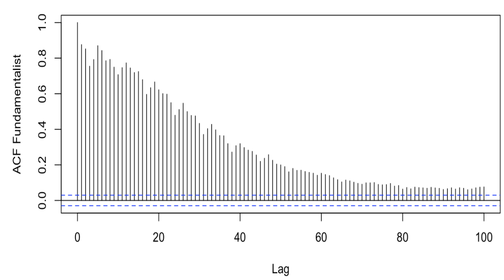
(e) Herd



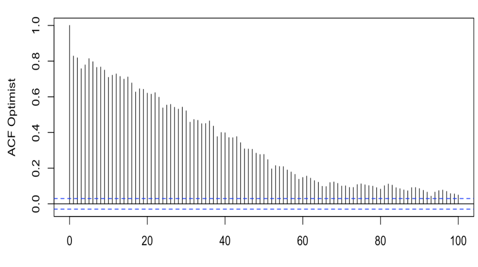
(f) Buyers



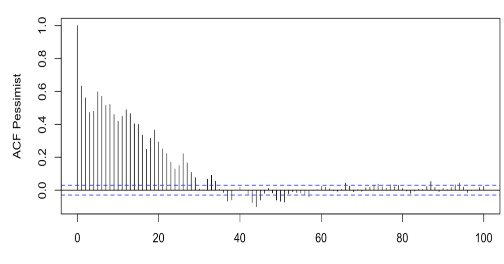
(g) Sellers



(h) Fundamentalists

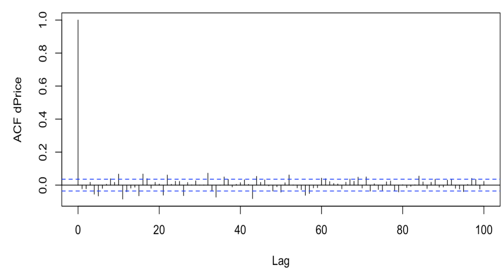


(i) Optimists

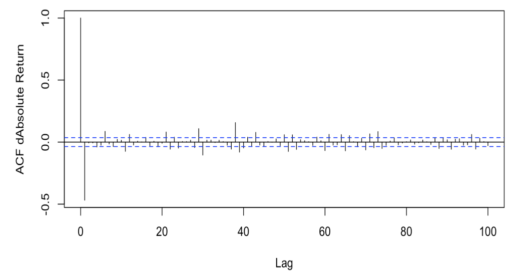


(j) Pessimists

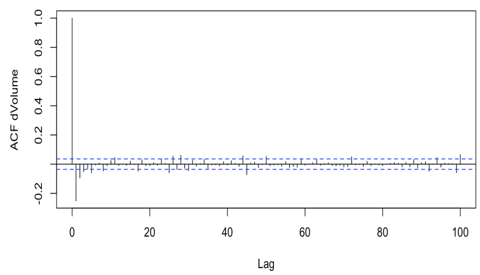
Figure 5.4: The autocorrelation plots for Model variables in their levels.



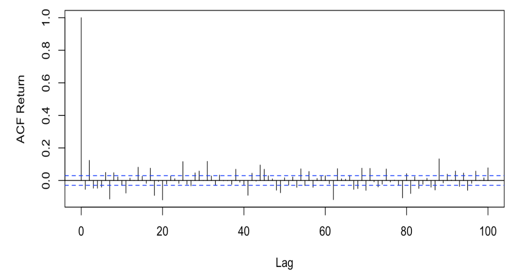
(a) dPrice



(b) dAbsolute returns



(c) dVolume



(d) Returns

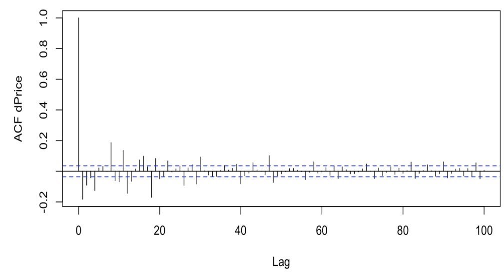
Figure 5.5: The autocorrelation plots for BAC variables in their first differences.

5.2 Trader's Characteristics Influence on Market Phenomena

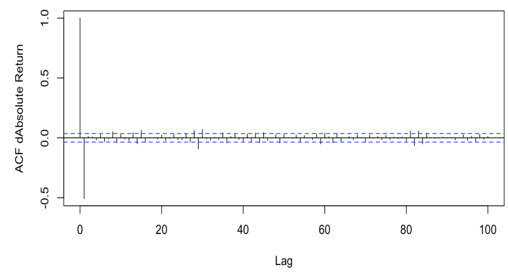
In this section, some of the parameters related to trader's choices and behaviors, and their effects on emergent properties of the market is explored.

5.2.1 Intensity of Belief

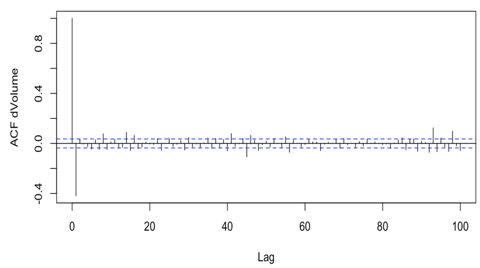
First, the intensity of beliefs for fundamentalists, optimists and pessimists are going to be analyzed. As it was explained in previous sections, to introduce more heterogeneity and get closer to real world, each fundamentalist, optimist and pessimist are unique in intensity of belief in their trading strategies. For fundamentalists, this specifies different opinions in the speed with which they expect the price to return to fundamental value (f). As for the chartists, it conveys different sensitivity levels to price changes and trends (c_o, c_p). Henceforth, at the beginning of each simulation, unique f, c_o and c_p are drawn from Gaussian distributions and assigned to each agent. It was illustrated in Chapter 4 that to achieve the emergence of stylized facts in the model, the shape of the aforementioned distributions are as follow:



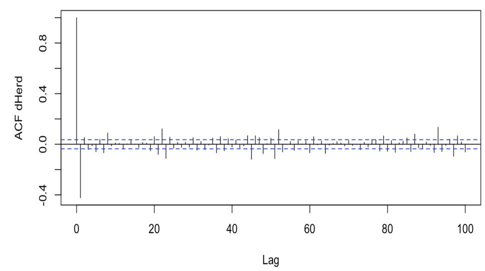
(a) dPrice



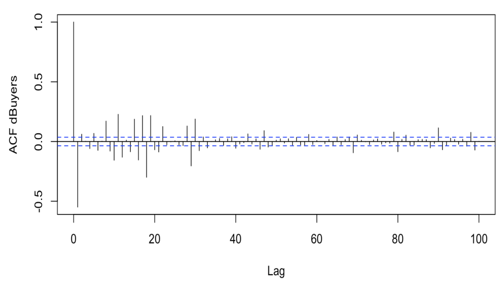
(b) dAbsolute returns



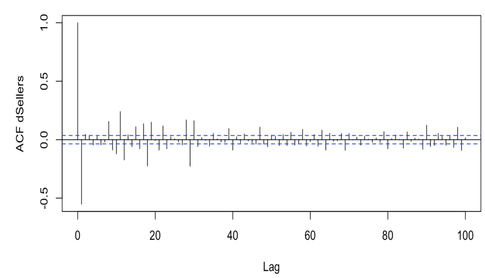
(c) dVolume



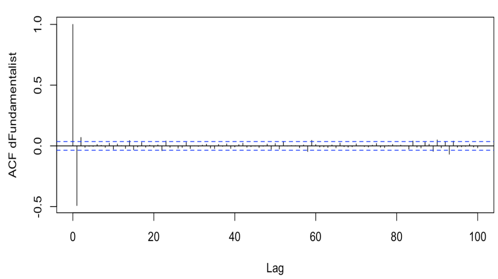
(d) dHerd



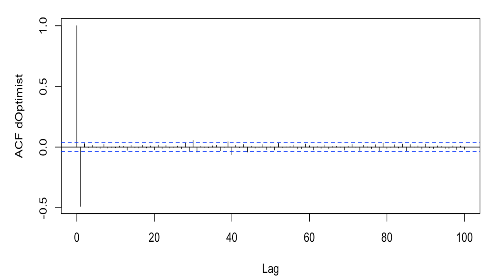
(e) dBuyers



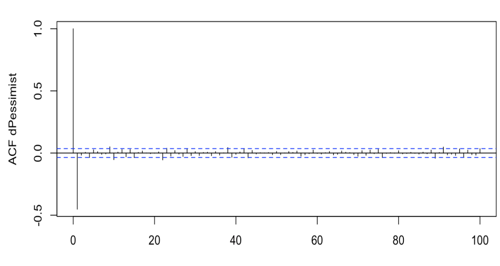
(f) dSellers



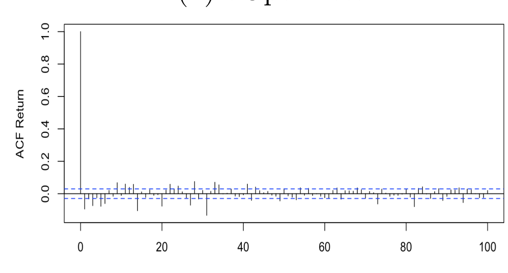
(g) dFundamentalists



(h) dOptimists



(i) dPessimists



(j) Return

Figure 5.6: The autocorrelation plots for model variables in their first differences.

- $f \sim N(0.11, 0.02)$
- $c_o \sim N(0.11, 0.03)$
- $c_p \sim N(-0.15, 0.04)$

Figure 5.7 displays the overlapping distributions of each trader's intensity of belief in their trading regime.

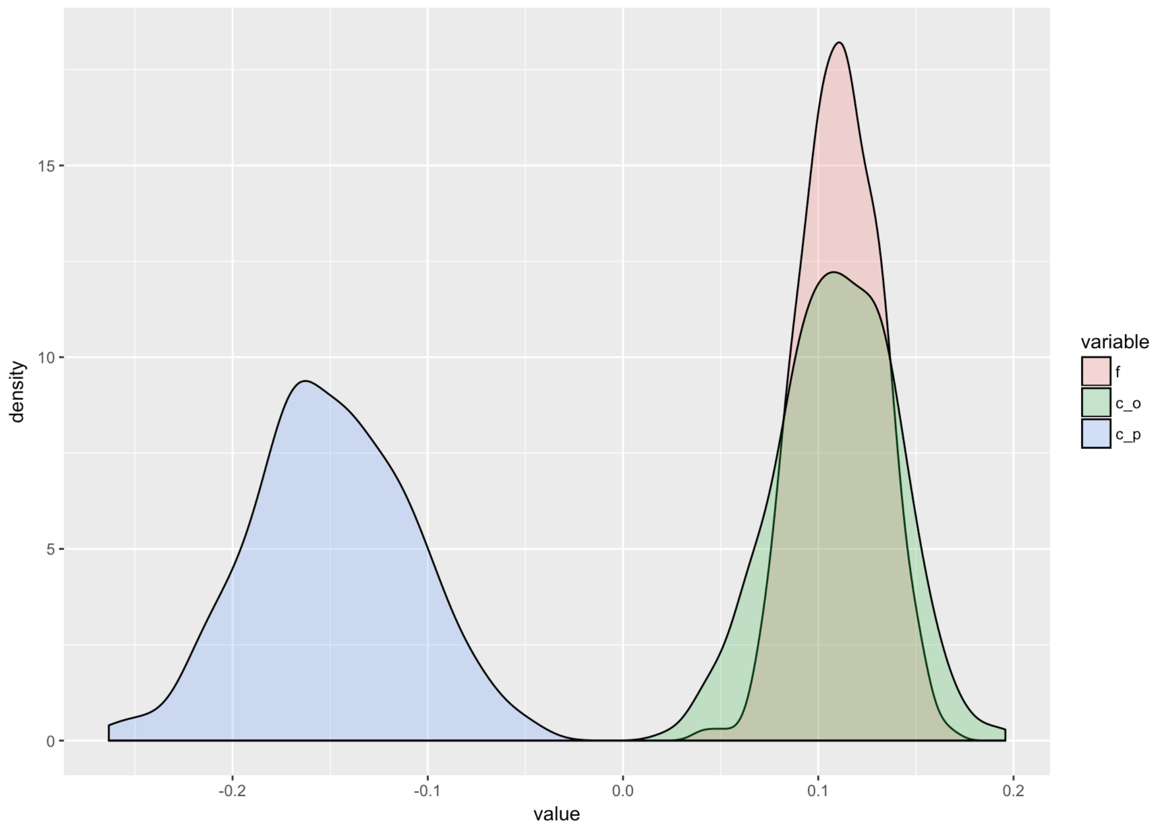


Figure 5.7: Traders intensity of belief.

5.2.2 Intensity of Choice

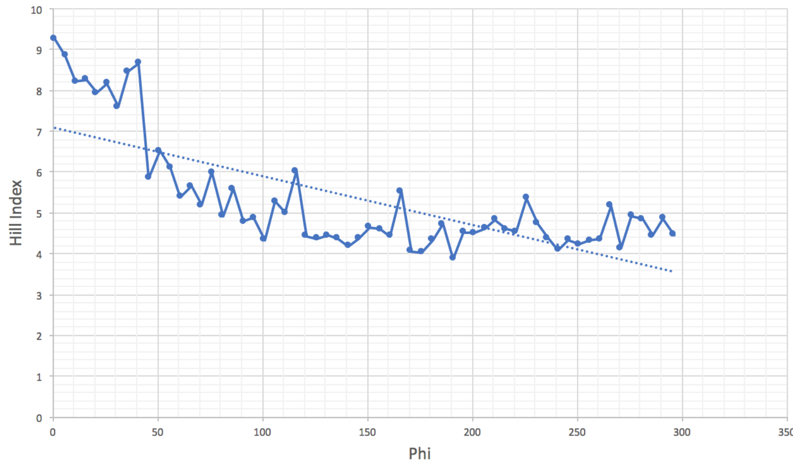
In the next set of simulations, the effects of trader's intensity of choice (λ), which measures how fast agents switch between different trading strategies, on generated stylized facts is studied. Figure 5.8 shows the changing simulation results for volatility clustering and fat tail of the returns as we increase the intensity of choice for traders.

The presented results are averaged over 10 simulation runs for each instance of λ to provide converged results.

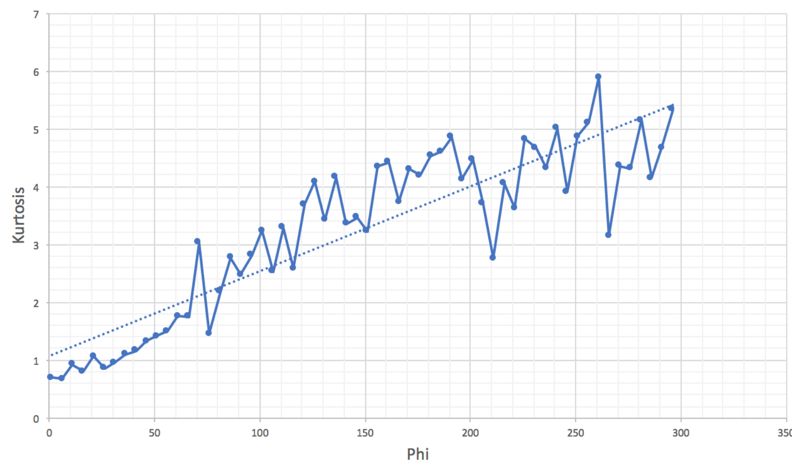
As it can be seen from the plots, $\lambda = 260$ provides the closest results with respect to the empirical statistics, which can be found in third column of Table 4.1, for all three cases.

5.2.3 Wealth Distribution

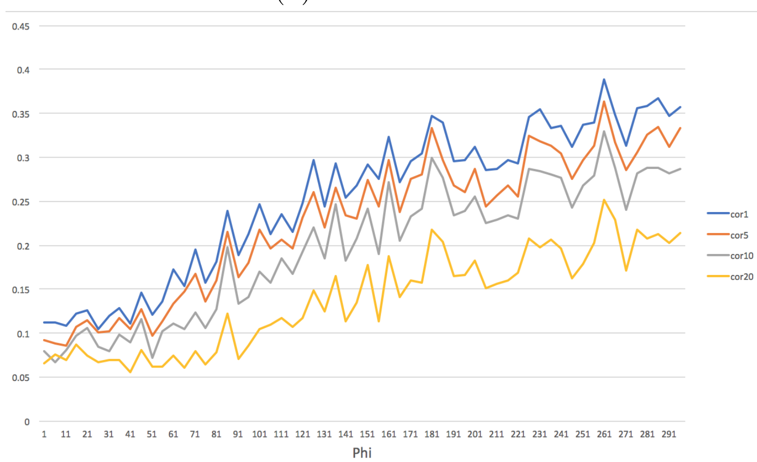
Next, the distribution of money held by agents in the beginning and end of the simulations are investigated. As was explained in Section 3.3.1 of Chapter 3, the distribution of money held by agents provides valuable information on how wealth is distributed. As per the discovery of power-law distribution of wealth across several countries [64, 65], the agents were initially assigned different amount of cash drawn from a power-law distribution. Figure 5.9 exhibits the cash distribution between agents at the beginning and end of simulation. As it can bee seen, both distributions demonstrate power-law behavior, meaning only a few number of agents hold most of the cash in model, even after the course of simulation. This is in agreement with wealth inequality of human population.



(a) Hill tail estimate of the return distribution.



(b) Kurtosis of returns.



(c) Autocorrelation of absolute return for lags 1,5,10,20(volatility clustering).

Figure 5.8: Observed stylized facts with increasing the intensity of choice (λ).

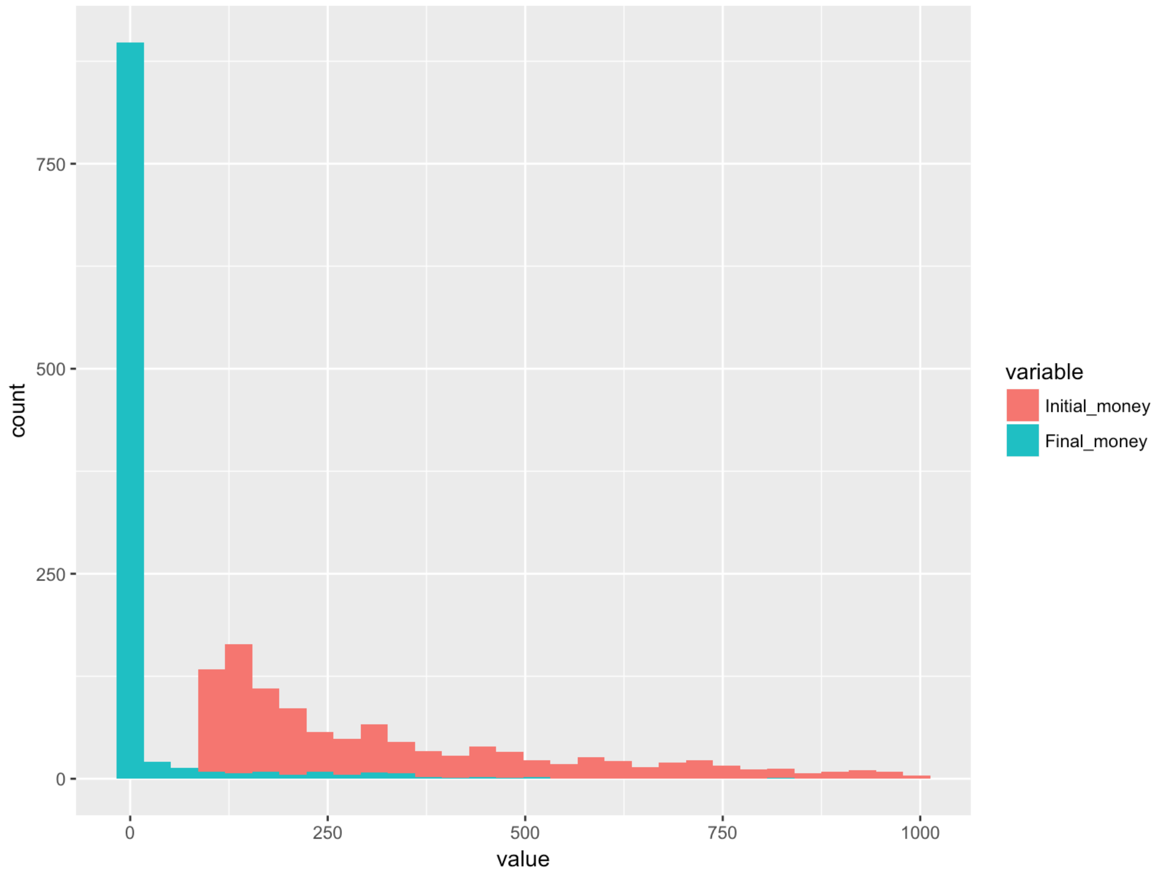


Figure 5.9: Wealth distribution.

5.2.4 Herding

The imitation threshold or herding threshold γ was explained in details in Section 3.4.2. Agents have connections with other agents on their eight neighboring patches. Traders observe the choices of their neighbors, and if less than γ percent have chosen the same action as them, disregard their information to follow the decision of the neighbors. In Chapter 4, GA has considered different imitation thresholds with respect to the fitness function. The results revealed that the best value for γ in order for model to more accurately replicate stylized facts and falls closer to the behavior of real markets was 55%. In other words, traders would change their decisions if more that 45% of their neighbors decide differently.

5.2.5 Trade Threshold

Fundamentalist, optimist and pessimist become active and trade (buy or sell) one share of the risky asset if the difference between the current price and the price they expect for the next period is above their trade threshold. These three thresholds were determined through model calibration by GA in Chapter 4. As it can be seen in Table 4.2, the fundamentalist, optimist and pessimist trade threshold were estimated to be 14%, 4% and 6% respectively. This is in agreement with real-world market traders, where the frequency at which fundamentalists trade is not as often as chartists or trend followers.

5.3 Causality Analysis

5.3.1 Definition and Method

The notion of causality used in this research is the one introduced by Granger [76] (Nobel Winner 2003). Although the concept of causality is, in many ways, a philosophical one, however, Granger proposed a practical and widely accepted definition of it between two or more time-series variables. This test goes beyond the classical regression model and correlation, which examine a relationship but not how one variable causes movement in another. Consider the process X, Granger causality (GC) assumes that only the past can help predict the present, therefore X can be written as an autoregressive process in which its current value can be explained (partially and under a certain lag length, l) by the past values:

$$X_t = \alpha + \sum_{j=1}^{l} \beta_j X_{t-j} + \epsilon_t \tag{5.3}$$

Another time-series variable, e.g. Y, is said to Granger cause X if the past values of Y offer additional information for the forecast of current value of X, beyond what was provided by past values of X alone.

$$X_t = \alpha + \sum_{j=1}^{l} \beta_j X_{t-j} + \sum_{j=1}^{l} \delta_j Y_{t-j} + \epsilon_t \tag{5.4}$$

where, X is the dependent variable and Y is the independent variable. The test is an F-test on the δ's being jointly equal to zero. If the null hypothesis is rejected, then Y is said to Granger cause X. It is also possible to do this test reversely, checking where X Granger causes Y.

As Stock and Watson [77] noted, Granger causality tests should only be undertaken on stationary variables since the existence of unit roots would typically result in nonstandard F-distributions and spurious correlation problem, where the variables that are not related seem to be correlated with each other. Therefor, the variables need to have time-invariant mean and variance and can be adequately represented by a linear autoregressive process $AR(l)$ of order l, with l chosen by an appropriate criterion.

5.3.2 Lag Length Selection

The choice of lag length l plays a critical role as causality test is very sensitive to it. There are different approaches to find the appropriate lag, from arbitrary, ad hoc methods, to testing VAR equation of (5.4) in both directions with different lag lengths, and some criteria such as Akaike Information Criterion and Bayesian Information Criterion (AIC/BIC) [78,79], where l is chosen by minimizing their functions. Here a combination of F-test from VAR models and minimizing AIC/BIC criteria was used to test and select the proper lag length for each model.

5.3.3 Experiments and Results

5.3.3.1 Causal Relationships Between Herding and Other Market Phenomena

In this section, the causes and effects of herding behavior in financial market is investigated. It should be noted that all these tests and results are from the model generated time series, and the causality analysis on BAC data will be presented later in this section. As it was explained in previous sections, herding is an important emergent property of financial markets, often leading to the creation of speculative bubbles that inevitably make markets unstable and prone to major crashes, which makes it crucial to understand the origins and driving forces of such phenomena. The causal models tested here are as follow:

- Does volatility Granger causes herding in the market? Does herd Granger cause volatility? Is the relationship bi-variate?

$$dherd_t = \alpha + \sum_{j=1}^{l^h} \beta_j^h dherd_{t-j} + \sum_{l^h}^{l^h} \delta_j^v dvolatility_{t-j} + \epsilon_t \qquad (5.5)$$

$$dvolatility_t = \alpha + \sum_{j=1}^{l^v} \beta_j^v dvolatility_{t-j} + \sum_{j=1}^{l^v} \delta_j^h dherd_{t-j} + \epsilon_t \qquad (5.6)$$

- Does herd Granger causes bull market? Does bull market Granger cause herd? Is the relationship bi-variate?

- Does herd Granger causes the volume, or vice versa? Is the relationship bi-variate?

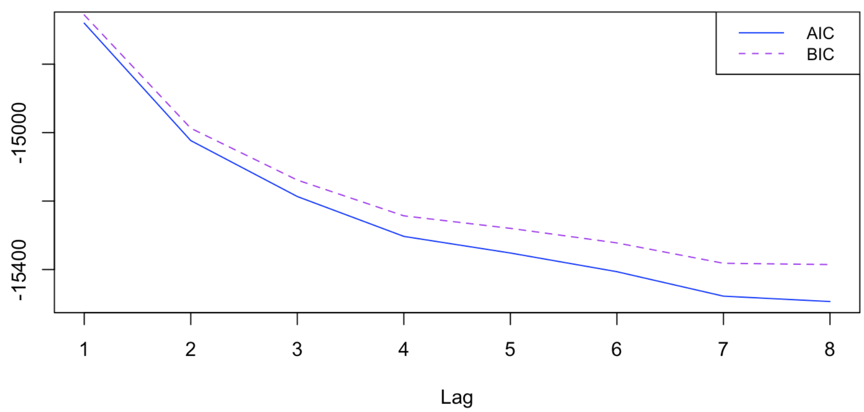
- Does herd Granger causes the number of fundamentalists, optimists or pessimists in market? Or vice versa? are the relationships bi-variate?

The above tests are conducted by the series of F-tests determining whether the lagged values of independent variables significantly contribute to explaining the current value of dependent variables. For any model that the null hypothesis of no Granger causality ($\delta_1 = \delta_2 = ... = \delta_l = 0$) can be rejected, it can be concluded that the independent variable Granger causes the dependent variable. The model structure of all the other tests are in the same fashion of first set of tests between *herd* and *volatility*, and each case is conducted in both directions.

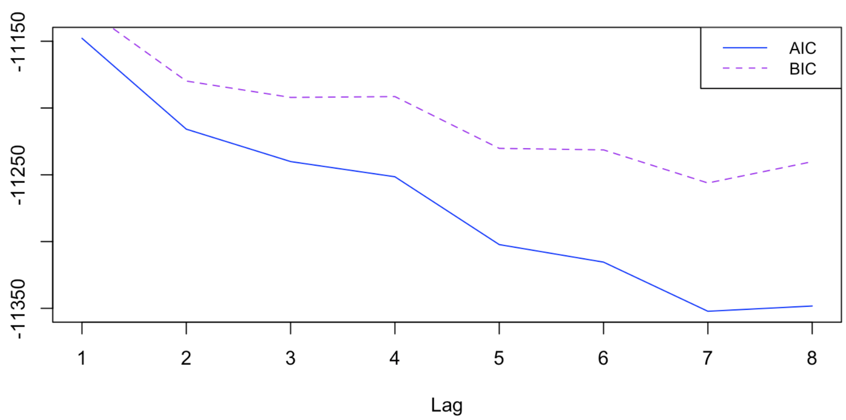
The lag-length (l) selection results based on different criteria are reported in Table 5.5. Numbers presented here are optimum lag lengths determined by the AIC, BIC, and a series of F-tests on decreasing lags. Figure 5.10 depicts a sample lag selection for herd and volatility relations by minimizing AIC and BIC functions. It can be observed that BIC usually selects a shorter lag structures than those selected by the AIC. For causality tests, the largest lag length from Table 5.5 is selected, and tests move on to a shorter length if no significant relationship is found. The results of the tests are presented in Table 5.6.

Table 5.5: Lag length selection for investigating variable *herd* by different criteria.

Dependent variable-Independent variable	AIC	BIC	F-test
herd-volatility	7	7	7
volatility-herd	8	8	8
herd-bull	7	7	7
bull-herd	1	1	1
herd-volume	7	7	7
volume-herd	8	8	8
herd-fundamentalist	8	7	8
fundamentalist-herd	8	6	8
herd-optimist	7	7	7
optimist-herd	8	8	8
herd-pessimist	8	7	8
pessimist-herd	8	6	8



(a) Herd GC volatility.



(b) Volatility GC herd.

Figure 5.10: AIC and BIC lag length selection for herd volatility causal relationship.

The results indicate that F-test fail to reject the hypothesis of volatility not cause herding at 5% significance level, while rejecting it at 1% significance the other way around. These results imply that herd behavior may cause and enhance volatility in the market, but not the other way around. The results also show no significant causal relationship between herding and bullish/bearish market states in any directions. On the other hand, there is a significant bi-directional causal relationship between the volume traded in the market and herding, indicating that the past values of volume in the market can help predict and cause the herding behavior, and vice versa. Finally, there are positive bi-directional causal relationships between number of fundamentalists, optimists and pessimists in the market and herding.

Table 5.6: Causality test between *herd* and other market variables of the model.

H_0	# of lags	F-value	p-value
Volatility does not cause herding	8	1.51	0.1486
Herding does not cause volatility	7	3.91	0.0003**
Bull market does not cause herding	1	1.50	0.2209
Herding does not cause bull market	7	0.43	0.8871
Volume does not cause herding	8	11.89	<.0001**
Herding does not cause volume	7	9.70	<.0001**
# of fundamentalists does not cause herding	8	14.56	<.0001**
Herding does not cause # of fundamentalists	8	30.40	<.0001**
# of optimists does not cause herding	8	5.02	<.0001**
Herding does not cause # of optimists	7	11.88	<.0001**
# of pessimists does not cause herding	8	7.60	<.0001**
Herding does not cause # of pessimists	8	25.96	<.0001**
** Significant at 99% * Significant at 95%			

5.3.3.2 Causal Relationships Between Volume and Price

Next, the causality between model generated price and trading volume is investigated, and the results are compared with those of empirical results for BAC time series. The relationship of stock price and volume has been getting noticeable attention in economics and finance, provided that investigating the dynamics of the price in conjunction with traded volume gives a more comprehensive picture of stock market nature [80]. The lag length selection based on AIC, BIC and a series of F-tests are presented in Table 5.7

Table 5.7: Lag length selection for investigating *volume* and *price* relationship for BAC and Model.

	Dependent variable-Independent variable	AIC	BIC	F-test
BAC	volume-price	8	8	8
	price-volume	5	1	5
Model	volume-price	7	7	7
	price-volume	8	8	8

The results of the tests are presented in Table 5.8. The tests are done in both directions using first differences, and as it can be seen, there appears be a bi-directional

Table 5.8: Causality test between *volume* and *price*.

H_0	# of lags	F-value	p-value
BAC price does not cause trading volume	8	3.10	0.0018**
BAC trading volume does not cause price	5	1.10	0.3575
BAC trading volume does not cause price	1	3.04	0.0816
Model price does not cause trading volume	7	7.42	<.0001**
Model trading volume does not cause price	8	6.77	<.0001**
** Significant at 99%			
* Significant at 95%			

causal relationship between price and volume in the model generated data. This result is significant in the sense that the data on past trading volume can be used to help predict stock returns, in agreement with an old Wall Street adage that says "it takes volume to make prices move". On the other hand, the direction of BAC causality shows that only return Granger causes volume, and not the other way around, not even for 1 lag. There are many empirical studies that have found the evidence for existence of either uni- or bi-directional causality between price and volume [81, 82].

5.3.3.3 Causal Relationships Between Volatility and Other Market Phenomena

In this section, the bi-directional causal relationships between volatility and bull/bear markets, returns, volume and number of different types of traders in the market are explored. Volatility is an important variable in financial markets, and it is one of the key inputs to many investment decision. Being able to predict volatility of prices and know the causes and effects of it would be of a great value for any market risk assessment. Table 5.9 presents the results of lag length selection using different criteria and Table 5.10 gives the results on bi-directional causality tests.

As before, the tests are first run with the highest number estimated lags, and if no significant causality was detected, they are repeated the the next largest number of lags.

The F-test and p-values for causal relationship between volatility and price confirm that the results of both BAC and model are in agreement, and point to a significant

Table 5.9: Lag length selection for investigating volatility causal relationships for BAC and model.

	Dependent variable-Independent variable	AIC	BIC	F-test
BAC	volatility-price	8	5	8
	price-volatility	8	8	8
	volatility-bull	1	1	5
	bull-volatility	8	8	8
	volatility-volume	8	8	8
	volume-volatility	8	8	8
Model	volatility-price	8	8	8
	price-volatility	8	8	8
	volatility-bull	1	1	1
	bull-volatility	8	8	8
	volatility-volume	7	7	7
	volume-volatility	8	8	8
	volatility-fundamentalist	6	4	6
	fundamentalist-volatility	8	8	8
	volatility-optimist	6	4	6
	optimist-volatility	8	8	8
	volatility-pessimist	8	6	8
	pessimist-volatility	8	8	8

bi-directional causal relationship between the two variables.

There are also bi-directional causalities between volatility and volume, for both BAC and model generated data.

One of the interesting results is the causality between bullish/bearish market and volatility. The test result for BAC stock show that bull market can cause or help predict volatility in the market at 5% significance level for 8 lags, while rejecting it for the other way around. However, there are no evidence of causal relationship between these two market variables in the model generated data, and F-tests fail to reject the null hypothesis of no causality.

Finally, the results indicate that there are significant bi-directional causal relationships between number of fundamentalists, optimists and pessimists in the market and volatility.

Table 5.10: Causality test between volatility and other market variables.

H_0	# of lags	F-value	p-value
BAC volatility does not cause price	8	10.25	<.0001**
BAC price does not cause volatility	8	7.38	<.0001**
BAC volatility does not cause bull market	1	0.10	0.7513
BAC bull market does not cause volatility	8	2.42	0.0134*
BAC volatility does not cause volume	8	3.61	0.0003**
BAC volume does not cause volatility	8	5.58	<.0001**
Model volatility does not cause price	8	3.05	0.0020**
Model price does not cause volatility	8	5.55	<.0001**
Model volatility does not cause bull market	1	2.23	0.1355
Model bull market does not cause volatility	8	0.68	0.7118
Model volatility does not cause volume	7	2.27	0.0201*
Model volume does not cause volatility	8	2.56	0.0088**
Model volatility does not cause # of fundamentalists	6	39.56	<.0001**
Model # of fundamentalists does not cause volatility	8	5.70	<.0001**
Model volatility does not cause # of optimists	6	2.12	0.0479*
Model # of optimists does not cause volatility	8	3.39	0.0007**
Model volatility does not cause # of pessimists	8	9.60	<.0001**
Model # of pessimists does not cause volatility	8	2.68	0.0061**
** Significant at 99% * Significant at 95%			

CHAPTER 6: CONCLUSIONS

In this research, we have developed an agent-based model of a stock market. As was argued in Chapter 1, financial markets are characterized by their complex and dynamic nature, making it nearly impossible for them to be modeled by traditional theoretical and experimental economic models. Much of these conventional models, with their top-down approach, assume all market participants are perfectly rational and homogeneous, and market price movements incorporate all information rationally and instantaneously. However, with the assumption of perfect rationality and information, they have lacked the means to properly explain or reproduce financial market in its complete dynamic complexity [6].

This research was built on the recognition that financial markets consist of boundedly rational participants, whose interactions and adaptive behaviors give rise to unique and unexpected (emergent) macroscopic properties. This has led us to Agent-based Modeling (ABM), which is the methodology followed in this work. We model a simple, yet rich, stock market in an incremental bottom-up fashion, as an evolving adaptive system of autonomous interacting agents.

We have aimed to provide an accurate, yet generic, representation of real financial markets and traders, with the goal of exploring market phenomena under different trading, decision-making and behavioral strategies. Since we believed the source of most emergent properties of financial markets is the behavior and interaction of its traders, a simple market mechanism with no strong assumptions is maintained throughout the design and implementation of the model.

To allow for a better tractability of the parameters, we started the model with the simple setting of two types of traders, fundamentalists and chartists. As the work proceeded, we have introduced further heterogeneity in various aspects of the model, and move on to N-typed design. Heterogeneity has shown to be one of the important properties for the model to be able to generate the desired empirical statistical properties that are often observed in financial time-series. This study suggests that local interactions, rational and irrational decision-making approaches and heterogeneity are among the key elements in modeling financial markets.

We have incorporated two processes of rational and irrational decision making in the model. In first case, agents try and predict the future of the price, relying on the information they have(either fundamental value of the stock, or past prices and trend) and their intensity of belief in their trading strategy and make a decision weather to buy, sell or hold. On the other hand, the irrational process is when they observe the choice of people around them, and if more than some percentage are doing the same action, decide to follow them. We have tested and observed that better results can be obtained when this imitation component is added to the model.

After calibrating the model with a standard Genetic Algorithm, we have conducted a series of statistical tests to investigate the behaviour of our market with respect to empirical benchmark and stylized facts. The model was able to reproduce some of the most important stylized facts of financial time series, such as heavy tail in distribution of returns, volatility clustering, and absence of autocorrelation in raw returns, solely from interactions, learning and evolution between traders. As a result, we arrived to a particular configuration that is considered the platform to further explore the characteristics of our market.

Volatility and volume are among the critical variables in financial markets, and play important roles in many risk assessments and investment decisions. The investigations on these parameters from both modeled generated and empirical time-series revealed that bullish market can cause or help predict volatility in the market. There are also significant bi-directional causal relationships between volatility and volume in the market. In addition, the significant causality between price and volume reveals that the data on past trading volume can be used to help predict stock returns, and its in agreement with an old Wall Street adage that says "it takes volume to make prices move". These results could prove highly relevant in achieving a better understanding of market structure and dynamics, and serve both academics and practitioners by understanding how volatility and volume are affected in the market.

This research also contributes to the body of literature that studies herding behavior, as it uniquely investigates the relationships of different herding measures with other market parameters and properties, and advances the understanding of the phenomenon. Our results suggest that herd behavior may cause and enhance volatility in the market, but not the other way around. The results also reveal significant bi-directional causal relationship between the volume traded in the market and herding, indicating that the past values of volume in the market can help predict and cause the herding behavior, and vice versa.

Although the financial data are available plentifully and accurately in different frequencies, there still are a number of limitations we are currently facing:

- In trying to distinguish herding in the market, one needs to know if the trader is discarding his private information to follow the crowd. However, there are no data on the private information available to traders and, therefore, it is difficult to know when traders decide not to follow it.

- It is very difficult, if not impossible, to differentiate between true and spurious herding due to unknown motives behind a trade and lack of data on the private information.

- The price in the model in determined as a function of excess demand in the market. However, in the long run, economic factors other than short-term excess demand may influence the evolution of the asset price, resulting in more complex types of behavior.

REFERENCES

[1] Thomas Lux and Michele Marchesi. Volatility clustering in financial markets: a microsimulation of interacting agents. *International journal of theoretical and applied finance*, 3(04):675–702, 2000.

[2] Shu-Heng Chen, Chia-Ling Chang, and Ye-Rong Du. Agent-based economic models and econometrics. *The Knowledge Engineering Review*, 27(02):187–219, 2012.

[3] Andrew W Lo. Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis. 2005.

[4] G William Schwert. Stock volatility and the crash of?87. *Review of financial Studies*, 3(1):77–102, 1990.

[5] Cars H Hommes. Heterogeneous agent models in economics and finance. *Handbook of computational economics*, 2:1109–1186, 2006.

[6] Leigh Tesfatsion. Agent-based computational economics: Growing economies from the bottom up. *Artificial life*, 8(1):55–82, 2002.

[7] Richard G Palmer, W Brian Arthur, John H Holland, Blake LeBaron, and Paul Tayler. Artificial economic life: a simple model of a stockmarket. *Physica D: Nonlinear Phenomena*, 75(1):264–274, 1994.

[8] Mark Rubinstein. Rational markets: yes or no? the affirmative case. *Financial Analysts Journal*, 57(3):15–29, 2001.

[9] John R Nofsinger. Social mood and financial economics. *The Journal of Behavioral Finance*, 6(3):144–160, 2005.

[10] Nicholas Barberis and Richard Thaler. A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128, 2003.

[11] Egle Samanidou, Elmar Zschischang, Dietrich Stauffer, and Thomas Lux. Agent-based models of financial markets. *Reports on Progress in Physics*, 70(3):409, 2007.

[12] Blake LeBaron. Evolution and time horizons in an agent-based stock market. *Macroeconomic Dynamics*, 5(02):225–254, 2001.

[13] Robert Axelrod. Agent-based modeling as a bridge between disciplines. *Handbook of computational economics*, 2:1565–1584, 2006.

[14] Eric Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7280–7287, 2002.

[15] Blake LeBaron. Agent-based computational finance. *Handbook of computational economics*, 2:1187–1233, 2006.

[16] Helen Allen and Mark P Taylor. Charts, noise and fundamentals in the london foreign exchange market. *The Economic Journal*, 100(400):49–59, 1990.

[17] Jeffrey A Frankel and Kenneth A Froot. Chartists, fundamentalists, and trading in the foreign exchange market. *The American Economic Review*, 80(2):181–185, 1990.

[18] William J Baumol. Speculation, profitability, and stability. *The Review of Economics and Statistics*, pages 263–271, 1957.

[19] Milton Friedman. *Essays in positive economics.* University of Chicago Press, 1953.

[20] Murray C Kemp. Speculation, profitability, and price stability. *The Review of Economics and Statistics*, pages 185–189, 1963.

[21] E Christopher Zeeman. On the unstable behaviour of stock exchanges. *Journal of mathematical economics*, 1(1):39–49, 1974.

[22] Richard H Day and Weihong Huang. Bulls, bears and market sheep. *Journal of Economic Behavior & Organization*, 14(3):299–329, 1990.

[23] William A Brock and Cars H Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic dynamics and Control*, 22(8):1235–1274, 1998.

[24] H Peter Boswijk, Cars H Hommes, and Sebastiano Manzan. Behavioral heterogeneity in stock prices. *Journal of Economic dynamics and control*, 31(6):1938–1970, 2007.

[25] Frank Westerhoff et al. Exchange rate dynamics: A nonlinear survey. *Handbook of research on complexity*, pages 287–325, 2009.

[26] Paul Milgrom and Nancy Stokey. Information, trade and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.

[27] Drew Fudenberg and Jean Tirole. Game theory, 1991. *Cambridge, Massachusetts*, 393, 1991.

[28] Thomas Lux. The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions. *Journal of Economic Behavior & Organization*, 33(2):143–165, 1998.

[29] Alan Kirman. Epidemics of opinion and speculative bubbles in financial markets. *Money and financial markets*, pages 354–368, 1991.

[30] Cars H Hommes. Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7221–7228, 2002.

[31] Bart Frijns, Thorsten Lehnert, and Remco CJ Zwinkels. Behavioral heterogeneity in the option market. *Journal of Economic Dynamics and Control*, 34(11):2273–2287, 2010.

[32] Charles Goodhart. The foreign exchange market: a random walk with a dragging anchor. *Economica*, pages 437–460, 1988.

[33] R Duncan Luce. *Individual choice behavior: A theoretical analysis*. Courier Corporation, 2005.

[34] Blake LeBaron et al. A builder?s guide to agent-based financial markets. *Quantitative Finance*, 1(2):254–261, 2001.

[35] W Brian Arthur, John H Holland, Blake LeBaron, Richard G Palmer, and Paul Tayler. Asset pricing under endogenous expectations in an artificial stock market. *Available at SSRN 2252*, 1996.

[36] Frank Schweitzer, Jörg Zimmermann, and Heinz Mühlenbein. Coordination of decisions in a spatial agent model. *Physica A: Statistical Mechanics and its Applications*, 303(1):189–216, 2002.

[37] Sushil Bikhchandani and Sunil Sharma. Herd behavior in financial markets. *IMF Staff papers*, pages 279–310, 2000.

[38] Mr Sunil Sharma and Sushil Bikhchandani. *Herd behavior in financial markets: a review*. Number 0-48. International Monetary Fund, 2000.

[39] David A Hirshleifer and Siew Hong Teoh. Thought and behavior contagion in capital markets. *Handbook of financial markets: Dynamics and evolution*, 2009.

[40] Marco Cipriani and Antonio Guarino. Estimating a structural model of herd behavior in financial markets. *The American Economic Review*, 104(1):224–251, 2014.

[41] Abhijit V Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, pages 797–817, 1992.

[42] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, pages 992–1026, 1992.

[43] André Orléan. Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion. *Journal of Economic Behavior & Organization*, 28(2):257–274, 1995.

[44] Rama Cont and Jean-Philipe Bouchaud. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic dynamics*, 4(02):170–196, 2000.

[45] Didier Sornette. *Why stock markets crash: critical events in complex financial systems*. Princeton University Press, 2009.

[46] Eugene N White and Peter Rappoport. The new york stock market in the 1920s and 1930s: Did stock prices move together too much? Technical report, National Bureau of Economic Research, 1994.

[47] Richard Topol. Bubbles and volatility of stock prices: effect of mimetic contagion. *The Economic Journal*, 101(407):786–800, 1991.

[48] Victor M Eguiluz and Martin G Zimmermann. Transmission of information and herd behavior: an application to financial markets. *Physical review letters*, 85(26):5659, 2000.

[49] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. 2001.

[50] Benoit B Mandelbrot. The variation of certain speculative prices. In *Fractals and scaling in finance*, pages 371–418. Springer, 1997.

[51] Bruce M Hill et al. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 3(5):1163–1174, 1975.

[52] Narayana R Kocherlakota. The equity premium: It's still a puzzle. *Journal of Economic literature*, 34(1):42–71, 1996.

[53] J Barkley Rosser Jr. Speculations on nonlinear speculative bubbles. *Nonlinear Dynamics, Psychology, and Life Sciences*, 1(4):275–300, 1997.

[54] Jun Muranaga and Makoto Ohsawa. Measurement of liquidity risk in the context of market risk calculation. *a BIS volume entitled The Measurement of Aggregate Market Risk*, 1997.

[55] Xavier Gabaix, Parameswaran Gopikrishnan, Vasiliki Plerou, and H Eugene Stanley. A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937):267–270, 2003.

[56] Stephen J Taylor. Asset price dynamics. *Volatility, and Prediction*, 2005.

[57] Ruey S Tsay. *Analysis of financial time series*, volume 543. John Wiley & Sons, 2005.

[58] Rama Cont. Volatility clustering in financial markets: empirical facts and agent-based models. In *Long memory in economics*, pages 289–309. Springer, 2007.

[59] Thomas Lux. Applications of statistical physics in finance and economics. Technical report, Kieler Arbeitspapiere, 2008.

[60] Maria Pacurar. Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. *Journal of Economic Surveys*, 22(4):711–751, 2008.

[61] Jeffrey R Russell, RF Engle, Y Ait-Sahalia, and LP Hansen. Analysis of high-frequency data. *Handbook of financial econometrics*, 1:383–426, 2009.

[62] Manfred Gilli and Peter Winker. A global optimization heuristic for estimating agent based models. *Computational Statistics & Data Analysis*, 42(3):299–312, 2003.

[63] Frank H Westerhoff. The use of agent-based financial market models to test the effectiveness of regulatory policies. *Jahrbücher für Nationalökonomie und Statistik*, 228(2-3):195–227, 2008.

[64] Vilfredo Pareto. *Cours d'économie politique*, volume 1. Librairie Droz, 1964.

[65] Adrian Drăgulescu and Victor M Yakovenko. Exponential and power-law probability distributions of wealth and income in the united kingdom and the united states. *Physica A: Statistical Mechanics and its Applications*, 299(1):213–221, 2001.

[66] Lukas Menkhoff. The use of technical analysis by fund managers: International evidence. *Journal of Banking & Finance*, 34(11):2573–2586, 2010.

[67] Andrea Devenow and Ivo Welch. Rational herding in financial economics. *European Economic Review*, 40(3):603–615, 1996.

[68] Johannes Voit. *The statistical mechanics of financial markets*. Springer Science & Business Media, 2013.

[69] John H Holland. Adaptation in natural and artificial systems. an introductory analysis with application to biology, control, and artificial intelligence. *Ann Arbor, MI: University of Michigan Press*, pages 439–444, 1975.

[70] Alison J Heppenstall, Andrew J Evans, and Mark H Birkin. Genetic algorithm optimisation of an agent-based model for simulating a retail market. *Environment and Planning B: Planning and Design*, 34(6):1051–1070, 2007.

[71] Benoît Calvez and Guillaume Hutzler. Automatic tuning of agent-based models using genetic algorithms. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation*, pages 41–57. Springer, 2005.

[72] Gary William Flake. *The computational beauty of nature: Computer explorations of fractals, chaos, complex systems, and adaptation*. MIT press, 1998.

[73] Yahoo finance.

[74] David A Dickey and Wayne A Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, pages 1057–1072, 1981.

[75] Adrian R Pagan and Kirill A Sossounov. A simple framework for analysing bull and bear markets. *Journal of Applied Econometrics*, 18(1):23–46, 2003.

[76] Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969.

[77] James H Stock and Mark W Watson. Interpreting the evidence on money-income causality. *Journal of Econometrics*, 40(1):161–181, 1989.

[78] Hirotugu Akaike. Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics*, 21(1):243–247, 1969.

[79] Hirotugu Akaike. A bayesian analysis of the minimum aic procedure. In *Selected Papers of Hirotugu Akaike*, pages 275–280. Springer, 1998.

[80] A Ronald Gallant, Peter E Rossi, and George Tauchen. Stock prices and volume. *The Review of Financial Studies*, 5(2):199–242, 1992.

[81] Craig Hiemstra and Jonathan D Jones. Testing for linear and nonlinear granger causality in the stock price-volume relation. *The Journal of Finance*, 49(5):1639–1664, 1994.

[82] Kemal Saatcioglu and Laura T Starks. The stock price–volume relationship in emerging stock markets: the case of latin america. *International Journal of forecasting*, 14(2):215–225, 1998.

APPENDIX A: AUGMENTED DICKEY-FULLER TEST RESULTS

Table A.1: Augmented Dickey-Fuller Unit Root Tests for Model Variable-price

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-0.9263	0.486	-1.39	0.1523		
	1	-0.9259	0.486	-1.53	0.1195		
Single Mean	0	-4.2753	0.5096	-1.9	0.3302	2.19	0.5084
	1	-4.0173	0.5378	-1.96	0.305	2.41	0.4522
Trend	0	-4.4101	0.8614	-1.31	0.885	1.81	0.8142
	1	-3.5759	0.9102	-1.16	0.9167	1.94	0.7894

Table A.2: Augmented Dickey-Fuller Unit Root Tests for Model Variable-absreturn

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-839.39	0.0001	-22.13	<.0001		
	1	-338.836	0.0001	-13.08	<.0001		
Single Mean	0	-1710.96	0.0001	-34.63	<.0001	599.77	0.001
	1	-902.455	0.0001	-21.29	<.0001	226.68	0.001
Trend	0	-1794.94	0.0001	-35.81	<.0001	641.03	0.001
	1	-974.801	0.0001	-22.11	<.0001	244.36	0.001

Table A.3: Augmented Dickey-Fuller Unit Root Tests for Model Variable-herd

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-212.283	0.0001	-10.64	<.0001		
	1	-91.9827	<.0001	-6.94	<.0001		
Single Mean	0	-224.99	0.0001	-10.94	<.0001	59.9	0.001
	1	-97.4081	0.0019	-7.11	<.0001	25.31	0.001
Trend	0	-252.328	0.0001	-11.57	<.0001	66.94	0.001
	1	-109.166	0.0001	-7.47	<.0001	27.94	0.001

Table A.4: Augmented Dickey-Fuller Unit Root Tests for Model Variable-buyers

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1514.81	0.0001	-31.98	<.0001		
	1	-663.168	0.0001	-18.39	<.0001		
Single Mean	0	-1582.42	0.0001	-32.92	<.0001	541.91	0.001
	1	-707.142	0.0001	-18.97	<.0001	179.92	0.001
Trend	0	-1705.34	0.0001	-34.63	<.0001	599.59	0.001
	1	-791.792	0.0001	-20.03	<.0001	200.61	0.001

Table A.5: Augmented Dickey-Fuller Unit Root Tests for Model Variable-sellers

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-2206.89	0.0001	-41.77	<.0001		
	1	-1211.3	0.0001	-24.6	<.0001		
Single Mean	0	-2255.08	0.0001	-42.49	<.0001	902.65	0.001
	1	-1262.61	0.0001	-25.11	<.0001	315.33	0.001
Trend	0	-2353.94	0.0001	-43.99	<.0001	967.53	0.001
	1	-1374.78	0.0001	-26.2	<.0001	343.3	0.001

Table A.6: Augmented Dickey-Fuller Unit Root Tests for Model Variable-fundamentalist

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-6.2327	0.0858	-1.8	0.0691		
	1	-2.2775	0.3001	-1.12	0.2385		
Single Mean	0	-418.756	0.0001	-15.02	<.0001	112.83	0.001
	1	-158.611	0.0001	-8.94	<.0001	39.98	0.001
Trend	0	-443.764	0.0001	-15.48	<.0001	119.87	0.001
	1	-168.641	0.0001	-9.2	<.0001	42.31	0.001

Table A.7: Augmented Dickey-Fuller Unit Root Tests for Model Variable-optimist

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-8.6962	0.0411	-2.07	0.0373		
	1	-2.8785	0.2441	-1.16	0.2236		
Single Mean	0	-543.725	0.0001	-17.29	<.0001	149.53	0.001
	1	-213.247	0.0001	-10.33	<.0001	53.4	0.001
Trend	0	-574.12	0.0001	-17.81	<.0001	158.63	0.001
	1	-226.697	0.0001	-10.65	<.0001	56.67	0.001

Table A.8: Augmented Dickey-Fuller Unit Root Tests for Model Variable-pessimist

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-9.4383	0.0331	-2.16	0.0296		
	1	-3.5002	0.1989	-1.3	0.1791		
Single Mean	0	-1140.9	0.0001	-26.55	<.0001	352.51	0.001
	1	-585.33	0.0001	-17.13	<.0001	146.75	0.001
Trend	0	-1172.3	0.0001	-26.99	<.0001	364.21	0.001
	1	-606.836	0.0001	-17.43	<.0001	151.86	0.001

Table A.9: Augmented Dickey-Fuller Unit Root Tests for Model Variable-bull

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1508.93	0.0001	-31.76	<.0001		
	1	-742.124	0.0001	-19.28	<.0001		
Single Mean	0	-3038.15	0.0001	-55.45	<.0001	1537.6	0.001
	1	-3033.39	0.0001	-38.92	<.0001	757.57	0.001
Trend	0	-3038.39	0.0001	-55.45	<.0001	1537.34	0.001
	1	-3034.23	0.0001	-38.92	<.0001	757.53	0.001

Table A.10: Augmented Dickey-Fuller Unit Root Tests for Model Variable-volume

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-205.269	0.0001	-10.46	<.0001		
	1	-89.8345	<.0001	-6.86	<.0001		
Single Mean	0	-217.844	0.0001	-10.76	<.0001	57.89	0.001
	1	-95.2415	0.0019	-7.03	<.0001	24.76	0.001
Trend	0	-243.915	0.0001	-11.36	<.0001	64.58	0.001
	1	-106.514	0.0001	-7.38	<.0001	27.27	0.001

Table A.11: Augmented Dickey-Fuller Unit Root Tests for Model Variable-dabsreturn

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4520.51	0.0001	-95.77	<.0001		
	1	-9001.84	0.0001	-67.11	<.0001		
Single Mean	0	-4520.51	0.0001	-95.76	<.0001	4584.52	0.001
	1	-9001.9	0.0001	-67.09	<.0001	2250.85	0.001
Trend	0	-4520.52	0.0001	-95.74	<.0001	4583.02	0.001
	1	-9002.19	0.0001	-67.08	<.0001	2250.18	0.001

Table A.12: Augmented Dickey-Fuller Unit Root Tests for Model Variable-return

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-3263.32	0.0001	-59.86	<.0001		
	1	-3638.68	0.0001	-42.74	<.0001		
Single Mean	0	-3264.08	0.0001	-59.87	<.0001	1792.15	0.001
	1	-3641.84	0.0001	-42.75	<.0001	913.75	0.001
Trend	0	-3266.06	0.0001	-59.9	<.0001	1794.04	0.001
	1	-3650.25	0.0001	-42.79	<.0001	915.69	0.001

Table A.13: Augmented Dickey-Fuller Unit Root Tests for Model Variable-dvolume

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4255.54	0.0001	-85.65	<.0001		
	1	-6011.19	0.0001	-54.99	<.0001		
Single Mean	0	-4255.57	0.0001	-85.64	<.0001	3666.95	0.001
	1	-6011.58	0.0001	-54.99	<.0001	1511.73	0.001
Trend	0	-4255.64	0.0001	-85.63	<.0001	3665.89	0.001
	1	-6012.45	0.0001	-54.98	<.0001	1511.51	0.001

Table A.14: Augmented Dickey-Fuller Unit Root Tests for Model Variable-dherd

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4267.77	0.0001	-86.1	<.0001		
	1	-5781.77	0.0001	-53.94	<.0001		
Single Mean	0	-4267.8	0.0001	-86.08	<.0001	3705.31	0.001
	1	-5782.15	0.0001	-53.93	<.0001	1454.41	0.001
Trend	0	-4267.87	0.0001	-86.07	<.0001	3704.24	0.001
	1	-5782.98	0.0001	-53.93	<.0001	1454.19	0.001

Table A.15: Augmented Dickey-Fuller Unit Root Tests for Model Variable-dbuyers

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4644.43	0.0001	-101.46	0.0001		
	1	-9550.22	0.0001	-69.44	<.0001		
Single Mean	0	-4644.44	0.0001	-101.44	0.0001	5145.23	0.001
	1	-9550.33	0.0001	-69.43	<.0001	2410.41	0.001
Trend	0	-4644.46	0.0001	-101.43	0.0001	5143.61	0.001
	1	-9550.55	0.0001	-69.42	<.0001	2409.62	0.001

Table A.16: Augmented Dickey-Fuller Unit Root Tests for Model Variable-dsellers

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4654	0.0001	-101.89	0.0001		
	1	-10176.3	0.0001	-71.88	<.0001		
Single Mean	0	-4654	0.0001	-101.87	0.0001	5188.99	0.001
	1	-10176.3	0.0001	-71.87	<.0001	2582.38	0.001
Trend	0	-4654	0.0001	-101.86	0.0001	5187.26	0.001
	1	-10176.3	0.0001	-71.85	<.0001	2581.56	0.001

Table A.17: Augmented Dickey-Fuller Unit Root Tests for Model Variable-dfundamentalist

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4471.6	0.0001	-93.71	<.0001		
	1	-7085.83	0.0001	-59.5	<.0001		
Single Mean	0	-4471.6	0.0001	-93.7	<.0001	4389.58	0.001
	1	-7085.91	0.0001	-59.49	<.0001	1769.33	0.001
Trend	0	-4471.62	0.0001	-93.68	<.0001	4388.16	0.001
	1	-7086.04	0.0001	-59.48	<.0001	1768.78	0.001

Table A.18: Augmented Dickey-Fuller Unit Root Tests for Model Variable-doptimist

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4463.07	0.0001	-93.36	<.0001		
	1	-7871.03	0.0001	-62.7	<.0001		
Single Mean	0	-4463.07	0.0001	-93.35	<.0001	4356.96	0.001
	1	-7871.07	0.0001	-62.69	<.0001	1965.11	0.001
Trend	0	-4463.07	0.0001	-93.33	<.0001	4355.51	0.001
	1	-7871.11	0.0001	-62.68	<.0001	1964.46	0.001

Table A.19: Augmented Dickey-Fuller Unit Root Tests for Model Variable-dpessimist

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4353.35	0.0001	-89.08	<.0001		
	1	-7698.93	0.0001	-61.99	<.0001		
Single Mean	0	-4353.35	0.0001	-89.07	<.0001	3966.45	0.001
	1	-7698.97	0.0001	-61.98	<.0001	1920.73	0.001
Trend	0	-4353.36	0.0001	-89.05	<.0001	3965.15	0.001
	1	-7699.03	0.0001	-61.97	<.0001	1920.11	0.001

Table A.20: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-price

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1.741	0.3628	-1.34	0.1673		
	1	-1.7203	0.3654	-1.35	0.1638		
Single Mean	0	-4.2468	0.5126	-1.77	0.3937	1.68	0.6414
	1	-4.1559	0.5225	-1.77	0.3944	1.68	0.6402
Trend	0	-2.6785	0.9497	-0.96	0.948	2.16	0.744
	1	-2.528	0.955	-0.92	0.952	2.23	0.7303

Table A.21: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-absreturn

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1109.42	0.0001	-26.08	<.0001		
	1	-536.848	0.0001	-16.38	<.0001		
Single Mean	0	-1603.31	0.0001	-33.06	<.0001	546.63	0.001
	1	-912.746	0.0001	-21.35	<.0001	227.95	0.001
Trend	0	-1650.46	0.0001	-33.73	<.0001	568.74	0.001
	1	-955.981	0.0001	-21.85	<.0001	238.82	0.001

Table A.22: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-bull

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1539.11	0.0001	-32.17	<.0001		
	1	-721.274	0.0001	-18.99	<.0001		
Single Mean	0	-3149.62	0.0001	-57.56	<.0001	1656.37	0.001
	1	-3139.22	0.0001	-39.58	<.0001	783.43	0.001
Trend	0	-3151.99	0.0001	-57.59	<.0001	1658.33	0.001
	1	-3147.02	0.0001	-39.62	<.0001	785.06	0.001

Table A.23: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-Volume

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-209.63	0.0001	-10.42	<.0001		
	1	-132.846	0.0001	-8.14	<.0001		
Single Mean	0	-455.386	0.0001	-15.7	<.0001	123.22	0.001
	1	-310.366	0.0001	-12.46	<.0001	77.62	0.001
Trend	0	-456.195	0.0001	-15.71	<.0001	123.45	0.001
	1	-311.038	0.0001	-12.47	<.0001	77.8	0.001

Table A.24: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-dprice

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-3064.11	0.0001	-55.96	<.0001		
	1	-3202.65	0.0001	-40	<.0001		
Single Mean	0	-3064.32	0.0001	-55.95	<.0001	1565.25	0.001
	1	-3203.33	0.0001	-40	<.0001	800.01	0.001
Trend	0	-3067.84	0.0001	-56.01	<.0001	1568.41	0.001
	1	-3214.93	0.0001	-40.07	<.0001	802.63	0.001

Table A.25: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-return

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-3028.52	0.0001	-55.29	<.0001		
	1	-2787.72	0.0001	-37.32	<.0001		
Single Mean	0	-3028.56	0.0001	-55.28	<.0001	1527.82	
	1	-2787.84	0.0001	-37.32	<.0001	696.26	0.001
Trend	0	-3030.58	0.0001	-55.31	<.0001	1529.35	0.001
	1	-2793.36	0.0001	-37.35	<.0001	697.4	0.001

Table A.26: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-dabsreturn

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4403.36	0.0001	-90.99	<.0001		
	1	-7974.77	0.0001	-63.12	<.0001		
Single Mean	0	-4403.36	0.0001	-90.97	<.0001	4137.86	0.001
	1	-7974.77	0.0001	-63.11	<.0001	1991.69	0.001
Trend	0	-4403.36	0.0001	-90.96	<.0001	4136.48	0.001
	1	-7974.79	0.0001	-63.1	<.0001	1991.04	0.001

Table A.27: Augmented Dickey-Fuller Unit Root Tests for BAC Variable-dVolume

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-3751.94	0.0001	-70.77	<.0001		
	1	-5262.22	0.0001	-51.28	<.0001		
Single Mean	0	-3751.94	0.0001	-70.76	<.0001	2503.17	0.001
	1	-5262.22	0.0001	-51.27	<.0001	1314.23	0.001
Trend	0	-3751.95	0.0001	-70.74	<.0001	2502.35	0.001
	1	-5262.29	0.0001	-51.26	<.0001	1313.81	0.001